

Studies in Computational Intelligence 998

Jitendra Kumar Verma
Sudip Paul *Editors*

Advances in Augmented Reality and Virtual Reality

 Springer

Studies in Computational Intelligence

Volume 998

Series Editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series “Studies in Computational Intelligence” (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

Indexed by SCOPUS, DBLP, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at <https://link.springer.com/bookseries/7092>

Jitendra Kumar Verma · Sudip Paul
Editors

Advances in Augmented Reality and Virtual Reality

 Springer

Editors

Jitendra Kumar Verma
IT Discipline
Indian Institute of Foreign Trade (under
Ministry of Commerce & Industry, Govt.
of India)
New Delhi, India

Sudip Paul
Department of Biomedical Engineering
North-Eastern Hill University
Shillong, Meghalaya, India

ISSN 1860-949X

ISSN 1860-9503 (electronic)

Studies in Computational Intelligence

ISBN 978-981-16-7219-4

ISBN 978-981-16-7220-0 (eBook)

<https://doi.org/10.1007/978-981-16-7220-0>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Editorial Advisory Board

Prof. D. P. Sharma, AMUIT MOSHEFDRE under UNDP and Research Adviser
VUB Belgium

Prof. Celia Shahnaz, Department of EEE, BUET, Bangladesh Chair, IEEE
Bangladesh Section

Prof. Jesús P. Zamora, UNED, Department of Logic and Philosophy of Science,
Spain

Prof. Adolfo Crespo, University of Seville, Department of Industrial Management,
School of Engineering, Spain

Prof. Vicente González-Prida Díaz, University of Seville, Department of Industrial
Management, School of Engineering, Spain

Prof. Shalini Bhaskar Bajaj, Head, Department of Computer Science, Amity
University, Gurugram, India

Dr. Vira Shendryk, Deputy Head of Computer Science Department, Head of the
Information Technology Section, Sumy State University, Ukraine

Dr. Deepak Saxena, Trinity College Dublin, Ireland

Technical Programme Committee

Abhishek Kaushik, ADAPT, Dublin City University, Ireland
María Carmen Carnero, University of Castilla-La Mancha, Spain
François Pérès, Université de Toulouse, Ecole Nationale d'Ingénieurs de Tarbes,
Laboratoire Génie de Production, France
Carmen Martin, Toulouse University, France
Fredy Kristjanpoller, University Federico Santa María, Valparaíso, Chile
Pablo Viveros, University Federico Santa María, Valparaíso, Chile
Vijay Verma, NIT Kurukshetra, Haryana, India
Sandeep Pawar Jogi, Amity University, Gurugram, India
Sunil Kumar Bharti, Noida Institute of Engineering and Technology, Greater Noida,
India
Dileep Kumar Yadav, Galgotias University, Greater Noida, India
Yogendra Meena, Kalindi College (University of Delhi), New Delhi, India
Utpal Shrivastava, G L Bajaj Group of Institution, Mathura, India
Krishan Kumar, Amity University Haryana, Gurugram, India
Akshat Agrawal, Amity University Haryana, Gurugram, India
Ankit Garg, Amity University Haryana, Gurugram, India
Vikas Thada, DPGITM Engineering College, Gurugram, India
Sunil Sikka, Amity University Haryana, Gurugram, India
Vivek Jaglan, Graphic Era Hill University, Dehradun, India
Rajesh Tyagi, Amity University Haryana, Gurugram, India
Tarun Agrawal, NIT Hamirpur, Himachal Pradesh, India
Amit Kumar, Galgotias University, Greater Noida, India
Ashok Kumar Yadav, Rajkiya Engineering College, Deogaon, Azamgarh, India
Suman Devi, Galgotias University, Greater Noida, India

Preface

This book recapitulates and outlines recent advances in computer-assisted environments, particularly those in augmented and virtual reality. While these technologies have been available for a quite some time now, they were mainly confined to research laboratories and science fiction. This book discusses recent advances by focussing upon the application of these technologies in various domains. The book is divided into four parts.

The first part focusses on understanding virtual environments. Chapter “[Recreating Reality: Classification of Computer-Assisted Environments](#)” discusses various classifications of computer-assisted environments, namely virtual reality (VR), augmented reality (AR), mixed reality (MR) and extended reality (XR). Chapter “[Impact of Different Field-of-Views on Visuospatial Memory and Cognitive Workload in a Complex Virtual Environment](#)” investigates influence of different field of views (FoV) on cognitive performance of the participants in a complex virtual environment. Chapter “[Free-Roam Virtual Reality: A New Avenue for Gaming](#)” discusses a business application of perceptions in such virtual environments, by discussing free-roam virtual games.

The second part discusses applications of AR and VR in engineering disciplines. Chapter “[Computer Simulation of Physical Processes in an Electric Circuit with Nonlinear Inductance](#)” outlines computer simulation of physical processes in an electric circuit with nonlinear inductance, particularly in the context of laboratory-based teaching and learning. Chapter “[A Comparative Motion Study of Mated Gears on AutoCAD and SOLIDWORKS](#)” moves to the discipline of mechanical engineering and presents a comparative motion study of mated gears on AutoCAD and SOLIDWORKS. Chapter “[An Overview of Different Approaches for Ternary Reversible Logic Circuits Synthesis Using Ternary Reversible Gates with Special Reference to Virtual Reality](#)” addresses different ternary reversible logic synthesis approaches in logic circuits and highlights its relevance for quantum computing.

The third part turns its gaze towards applications of AR and VR technologies in agriculture and medical fields. Chapter “[Tomato’s Disease Identification Using Machine Learning Techniques with the Potential of AR and VR Technologies for Inclusiveness](#)” applies machine learning techniques for tomato’s disease identification, while discussing the potential of AR and VR technologies for inclusiveness. Chapter “[Augmented Reality and Virtual Reality Creating Ripple in Medical and Pharmaceutical World](#)” highlights a number of AR and VR applications in the medical field, such as in phantom limb pain management or in the treatment of phobias. Chapter “[Autonomous Drones for Medical Assistance Using Reinforcement Learning](#)” focuses on the navigation of autonomous drones for medical assistance by following the bike lanes present on the roads.

The fourth part links AR and VR with another set of technological revolutions by integrating virtual/augmented environments with machine and reinforcement learning approaches. Chapter “[Autonomous Navigation of Drones Using Reinforcement Learning](#)” employs reinforcement learning for autonomous navigation of drones. Chapter “[Human Face Detection and Recognition Protection System Based on Machine Learning Algorithms with Proposed AR Technology](#)” proposes a human face detection and recognition protection system based on machine learning algorithms. Finally, Chapter “[A Reinforcement Learning Approach for Shortest Path Navigation in Automated Guided Vehicles for Medical Assistance](#)” discusses reinforcement learning approach for obstacle detection and shortest path navigation for automated guided vehicles.

New Delhi, India
Shillong, India

Jitendra Kumar Verma
Sudip Paul

Contents

Understanding Virtual Environments

Recreating Reality: Classification of Computer-Assisted Environments	3
Deepak Saxena and Jitendra Kumar Verma	

Impact of Different Field-of-Views on Visuospatial Memory and Cognitive Workload in a Complex Virtual Environment	11
Akash K. Rao, Ronnie V. Daniel, Vishal Pandey, Sushil Chandra, and Varun Dutt	

Free-Roam Virtual Reality: A New Avenue for Gaming	29
Francesco Rega and Deepak Saxena	

Application in Engineering

Computer Simulation of Physical Processes in an Electric Circuit with Nonlinear Inductance	37
Vita Ogar, Andrii Perekrest, Oleksii Kravets, and Oleksandr Bilyk	

A Comparative Motion Study of Mated Gears on AutoCAD and SOLIDWORKS	57
Parth Patpatiya	

An Overview of Different Approaches for Ternary Reversible Logic Circuits Synthesis Using Ternary Reversible Gates with Special Reference to Virtual Reality	73
P. Mercy Nesa Rani and Phrangboklang Lyngton Thangkhiew	

Applications in Agriculture and Medical

Tomato’s Disease Identification Using Machine Learning Techniques with the Potential of AR and VR Technologies for Inclusiveness 93

Md. Sadik Tasrif Anubhove, S. M. Masum Ahmed, Mohammad Zeyad, Md. Abul Ala Walid, Nawreen Ashrafi, and Ahmed Mortuza Saleque

Augmented Reality and Virtual Reality Creating Ripple in Medical and Pharmaceutical World 113

Siddharth Singh, Pankaj Kumar, Hema Chaudharya, Riya Khurana, and Ankit Shokeen

Autonomous Drones for Medical Assistance Using Reinforcement Learning 133

Billy Jacob, Abhishek Kaushik, Pankaj Velavan, and Mahak Sharma

Integration with Artificial Intelligence

Autonomous Navigation of Drones Using Reinforcement Learning 159

Billy Jacob, Abhishek Kaushik, and Pankaj Velavan

Human Face Detection and Recognition Protection System Based on Machine Learning Algorithms with Proposed AR Technology 177

Md. Imran Hossain Chowdhury, Newaz Mahmud Sakib, S. M. Masum Ahmed, Mohammad Zeyad, Md. Abul Ala Walid, and Golam Kawcher

A Reinforcement Learning Approach for Shortest Path Navigation in Automated Guided Vehicles for Medical Assistance 193

Pankaj Velavan, Abhishek Kaushik, Billy Jacob, and Mahak Sharma

Editors and Contributors

About the Editors

Dr. Jitendra Kumar Verma is Assistant Professor (Senior Grade) of Indian Institute of Foreign Trade (IIFT), New Delhi in IT Discipline. Prior to joining IIFT, he served at Amity University Haryana, Galgotias University and Deen Dayal Upadhyaya College (University of Delhi). He received his degree of M.Tech and Ph.D. in computer science from Jawaharlal Nehru University, New Delhi in 2013 and 2017, respectively. He obtained his degree of B.Tech in Computer Science & Engineering from Kamla Nehru Institute of Technology (KNIT), Sultanpur in 2008. He is awardee of prestigious DAAD “A new Passage to India” Fellowship (2015–16) funded by Federal Ministry of Education and Research—BMBF, Germany and German Academic Exchange Service. He worked at JULIUS-MAXIMILIAN UNIVERSITY OF WÜRZBURG, GERMANY as a Visiting Research Scholar.

Over his short career, he has published numerous papers with reputed peer-reviewed International Journals, books, conference proceedings papers, book chapters with reputed publishers such as IEEE, Springer Nature, Elsevier, and CRC Press (Taylor & Francis Group). He has organised a numerous International conferences, seminars, FDPs, MDPs and workshops. He has delivered several invited talks. He is associate editor and guest editor to many International Journals. He is a member of different Societies and professional bodies including ACM, IEEE Industrial Applications Society, and IEEE. He serves as a reviewer for various International Journals, conferences, and workshops.

Dr. Sudip Paul is currently Assistant Professor in Department of Biomedical Engineering, School of Technology, North-Eastern Hill University (NEHU), Shillong, India since 2012. He completed his postdoctoral research at School of Computer Science and Software Engineering, The University of Western Australia, Perth. He was one of the most precious fellowship awardee (Biotechnology Overseas Associateship for the scientists working in North-Eastern States of India: 2017–18 supported by Department of Department of Biotechnology, Government of India). He received

his Ph.D. degree from Indian Institute of Technology (Banaras Hindu University), Varanasi, with specialization in electrophysiology and brain signal analysis. He has many credentials in his credit out of which his first prize in Sushruta Innovation Award 2011 sponsored by Department of Science and Technology, Government of India, and also he also organized many workshops and conferences out of which most significant are the 29th annual meeting of the Society for Neurochemistry, India, and IRBO/APRC Associate School 2017. He published more than 90 international journal and conference papers and also filled four patents. Recently, he completed three book projects and two are ongoing as Editor. He is Member of different societies and professional bodies, including APSN, ISN, IBRO, SNCI, SfN, IEEE and many more since 2009, 2010 and 2011 onwards. He received many awards specially World Federation of Neurology (WFN) travelling fellowship, Young Investigator Award, IBRO Travel Awardee and ISN Travel Awardee. He also served as Editorial Board Member for some international journals. He has presented his research accomplishments in USA, Greece, France, South Africa and Australia.

Contributors

Md. Abul Ala Walid Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU), Gopalganj, Bangladesh

Nawreen Ashrafi American International University-Bangladesh (AIUB), Dhaka, Bangladesh

Oleksandr Bilyk Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine

Sushil Chandra Department of Biomedical Engineering, Institute of Nuclear Medicine and Allied Sciences, Defence Research and Development Organization, Delhi, India

Hema Chaudharya Faculty of Pharmaceutical Sciences, PDM University, Bahadurgarh, Haryana, India

Md. Imran Hossain Chowdhury American International University-Bangladesh (AIUB), Dhaka, Bangladesh

Ronnie V. Daniel Department of Biomedical Engineering, Institute of Nuclear Medicine and Allied Sciences, Defence Research and Development Organization, Delhi, India

Varun Dutt Applied Cognitive Science Laboratory, Indian Institute of Technology, Mandi, Himachal Pradesh, India

Billy Jacob Dublin Business School, Dublin, Ireland

Abhishek Kaushik Adapt Centre, Dublin City University, Dublin, Ireland

Golam Kawcher University of Siegen, Siegen, Germany

Riya Khurana Faculty of Pharmaceutical Sciences, PDM University, Bahadurgarh, Haryana, India

Oleksii Kravets Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine

Pankaj Kumar Department of Applied Sciences, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

S. M. Masum Ahmed Heriot-Watt University, Edinburgh, Scotland, UK; Universidad del País Vasco/Euskal Herriko Unibertsitatea, Bilbao, Spain; Energy and Technology Research Division, Advanced Bioinformatics, Computational Biology and Data Science Laboratory, Bangladesh, Chattogram, Bangladesh

P. Mercy Nesa Rani College of Post-Graduate Studies in Agricultural Sciences, Umiam, Meghalaya, India

Vita Ogar Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine

Vishal Pandey Department of Biomedical Engineering, Institute of Nuclear Medicine and Allied Sciences, Defence Research and Development Organization, Delhi, India

Parth Patpatiya School of Automation, Rajasthan, India

Andrii Perekrest Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine

Akash K. Rao Applied Cognitive Science Laboratory, Indian Institute of Technology, Mandi, Himachal Pradesh, India

Francesco Rega Trinity College Dublin, Dublin, Ireland

Md. Sadik Tasrif Anubhove Energy and Technology Research Division, Advanced Bioinformatics, Computational Biology and Data Science Laboratory, Bangladesh, Chattogram, Bangladesh; American International University-Bangladesh (AIUB), Dhaka, Bangladesh

Newaz Mahmud Sakib American International University-Bangladesh (AIUB), Dhaka, Bangladesh

Ahmed Mortuza Saleque The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Deepak Saxena Birla Institute of Technology and Science Pilani, Pilani, India

Mahak Sharma The Maharaja Sayajirao University of Baroda, Baroda, India

Ankit Shokeen Faculty of Pharmaceutical Sciences, PDM University, Bahadurgarh, Haryana, India

Siddharth Singh School of Health Sciences, University of Petroleum and Energy Studies, Dehradun, India

Phrangboklang Lyngton Thangkhiew School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Chennai, India

Pankaj Velavan Dublin Business School, Dublin, Ireland

Jitendra Kumar Verma IT Discipline, Indian Institute of Foreign Trade (under Ministry of Commerce & Industry, Govt. of India), New Delhi, India

Md. Abul Ala Walid Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU), Gopalganj, Bangladesh

Mohammad Zeyad Heriot-Watt University, Edinburgh, Scotland, UK;
Universidad del País Vasco/Euskal Herriko Unibertsitatea, Bilbao, Spain;
Energy and Technology Research Division, Advanced Bioinformatics, Computational Biology and Data Science Laboratory, Bangladesh, Chattogram, Bangladesh

Understanding Virtual Environments

Recreating Reality: Classification of Computer-Assisted Environments



Deepak Saxena and Jitendra Kumar Verma

Abstract In everyday life, we understand actual physical reality through our sense perception. In the recent years, technology has added many new elements to the perception of reality around us. In this regard, many terms are in use—virtual reality (VR), augmented reality (AR), mixed reality (MR), and more recently, extended reality (XR). In general, these are computer-assisted environments that attempt to recreate reality for the users using cameras, sensors, display and projection devices. However, there are some subtle different among these terms depending upon their relationship with actual physical reality. This chapter makes sense of these terms, with suitable examples in order to present a clear picture for academics and practitioners.

Keywords Augmented reality · Extended reality · Mixed reality · Virtual reality · AR · MR · VR · XR

1 Introduction

Reality may be understood as how we perceive our environment. In everyday life, we understand reality through our sense perception, mediated by our mental schema. For instance, 25 °C may be considered cold, hot, or pleasant depending upon whether you are measuring it in Mali, Russia, or India. The measurement remains the same but our mental schema defines our interpretations. In the recent years, computers have also become a factor in our reality perception. Computer-assisted environments have become more common across many settings such as gaming or healthcare. In this regard, many terms are in use—virtual reality, augmented reality, mixed reality, and more recently, extended reality. In general, these are computer-assisted environments

D. Saxena (✉)
Birla Institute of Technology and Science, Pilani, India
e-mail: saxenad@tcd.ie

J. K. Verma
IT Discipline, Indian Institute of Foreign Trade (under Ministry of Commerce & Industry, Govt. of India), New Delhi, India
e-mail: jk-verma@ieee.org

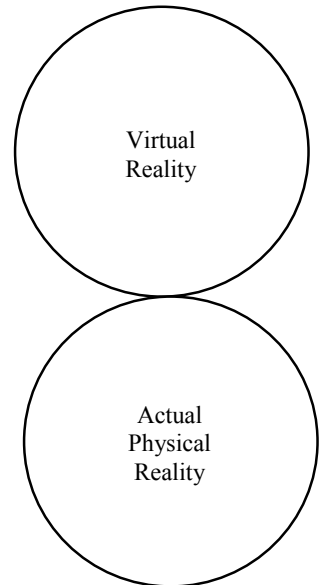
that attempt to recreate reality for the users using devices such as cameras, sensors, display, and projection devices. However, there are some subtle differences among these terms. This chapter makes sense of these terms, with suitable examples in order to present a clear picture for academic and practitioners. We do this by explaining their relationship with *actual physical reality* (which we define as external reality that is perceptible to users with help of their basic senses, e.g., via sight, audio, or touch) and by outlining the use of each type of computer-assisted environment in various domains such as medicine and tourism.

2 Virtual Reality

As the name suggests, virtual reality (VR) refers to computer-generated environment that allows for full immersion in an artificially created world [1]. The artificially created world may be completely imaginary such as in gaming or may be based on real situations such as a tourist spot. By using devices such as VR headsets or Google Cardboard, VR technology attempts to provide perceptual stimulation similar to the one gained in real situations [2], providing the users a feeling of telepresence in a virtual environment [3]. Steffen et al., [4] suggest that VR is significantly different from other similar technologies in terms of depicting the nonexistent and in overcoming space–time linearity.

Figure 1 shows the relationship of VR with actual physical reality. In its strict sense, VR is distinct from actual physical reality, with only a minor connection

Fig. 1 VR in relation to actual physical reality



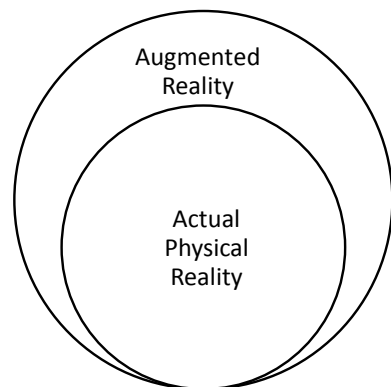
between the two in terms of users' sense of wearing a device. VR is increasingly being used for the assessment, understanding, and treatment of mental health disorders [5, 6] since it can easily simulate intangibles. Similarly, it can facilitate virtual tourism [7, 8] when coronavirus pandemic has resulted in travel restrictions across the world. However, perhaps the most popular application of VR is in the area of gaming [9, 10]. Studies suggest that VR enhances gamers' experience in terms of enjoyment, immersion, and overall satisfaction [11, 12]. It is no wonder that major gaming companies focus on developing and selling VR gears such as Oculus Quest, Sony PS VR, and HTC Vive Pro.

3 Augmented Reality

In contrast to VR that attempts to create an artificial world, augmented reality (AR) draws from the real environment of the user and adds additional details to complement/augment their experience. Thus, AR provides a semi-immersive experience by superimposing computer-generated content (text, images, or animation) on the features of user's actual environments [13]. Pokemon Go or Google Glass are some better-known examples of AR technology. Both take real-time video of users' surrounding and complement the video with virtual creatures (Pokemon Go) or complementary information (Google Glass).

Figure 2 shows the relationship between the AR and actual physical reality. Unlike VR that is distinct from actual physical reality, AR builds on it and complements it with added features. Due to its closer connection with actual reality, AR is now widely being used in the field of medicine [14]. With help of AR technology, the surgeon can view superimposed virtual organs to have a more realistic perception of the operation procedure [15]. Studies conducted with nursing students [16, 17] suggest that, with use of AR visualization, students get a better understanding of the functioning of internal organs, when compared with traditional methods such as

Fig. 2 AR in relation to actual physical reality



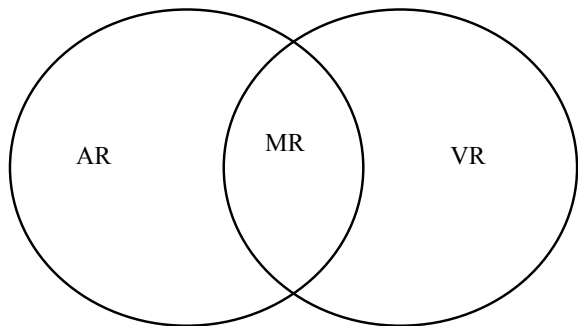
pictures, videos, or models. In fact, this may be true not only for medical education but also for any field of education in the wake of coronavirus pandemic [18]. Tourism is also a suitable avenue for AR that can complement tourist experience by providing interesting trivia or historical background when they are able to visit a place [8, 19].

4 Mixed Reality

Mixed reality (MR) is a hybrid between the VR and AR (Fig. 3). Similar to AR, MR complements external reality; while also adding virtual elements as in VR. However, in terms of computational processing and the level of interactivity, MR is much complex than AR or VR alone. While AR primarily complements or augments external reality, MR allows for the interaction between the objects coexisting in external and virtual worlds [20, 21]. MR is more complex than VR as well since it requires continuous processing of elements from the external and virtual worlds. MR provides the users a truly integrated experience using voice/gesture recognition via headset and/or motion controllers. Although the technology is yet to mature, some of the products from this relatively new domain are Microsoft's HoloLens and Samsung Odyssey.

Once again, medical science is at the forefront of using mixed reality technology. Tepper et al. [22] outline the use of Microsoft HoloLens in the operation theater, resulting in improved surgical workflow and better decision-making. However, since the technology is yet to mature, there remain both opportunities and challenges in using MR in the field of medicine [23]. Apart from medicine, the world of business is also warming up for the use of MR. The use of MR has been suggested for creating unique customer experience [20, 21] and in facilitating group decision-making [24, 25]. In the field of education, teachers' training using MR has shown to result in increased confidence in applying skills in the classroom, though challenges associated with suspension of disbelief remain [26].

Fig. 3 MR in relation to AR and VR



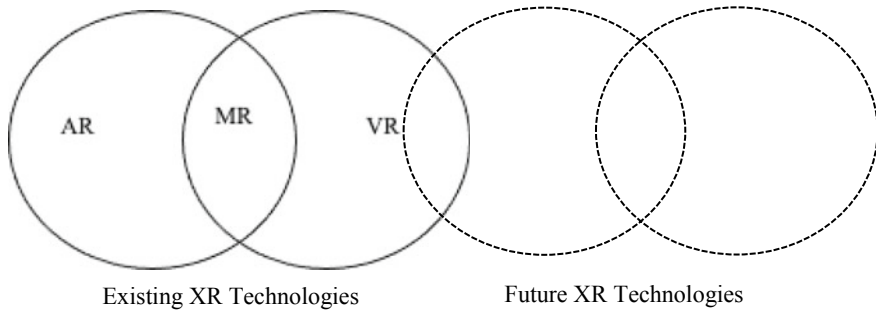


Fig. 4 XR as an umbrella term

5 Extended Reality

Extended reality (XR) is an umbrella term that not only subsumes VR, AR, and MR but also includes any future technology as well [27]. Thus, XR is a general term that indicates the recreation of use's reality in a specific context. XR offers a unique sense of presence to the users by extending their reality in a way that computer-generated reality becomes indistinguishable from the actual physical reality [28]. Figure 4 illustrates the relationship between different forms of realities discussed in the chapter. Since XR is an umbrella term, it is advised that the researchers use specific terms (AR, VR, or MR) for clarity in their writing, unless they are referring to all three.

6 Conclusion

Technology has added many novel elements to the perception of reality around us. In this regard, diverse terms are used to denote the extent of their relationship with actual physical reality. This chapter helps researchers and practitioners understand the meaning of different terms in this regard—virtual reality, augmented reality, mixed reality, and extended reality. The chapter also differentiates these terms by highlighting their use in different applications. It is hoped that this chapter helps researchers and practitioners in correctly using the appropriate term in the relevant context and contributes to the clarity of communication in the field of computer-assisted environments.

References

1. Guttentag, D. A. (2010). Virtual reality: Applications and implications for tourism. *Tourism Management*, 31(5), 637–651.

2. Diemer, J., Alpers, G. W., Peperkorn, H. M., Shibani, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology*, 6, 26.
3. Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4), 73–93.
4. Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L., & Wolman, I. (2019). Framework of affordances for virtual reality and augmented reality. *Journal of Management Information Systems*, 36(3), 683–729.
5. Emmelkamp, P. M., & Meyerbröker, K. (2021). Virtual reality therapy in mental health. *Annual Review of Clinical Psychology*, 17.
6. Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 47(14), 2393–2400.
7. Beck, J., Rainoldi, M., & Egger, R. (2019). Virtual reality in tourism: A state-of-the-art review. *Tourism Review*, 74(3), 586–612.
8. Wei, W. (2019). Research progress on virtual reality (VR) and augmented reality (AR) in tourism and hospitality: A critical review of publications from 2000 to 2018. *Journal of Hospitality and Tourism Technology*, 10(4), 539–570.
9. Rega, F., & Saxena, D. (2022). Free-roam virtual reality: A new avenue for gaming. In Verma, J. K., & Paul, S. (Eds.), *Advances in Augmented Reality and Virtual Reality*. Springer Nature.
10. Liszio, S., & Masuch, M. (2016). Designing shared virtual reality gaming experiences in local multi-platform games. In *International Conference on Entertainment Computing* (pp. 235–240). Springer, Cham.
11. Evans, L., & Rzeszewski, M. (2020). Hermeneutic relations in VR: Immersion, embodiment, presence and HCI in VR gaming. In *International Conference on Human-Computer Interaction* (pp. 23–38). Cham: Springer.
12. Shelstad, W. J., Smith, D. C., & Chaparro, B. S. (2017). Gaming on the rift: How virtual reality affects game user satisfaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61(1), pp. 2072–2076). Sage CA: Los Angeles, CA: SAGE Publications.
13. Danado, J., Dias, E., Romão, T., Correia, N., Trabuco, A., Santos, C., & Câmara, A. (2005). Mobile environmental visualization. *The Cartographic Journal*, 42(1), 61–68.
14. Eckert, M., Volmerg, J. S., & Friedrich, C. M. (2019). Augmented reality in medicine: Systematic and bibliographic review. *JMIR mHealth and uHealth*, 7(4), e10967.
15. Ha, H. G., & Hong, J. (2016). Augmented reality in medicine. *Hanyang Medical Reviews*, 36(4), 242–247.
16. Rahn, A., & Kjaergaard, H. W. (2014). Augmented reality as a visualizing facilitator in nursing education. In *INTED 2014 Valencia: 8th International Technology, Education and Development Conference*. IATED.
17. Vaughn, J., Lister, M., & Shaw, R. J. (2016). Piloting augmented reality technology to enhance realism in clinical simulation. *CIN: Computers, Informatics, Nursing*, 34(9), 402–405.
18. Elmqaddem, N. (2019). Augmented reality and virtual reality in education. Myth or reality? *International journal of emerging technologies in learning*, 14(3).
19. Cranmer, E. E., tom Dieck, M. C., & Fountoulaki, P. (2020). Exploring the value of augmented reality for tourism. *Tourism Management Perspectives*, 35, 100672.
20. Farshid, M., Paschen, J., Eriksson, T., & Kietzmann, J. (2018). Go boldly!: Explore augmented reality (AR), virtual reality (VR), and mixed reality (MR) for business. *Business Horizons*, 61(5), 657–663.
21. Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, 100, 547–560.
22. Tepper, O. M., Rudy, H. L., Lefkowitz, A., Weimer, K. A., Marks, S. M., Stern, C. S., & Garfein, E. S. (2017). Mixed reality with HoloLens: Where virtual reality meets augmented reality in the operating room. *Plastic and Reconstructive Surgery*, 140(5), 1066–1070.

23. Chen, L., Day, T. W., Tang, W., & John, N. W. (2017). Recent developments and future challenges in medical mixed reality. In *2017 IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 123–135). IEEE.
24. Ens, B., Lanir, J., Tang, A., Bateman, S., Lee, G., Piumsomboon, T., & Billinghamurst, M. (2019). Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, *131*, 81–98.
25. Piumsomboon, T., Day, A., Ens, B., Lee, Y., Lee, G., & Billinghamurst, M. (2017). Exploring enhancements for remote mixed reality collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics and Interactive Applications* (pp. 1–5).
26. Dalinger, T., Thomas, K. B., Stansberry, S., & Xiu, Y. (2020). A mixed reality simulation offers strategic practice for pre-service teachers. *Computers and Education*, *144*, 103696.
27. Kwok, A. O., & Koh, S. G. (2020). COVID-19 and extended reality (XR). *Current Issues in Tourism*, Ahead-of-Print, 1–6.
28. Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, *86*, 77–90.

Impact of Different Field-of-Views on Visuospatial Memory and Cognitive Workload in a Complex Virtual Environment



Akash K. Rao, Ronnie V. Daniel, Vishal Pandey, Sushil Chandra, and Varun Dutt

Abstract The evolution of state-of-the-art visual interfaces incorporating 360° video feeds has paved the way in expanding the human's capacity to sense complex environments exponentially. As the horizontal field-of-view (FoV) of an adult healthy human being is limited, it is imperative to evaluate the human-perceptual cognitive functions (like visuospatial working memory) as a function of FoVs offered by different visual interfaces. This research's primary objective is to address this need and to investigate the influence of different FoVs on cognitive performance in a complex virtual environment. In a laboratory-based experiment, 60 healthy subjects were equally and randomly divided into three between-subject conditions: 180° × 2 FoV (interface divided into an upper section projecting a front view of 180° and a lower section projecting a back view of 180°); 360° × 1 FoV (panoramic); and the 360° × 90° FoV (panoramic view in the lower section of the interface and 90° view of the region of interest in the upper section, which could be panned as per the participant's convenience). Participants were initially asked to place six mines at six different locations of their choice within 5 min. Starting from a different initiation point in the virtual environment, participants were then asked to specify the mines' approximate location within 5 min. Behavioral and cognitive workload results revealed that participants performed better and perceived lesser workload requirements in the 180° × 2 FoV and the 360° × 90° FoV condition compared to the 360° × 1 FoV condition. We highlight the implications of this research for creating compelling and reliable virtual environments for training personnel.

Keywords 360° visual interfaces · Fields-of-view · Visuospatial memory · Cognitive workload · Decision making · Virtual environments

A. K. Rao (✉) · V. Dutt

Applied Cognitive Science Laboratory, Indian Institute of Technology, Mandi, Himachal Pradesh, India

R. V. Daniel · V. Pandey · S. Chandra

Department of Biomedical Engineering, Institute of Nuclear Medicine and Allied Sciences, Defence Research and Development Organization, Delhi, India

1 Introduction

The rapid evolution of technology has empowered our propensity to sense complex, three-dimensional environments at a distance. This augmentation of our capacity is very important, especially for facilitating cognitive processes like visual perception, visual encoding, visual information processing, and subsequent action translation. But the meticulous and systematic extension of such sensory augmentations is crucial, especially in the military context. The design and development of 360° visual interface through image stitching/video transformation has been one such modern augmentation for empowering our perceptual capabilities. According to [1–3], modern 360° visual interfaces enable better spatial allocation of attention and better situational awareness. However, the average horizontal field-of-view (FoV) of a healthy adult human ranges from 120 to 200° [4]. Hence, viewing the entirety of 360° requires stitching multiple views together or compressing the display horizontally [1, 4]. But, compressing the 360° visual interface horizontally, primarily through image transformation, leads to information loss, negatively influencing the spatial allocation of attention and information perception [5]. Besides, the resulting distortion from compressing the 360° visual interface horizontally interferes with the user's capacity to efficiently perceive the visuospatial relationship between multiple three-dimensional structures in a complex environment [1]. Therefore, stitching various, mutually exclusive FoVs together has been recommended by most researchers [4, 5]. Generic techniques, like combining views from multiple cameras [6] and two-dimensional projective transformations, have been used for this purpose [4]. Preliminary research has evaluated numerous ways to incorporate video from various cameras to produce a continuous 360° visual interface [4, 7–9]. For example, a visual interface with a 90° × 4 FoV was compared to an interface with a 180° × 2 FoV in a dynamic decision-making task [4]. Similarly, researchers in [10] investigated the visual interface's efficacy with a 90° × 4 FoV compared to an interface with a 180° × 2 FoV in teleoperated search-and-rescue task [10]. In both cases, it was found out that the visual interface with the 90° × 4 FoV induced high cognitive workload requirements and below-par performance from the participants [4, 10]. Other researchers have also evaluated the efficacy of a panoramic 360° visual interface for intelligence, reconnaissance, and surveillance [11, 12]. According to recent studies, panoramic 360° visual interfaces and visual interfaces with a 180° × 2 FoV are the most viable, reliable methods for conveying 360° information for enhanced cognitive performance. But most of the research concerning these interfaces has been either focusing on target detection [1], formulation of cognitive maps [7], or search-and-rescue [10]. Recent research has also suggested designing and evaluating a visual interface with a 360° × 90° FoV, where the 360° panoramic view is displayed in the lower section of the interface and a 90° FoV of the region of interest is conveyed in the upper section of the interface [8, 13]. This 90° FoV can be panned as per the user's convenience [8, 13]. Even though the researchers have claimed that this visual interface could be the most optimal 360° visual interface designed till present [13], research concerning the perceptual, behavioral, and workload-based ramifications

of such an interface has hardly been investigated in the literature. Besides, research evaluating perceptual-cognitive functions (like visuospatial memory) and the relationship between visuospatial memory and cognitive workload as a function of FoV is lacking and much needed in literature.

Since the horizontal FoV of a healthy, adult human being ranges from 120 to 200° [4], the amount of information acquired through 360° visual interfaces could be perceptually and cognitively intimidating, especially to a novice user [1, 4]. This can be corroborated by theories in perceptual-cognitive performance and decision making like visuospatial allocation of attention theory [14] and instance-based learning theory (IBLT) [15, 16]. According to the visuospatial allocation of attention theory, the visuospatial distribution of attention is more efficient if the 360° visual interface provides easily discernible visual landmarks. This leads to FoV peripheries being directed left or right concerning the participant's heading in the virtual environment, subsequently allocating visuospatial attention better [14]. According to [14], this also leads to optimal utilization of information processing resources in the brain due to the user's propensity to adapt to a more pronounced distribution of visuospatial attention [14]. According to IBLT [15, 16], a well-researched theory on how humans makes decisions from experience, decision making is usually a five-step process. It involves recognizing the situation, the judgment based on the user's experience, the user's choices based on the judgments, translation of the chosen action, and feedback to those experiences that led to the chosen action [15, 16]. Hence, as per IBLT and visuospatial allocation of attention theory, users should find 360° visual interfaces with a $180^\circ \times 2$ FoV and a $360^\circ \times 90^\circ$ FoV easier to negotiate with optimal cognitive resource utilization compared to the visual interface with $360^\circ \times 1$ FoV, even though all the interfaces mathematically convey the same amount of information. However, research on the effects of different FoVs offered by 360° visual interfaces on visuospatial memory and cognitive workload in a complex virtual environment is lacking and much needed in literature.

The primary objective of this research is to address this literature gap. Specifically, we design an experiment to evaluate the impact of different 360° visual interface FoVs ($180^\circ \times 2$ FoV, $360^\circ \times 1$ FoV, and $360^\circ \times 90^\circ$ FoV) on visuospatial memory and cognitive workload in a complex virtual environment. In what follows, we provide a brief overview of research involving 360° visual interfaces and perceptual-cognitive functions. Next, we evaluate three 360° visual interfaces that differed in their FoVs and evaluated human performance and cognitive workload in an experiment. We then discuss the results in the experiment and highlight the inference of our results on visuospatial memory and cognitive workload in tasks involving 360° visual interfaces.

2 Background

A few studies have evaluated higher-level cognitive functions against 360° visual interfaces with variable FoVs in generic virtual environments. For instance,

researchers in [1] assessed the efficacy of three different 360° interfaces in target detection and remote-control task. Participants were made to locate some pre-specified targets in a complex, three-dimensional terrain using a two-dimensional 360° visual interface. Participants initially perceived their respective egocentric directions to targets and later placed them on a digitally reconstructed overhead view of the complex terrain. The results indicated that non-seamless interfaces, i.e., interfaces with discernible visual boundaries, facilitated better spatial understanding.

Similarly, researchers in another study [4] investigated the influence of two different FoVs ($90^\circ \times 4$ FoV and $180^\circ \times 2$ FoV) and two different proportions of targets and distractors (80–20% and 20–80%) in a search-and-shoot task. Twenty-five participants executed all four different FoVs and proportions of targets and distractors scenario in random order. Results suggested that participants performed better and perceived lesser cognitive workload in the $180^\circ \times 2$ FoV condition compared to the $90^\circ \times 4$ FoV condition regardless of the proportions of targets and distractors in the virtual environment.

Researchers in [17] highlighted the importance of investigating the relationship between operator capabilities, operational needs, and the operator's inherent cognitive ability to retrieve and process complex information coming from visual interfaces with 360° sensors. They investigated a set of evaluation criteria crucial for the maintenance of 360° situational awareness in complex military environments. They also highlighted the need for the 360° visual interface designs to conform to the individual's neurocognitive processing capability, which would subsequently enhance human performance.

Researchers in [13, 18] designed and proposed the possibility of a scalable, high-fidelity visual interface with a $360^\circ \times 90^\circ$ FoV for a manned ground vehicle for maintaining optimal local area awareness in complex, three-dimensional terrains. As explained in Sect. 1, this design consisted of a 360° panoramic view displayed in the lower section of the interface and a 90° FoV of the region of interest displayed in the upper section of the interface [13, 18]. This 90° FoV can be panned as per the user's convenience. The researchers in [18] then evaluated this visual interface's efficacy by comparing it to the traditional $90^\circ \times 4$ interface. The visual interface with the $360^\circ \times 90^\circ$ FoV significantly outperformed the conventional $90^\circ \times 4$ interface on all accounts (behavioral, perceptual, and workload). But an evaluation of its efficacy with more optimal 360° interfaces ($180^\circ \times 2$ FoV, $360^\circ \times 1$ panoramic) has hardly been documented and much needed in literature.

As per the visuospatial allocation of attention theory [14] and IBLT [15, 16], we expect the participants to perform better in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition compared to the $360^\circ \times 1$ FoV condition. Also, owing to its FoV, which is comparable to the healthy adult horizontal FoV, we expect the participants to perceive lesser cognitive workload requirements in the $180^\circ \times 2$ FoV condition.

3 Materials and Methods

3.1 Participants

A total of 60 participants (21 females and 39 males, mean age = 21.4 years, standard deviation = 2.1 years, minimum age = 19 years, maximum age = 26 years) at the Department of Biomedical Engineering, Institute of Nuclear Medicine and Allied Sciences, Delhi, India and the Indian Institute of Technology Mandi, Himachal Pradesh, India, participated in this study. An institutional ethical committee approved the study at the Institute of Nuclear Medicine and Allied Sciences and the Indian Institute of Technology Mandi. Participation in the study was voluntary, and the subjects gave written consent before participating in the study. Out of 60 participants, 56 participants were right-handed, 3 participants were left-handed, and 1 was ambidextrous. None of the participants reported neurological/psychological/mental history of any kind. All the participants hailed from a science/engineering/technology/mathematics (STEM) background. None of the participants had experienced a 360° visual interface before. All the participants received a flat payment of INR 50 irrespective of their performance in the study. The participants were also paid an additional INR 50 if they completed all the objectives in the study within the given time frame, i.e., 5 min.

3.2 The Virtual Environment

A ground-based virtual environment was designed for the desktop screen using Unity3D version 2019.2.3 *f1* [19]. Some of the 3D avatars incorporated in the virtual environment were created using Blender3D version 2.80 [20], and some of them were directly imported from the Unity3D asset store. Three different virtual environment versions were designed: 180° × 2 FoV, 360° × 90° FoV, and 360° × 1 FoV. A detailed explanation of each FoV condition is given in Sect. 3.1. The virtual environment consisted of two phases: the mines placing phase and the mines locating phase. In the mines placing phase, the participants were asked to place six mines at distinctly different locations (of their choice) in the virtual environment. The participants were not allowed to place the mines very close to each other. The virtual environment was programmed to indicate through an audio-visual stimulus (a beep along with a red light) if the Euclidean distance between two mines (i.e., a mine placed and a mine which was currently being placed by the participant) was less than 50 Unity3D units (1 Unity3D unit could be roughly approximated to 1 m in the real world). All the participants were given 5 min to complete the mines placing task. However, participants could switch to the mines locating task as soon as they finished the mines placing task. Several 3D objects (like cows, tents, barrels, and ropes) were placed randomly in the virtual environment so that participants could use them as landmarks (see Fig. 1). In the mines locating phase, all participants would start from a different

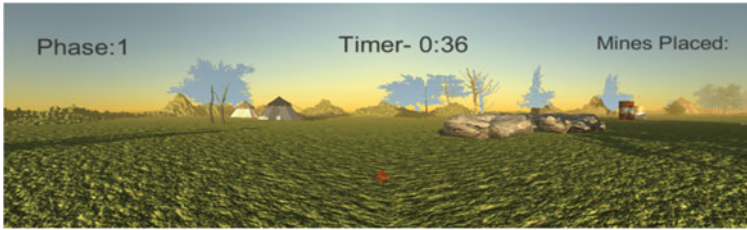


Fig. 1 Front view of the virtual environment. The number of mines placed was indicated in the top-right corner of the interface. The timer (in seconds) was placed in the top-central corner of the interface. The virtual environment phase was indicated in the top-left corner of the interface

initiation point (i.e., a point in the virtual environment that was randomly decided and one that was different from the point of beginning of the mines placing task). Participants were asked to point to the approximate location of the mines they had initially placed in no particular order. Participants were given 5 min to complete the mines locating phase. If participants could not complete the task within the stipulated time frame, the virtual environment would automatically terminate at the end of 5 min. The virtual environment was executed on a 24 inch workstation computer at a resolution of 1920×1080 pixels. Participants used a Logitech Extreme 3D Pro joystick to navigate the virtual environment and to place/locate the mines. The study was conducted in an isolated room, devoid of any external interference/noise.

3.3 Experiment Design

In a laboratory-based experiment, 60 healthy participants were divided equally and randomly into three between-subject conditions: the $180^\circ \times 2$ FoV condition, the $360^\circ \times 1$ FoV condition; and the $360^\circ \times 90^\circ$ FoV condition (see Fig. 2). As shown in Fig. 2, all the participants undertook an acclimatization session for approximately 5 min, where they got familiar with the immersion, resolution, and navigating in a dummy virtual environment using the 3D controller (see Fig. 2).

In the $180^\circ \times 2$ FoV condition, the interface was divided into an upper section projecting a front view of 180° and a lower section projecting a back view of 180° (see Fig. 3a)). This FoV condition was designed based on the approximate horizontal FoV of a healthy human, as suggested by [4]. In the $360^\circ \times 1$ FoV condition, the interface was a continuous panoramic image (see Fig. 3b). According to [3, 4], this interface reduced the load on human visuospatial processing since two 180° views are stitched together in the interface. In the $360^\circ \times 90^\circ$ condition, the interface was a panoramic view in the lower section of the interface and a 90° view of the region of interest in the upper section, which could be panned as per the participant's convenience (see Fig. 3c). This FoV condition was based on the design initially proposed by [13, 18]. This condition is motivated from the fact that for achieving complete situational

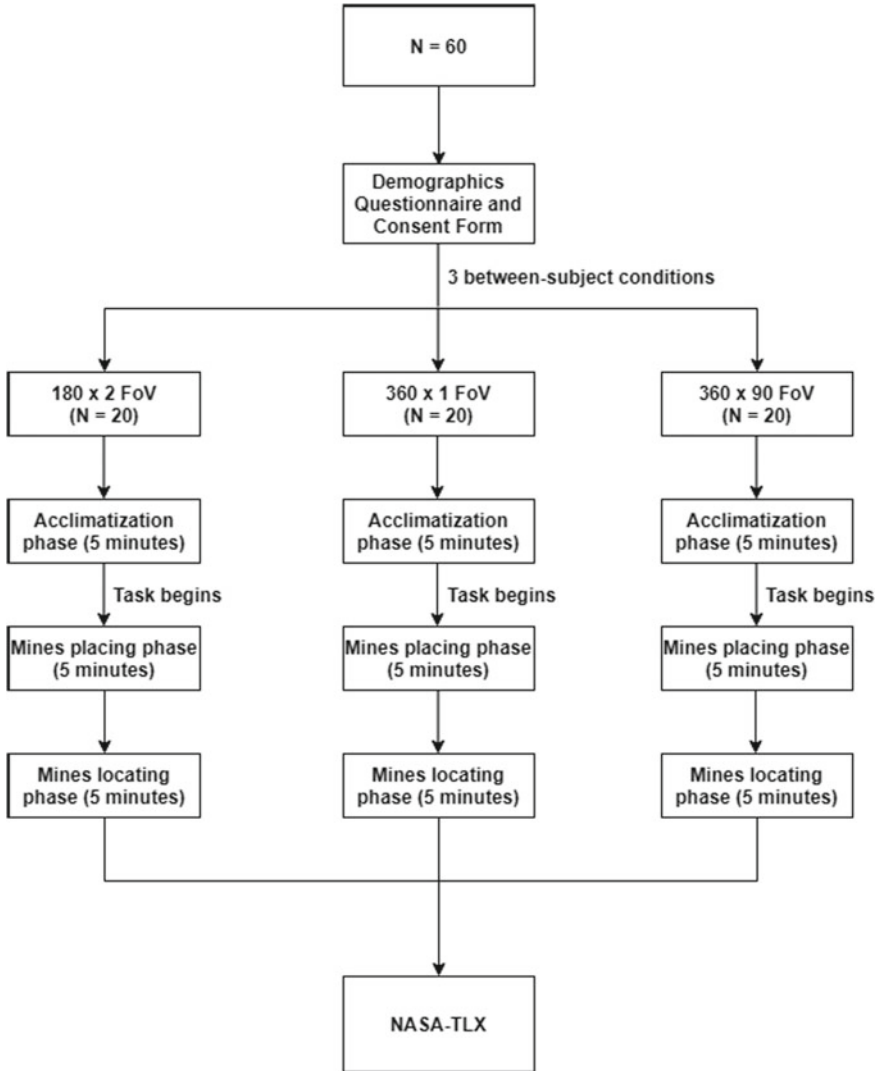
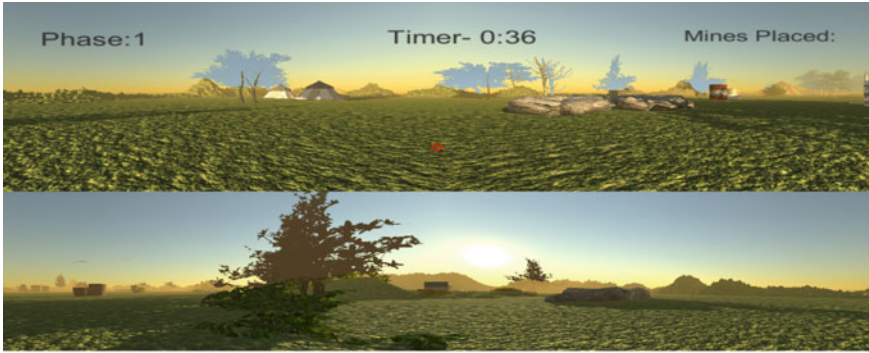
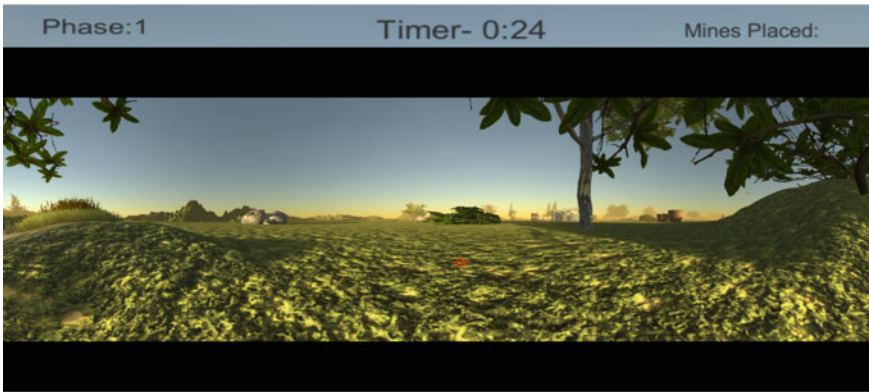


Fig. 2 Experimental design of the study

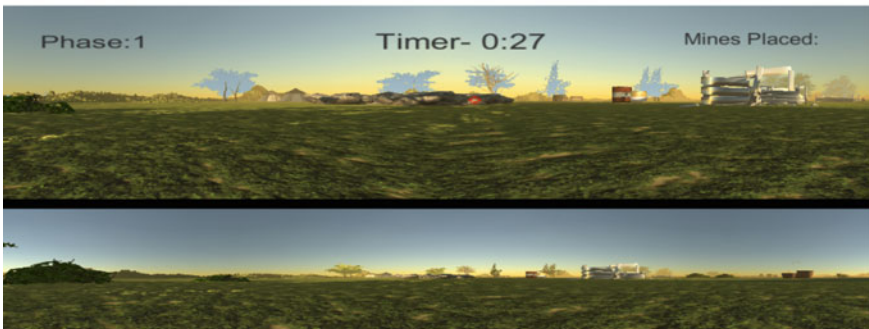
awareness, the interface should be able to provide a wholesome picture of the environment and specific regions of interest [18]. The regions of interest could be panned as per the user’s convenience. During the task, various performance measures like percentage of mines correctly located, the average time taken (in seconds), and the average Euclidean distance between mines placed and mines located (in Unity3D units, where 1 Unity3D unit equaled 1 m in the real world) were calculated. The percentage of mines correctly located was calculated by dividing the number of mines correctly located by 6 (the total number of mines placed) and multiplying it by



(a)



(b)



(c)

Fig. 3 Three different FoV conditions used in the virtual environment. **a** The $180^\circ \times 2^\circ$ FoV condition. **b** The $360^\circ \times 1^\circ$ FoV condition and **c** the $360^\circ \times 90^\circ$ FoV condition

100. Each mine located was logged as ‘located correctly’ if the distance between the mines placed in the first phase and the corresponding mine located in the second phase was less than 10 Unity3D units (~10 m in the real world). The average Euclidean distance (AED) between mines placed and mines located was calculated as follows:

$$\text{AED} = \sqrt{(x_{\text{minesplaced}} - x_{\text{mineslocated}})^2 + (y_{\text{minesplaced}} - y_{\text{mineslocated}})^2} \quad (1)$$

Besides, participants also undertook a computerized version of the NASA-Task Load Index (TLX) [21]. The NASA-TLX is a widely used self-reported scale for measuring participants’ perceived workload on a 100-point Likert scale, with increments of 5 [5]. The NASA-TLX was administered after participants executed both the phases, as shown in Fig. 2. The NASA-TLX was divided into six subscales: mental demand, physical demand, temporal demand, performance satisfaction, effort, and frustration level. We carried out one-way ANOVAs to evaluate the effect of different FoV conditions ($180^\circ \times 2$, $360^\circ \times 1$, and $360^\circ \times 90^\circ$) on various performance and cognitive/workload measures. The alpha level was set to 0.05, and the power was set to 0.8. Overall, on account of IBLT [15, 16] and visuospatial allocation of attention theory [14], we expected the participants in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition to perform better compared to the $360^\circ \times 1$ condition. Also, on account of IBLT [15, 16], effective spatial allocation theory [14], and cognitive load theory [22], we expected the $180^\circ \times 2$ FoV condition to yield lower mental demand and effort requirements and higher-performance satisfaction compared to the $360^\circ \times 90^\circ$ FoV condition and the $360^\circ \times 1$ FoV condition. Since the study did not involve vigorous physical movement/locomotion of any kind and did not incorporate particularly annoying/controversial/irritating stimuli, we did not expect the self-reported physical demand and frustration level to reach statistical significance.

3.4 Procedure

Before the participants began the experiment, they were briefed about the objectives to be achieved. The participants were given detailed instructions on executing both the phases of the experiment: the mines placing phase and the mines locating phase. Participants were also informed that the mines they would place in the mines placing phase would have to be subsequently located in the mines locating phase, beginning the phase from a different initiation point (as described in Sect. 3.2). All the participants then undertook a basic demographics questionnaire and gave their written consent to participate in the study (see Fig. 2).

4 Results

4.1 Performance Measures

Percentage of Mines Correctly Located

Figure 4 shows the percentage of mines correctly located across all FoV conditions. As shown in Fig. 4, the percentage of mines correctly located was significantly different across all FoV conditions ($F(2.59) = 9.61, p < 0.05, \eta_p^2 = 0.74$). Bonferroni post-hoc test revealed that the percentage of mines correctly located was similar in the $180^\circ \times 2$ and $360^\circ \times 2$ FoV conditions; however, there was a significant difference between the $360^\circ \times 1$ and $180^\circ \times 2$ FoV condition and the $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition ($180^\circ \times 2$ FoV: $\mu = 62.5\% \sim 360^\circ \times 90^\circ$ FoV: $\mu = 61.67\%$ ($p = 0.99$); $360^\circ \times 1$ FoV: $\mu = 31.67\% < 180^\circ \times 2$ FoV: $\mu = 62.5\%$ ($p < 0.05$); $360^\circ \times 1$ FoV: $\mu = 31.67\% < 360^\circ \times 90^\circ$ FoV: $\mu = 61.67\%$ ($p < 0.05$)). Overall, as per our expectations, participants in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition correctly located more mines compared to the $360^\circ \times 1$ FoV condition.

Average Time Taken

Figure 5 shows the average time taken (in seconds) across all FoV conditions. As shown in Fig. 5, the average time taken was significantly different across different FoV conditions ($F(2.59) = 5.08, p < 0.05, \eta_p^2 = 0.84$). Bonferroni post-hoc test revealed that the average time taken was similar in the $180^\circ \times 2$ and $360^\circ \times 90^\circ$ FoV condition and the $180^\circ \times 2$ and $360^\circ \times 1$ FoV condition; however, there was a significant difference between the $360^\circ \times 1$ and the $360^\circ \times 90^\circ$ FoV conditions ($180^\circ \times 2$ FoV: $\mu = 246.55 \text{ s} \sim 360^\circ \times 90^\circ$ FoV: $\mu = 257.95 \text{ s}$ ($p = 0.99$); $180^\circ \times 2$ FoV: $\mu = 246.55 \text{ s} \sim 360^\circ \times 1$ FoV: $\mu = 216.95 \text{ s}$ ($p = 0.09$); $360^\circ \times 1$ FoV: μ

Fig. 4 Means and 95% confidence intervals of the percentage of mines correctly located across different FoV conditions. The error bars show 95% CI across point estimates

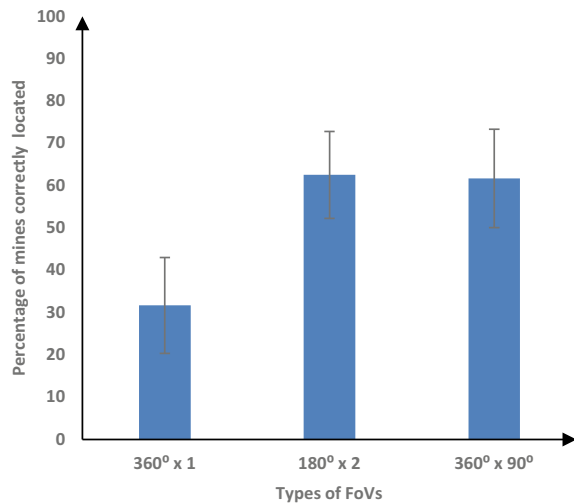
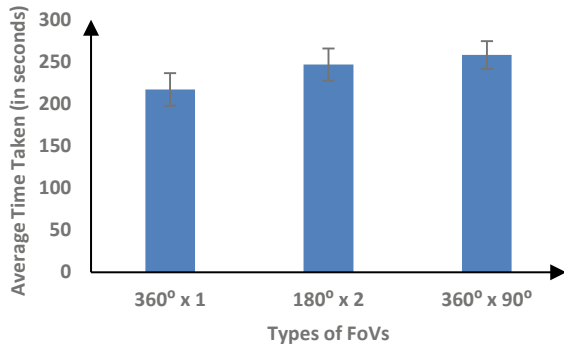


Fig. 5 Means and 95% confidence intervals of the average time taken (in seconds) across different FoV conditions. The error bars show 95% CI across point estimates



= 216.95 s < 360° × 90° FoV: $\mu = 257.95$ s ($p < 0.05$)). Overall, participants in the 360° × 1 FoV condition took significantly lesser time compared to those in the 180° × 2 and 360° × 1 FoV conditions.

Average Euclidean Distance Between Mines Placed and Mines Located

Figure 6 shows the average Euclidean distance between mines placed and mines located (in Unity3D units) across different FoV conditions. As shown in Fig. 6, the average Euclidean distance between mines placed and mines located was significantly different across all FoV conditions ($F(2.59) = 13.81$, $p < 0.05$, $\eta_p^2 = 0.67$). Bonferroni post-hoc test revealed that the average Euclidean distance between mines placed and mines located was similar in the 180° × 2 and 360° × 90° FoV condition; however, there was a significant difference between the 360° × 1 and 180° × 2 FoV condition and the 360° × 1 and 360° × 90° FoV condition (180° × 2 FoV: $\mu = 14.3$ ~ 360° × 90° FoV: $\mu = 17.25$ ($p = 0.99$); 360° × 1 FoV: $\mu = 31.7$ > 180° × 2 FoV: $\mu = 14.3$ ($p < 0.05$); 360° × 1 FoV: $\mu = 31.7$ > 360° × 90° FoV: $\mu = 17.25$ ($p < 0.05$)). Overall, as per our expectations, the participants in the 180° × 2 FoV condition and the 360° × 90° FoV condition recorded a significantly lesser Euclidean distance between the mines placed and the mines located compared to the 360° × 1 FoV condition.

4.2 Workload Measures

Mental Demand

Figure 7 shows the self-reported mental demand across all FoV conditions. As shown in Fig. 7, the self-reported mental demand was significantly different across all FoV conditions ($F(2.59) = 8.54$, $p < 0.05$, $\eta_p^2 = 0.76$). Bonferroni post-hoc test revealed that the self-reported mental demand was significantly different in the 180° × 2 and 360° × 90° FoV condition and the 180° × 2 and 360° × 1 FoV condition; however, the self-reported mental demand was similar across 360° × 1 and 360° × 90° FoV condition (180° × 2 FoV: $\mu = 34.75$ < 360° × 1 FoV: $\mu = 51$ ($p < 0.05$); 180° × 2

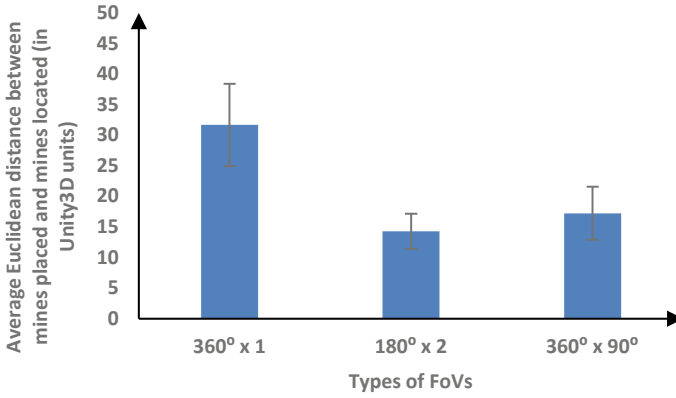
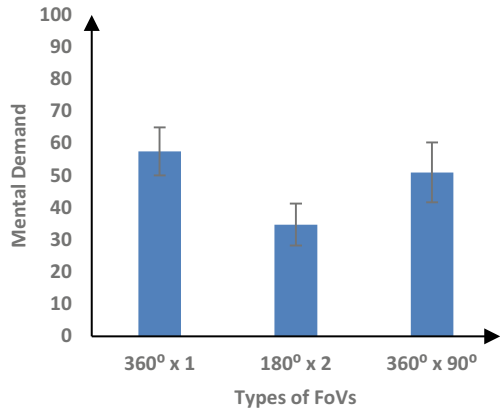


Fig. 6 Means and 95% confidence intervals of the average Euclidean distance between mines placed and mines located (in Unity3D units) across different types of FoV conditions. The error bars show 95% CI across point estimates

Fig. 7 Means and 95% confidence intervals of the self-reported mental demand across different types of FoV conditions. The error bars show 95% CI across point estimates

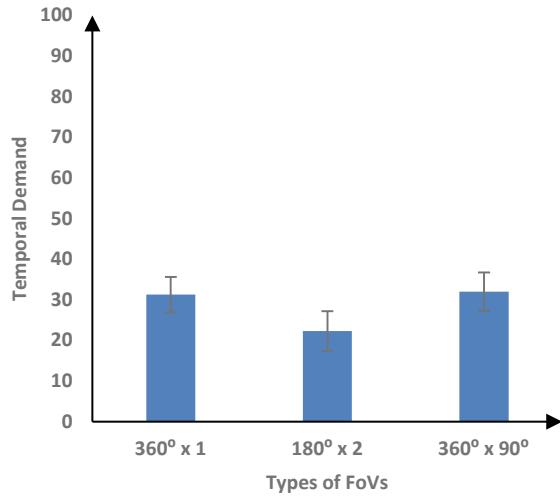


FoV: $\mu = 34.75 < 360^\circ \times 90^\circ$ FoV: $\mu = 57.5$ ($p < 0.05$); $360^\circ \times 1$ FoV: $\mu = 57.5 \sim 360^\circ \times 90^\circ$ FoV: $\mu = 51$ ($p = 0.77$). Overall, as per our expectations, the $180^\circ \times 2$ FoV condition yielded significantly lesser mental demand requirement compared to the $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition.

Physical Demand

The self-reported physical demand was not significantly different across different FoV conditions ($F(2.59) = 0.31, p = 0.74, \eta_p^2 = 0.07$). Since this task did not involve significant muscle/body movement of any kind, the self-reported physical demand's failure to reach statistical significance was conforming to our expectations.

Fig. 8 Means and 95% confidence intervals of the self-reported temporal demand across different types of FoV conditions. The error bars show 95% CI across point estimates



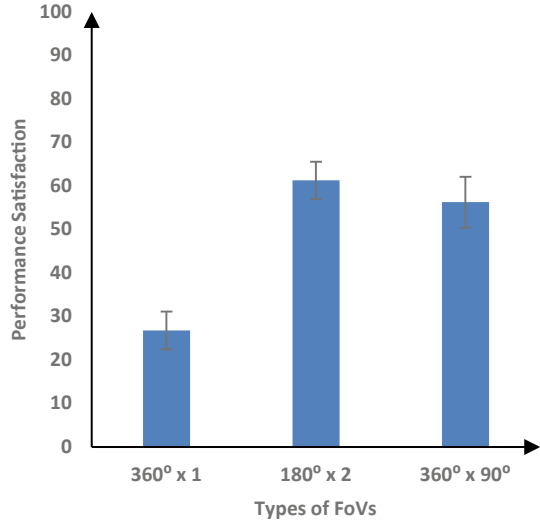
Temporal Demand

Figure 8 shows the self-reported temporal demand across all FoV conditions. As shown in Fig. 8, the self-reported temporal demand was significantly different across all FoV conditions ($F(2.59) = 5.13, p < 0.05, \eta_p^2 = 0.74$). Bonferroni post-hoc test revealed that the self-reported temporal demand was significantly different in the $180^\circ \times 2$ and $360^\circ \times 90^\circ$ FoV condition and the $180^\circ \times 2$ and $360^\circ \times 1$ FoV condition; however, the self-reported temporal demand was similar across $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition ($180^\circ \times 2$ FoV: $\mu = 22.25 < 360^\circ \times 1$ FoV: $\mu = 32$ ($p < 0.05$); $180^\circ \times 2$ FoV: $\mu = 22.25 < 360^\circ \times 90^\circ$ FoV: $\mu = 31.25$ ($p < 0.05$); $360^\circ \times 1$ FoV: $\mu = 31.25 \sim 360^\circ \times 90^\circ$ FoV: $\mu = 32$ ($p = 0.99$)). Contrary to the results obtained in average time taken (in seconds) as shown in Fig. 2, as per our expectations, participants in the $180^\circ \times 2$ FoV condition yielded significantly lesser time management requirement compared to the $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition.

Performance Satisfaction

Figure 9 shows the self-reported performance satisfaction across all FoV conditions. As shown in Fig. 9, the self-reported performance satisfaction was significantly different across all FoV conditions ($F(2.59) = 55.29, p < 0.05, \eta_p^2 = 0.34$). Bonferroni post-hoc test revealed that the self-reported performance satisfaction was similar in the $180^\circ \times 2$ and $360^\circ \times 90^\circ$ FoV condition; however, there was a significant difference between the $360^\circ \times 1$ and $180^\circ \times 2$ FoV condition and the $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition ($180^\circ \times 2$ FoV: $\mu = 61.25 \sim 360^\circ \times 90^\circ$ FoV: $\mu = 56.25$ ($p = 0.49$); $360^\circ \times 1$ FoV: $\mu = 26.25 < 180^\circ \times 2$ FoV: $\mu = 61.25$ ($p < 0.05$); $360^\circ \times 1$ FoV: $\mu = 26.25 < 360^\circ \times 90^\circ$ FoV: $\mu = 56.25$ ($p < 0.05$)). Overall, as per our expectations, the participants in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition yielded significantly more performance satisfaction compared to the $360^\circ \times 1$ FoV condition.

Fig. 9 Means and 95% confidence intervals of the self-reported performance satisfaction across different types of FoV conditions. The error bars show 95% CI across point estimates



Effort

Figure 10 shows the self-reported effort across all FoV conditions. As shown in Fig. 10, the self-reported effort was significantly different across all FoV conditions ($F(2.59) = 12.08, p < 0.05, \eta_p^2 = 0.69$). Bonferroni post-hoc test revealed that the self-reported effort was significantly different in the $180^\circ \times 2$ and $360^\circ \times 90^\circ$ FoV condition and the $180^\circ \times 2$ and $360^\circ \times 1$ FoV condition; however, the self-reported effort was similar across $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition ($180^\circ \times 2$ FoV: $\mu = 38.5 < 360^\circ \times 1$ FoV: $\mu = 64.75 (p < 0.05)$; $180^\circ \times 2$ FoV: $\mu = 38.5 < 360^\circ \times 90^\circ$ FoV: $\mu = 54.5 (p < 0.05)$; $360^\circ \times 1$ FoV: $\mu = 64.75 \sim 360^\circ \times 90^\circ$ FoV: $\mu = 54.5 (p = 0.19)$). Overall, as per our expectations, the $180^\circ \times 2$ FoV condition yielded significantly lesser effort requirement compared to the $360^\circ \times 1$ and $360^\circ \times 90^\circ$ FoV condition.

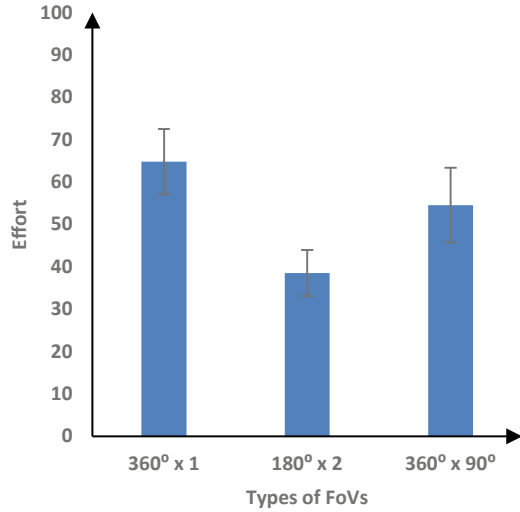
Frustration Level

The self-reported frustration level was not significantly different across different FoV conditions ($F(2.59) = 0.35, p = 0.7, \eta_p^2 = 0.09$). Since the experiment did not contain irritating and annoying stimuli of any kind which could potentially upset the participants, the non-significance of the self-reported frustration level was as per our expectations.

4.3 Correlation Analysis

In addition to one-way ANOVA, we also performed a two-tailed Pearson correlation analysis to investigate the relationship between different performance and cognitive measures in each of the three conditions. In the $360^\circ \times 1$ FoV condition, the average

Fig. 10 Means and 95% confidence intervals of the self-reported effort across different types of FoV conditions. The error bars show 95% CI across point estimates



time taken (in seconds) was significantly positively correlated with the self-reported performance satisfaction ($r(18) = 0.47, p < 0.05$). In the $180^\circ \times 2$ FoV condition, the average time taken was significantly positively correlated with the self-reported mental demand ($r(18) = 0.47, p < 0.05$). In the $360^\circ \times 90^\circ$ FoV condition, the average time taken was significantly negatively correlated with the self-reported physical demand ($r(18) = -0.67, p < 0.01$).

In the $360^\circ \times 1$ FoV condition, the average Euclidean distance between mines placed and mines located was significantly negatively correlated with the self-reported physical demand ($r(18) = -0.5, p < 0.05$). Also, in the $360^\circ \times 1$ FoV condition, the average Euclidean distance between mines placed and mines located was significantly negatively correlated with the self-reported effort ($r(18) = -0.45, p < 0.05$). Other correlations were not significant.

5 Discussion and Conclusions

This experiment evaluated the impact of 360° visual interfaces with different FoVs on visuospatial memory and cognitive workload in a complex virtual environment. Results revealed that the participants performed better in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition compared to the $360^\circ \times 1$ FoV condition. Results also indicated that participants generally perceived significantly lesser cognitive workload requirements in the $180^\circ \times 2$ FoV condition compared to the $360^\circ \times 90^\circ$ FoV condition and the $360^\circ \times 1$ FoV condition. Correlation analysis also revealed relationships between performance measures and cognitive workload measures in all the FoV conditions. Overall, most of the results were consistent with the visuospatial allocation of attention theory [14], IBLT [15, 16], and the cognitive load theory [22].

First, we found a significantly higher percentage of mines correctly located in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition compared to the $360^\circ \times 1$ FoV condition. Also, we found the participants in the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition recorded significantly lesser average Euclidean distance between mines placed and mines located compared to the $360^\circ \times 1$ FoV condition. These results were likely because the discernible visual landmarks and significantly pronounced image peripheries in both the $180^\circ \times 2$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition led to better visuospatial distribution of attention [14]. In accordance with IBLT [15, 16], the participants were able to create a better mental model of the three-dimensional landmarks present in the virtual environment, which eventually led to better creation of instances in memory concerning the mines placed in the first phase.

Second, we found the average time taken in the $360^\circ \times 90^\circ$ FoV condition to be significantly higher compared to the $180^\circ \times 2$ FoV and the $360^\circ \times 1$ FoV conditions. This result agrees with the performance results in these conditions, where spending little time deteriorated performance in $360^\circ \times 1$ FoV condition compared to that in the $360^\circ \times 90^\circ$ FoV condition. In fact, the correlation analyses supported lesser performance satisfaction in the $360^\circ \times 1$ FoV condition and more mental demand in the $360^\circ \times 90^\circ$ FoV condition.

We also found that participants in the $180^\circ \times 2$ FoV condition perceived significantly lesser cognitive workload requirements (lesser mental demand, lesser effort, lesser temporal demand, and more performance satisfaction) compared to the $360^\circ \times 1$ FoV condition and the $360^\circ \times 90^\circ$ FoV condition. This result was in conjunction with the cognitive load theory [22] and the visuospatial allocation of attention theory [14], where it was argued that the $180^\circ \times 2$ FoV condition was close to the horizontal FoV of an adult, healthy human being ($120\text{--}200^\circ$) [4]. This likely resulted in participants perceiving complex information from the virtual environment with optimal cognitive resources utilization in the brain. Even though the participants had performed equally well in the $360^\circ \times 90^\circ$ FoV condition, the presence of a panoramic 360° view of the environment in addition to the 90° region of interest created significantly more instances in the memory, leading to higher information processing and workload requirements.

We also found out that the average Euclidean distance between mines placed and mines located was negatively correlated with both self-reported physical demand and effort in the $360^\circ \times 1$ FoV condition. This result meant that participants had to exert significantly higher physical demand and exert considerably more effort to reduce the average Euclidean distance between mines placed and mines located which could eventually lead to a higher number of mines being correctly located. This result agreed with the visuospatial allocation of attention theory [14], where the absence of clearly distinguishable landmarks and peripheries in the $360^\circ \times 1$ FoV condition led to participants finding it difficult to distribute their attentional resources uniformly in the virtual environment.

Overall, our results indicated that visual interface with a $180^\circ \times 2$ FoV performed significantly better for training participants in complex visuospatial memory-based tasks. However, participants in the novel $360^\circ \times 90^\circ$ FoV condition also performed

well. The higher cognitive workload and information processing requirements in the $360^\circ \times 90^\circ$ FoV could probably be nullified to an extent by imparting heterogeneous longitudinal training to the participants.

The results obtained in this research work could be replicable. However, these results may need to be corroborated and cross-validated by appointing personnel to execute complex tasks in real-world terrain. Thus, as part of future work, we plan to integrate the results obtained in this research to test the efficacy of the transfer of virtual environment-based training in real-world settings. We also plan to evaluate different 360° visual interfaces based on various physiological, neuro-physiological, ergonomic, behavioral, and cognitive parameters.

References

1. Boonsuk, W., Gilbert, S. B., & Kelly, J. W. (2012). The impact of three interfaces for 360-degree video on spatial cognition. In *Conference on Human Factors in Computing Systems—Proceedings* (pp. 2579–2588). New York, USA: ACM Press. <https://doi.org/10.1145/2207676.2208647>.
2. Boonsuk, W. (2011). Evaluation of desktop interface displays for 360° video. <https://lib.dr.ias.tate.edu/etd/10249/>. <https://doi.org/10.31274/etd-180810-2797>.
3. Rao, A. K., Dhankar, U., Satyarthi, C., Chandra, S., & Dutt, V. (2018). Influence of different fields of view on decision-making in a search-and-shoot scenario. In *Proceedings—2017 International Conference on Machine Learning and Data Science, MLDS 2017* (pp. 61–67). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/MLDS.2017.23>.
4. Rao, A., Satyarthi, C., Dhankar, U., Chandra, S., & Dutt, V. (2019). Indirect visual displays: Influence of field-of-views and target-distractor base-rates on decision-making in a search-and-shoot task. In *Proceedings—2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018* (pp. 4326–4332). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/SMC.2018.00731>.
5. Foote, J., & Kimber, D. (2000). FlyCam: Practical panoramic video and automatic camera control. In *IEEE International Conference on Multi-Media and Expo*. (pp. 1419–1422). <https://doi.org/10.1109/icme.2000.871033>.
6. Stratton, G. (1896). Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review*, 3, 611–617. <https://doi.org/10.1037/h0072918>
7. Kadous, M. W., Sheh, R. K. M., & Sammut, C. (2006). Effective user interface design for rescue robotics. In *HRI 2006: Proceedings of the 2006 ACM Conference on Human-Robot Interaction* (pp. 250–257). New York, USA: Association for Computing Machinery. <https://doi.org/10.1145/1121241.1121285>.
8. Meguro, J. I., Hashizume, T., Takiguchi, J. I., & Kurosaki, R. (2005). Development of an autonomous mobile surveillance system using a network-based RTK-GPS. In *Proceedings—IEEE International Conference on Robotics and Automation* (pp. 3096–3101). <https://doi.org/10.1109/ROBOT.2005.1570586>.
9. Greenhill, S., & Venkatesh, S. (2006). Virtual observers in a mobile surveillance system. In *Proceedings of the 14th Annual ACM International Conference on Multimedia, MM 2006* (pp. 579–588). New York, USA: ACM Press. <https://doi.org/10.1145/1180639.1180759>.
10. Cosenzo, K., Metcalfe, J., Gordon, S., Johnson, T., Feng, T., Brumm, B., Evans, A., & Capstick, E. (2001). Impact of 360° Sensor Information on Vehicle Commander Performance. Undefined.
11. Rao, A. K., Pramod, B. S., Chandra, S., & Dutt, V. (2019). Influence of indirect vision and virtual reality training under varying manned/unmanned interfaces in a complex search-and-shoot simulation. In *Advances in Intelligent Systems and Computing* (pp. 225–235). Springer Verlag. https://doi.org/10.1007/978-3-319-94223-0_21.

12. Cowan, L. V., Lyons, A., Babington, J., Carles Santacana, G., Wood, A.P., & Harvey, A.R. (2020). Multi-aperture 360° panoramic imaging for enhanced situational awareness. In R. Wu, O. Matoba, Y. Wang, & T. E Kidger (Eds.). *Optical Design and Testing X* (p. 2). SPIE. <https://doi.org/10.1117/12.2574802>.
13. Southworth, M., & Paytan, R. (2017). Delivering a true 360° view of an operating environment on any platform
14. Gobell, J. L., Tseng, C. H., & Sperling, G. (2004). The spatial distribution of visual attention. *Vision Research*, 1273–1296. <https://doi.org/10.1016/j.visres.2004.01.012>.
15. Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, 25, 143–153. <https://doi.org/10.1002/bdm.722>
16. Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, 118, 523–551. <https://doi.org/10.1037/a0024558>
17. Metcalfe, J. S., Mikulski, T., & Dittman, S. (2011). Accounting for human neurocognitive function in the design and evaluation of 360-degree situational awareness display systems. In *Display Technologies and Applications for Defense, Security, and Avionics V; and Enhanced and Synthetic Vision 2011* (p. 80420F). SPIE. <https://doi.org/10.1117/12.885051>
18. Metcalfe, J., Cosenzo, K., Johnson, T., Brumm, B., Manteuffel, C., Evans, A., & Tierney, T.M. (2010). Human dimension challenges to the maintenance of local area awareness using a 360 degree indirect vision system
19. Unity Real-Time Development Platform | 3D, 2D VR & AR Engine, <https://unity.com/>. Last accessed 2020/12/29.
20. blender.org—Home of the blender project—Free and open 3D creation software, <https://www.blender.org/>. Last accessed 2020/12/29.
21. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139–183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
22. Sweller, J. (2011). *Cognitive Load Theory*. <https://doi.org/10.1016/B978-0-12-387691-1.00002-8>

Free-Roam Virtual Reality: A New Avenue for Gaming



Francesco Rega and Deepak Saxena 

Abstract Recent coronavirus pandemic has restricted people's movement. This has resulted in people spending more time indoors, and quite often, in pursuits like streaming and/or gaming. Traditional gaming often requires the users to sit for hours in front of a screen. This makes the practice unhealthy, and at the same time, makes it relatively less social. Free-roam gaming is a new form of gaming that utilizes virtual reality (VR) techniques to create an environment in which the user feels present. In free-roam gaming, a camera captures users' movements and converts into gaming actions in the game. Generally such free-roam games are conducted in large spaces to allow for free movement, and thus, allow a group of users to play the same simultaneously. This chapter explores the free-roam VR games via a case study of Zero Latency and the potential of such VR games in dealing with the isolation caused by the coronavirus pandemic.

Keywords Free-roam · Gaming · Virtual reality · VR

1 Introduction

Recent coronavirus pandemic has restricted people's movement. This has resulted in people spending more time indoors, and quite often, in pursuits like streaming and/or gaming [1, 2] at home. Despite the advent of multiplayer online games and growing popularity of e-Sports due to their social nature [3], restricted living has also resulted in people feeling isolated at their homes and experiencing a deterioration in mental health [4, 5]. In such situations, virtual reality-based free-roam could be a possible solution since it simultaneously allows for leisure and social interaction at the same time [6].

F. Rega
Trinity College Dublin, Dublin, Ireland

D. Saxena (✉)
Birla Institute of Technology and Science Pilani, Pilani, India
e-mail: saxenad@tcd.ie

This chapter explores free-roam VR games via a case study of Zero Latency, based in Dublin Ireland and the potential of such VR games in dealing with the isolation caused by the coronavirus pandemic. The next section defines the term ‘virtual reality’ as understood in this chapter and distinguishes it with other terms such as augmented reality, mixed reality, virtual worlds, and 360° videos. This is followed by a discussion on free-roam gaming and how it differentiates itself from traditional gaming. Thereafter, the case study of a company that offers free-roam gaming is presented to illustrate the potential of such endeavours. Finally, the last section discusses the challenges and opportunities thrown-in by the coronavirus pandemic for such VR gaming avenues.

2 Defining Virtual Reality

Biocca [7] defines VR as ‘an environment created by a computer or other media, an environment in which the user feels present’ (p. 5–6). When a user is immersed in a human–computer environment that can be perceived or interacted with, they then find themselves in a VR system [8]. VR is a technology capable of replicating or simulating lifelike experiences [9, 10], imparting the users the feeling to be in another place and time. While there are slight variations in the definitions included above, almost all scholars agree on a common set of core components that define VR. Two major elements with pivotal roles in VR experience are *vividness* and the level of possible *interaction* [11]. Vividness provides a wide range of sensory breadth and depth [12] and refers to the simulation’s capacity to optimally replicate the real world. Interactivity refers to the provided capacity of manipulating virtual objects and opportunity to observe the consequent results in real time. In short, it represents the capacity to replicate what would happen in the real world with similar actions [7, 8].

A common enough mistake is to conflate VR with virtual worlds or with augmented reality. Virtual worlds, with Second Life being a good example, are commonly defined as 3D Internet-based simulated environments, with more or less rules and patterns to follow, and where the user is generally represented by an avatar [13]. Augmented reality (AR), as the name suggests, displays virtual components within the actual context of a user’s surroundings [10, 14], thus augmenting our own reality with new elements. This is often done using 360° videos that capture every direction at the same time, using specific cameras and techniques. Finally, mixed reality can be explained as a continuum between VR and AR [10], allowing virtual elements to interact in an environment (Fig. 1).

VR is becoming common in many markets and industries, with new and more powerful devices being launched regularly. This has prompted audiences to be gradually more receptive towards the technology and has opened the avenues for many applications [13]. Free-roam gaming is one such application that is discussed in the next section.

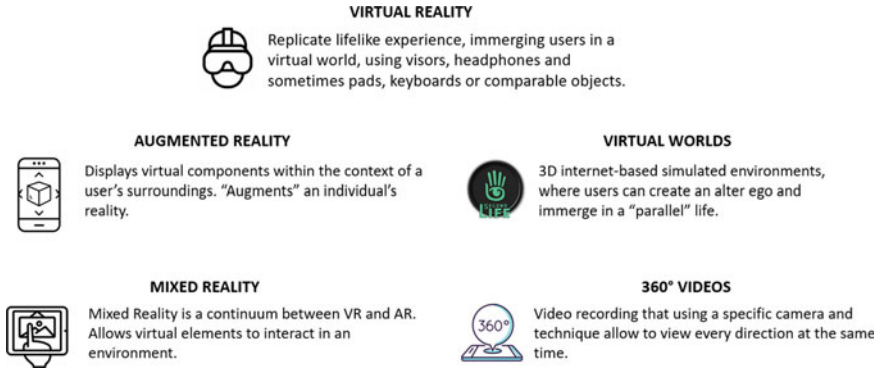


Fig. 1 Virtual reality and differences from other similar technologies

3 Free-Roam Gaming

Traditional gaming requires the users to sit for hours in front of a screen. This makes the practice unhealthy and at the same time makes it relatively less social. Multiplayer online games and e-Sports provide some sense of social interaction, and their benefits are still unclear [15]. Gaming consoles at home, while allowing for some physical activity, they are still constrained in terms of free space at users' home. In contrast, free-roam gaming is a new form of gaming that utilizes VR techniques to create an environment in which the user feels present. In contrast to traditional linear gameplay usually played within an online environment, free-roam games are more nonlinear and played in vast open spaces (typically from 200 to 400 square meters) to allow for roaming and exploring the virtual reality. Generally such free-roam games are conducted in large spaces to allow for free movement and thus allow a group of users to play the same simultaneously. Therefore, free-roam gaming not only allows for a physical activity, it also allows for social gaming without compromising the richness of the game. In a sense, it provides the affordances for preserving the good features of physical reality and enhancing the features of a virtual reality [14].

As shown in Fig. 2, a free-roam gaming avenue is divided into grids with each grid having a motion capture camera. The motion capture camera captures users' movements, which is then converted into actions in the game. The users experience the game through a VR headset and have a powerful laptop in a backpack, and sometimes a controller depending on the nature of the game [16]. The VR headset provides realistic visuals and sound and allows for communication with other team members. The backpack laptop is mainly used to process the input from and transfer the output to the user headset. This also makes the user movements free since there is no cable attached with an external machine. Depending on the need of the game (e.g. a shooting game), the users get the pinpoint accuracy within the game. Finally, all this is complemented with powerful cameras, storage devices and back-end processors,

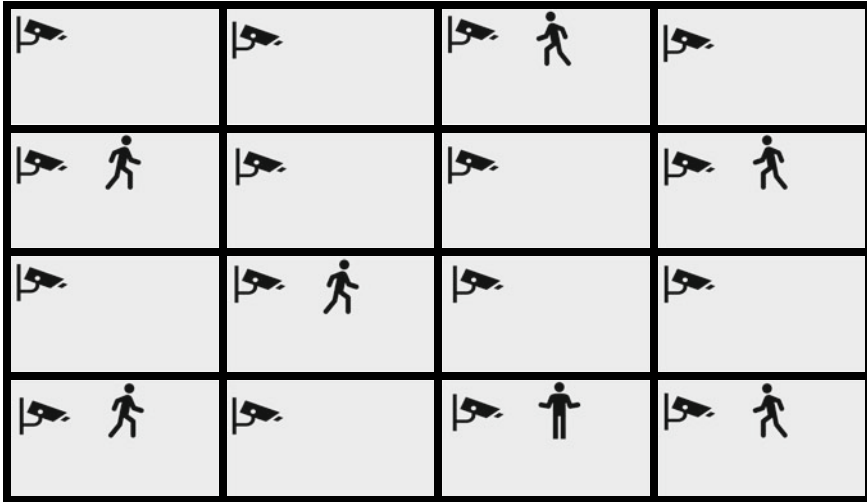


Fig. 2 An illustration of a free-roam gaming venue

thus making the establishment of free-roam virtual gaming venue a costly exercise from the infrastructure perspective.

4 The Case of Zero Latency

Zero Latency is a technology company, originally founded in 2013, in Melbourne, Australia. After a successful fundraising campaign bolstered by both crowdfunding and corporate investors, Zero Latency gained international recognition and partnered with Alienware, Dell's gaming division, to further improve their technology. Their business model evolved over time and now operates as a franchise model, with venues located across Asia, Australia, Europe, North America and South America. The franchisees operate fall under the larger Zero Latency umbrella, which requires franchisees to utilize the provided hardware and software. While this model provides required expertise to new franchisee, it also leaves little room for agency over game choices and decisions regarding the gaming equipment.

Zero Latency's mission is to offer world class entertainment to clients by providing them with an unforgettable experience that cannot be found in traditional gaming. Their primary vision is multi-faceted and aims to normalize the perception of virtual reality, with an overall push to help VR become a mainstream entertainment choice that allows customers to engage in a more social style of gaming. Their core values are heavily centred on customer experience, which is demonstrated through staff training focused on soft skill development, thoughtful facility design and attentive venue and equipment maintenance. As a young business, their long-term goals focus on both

expansion and improvement. Zero Latency hopes to achieve these benchmarks by maximizing the number of available spots used in their venue and by growing the number of repeat customers, an initiative they are already actively pursuing.

It may also be noted that free-roam gaming avenues are not just a draw for a group of gaming enthusiasts. VR-based free-roam games are also being used by corporate clients for team building exercises and leadership development, thus creating an additional revenue stream for the company. Considering the benefits of such gaming exercises in terms of productivity and leadership development [17, 18], corporate gaming sessions are increasingly becoming common. At the same time, VR games may also be used for in-game advertising as an additional revenue stream to keep the price low for individual gamers.

5 Coronavirus: Challenges and Opportunities

Coronavirus pandemic, starting in late 2019 and continuing throughout 2020, resulted in restriction of movement across the globe. Governments across the world imposed severe restrictions on the population to contain the spread of the virus. Due to their extended stay at home, people resorted to entertainment avenues like video streaming or gaming within their homes [1, 2]. Initiatives like #PlayApartTogether attempt to raise a sense of community among those who are playing games within the confines of their homes. There is also a rise in the popularity of multiplayer online games and e-Sports due to their social nature [3]. However, extended stay within confined spaces is found to have a deteriorating effect on mental health, especially if one is staying alone [4, 5].

In such situations, free-roam gaming in a group could be a potential solution [6] since it simultaneously allows for psycho-somatic activity while observing social distancing. In other words, it provides the desired affordances (i.e. social distancing) within physical reality and enhancing the features (i.e. leisure and entertainment) of a virtual reality [14]. While many free-roam gaming venues found it difficult to sustain themselves at the beginning of the pandemic [19], with adequate hygiene measures such venues can still provide a stimulating yet safe service to their consumers. As Zero Latency defines itself on its twitter handle (Zero [20], ‘social multiplayer, free-roam, epic scale virtual reality entertainment’, VR-based free-roam gaming can tackle some of the isolation and mental health issues caused by the pandemic for gaming enthusiasts and corporate workers alike.

References

1. King, D. L., Delfabbro, P. H., Billieux, J., & Potenza, M. N. (2020). Problematic online gaming and the COVID-19 pandemic. *Journal of Behavioral Addictions*, 9(2), 184–186.

2. López-Cabarcos, M. Á., Ribeiro-Soriano, D., & Piñeiro-Chousa, J. (2020). All that glitters is not gold. The rise of gaming in the COVID-19 pandemic. *Journal of Innovation and Knowledge*, 5(4), 289–296.
3. Hall, S. (2020). How COVID-19 is taking gaming and esports to the next level. World Economic Forum. [Online] <https://www.weforum.org/agenda/2020/05/covid-19-taking-gaming-and-esports-next-level/>
4. Kellen, M. & Saxena, D. (2020). Calm my headspace: Motivations and barriers for adoption and usage of meditation apps during times of crisis. In *Proceedings of the 20th International Conference on Electronic Business (ICEB 2020)* (pp. 5–6). University of Hong Kong
5. Kumar, A., & Nayar, K. R. (2020). COVID 19 and its mental health consequences. *Journal of Mental Health*, p. 1–2. [ahead of print] <https://doi.org/10.1080/09638237.2020.1757052>.
6. Jones, C., Scholes, L., Johnson, D., Katsikitis, M., & Carras, M. C. (2014). Gaming well: Links between videogames and flourishing mental health. *Frontiers in Psychology*, 5, 260.
7. Biocca, F. (1992). Communication within virtual reality: Creating a space for research. *Journal of Communication*, 42(4), 5–22.
8. Borsci, S., Lawson, G., Jha, B., Burges, M., & Salanitri, D. (2016). Effectiveness of a multidevice 3D virtual environment application to train car service maintenance procedures. *Virtual Reality*, 20(1), 41–55.
9. Berg, L., & Vance, J. (2017). Industry use of virtual reality in product design and manufacturing: A survey. *Virtual Reality*, 21(1), 1–17.
10. Saxena, D., & Verma J. K. (2022). Recreating Reality: Classification of Computer-Assisted Environments. In Verma, J. K., & Paul, S. (Eds.), *Advances in Augmented Reality and Virtual Reality*. Springer Nature.
11. Coyle, J. R., & Thorson, E. (2001). The effects of progressive levels of interactivity and vividness in web marketing sites. *Journal of Advertising*, 30(3), 65–77.
12. Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4), 73–93.
13. Van Kerrebroeck, H., Brengman, M., & Willems, K. (2017). When brands come to life: Experimental research on the vividness effect of Virtual Reality in transformational marketing communications. *Virtual Reality*, 21(4), 177–191.
14. Steffen, J. H., Gaskin, J. E., Meservy, T. O., Jenkins, J. L., & Wolman, I. (2019). Framework of affordances for virtual reality and augmented reality. *Journal of Management Information Systems*, 36(3), 683–729.
15. Cole, D. A., Nick, E. A., & Pulliam, K. A. (2020). Are massively multiplayer online role-playing games healthy or not and why? Preliminary support for a compensatory social interaction model. *Computers in Human Behavior*, 102, 57–66.
16. Zero Latency (n.d.). Free-roam VR. [online] <https://zerolatenessvr.com/experience/>.
17. Keith, M. J., Anderson, G., Gaskin, J., & Dean, D. L. (2018). Team video gaming for team building: Effects on team performance. *AIS Transactions on Human-Computer Interaction*, 10(4), 205–231.
18. Nuangjumnong, T. (2016). The influences of online gaming on leadership development. In *Transactions on Computational Science XXVI* (pp. 142–160). Berlin, Heidelberg: Springer
19. Roettgers, J. (2020). The VR gaming centers of the future may not survive the crisis of today. Protocol, [online] <https://www.protocol.com/location-based-vr-covid-19>.
20. Zero Latency (2020). Zero Latency Dublin. [online] <https://twitter.com/ZeroLatencyDub>.

Application in Engineering

Computer Simulation of Physical Processes in an Electric Circuit with Nonlinear Inductance



Vita Ogar , Andrii Perekrest , Oleksii Kravets , and Oleksandr Bilyk

Abstract In modern society, the use of the latest technologies in any field of human activity is becoming increasingly important. The use of computerized equipment in laboratory classes for technical disciplines allows both to increase interest in learning the subject and to enhance the cognitive activity of the student. In most electrical devices with nonlinear elements, not only do phenomena occur that are fundamentally impossible in linear circuits, but also the principle of device operation is based on the circuit nonlinearity. The paper presents a computerized system that allows you to study the processes occurring in the AC circuit in the presence of a nonlinear element. As a result of the study, the user interface of a computerized complex, which reproduces an alternating current circuit with nonlinear inductance was developed. The universal LabVIEW package was used to build the mnemonic diagram, which contains a library of elements designed for the development of physical devices virtual interfaces and laboratory research facilities. The elements parameters that make up the circuit can be changed. The nonlinearity of the inductor is given in the form of an approximation dependence. As a result, on the basis of the created complex, with the help of computer modeling, the influence of nonlinearity degree of the element characteristic on the level of non-sinusoidality of the signal was determined. The developed complex can be used for laboratory workshops in the study of “Nonlinear AC circuits” in classes on “Electrical Engineering.”

Keywords Computerized complex · Physical processes · Electric circuit · Nonlinear element

1 Introduction

In the era of post-industrial society, the instrumental provision of information technology provides increased efficiency of processes in various sectors of the economy.

V. Ogar · A. Perekrest (✉) · O. Kravets · O. Bilyk
Kremenchuk Mykhailo Ostrohradskyi National University, Pershotravneva Street 20,
Kremenchuk 39600, Ukraine
e-mail: a.perekrest@icloud.com

The issues of using augmented and virtual reality technologies in the study of complex technical issues in electronics, electrical engineering, automation, etc., deserve special attention. Their use through new, innovative methods of implementing digital technologies provides a sense of user presence in objective reality with a high degree of realism (virtual reality) and provides additional opportunities for research planning and comparison with real processes (augmented reality).

The purpose of this work—is to develop a computer model of physical processes in an electric circuit with nonlinear inductance based on the study of methods of analysis, methods of software implementation and with direct implementation in the form of a computerized complex.

2 Characterization of Methods for the Analysis of Nonlinear Circuits with Steel

When solving most engineering problems, there is often a difficulty associated with the analysis of nonlinear circuits. Nonlinear circuits contain elements that can be described using constant coefficients, and the characteristics are nonlinear functions of one or more variables [1]. Since the mathematical apparatus for analyzing linear circuits is reduced to solving simple algebraic equations, researchers often neglect the nonlinearity of circuits. However, in such a consideration of the issue, important information is lost, and in some cases, the entire physical meaning of the phenomenon.

Currently, the following methods are used in the study of nonlinear circuits:

- methods of small parameter and conditional linearization [1]. For the calculations of nonlinear AC circuits, non-sinusoidal currents are replaced by equivalent sinusoidal ones, and for analysis, a complex calculation method is used, taking into account the nonlinear relationship between the effective values and the phases of the equivalent sinusoidal current and voltage. A particular case of the described method is the method of harmonic analysis, where in the calculations, it is assumed that the amplitudes of the harmonic components change [2]. In this case, there is no need to take into account the entire spectrum of harmonics associated with a change in amplitude. This simplification, in essence, is a replacement for the nonlinear dependence by the linear one, which is valid only for a certain value of the current and voltage amplitudes;
- method of analytical approximation of nonlinear characteristics [3]. The method consists of obtaining some approximation dependence of the characteristic. Depending on how well the approximation dependence is selected, the problem will be solved so easily;
- the method of piecewise linear approximation and fitting of linear solutions [3]. In this case, the nonlinear characteristic is replaced by a broken line, and the solutions obtained for each of its sections are added. This method is widely applicable for solving various problems;

- iterative method [2]. It is based on the principles of iteration, i.e., the approximate solution is refined by substituting it into the original equation until the specified accuracy is achieved;
- graphical method [2]. In this case, differential equations are transformed into nonlinear ones, which are solved using graphical constructions;
- method of successive intervals [3]. The essence of the method consists of replacing the differential equations with algebraic ones, which contain increments of the investigated quantities at the appropriate time intervals. As a result, complex dependencies are transformed into expressions, which are calculated using a computer.

Each of the methods mentioned has its own disadvantages inherent in the chosen assumptions. Therefore, when solving complex engineering problems, a multi-criteria analysis and the use of an integrated approach are necessary, i.e., the use of a certain set of methods to simplify the solution of the problem.

3 Determination of the Magnetization Curve of the Inductor with Saturation

Nonlinear circuits are classified by the type of nonlinearity as follows [4]:

- parametric or linear with variable parameters—these are electrical circuits described by linear differential equations with variable coefficients. Such a case takes place if the value of at least one of the parameters of the electrical circuit depends on time. In parametric circuits, the principle of superposition takes place;
- nonlinear circuits described by nonlinear differential equations. The equation of an electric circuit turns out to be nonlinear in the case when any nonlinear elements are used in the circuit, i.e., elements, the parameters of which depend on the current or voltage. One of the most important features of nonlinear circuits is that the principle of superposition is not fulfilled in them;
- nonlinear-parametric, which include circuits described by nonlinear equations with variable coefficients. Circuits are nonlinear parametric if they contain nonlinear and parametric elements. Their features: inapplicability of the principle of superposition, enrichment of the harmonic components of the response spectrum in comparison with the spectrum of input signals.

The characteristic of nonlinear inductance is set by the relationship between the flux linkage and the current of the coil called the Weber-ampere characteristic [5–7]:

$$\psi = f(i). \quad (1)$$

Unlike linear inductance, in which the relationship between flux linkage and current is linear, expression (1) is a nonlinear function.

Physically nonlinear inductance is a coil or several coils located on a closed ferromagnetic core.

If we take the magnetic field in the core to be uniform, then, taking into account the law of total current, the weber-ampere characteristic of the inductive element can be written as:

$$\psi = BS w = \psi(H l_{\mu} / w) = \psi(i), \quad (2)$$

where H —magnetic field strength,

l_{μ} —average magnetic line length,

w —number of coil turns.

From relation (2), it follows that the characteristic of nonlinear inductance is determined by the dependence of induction on the field strength: $B = f(H)$, i.e., the characteristic of magnetization of the ferromagnetic material (steel) of the core.

All known methods of analysis and calculation of nonlinear circuits can be subdivided into graphical and analytical. The advantages of analytical methods include the ability to analyze the problem in general form, and not for particular values of parameters, but their common disadvantage is that it is necessary to analytically express the characteristics of nonlinear resistances (this is not required in graphical methods), which is always associated with—sins. Analytical methods of calculation are proposed in [8, 9] and experimental methods of calculation in [10, 11]. But when solving the problem, special attention should be paid to the calculation of a nonlinear electric circuit by reducing it to a circuit with time-variable parameters, while the nonlinear inductance is replaced by inductance that is clearly time dependent [12].

Let's consider how such a transition is carried out. The voltage across the nonlinear inductance is:

$$u_L = \frac{d\psi}{dt} = \frac{d\psi}{di} \times \frac{di}{dt} = \frac{d}{dt} \{L(i) \times i(t)\}. \quad (3)$$

In (3) $L(i) = \frac{d\psi}{di}$ —dynamic differential inductance.

Let the current $i(t)$, flowing through a circuit with a nonlinear inductance changes according to a sinusoidal law in time: $i(t) = i_m \sin(\omega t)$ (Fig. 1). Knowing addition $L = f(i)$ and $i(t)$, on the same figure we will plot the dependence $L = f(t)$.

As seen from Fig. 1, $L(t)$ is a periodic function that can be expanded in a Fourier series into a constant component, as well as cosine and sine components of even harmonics:

$$L(t) = l_0 + l_{2a} \cos(2\omega t) + l_{2b} \sin(2\omega t) \\ + l_{4a} \cos(4\omega t) + l_{4b} \sin(4\omega t) + \dots$$

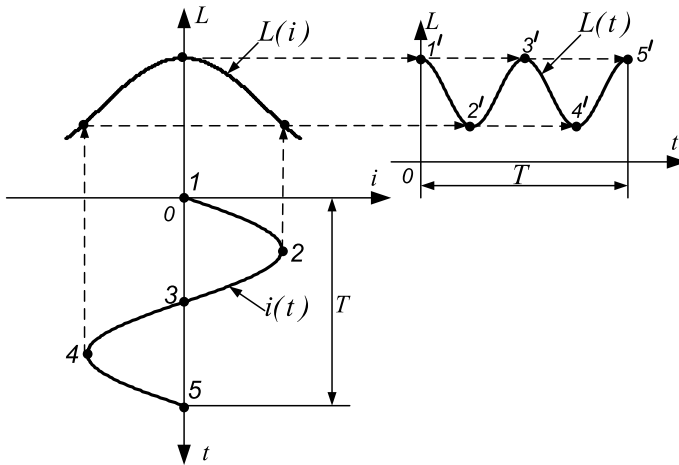


Fig. 1 Dependences of current and inductance on time

It should be noted that if the current curve is symmetrical about the time axis, then the frequency of the first harmonic is twice the frequency of the first harmonic of the current.

Analyzing the approaches to solving problems associated with finding the nonlinear characteristics of elements, special attention should be paid to the analysis of circuits using the energy method. Let us consider the formation of a mathematical apparatus based on the power balance when solving the problem of determining the dependence $L(I)$ coil based on equivalent circuit analysis (Fig. 2), where r —active resistance, x_{ns} —nonlinear inductive reactance.

Let us assume that in the equivalent circuit (Fig. 2) is the inductive reactance x_{ns} , equal to value a_0 in an unsaturated state, and the phenomenon of saturation will be associated with the appearance of an additional e.m.f in the loop $e(t)$, which allows to talk about a decrease in inductive resistance x_{ns} circuit, which is observed during saturation phenomenon.

In view of the above, we turn to the scheme presented in Fig. 3.

Let us take the dependences of voltage, current and power in the following form, respectively:

Fig. 2 Inductor coil equivalent circuit

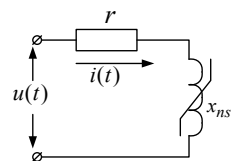
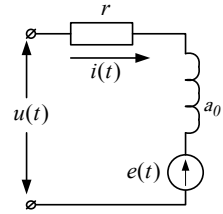


Fig. 3 Equivalent circuit for an inductor with e.m.f. $e(t)$ in the circuit



$$u(t) = \sum_{n=1}^{\infty} (u_{na} \cos(n\omega t) + u_{nb} \sin(n\omega t)), \tag{4}$$

$$i(t) = \sum_{m=1}^{\infty} (i_{ma} \cos(m\omega t) + i_{mb} \sin(m\omega t)), \tag{5}$$

$$p(t) = p_0 + \sum_{k=1}^{\infty} (p_{ka} \cos(k\omega t) + p_{kb} \sin(k\omega t)), \tag{6}$$

where the subscript “0” means the constant component of the signal;

“b”—sine component (Fig. 4);

n, m, k —harmonics of voltage, current and power, respectively.

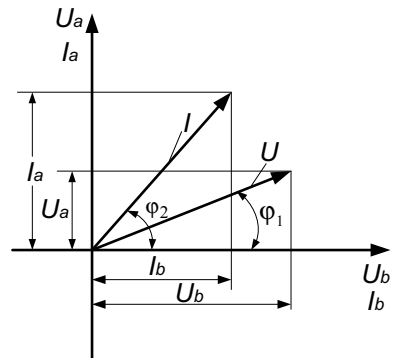
We represent the inductance dependence as:

$$l(t) = a_0 + l_{\text{var}}(t), \tag{7}$$

where a_0 —unsaturated inductance;

$l_{\text{var}}(t)$ —time dependence of the variable component of inductance.

Fig. 4 Vector diagram of voltage and current



In connection with the above, the equivalent circuit can be converted to the form (Fig. 5):

Due to the fact that the variable component of the inductance at the initial moment of time ($t = 0$) must be equal to zero, then it can be represented by the trigonometric polynomial of the following form:

$$l_{var}(t) = - \sum_{p=1}^{\infty} a_{2pa} + \sum_{p=1}^{\infty} [a_{2pa} \cos(2p\omega t) + a_{2pb} \sin(2p\omega t)], \quad (8)$$

where p —harmonic inductance number.

Instantaneous power of each of the circuit elements (Fig. 5) will be, respectively, equal to:

- source:

$$p_i(t) = u(t)i(t); \quad (9)$$

- active resistance:

$$p_r(t) = i^2(t)r; \quad (10)$$

- nonlinear inductance:

$$p_L(t) = i(t) \left\{ i(t) \frac{dl(t)}{dt} + l(t) \frac{di(t)}{dt} \right\}. \quad (11)$$

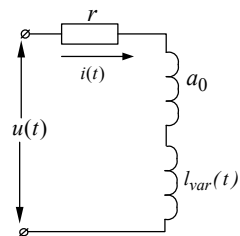
On the other hand, the instantaneous power of the source is equal to the sum of the instantaneous powers of each of the elements included in the circuit, i.e.,

$$p_i(t) = p_r(t) + p_L(t). \quad (12)$$

From expression (12), the power on nonlinear inductance $p_L(t)$ will be:

$$p_L(t) = p_i(t) - p_r(t). \quad (13)$$

Fig. 5 Equivalent circuit of nonlinear inductance, taking into account its decomposition into constant and variable components



Substituting into (13) the dependence describing the power on the nonlinear inductance, we obtain:

$$p_i(t) - p_r(t) = i(t) \cdot \left\{ i(t) \cdot \frac{dl(t)}{dt} + l(t) \cdot \frac{di(t)}{dt} \right\}. \quad (14)$$

Taking into account expression (7), according to which the nonlinear inductance can be decomposed into constant and variable components (Fig. 5), dependence (14) can be represented in the following form:

$$p_i(t) - p_r(t) - \Delta p_L(t) = i(t) \left\{ i(t) \frac{dl_{\text{var}}(t)}{dt} + l_{\text{var}}(t) \frac{di(t)}{dt} \right\}, \quad (15)$$

where $\Delta p_L(t) = a_0 \frac{di(t)}{dt} i(t)$ —power of the constant component of inductance.

Based on expressions (9)–(10), dependence (15) will be:

$$u(t)i(t) - i^2(t)r - a_0 \frac{di(t)}{dt} i(t) = i(t) \left\{ i(t) \frac{dl_{\text{var}}(t)}{dt} + l_{\text{var}}(t) \frac{di(t)}{dt} \right\}. \quad (16)$$

For the convenience of further mathematical calculations, we introduce the following notation for the power on a variable inductance as $p(t)$:

$$p(t) = u(t)i(t) - i^2(t)r - a_0 \frac{di(t)}{dt} i(t), \quad (17)$$

then expression (15) takes the final form:

$$p(t) = i(t) \left\{ i(t) \frac{dl_{\text{var}}(t)}{dt} + l_{\text{var}}(t) \frac{di(t)}{dt} \right\}. \quad (18)$$

Taking into account expressions (5), (8), on the basis of dependence (18), we compose a general expression for finding the approximation coefficients of variable inductance.

$$\begin{aligned} & p_0 + p_{ka} \cos(k\omega t) + p_{kb} \sin(k\omega t) \\ &= \omega \sum_{m=1}^{\infty} \sum_{p=1}^{\infty} \left\{ \frac{1}{2} m l_{2pa} i_{ma}^2 \sin(2m\omega t) \right. \\ & \quad - m l_{2pa} i_{mb} i_{ma} \cos(2m\omega t) - \frac{1}{2} m l_{2pa} i_{mb}^2 \sin(2m\omega t) \\ & \quad - \frac{1}{4} m l_{2pa} i_{ma}^2 (\sin([2m - 2p]\omega t) + \sin([2m + 2p]\omega t)) \\ & \quad \left. - \frac{1}{2} m l_{2pa} i_{mb} i_{ma} (\cos([2m + 2p]\omega t) + \cos([2p - 2m]\omega t)) \right\} \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{4}ml_{2pa}i_{mb}^2(\sin([2m - 2p]\omega t) + \sin([2m + 2p]\omega t)) \\
& - \frac{1}{4}ml_{2pb}i_{ma}^2(\cos([2p - 2m]\omega t) - \cos([2p + 2m]\omega t)) \\
& + \frac{1}{2}ml_{2pb}i_{mb}i_{ma}(\sin([2p - 2m]\omega t) + \sin([2p + 2m]\omega t)) \\
& + \frac{1}{4}ml_{2pb}i_{mb}^2(\cos([2p - 2m]\omega t) - \cos([2p + 2m]\omega t)) \Big\} \\
& + \omega \sum_{m=1}^{\infty} \sum_{p=1}^{\infty} \left\{ -pl_{2pa}i_{ma}^2 \sin(2p\omega t) + pl_{2pb}i_{mb}^2 \cos(2p\omega t) \right. \\
& - pl_{2pa}i_{mb}^2 \sin(2p\omega t) + l_{2pb}i_{mb}^2 \cos(2p\omega t) \\
& - \frac{1}{2}pl_{2pa}i_{ma}^2(\sin([2p - 2m]\omega t) + \sin([2p + 2m]\omega t)) \\
& + \frac{1}{2}pl_{2pb}i_{ma}^2(\cos([2m - 2p]\omega t) + \cos([2m + 2p]\omega t)) \\
& - pl_{2pa}i_{ma}i_{mb}(\cos([2m - 2p]\omega t) + \cos([2m + 2p]\omega t)) \\
& + pl_{2pb}i_{ma}i_{mb}(\sin([2m - 2p]\omega t) + \sin([2m + 2p]\omega t)) \\
& + \frac{1}{2}pl_{2pa}i_{mb}^2(\sin([2p - 2m]\omega t) + \sin([2p + 2m]\omega t)) \\
& \left. + \frac{1}{2}pl_{2pb}i_{mb}^2(\cos([2m - 2p]\omega t) + \cos([2m + 2p]\omega t)) \right\}. \quad (19)
\end{aligned}$$

The system of equations for finding the harmonic components of the variable inductance $l_{\text{var}}(t)$ we will form on the basis of the following provisions:

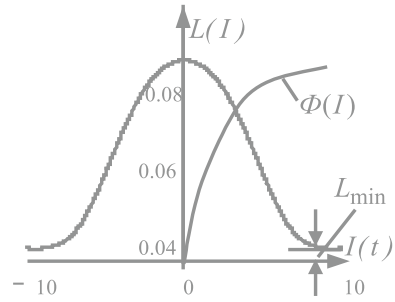
- the constant component of the power is formed from the terms of expression (19) that do not have a factor “cos” or “sin”;
- the cosine component is formed from the terms of expression (19) with the factor “cos” when the frequencies of the right and left sides are equal: $k = |2m \pm 2p|$, $k = 2m$, $k = 2p$;
- the sine component is formed from the terms of expression (19) with the factor “sin” when the frequencies of the right and left sides are equal $k = |2m \pm 2p|$, $k = 2m$, $k = 2p$.

For inductor with active resistance $r = 0.312 \Omega$ and inductance in an unsaturated state $a_0 = 0.0874 \text{ H}$ calculate the nonlinear characteristic $L'(i)$. The correctness of the method will be judged by the coincidence of the calculated nonlinear dependence $L'(i)$ with a given $L(i)$. The analysis of the processes occurring in a circuit containing a nonlinear element will be carried out taking into account the circuits in Figs. 2, 3 and 5.

Let us assume that the given characteristic $L(i)$ (for the equivalent circuit Fig. 2) has the form (Fig. 6).

Dependence $L(i)$ is represented by a power polynomial of the form:

Fig. 6 Specified nonlinear dependence of inductance on current



$$L(i) = a_0 + a_1i^2(t) + a_2i^4(t) + a_3i^6(t) + a_4i^8(t) + a_5i^{10}(t), \quad (20)$$

where a_0 – a_5 —coefficients equal to $a_0 = 0.0874$;

For the equivalent circuit shown in Figs. 2 and 3, the voltage on her clamps will be:

$$u(t) = i(t)r + L(i)\frac{di(t)}{dt} + i(t)\frac{dL(i)}{dt} \quad (21)$$

From expression (21), the analytical dependence for the current flowing through the circuit, taking into account the specified nonlinear dependence of inductance on current (20), will be:

$$\frac{di(t)}{dt} = \frac{u(t) - i(t) \cdot r}{L(i) + 2i^2(t) \cdot [a_1 + 2a_2i^2(t) + 3a_3i^3(t) + 4a_4i^4(t) + 5a_5i^5(t)]}. \quad (22)$$

For certain voltage levels, we calculate the current flowing through the circuit. So, for voltage, the dependence of the current obtained by simulating a circuit with inductance (20) has the form (Fig. 7).

It is known [13] that the presence of nonlinearity leads to the appearance of higher harmonics. Using the Fourier analysis, the harmonic components of the current necessary for the calculation were obtained (Fig. 8).

Fig. 7 Dependence of the current flowing through the circuit

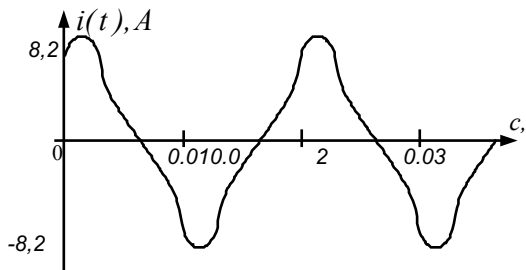
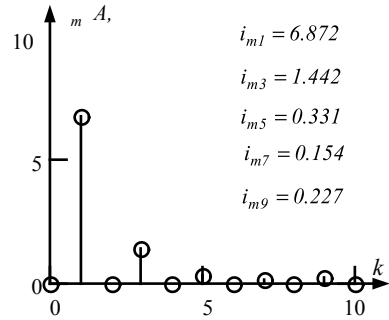


Fig. 8 Harmonic composition of current



Based on expressions (9), (10) and taking into account the dependence of the current flowing through the circuit, the power balance equation was compiled and the harmonic components of the power on the nonlinear inductance of the circuit were calculated.

Based on Eq. (19), we will compose a system of equations for finding the approximation coefficients taking into account the first, third, fifth, seventh, ninth harmonics of the current, the second, fourth, sixth and eighth harmonics of inductance.

The solution of the system of equations is the approximation coefficients of the dependence.

In Figs. 9 and 10, the dependence of the variable inductance component as well as the total inductance component on time is shown.

Fig. 9 Time dependence of the variable component of inductance

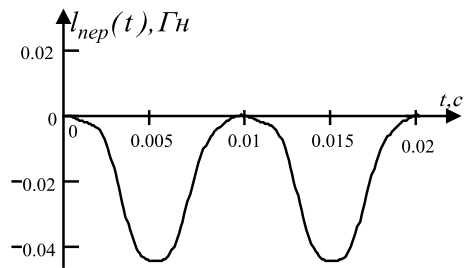


Fig. 10 Dependence of total inductance on time

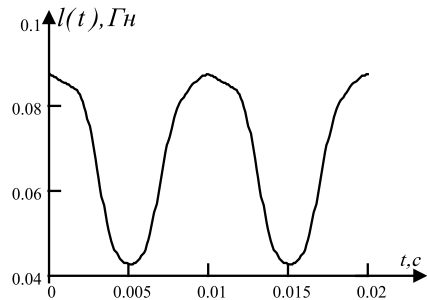
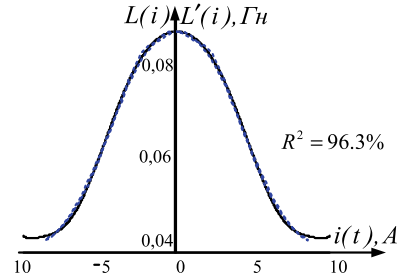


Fig. 11 Dependence of inductance on current



To check the adequacy of the model, an estimate of the coefficient of determination was used, which shows the percentage of explanation for the share of variation in this model and is about 96%. For the given dependence of inductance on current, the coefficient of determination is 96.3% taking into account five harmonics of current (1, 3, 5, 7, 9) and four harmonics of inductance (2, 4, 6, 8).

The coincidence of the given and obtained nonlinear characteristics (Fig. 11) testifies to the correctness of the proposed mathematical apparatus based on the power balance equations in determining the nonlinear dependence of the inductive element.

4 A Computerized Complex for Studying the Characteristics of an Electrical Circuit Using the LabVIEW Application Package

Analysis of existing software has shown that the development of a computerized system for studying the characteristics of an AC circuit with a nonlinear element is possible on the basis of universal software packages used in a wide range of subject areas (e.g., MATLAB, NI LabVIEW, etc.). Such universal packages contain a huge number of libraries of elements designed for the development of virtual interfaces of physical devices and laboratory research facilities.

The main criteria for choosing a software development environment were: reduction of development time, minimization of difficulties associated with the debugging and modeling of mathematical algorithms. According to these criteria, the most convenient graphical programming environment is NI LabVIEW.

LabVIEW is based on the graphical programming language G [14]. In addition to the programming capability itself, LabVIEW provides the user with a wide range of tools and libraries: from interactive setup wizards and user interfaces to built-in compiler, linker and debugging tools.

LabVIEW implements the concept of G graphical programming, so the source code is a block diagram (connected icons of language elements), which is then compiled into machine code. In this regard, the main advantage of programming in G is the ability to intuitively use graphical tools.

In addition to the direct choice of software environment, it is necessary to study [10, 13] and take into account [11, 15] the physical processes occurring in the consumer and take into account best practices in introducing innovative technologies in laboratory workshops for technical students [16, 17], the use of virtual designers hydraulic transport systems [18] according to certain automatic procedures [19], the use of information and analytical solutions to improve the efficiency of energy management in municipalities [20, 21] based on the analysis of primary data [22] using different efficiency criteria [23–27] and optimization [28, 29]. The study of [10, 13] is important and taking into account [11, 15] of physical processes occurring in the consumer.

To study the transients occurring in an electrical circuit containing an inductive element with a nonlinear characteristic, which is determined by the method proposed in the study, in LabVIEW a mnemonic of a computerized system was developed (Fig. 12).

The mnemonic diagram contains:

- electrical circuit, which includes a source of alternating voltage (Fig. 13) with the possibility of changing the amplitude value and frequency of the voltage, a resistor (Fig. 14), the slider of which allows to change the value of resistance,

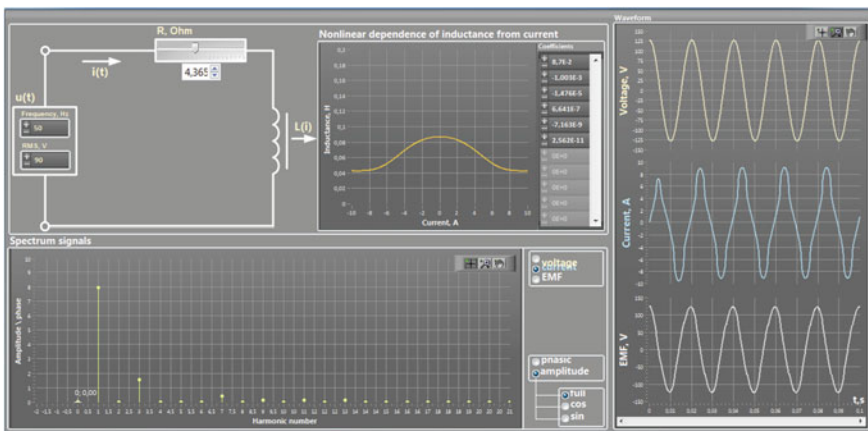


Fig. 12 Mnemonic diagram of the computerized system

Fig. 13 Source of the alternating voltage

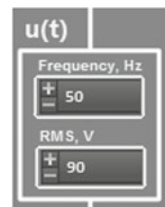


Fig. 14 Resistance

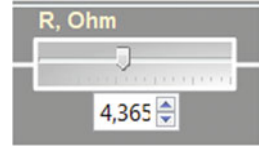


Fig. 15 Nonlinear inductivity



nonlinear inductance (Fig. 15); nonlinear dependence of inductance on current (Fig. 16);

- characteristics of current (Fig. 17), voltage (Fig. 18), e.m.f. (Fig. 19), changing over time;
- harmonic analysis of signals (Fig. 20).

In the block “Coefficients” (Fig. 21), it is possible to set different kind of nonlinear dependence of inductance on current with use of approximation coefficients.

In the computerized complex, the possibility of problem of different kind of approximation dependence of nonlinear characteristic is implemented. Analysis of

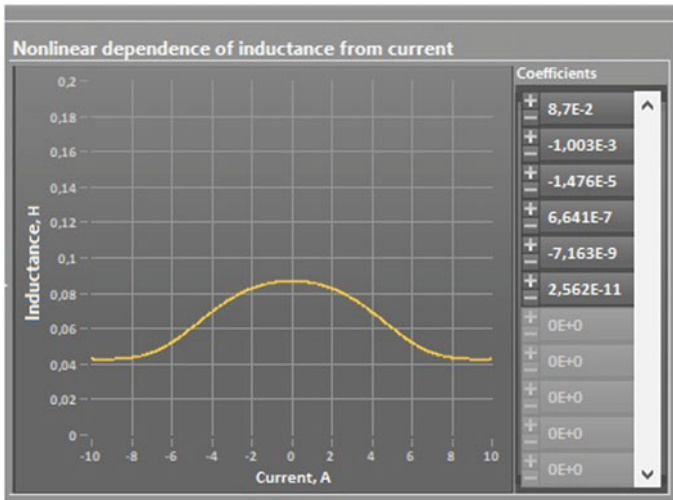


Fig. 16 Nonlinear dependence of inductance on current

Fig. 17 Time dependence of current

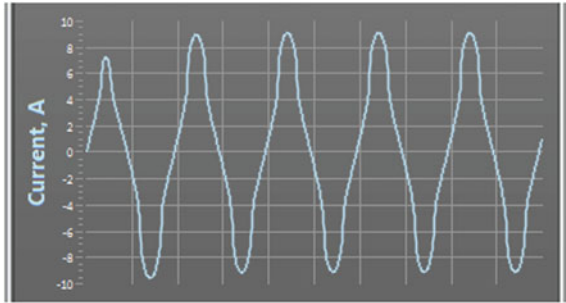
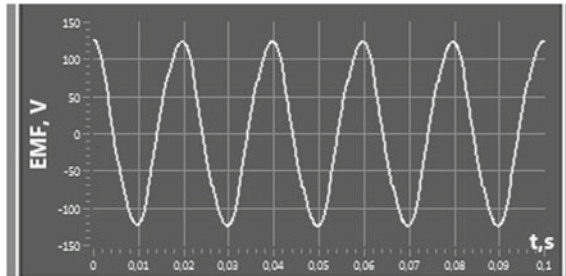


Fig. 18 Time dependence of voltage

Fig. 19 Time dependence of EMF



the harmonic components of the signal can be analyzed for both cosine and sine components.

In Figs. 22 and 23, a block diagram of the computerized system user interface developed in the G language is shown.

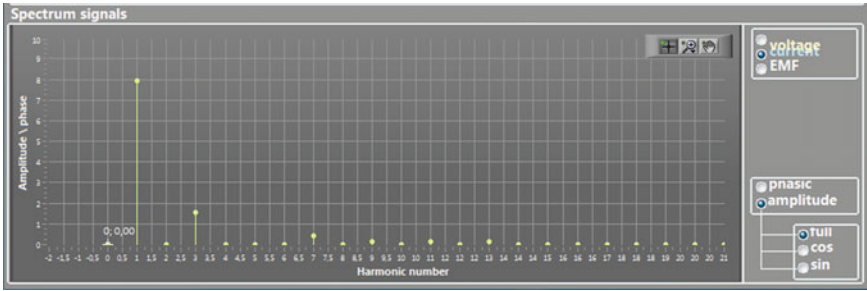


Fig. 20 Harmonic analysis of signals

Fig. 21 Block “coefficients”

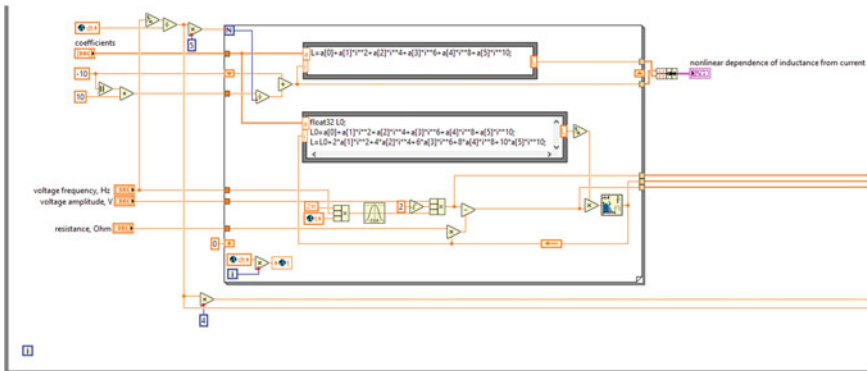


Fig. 22 Fragment of a block diagram for determining current and electromotive force signals in an electrical circuit with nonlinear inductance

5 Conclusions

The study presents a computerized laboratory complex that allows to study the processes occurring in a circle with a nonlinear element.

The paper proposes a method for calculating the nonlinear dependence of inductance on current, which is the basis of the mathematical apparatus for the analysis of

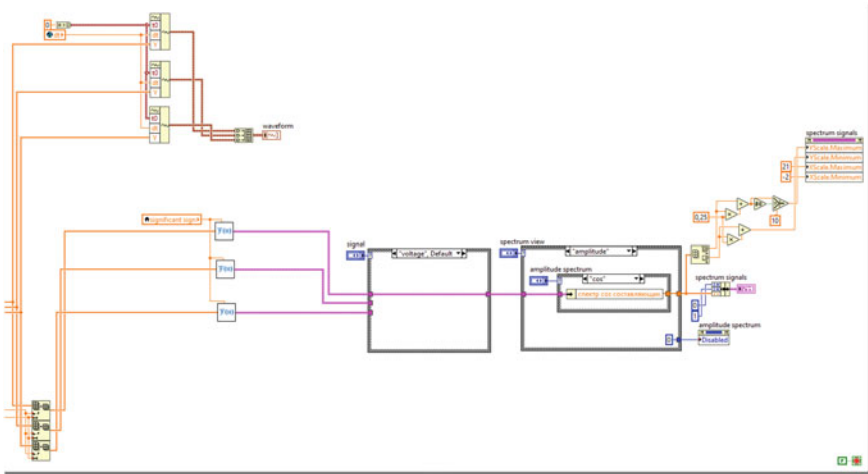


Fig. 23 Fragment of a block diagram for determining and displaying the spectral composition of current and electromotive force signals in an electric circuit with nonlinear inductance

circuits with nonlinear elements. Its main feature is to take into account the harmonic components of the current that appear in the circuit as a result of the presence of a nonlinear element.

Analysis of the characteristics of the alternating current circuit with nonlinear inductance showed that the nonlinearity of the element affects the current flowing in the circuit. As a result of the presence of a nonlinear element in the circuit, a non-sinusoidal current begins to flow. The level of non-sinusoidal signals depends on the degree of nonlinearity of the element characteristics.

The developed complex can be used for laboratory workshops during the study of disciplines of electrical engineering.

References

1. Shidlovska, N. (1992). *Analysis of nonlinear electric circuits by the method of small parameter* (p. 192). Kyiv: Euroindex.
2. Cunningham, W. (1962). *Introduction to the theory of nonlinear systems* (p. 456). M.-L.: Gosenergoizdat.
3. Zeveke, G. V. (1975). *Fundamentals of circuit theory: Textbook (for universities)*. In G. V. Zeveke, P. A. Ionkin, A. V. Netushil, S. V. Strakhov (Eds.) (4th ed., p. 752). M.: Energiya
4. Martynenko, V. A. (2003). Analysis of a nonlinear chain with steel using energy criteria. *Bulletin of the Kremenchutsk State Polytechnic University: Science Pratsi*, 2(19), 167–170. Kremenchuk: KDPU
5. Matkhanov, P. N. (1977). Fundamentals of analysis of electrical circuits. Nonlinear chains (p. 260). M.: Higher school.
6. Voldek, A. I., & Popov, V. V. (2008). Electric machines. *AC Machines: A Textbook for High Schools* (p. 350). SPb: Peter.

7. Zagirnyak, M., Ogar, V., & Chenchevoi, V. (2014). Analysis of induction motors features taking into account change of iron properties. *Acta Technica CSAV (Ceskoslovensk Akademie Ved)*, 59(1), 25–47.
8. d' Aquino, M., Serpico, C., Visone, C., & Adly, A. A. (2003). A new vector model of magnetic hysteresis based on a novel class of play hysterons. *IEEE Transactions on magnetics*, 39(5), 2537–2539.
9. Machado, V. M. (2002). Instantaneous magnetic field in non-linear inductive circuits. In *XVII Symposium Electromagnetic Phenomena in Non-Linear Circuits* (pp. 51–54). Porman.
10. Zagirnyak, M., Kalinov, A., Ogar, V. & Lotous, V. (2017). Experimental assessment of the accuracy of the method for determination the power on an induction motor shaft. In *2017 18th International Conference on Computational Problems of Electrical Engineering (CPEE)* (pp. 1–4). Kutna Hora. <https://doi.org/10.1109/CPEE.2017.8093054>.
11. Zagirnyak, M., Chenchevoi, V., Ogar, V., Chencheva, O., & Yatsiuk, R. (2020). Refining induction machine characteristics at high saturation of steel. *Przegląd Elektrotechniczny*, 96(11), 119–123. <https://doi.org/10.15199/48.2020.11.24>
12. Bessonov, L. A. (1964). *Nonlinear electrical circuits* (p. 430). M.: Higher School
13. Chenchevoi, V., Zachepa, I., Ogar, V., & Chencheva, O. (2019). Research on harmonic composition of voltage and current of induction generator with high saturation magnetic system. In *2019 IEEE International Conference on Modern Electrical and Energy Systems (MEES)* (pp. 170–173). Kremenchuk, Ukraine. <https://doi.org/10.1109/MEES.2019.8896498>.
14. Peich, L. I. (2004). Labview for beginners and specialist. In L. I. Peich, D. A. Tochilin, B. P. Pollak (Eds.) (p. 384). Hotline: Telecom.
15. Zagirnyak, M., Ogar, V., Chenchevoi, V. (2014). Analysis of induction motors features taking into account change of iron properties. *Acta Technica CSAV (Ceskoslovensk Akademie Ved)*, 59(1), 25–47
16. Zagirnyak, M., Serhiienko, S., & Chorni, O. (2017). Innovative technologies in laboratory workshop for students of technical specialties. In *Proceedings of 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)* (pp. 1216–1220). Kyiv, Ukraine, IEEE Catalog Number: CFP17K03-USB.
17. Chorni, O., Serhiienko, S., Yudyna, A., & Sydorenko, V. (2017). The analysis of the process of the laboratory practicum fulfillment and the assessment of its efficiency on the basis of the distance function. In *Proceedings of the 2017 IEEE International Conference on Modern Electrical and Energy System (MEES)* (pp. 328–331). <https://doi.org/10.1109/MEES.2017.8248924>.
18. Zagirnyak, M., Korenkov, E., Kravets, A., & Korenkova, T. (2013). Virtual constructor used for research of hydrotransport system operation conditions. In *Proceedings of International IEEE Conference EUROCON 2013* (pp. 1193–1200). Zagreb, Croatia. <https://doi.org/10.1109/EUROCON.2013.6625132>.
19. Zagirnyak, M., Kovalchuk, V., Korenkova, T. (2018) The automation of the procedure of the electrohydraulic complex power harmonic analysis. *Przegląd Elektrotechniczny*, 1, 1–4.
20. Perekrest, A., Shendryk, V., Pijarski, P., Parfenenko, Y., & Shendryk, S. (2017). Complex information and technical solutions for energy management of municipal energetics. In *Proceeding SPIE 10445, Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments* (p. 1044567). <https://doi.org/10.1117/12.2280962>.
21. Zagirnyak, M., Perekrest, A., Ogar, V., Chebotarova, Y., & Mur, O. (2021). Segmentation of heat energy consumers based on data on daily power consumption, *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, 2, 88–96.
22. Perekrest, A., Chorni, O., Mur, O., Kuznetsov, V., Kuznetsova, Y., & Nikolenko, A. (2018). Preparation and preliminary analysis of data on energy consumption by municipal buildings. *Eastern-European Journal of Enterprise Technologies*, 6(896), 32–42. <https://doi.org/10.15587/1729-4061.2018.147485>
23. Perekrest, A., Konokh, I., & Kushch-Zhyrko, M. (2019). Administrative buildings heating automatic control based on maximum efficiency criterion. In *2019 IEEE International Conference on Modern Electrical and Energy Systems (MEES)* (pp. 202–205). Kremenchuk, Ukraine. <https://doi.org/10.1109/MEES.2019.8896517>

24. Perekrest, A., Kushch-Zhyrko, M., Ogar, V., Zalunina, O., Bilyk, O. & Chebotarova, Y. (2020). Key performance indicators assessment methodology principles adaptation for heating systems of administrative and residential buildings. In *2020 IEEE Problems of Automated Electrodrive. Theory and Practice (PAEP)* (pp. 82–86). Kremenchuk, Ukraine. <https://doi.org/10.1109/PAEP49887.2020.9240784>.
25. Perekrest, A., Chebotarova, Y., & Herasimenko, O. (2017). Information and analytical set of tools for assessing efficiency of the civil buildings heating modernization. In *2017 International Conference on Modern Electrical and Energy Systems (MEES)* (pp. 216–219). Kremenchuk. <https://doi.org/10.1109/MEES.2017.8248893>.
26. Perekrest, A., Chebotarova, A. Y., Al-Issa, H. A. (2019). Principles of designing and functioning of the expert system of assessing the efficiency of introducing energy conservation measures. In *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering, (UKRCON)* (pp. 871–875). Lviv. <https://doi.org/10.1109/UKRCON.2019.8879825>.
27. Chebotarova, Y., Perekrest, A., & Ogar, V. (2019). Comparative analysis of efficiency energy saving solutions implemented in the buildings. In *2019 IEEE International Conference on Modern Electrical and Energy Systems (MEES)* (pp. 434–437). Kremenchuk, Ukraine. <https://doi.org/10.1109/MEES.2019.8896691>.
28. Zalunina, O., Perekrest, A., Ogar, V., & Zbyrannyk, O. (2019). Formation of informational support in construction for the implementation of energy saving measures. In *2019 IEEE International Conference on Modern Electrical and Energy Systems (MEES)* (pp. 526–529). Kremenchuk, Ukraine. <https://doi.org/10.1109/MEES.2019.8896622>.
29. Zalunina, O., Kasych, A., Ogar, V., Perekrest, A., Serhiienko, S., & Kushch-Zhyrko, M. (2020). Energy system control optimization criterion development. In *2020 IEEE Problems of Automated Electrodrive. Theory and Practice (PAEP)* (pp. 191–196). Kremenchuk, Ukraine. <https://doi.org/10.1109/PAEP49887.2020.9240869>.

A Comparative Motion Study of Mated Gears on AutoCAD and SOLIDWORKS



Parth Patpatiya

Abstract Motion is a tool for exploring the kinematic and dynamic performance of a mechanical system. For the creation of motion of assembled entities, motion is fully incorporated and embedded in CAD software as an add-in module. This paper acquaints us with the multidisciplinary comparison of motion study performed on two mated gears of similar aspects on AutoCAD and SOLIDWORKS designed as a part model. During the case study as a part of product development, two solid models of gears have been mated on both the software and a motion study is performed. Modeling has been done keeping the designing parameters the same on both the software. AutoCAD involves modeling and assembling on the same platform, whereas SOLIDWORKS generates two separate files for designing and assembly. During the motion study, a series of commands have been executed to study the motion. Rscript command has been utilized for generating and running the loop in AutoCAD, whereas motion manager is utilized for creating and playing the animation in SOLIDWORKS. A 3D modeling comparison matrix has been developed which compares the looping feature of AutoCAD and SOLIDWORKS based on multiple disciplines. Further, quantitative comparison has been made for comparing the effectiveness of the process of product modeling to motion. The developed comparison may channelize the user to redesign the products and obtain a shorter and faster lead time to analyze and develop the motion.

Keywords AutoCAD · SOLIDWORKS · Motion analysis

1 Introduction

The beginning of the popular CAD systems, today, is widely accepted throughout the industry can be traced to the late twentieth-century when PRONTO, the very first commercial numerical control programming system was developed by Dr. Patrick J. Hanratty. The first CAD systems served as mere replacements of drawing boards,

P. Patpatiya (✉)

School of Automation, Banasthali Vidyapith, Rajasthan 304022, India

but gradually, CAD software and hardware became affordable, user friendly, and also grew in functionality. Syntha Vision by MAGI in 1969 introduced the enhanced 3D capabilities of CAD systems and in 1989 CAD's based on parametric engines by T-Flex made modeling more intuitive and precise. Over years, the development of open-source CAD has revolutionized the designing industry, and today, many commercial and proprietary programs have their strong open-source alternatives. Yet currently, there is no viable strong open-source competitor to commercial CAD systems like AutoCAD or SOLIDWORKS. AutoCAD, today, available in thirteen languages and usable on any device was introduced in 1982 as AutoCAD 360 and AutoCAD LT; the former being used for 2D, 3D, and isometric view, whereas the latter only for 2D. Features such as cloud storage connectivity, enhanced DWG compare, purge redesign, and more include industry-specific features and intelligent objects. AutoCAD is quite famous in all fields of design. SOLIDWORKS, the most popular CAD software according to the Sheffield Telegraph released its first product SOLIDWORKS 95 in Nov 1995. It is a solid modeler and utilizes the parametric features-based approach with additional mating features such as gear and CAM follower mates. For communicating design intents and in the transmission of knowledge, 3D modeling has proven to be a method superior to others. One of the main methods [1] prevailing at present for the improvement in the exactness of gear production is the beginning of electronic technologies into their design and manufacture. The 3D model is the most valuable design method and geometrically accurate enables us to make strength and kinetic calculations and to evaluate and check operations of the future mechanisms. Generation of solid models of the real body [2] in different variants is proposed using CAD systems such as AutoCAD and SOLIDWORKS. The generation of solid 3D models of equal variants is possible using both CAD systems. The AutoCAD package enables us to form a realistic model of the gear set. Unfortunately, the model did not fit in the initial three (the highest) exactness degrees. That is the reason [3–5] why the modeling was performed in the SOLIDWORKS via 3D parameterization. Wireframe, surface solid models are types of design layouts. Real body models are generated using the 3D modeling methods—solid modeling and feature parametric modeling [6]. Complex solid shapes are easier for generation and reduction as compared to the wireframe and solid models. Hence, solid modeling has been improved in modern CAD systems. Yet when huge files are created [7] in CAD, certain problems are encountered by the user while exporting these files. Creation of 3D solid models from 2D vector [8], drawing using the appropriate commands increases the file size manifolds. Therefore, certain commands [9] in AutoCAD files like AUDIT, RECOVERALL, -PURGE, OVERKILL, and WBLOCK are used to reduce the file size. The 3D model is generated [10–12] using the 2D vectors recorded in a file and also the experimental verification of the 3D model is generated from the 2D vector record.

Motion study not only eliminates the wastage of time and labor but also standardizes the method obtained after conducting the study. Every software developed aims at enhancing the efficiency of our work. Thus to determine the optional method to perform a task, the motion study is compared. This paper compares the total coefficients of change during the process of product modeling to motion while modeling using SOLIDWORKS and AutoCAD. A 3D model of the gear is developed which corresponds to the third degree (the highest) of accuracy in the SOLIDWORKS package. The method of product modeling was split into four phases, with each stage analyzed, and its complexity illuminated. Further, a matrix comparing the looping capabilities of AutoCAD and SOLIDWORKS across multiple disciplines has been established evaluating the effectiveness of the product modeling method in reference to motion.

2 Problem Description

The objective of this work is to analyze the motion study performed on two mated gears of similar aspects on AutoCAD and SOLIDWORKS designed as a part model. A quantitative comparison is done to predict the effectiveness of the process of product modeling to motion. The study is divided into several steps involving part modeling, designing of the gear, interlocking of gears, and study.

3 Methodology

The gears are modeled on AutoCAD and SOLIDWORKS using various part modeling both the softwares.

3.1 *AutoCAD*

We open a new drawing Window (Fig. 1) in AutoCAD and draw a circle. Then, we draw a tooth on one of its quadrants and polar array it (Fig. 2). We trim the extra lines and curves using the TRIM command and then we join all the required curves and lines (Fig. 3). We draw a circle at the center and then extrude both the geometries (Fig. 5). Subtracting the center circle from the outer geometry, we make another copy of the gear, and using rotate and move commands we interlock their teeth (Fig. 6). Grouping the two gears into different groups using the GROUP (Fig. 7) command and naming them differently makes the work easier. We open Notepad in AutoCAD and write the Rscript to perform the desired operation (Fig. 8). R is an integrated suite of software facilities for data manipulation, calculation, and graphical display. In addition to an effective data handling and storage facility, it also produces a large,

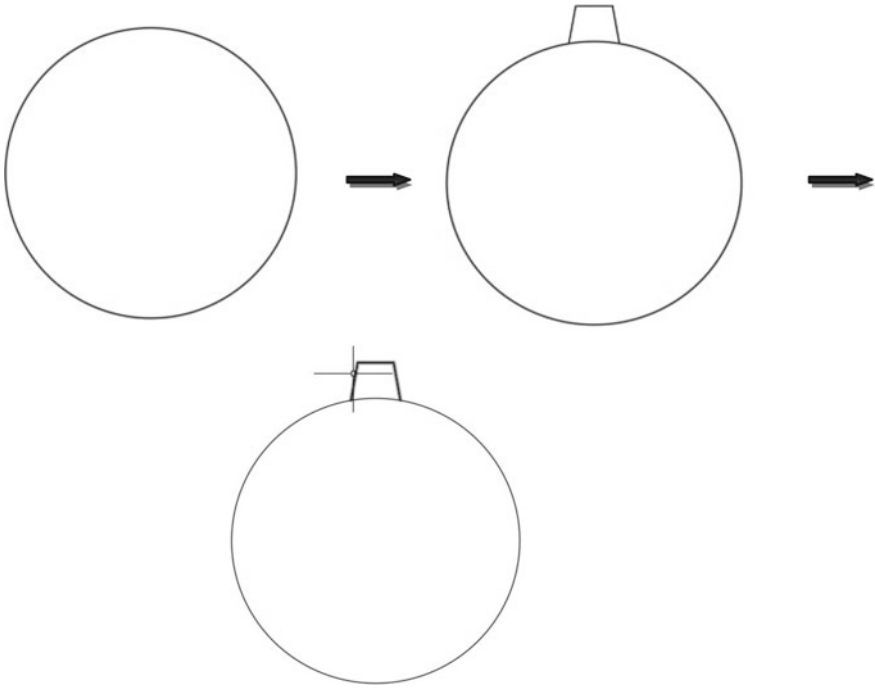


Fig. 1 Tooth on the quadrant of the circle

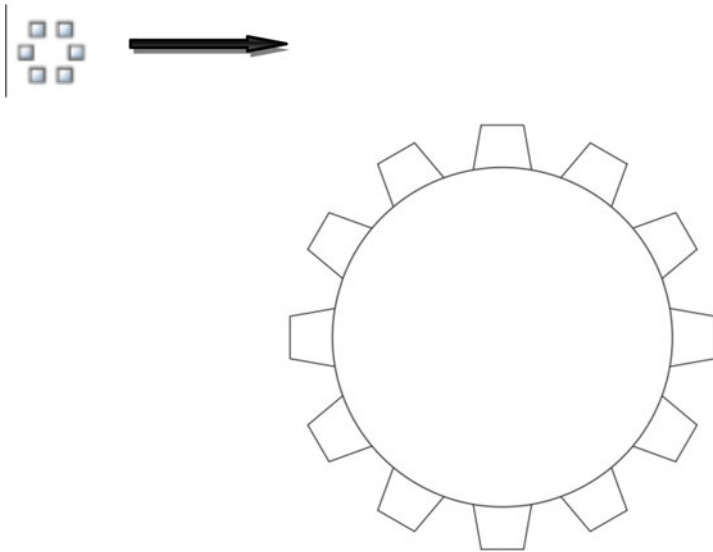


Fig. 2 Polar array of tooth

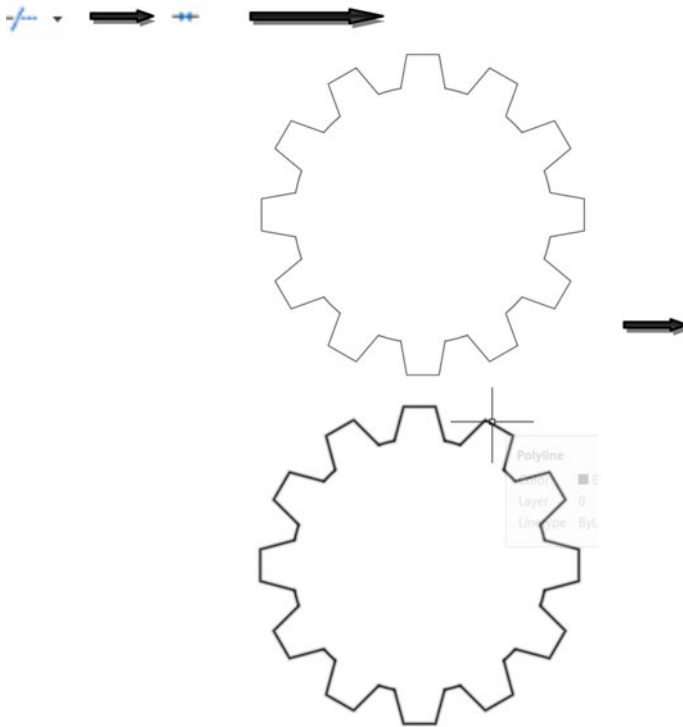


Fig. 3 Trim and join

coherent, and integrated collection of intermediate tools for data analysis. When performing a long analysis, it is better to type it into a script. Rscript is a series of commands that can be executed at one go and saves a lot of time and can be prepared on any text editor. Using the ID command, we find the coordinates of the center (Fig. 9) of gear. The coordinates are mentioned in Rscript as the rotating point around which the gears will rotate. We save the Rscript with a .scr extension (Fig. 10). Choosing the Manage option from the control panel we select the Run Script option (Fig. 11). A window appears and we select the required script to run (Fig. 13). We are using ISO spur gears as the rotating elements and to get it, we go into the design library's Toolbox and select ISO. Several options are displayed and we select power transmission among these options, and then further, we choose gears where we get spur gears option (Fig. 13). A property window appears on the left and we make the required modifications in the properties (Figs. 12 and 14).

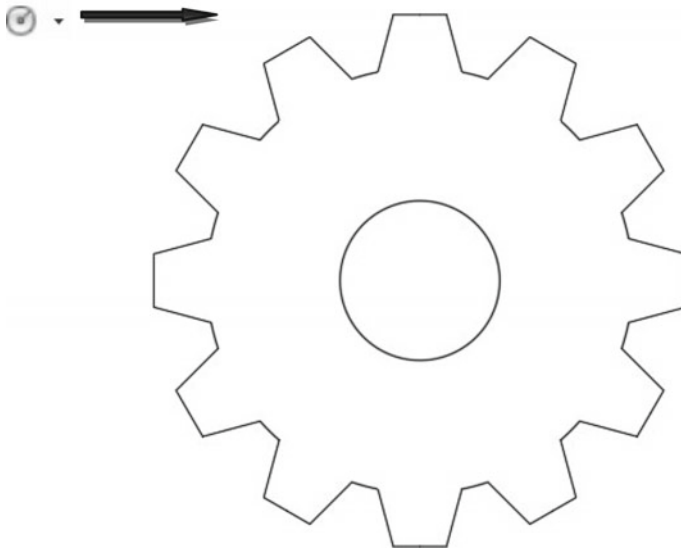


Fig. 4 Creating bored hole of the gear

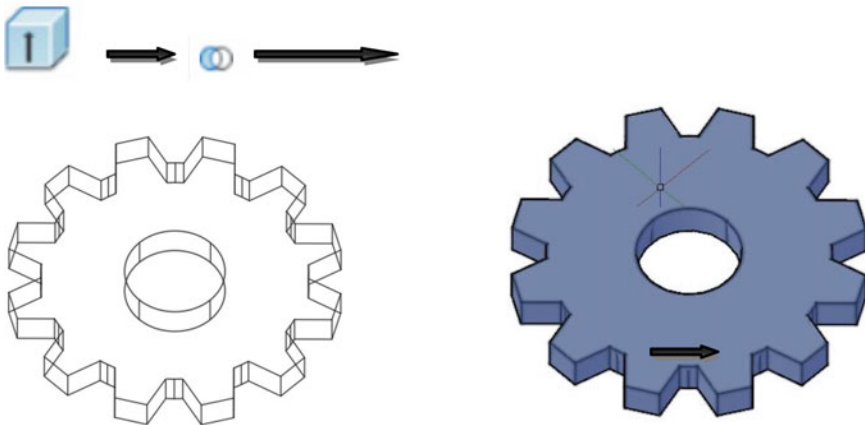


Fig. 5 Extrude and subtract bored hole (conceptual visual style)

3.2 Solidworks

After selecting Assembly as a new SOLIDWORKS document, we choose the Front Plane for designing our gears. We draw a line using mid-point line and taking its two halves as radius, we draw two circles (Fig. 12). Choosing the gear option from

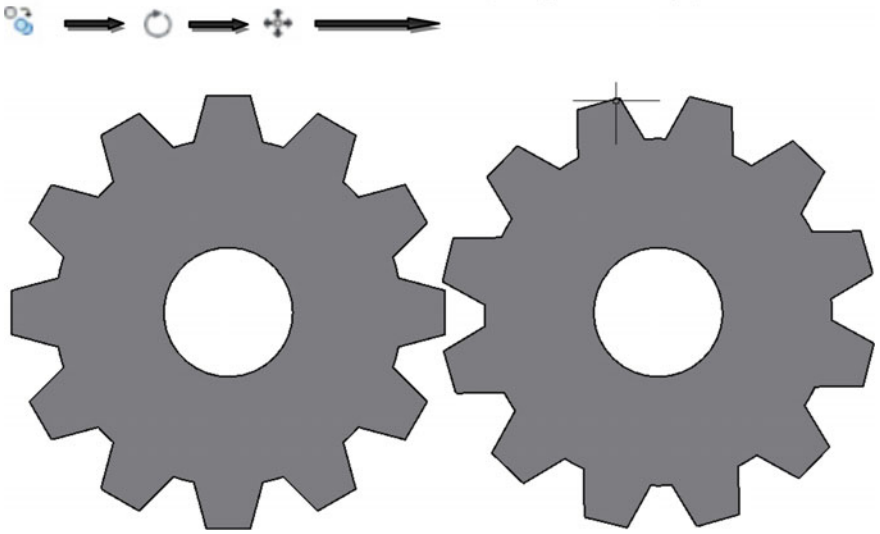


Fig. 6 Interlocking of gears

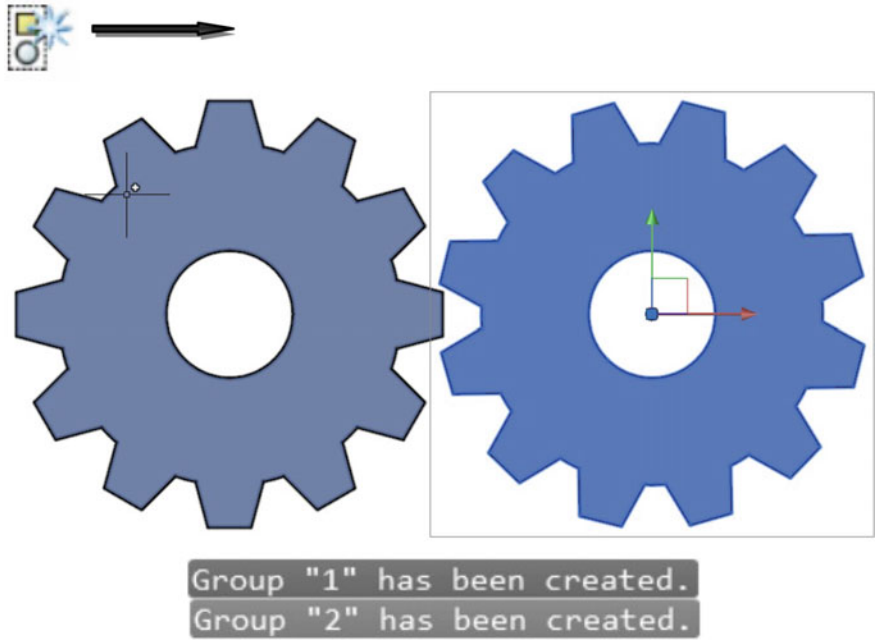


Fig. 7 Grouping of geometries



Fig. 8 Rscript on notepad

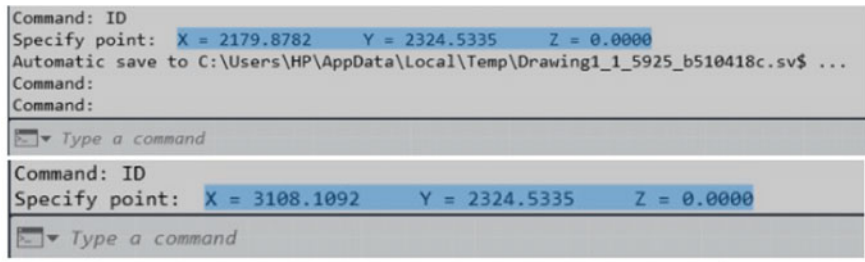


Fig. 9 IDs of the center of the gears using ID command

the Mate property of SOLIDWORKS, we mate the two gears is that we can apply the basic principle of the working of gear on them, i.e., the cut tooth of one gear is meshed with another toothed part to transmit torque. Using the Mate property of SOLIDWORKS, we mate the center of gears with the center of circles. Now for performing motion study, we place a rotary motor on one of the gear and we examine the working of the gears (Figs. 15, 16, 17 and 18).

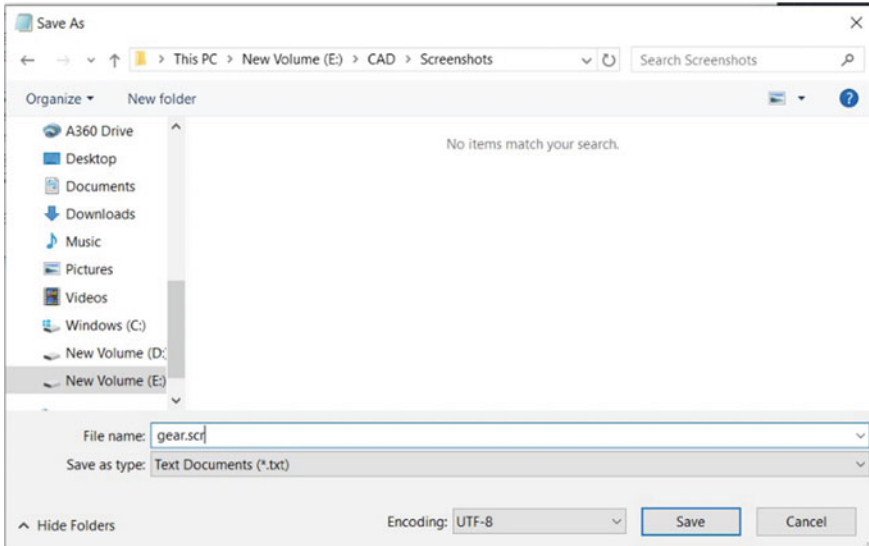







Fig. 10 Saving Rscript with .scr extension













4 Conclusion

Motion simulation software enables almost instantaneous simulation of the inverted slider’s motion, utilizing data already existing in the CAD assembly model. The study compares the motion study of two mated gears on the two CAD software. The motion comparison matrix reveals that it is easy to generate the motion of the mated gears in SOLIDWORKS as compared to AutoCAD. Further, quantitative comparison has been made for comparing the effectiveness of the process of product modeling to motion. The product modeling has been split up into four steps, i.e., geometry, designing of gear, interlocking of gears, and motion study. Since AutoCAD is nonparametric software, hence, the design is built piece by piece which amplifies the lead time of the designer from product development to motion analysis. Motion generation also requires additional effort for running the Rscript. Hence, SOLIDWORKS provides an easy platform for analyzing the motion.

Steps	Motion study on AutoCAD		Motion study on SOLIDWORKS	
Geometry	1. Start drawing		1. Assembly	
			2. Front plane	
Designing gear	1. Circle		1. Center line	
	2. Tooth on the quadrant of the circle		2. Circles	

(continued)

(continued)

Steps	Motion study on AutoCAD	Motion study on SOLIDWORKS
	3. Polar array of tooth 	3. Spur gears from design library 
	4. Trim extra curves 	4. Properties of gears 
	5. Join teeth 	
	6. Creating bored hole	
	7. Extrude 	
	8. Subtract bored hole 	
Interlocking of gears	1. Copy Gear 	
	2. Rotate 	2. Gear mating 
	3. Move 	
Motion study	1. Grouping	1. Placing rotary motor 
	2. IDs	
	3. Notepad (Rscript)	
	4. Save Rscript with .scr extension	
	5. Run script	

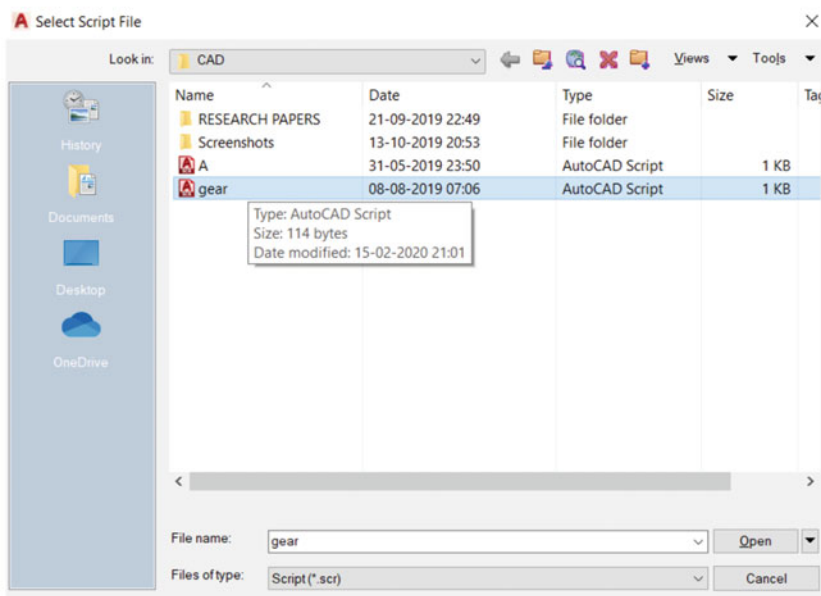
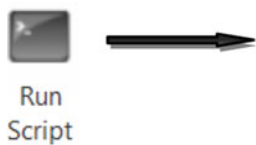


Fig. 11 Run script

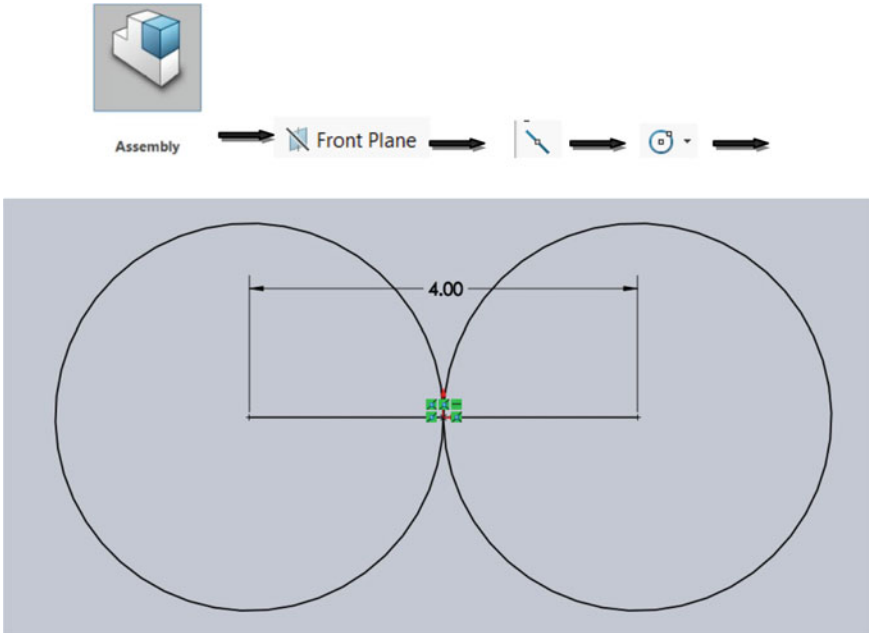


Fig. 12 Drawing circles on the front plane

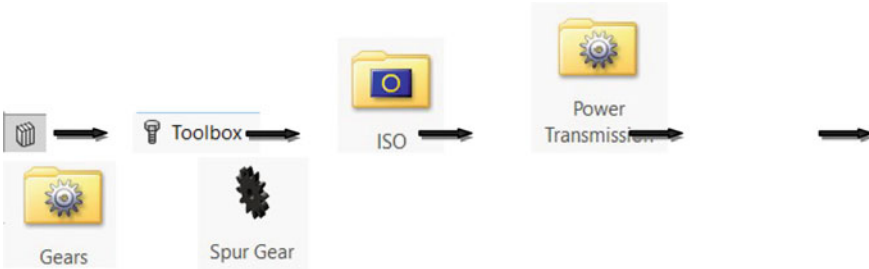


Fig. 13 Taking ISO spur gears from design library

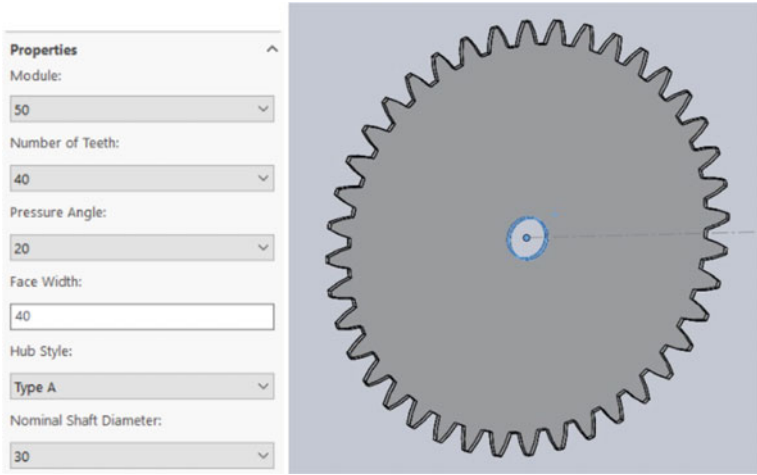


Fig. 14 Properties of gears

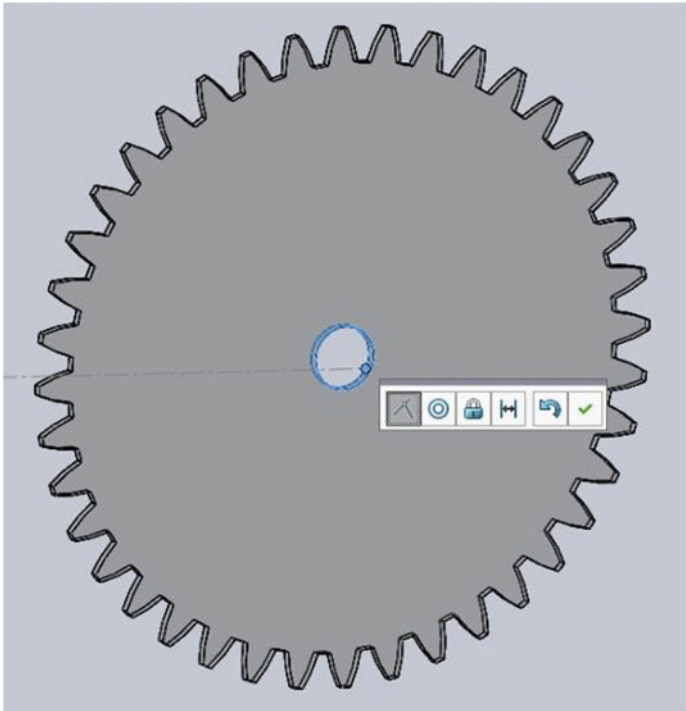


Fig. 15 Mating the center of gears with the center of circles

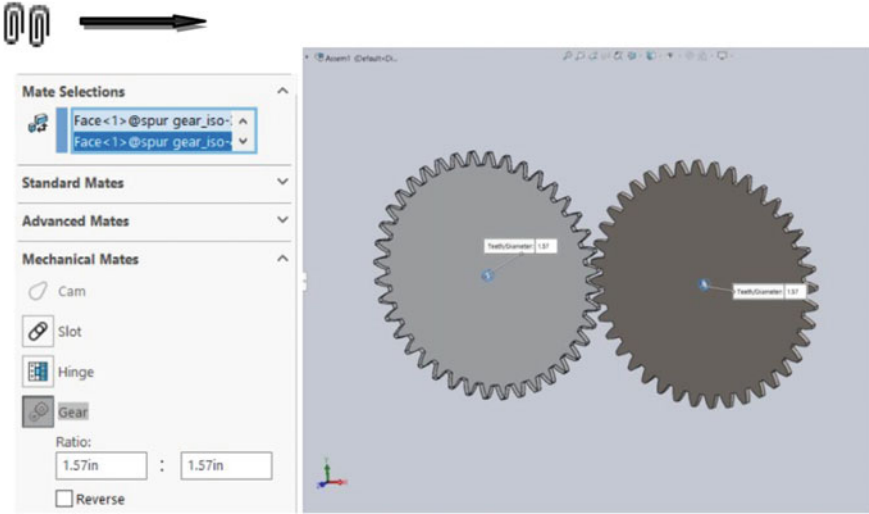


Fig. 16 Gear mating

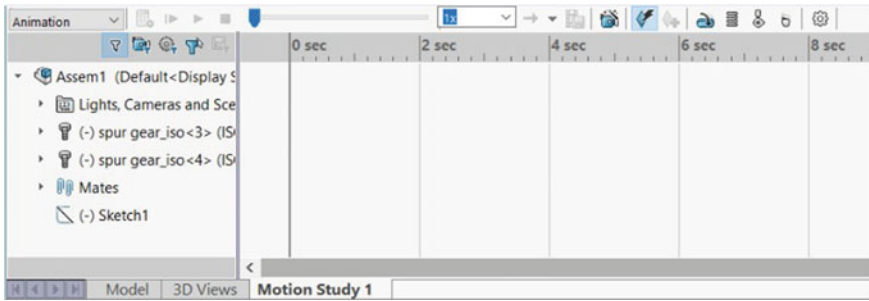


Fig. 17 Motion study

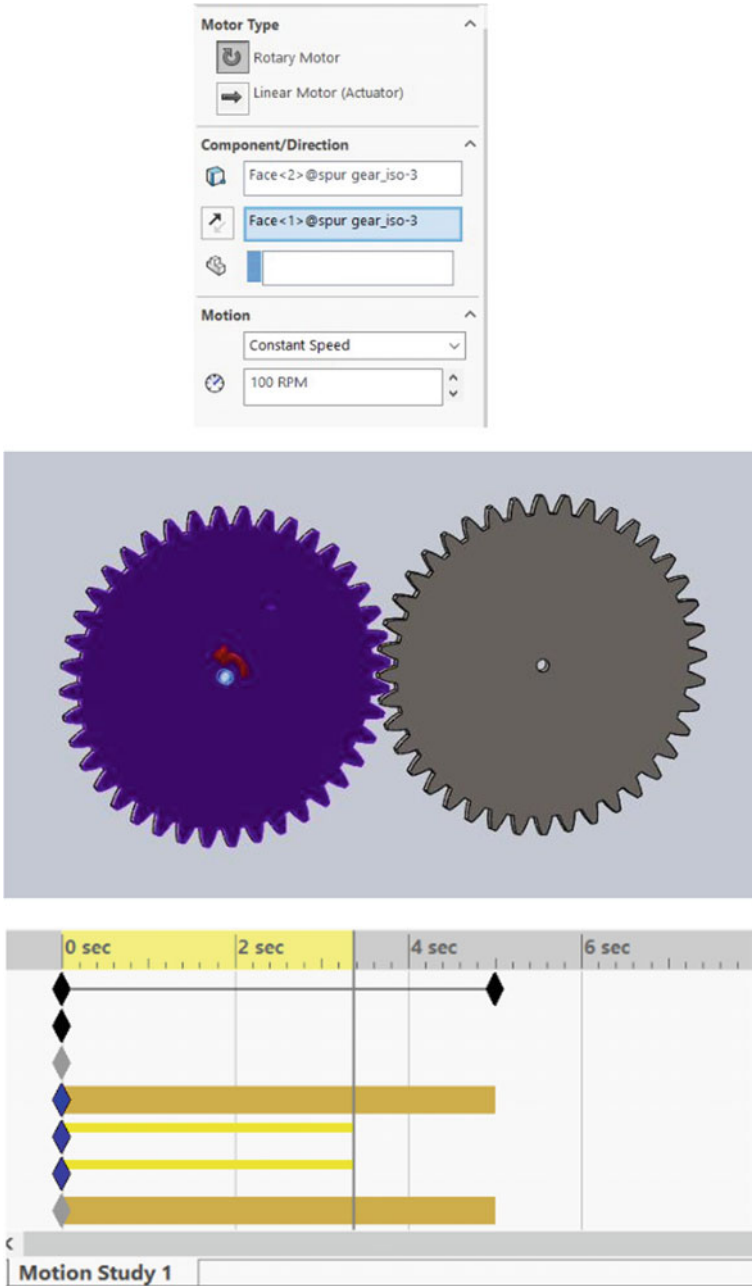


Fig. 18 Placing motor at Gear 1

References

1. Penev, M., Petkova, D., & Vasileva, V. (2005). Comparison at construction of 3D models in CAD systems. *Proceedings of International Science Conference, Series*, 2(44), 534–539.
2. Loginovsky, A. N., & Khmarova, L. I. (2016). 3D model of geometrically accurate helical-gear set. In *International Conference on Industrial Engineering, ICIE 2016 Procedia Engineering* (Vol. 150, pp. 734–741).
3. Kheifets, A. L., Loginovsky, A. N., Butorina, I. V., & Vasilieva, V. N. (2015). *Engineering 3D computer graphics: Textbook and tutorial* (3rd edn.) Yurait.
4. Tickoo, Sh. (2005). *Solid works for designers release 2004*. Peter.
5. Smirnov, A. A. (2008). *3D geometric modeling: Textbook for the course basics of CAD*. Publishing House of Bauman Moscow State Technical University.
6. Sandalski, B., et al. (2006). *Applied geometry and engineering graphics*. Sofia
7. Staneva, N. N. (2008). Approaches for generating 3D solid models in Autocad and solid works. *Journal of Engineering Annals TOME, IV*. ISSN 1584-2673 (FASCICULE 3).
8. Patpatiya, P., Sharma, S., Bhatnagar, V., Tomar, J., & Shalu, J. K. (2019) Approaches for concising AutoCAD files. In *International Conference on Advancement in Computing and Management (ICACM 2019)*.
9. Kheifets, A. L., & Loginovskiy, A. N. (2008). 3D-modeling of gear and worm gears in the package AutoCad. In *Improved training of pupils and students in the field of graphic design and standardization* (pp. 13–19). In Hi. Sat. Scien. tr., Saratov, Publishing House of Saratov State, Techn. University Press.
10. Vaský, J., Eliáš, M., Palaj, J. (2005). 3D solid model generation from 2D vector drawing. In: *CO-MAT-TECH 2005: Proceedings. 13th International Scientific Conference: Trnava, Slovak Republic, 2005* (pp. 1306–1311). STU v Bratislave. ISBN 80-227-2286-3
11. Vaský, J., Eliáš, M., & Palaj, J. (2004). Experimental verification of 3d model generation from 2D vector record. In: *CO-MAT-TECH 2004: Proceedings. Bratislava: STU v Bratislave, 2004* (pp. 1462–1466). ISBN 80-227-2117-4.
12. Kostic, Z., Radakovic, D., Cvetkovic, D., Trajkovic, S., & Jevremovic, A. (2012). Comparative study of CAD software, web 3D technologies and existing solutions to support distance-learning students of engineering profile. *IJCSI International Journal of Computer Science Issues*, 9(4, No 2). ISSN (Online): 1694-0814.

An Overview of Different Approaches for Ternary Reversible Logic Circuits Synthesis Using Ternary Reversible Gates with Special Reference to Virtual Reality



P. Mercy Nesa Rani and Phrangboklang Lyngton Thangkhiew

Abstract Quantum computing has been projected as an alternative to classical computing and promises to offer feasible solutions to problems that are considered intractable today. It has also been observed that information storage, processing, and communication can be improved to a great extent by the use of quantum mechanical effects like non-classical correlation and superposition of states. In a reversible circuit, we have a cascade of reversible gates each of which carry out some bijective mapping between inputs and outputs. They are also structurally different from conventional circuits in the sense that fanout and feedback connections are not permitted. In a binary quantum system expressed as a quantum circuit, the number of circuit lines represent the basic unit of information being processed (called *quantum bits* or *qubits*). Specifically, *qutrit* is the unit of information for a three-valued quantum system and is represented in one of the states such as $|0\rangle$, $|1\rangle$, and $|2\rangle$. As we know, qutrits are very much expensive resources in the synthesis process, and thus reducing the number of qutrits plays a major role during synthesis. The idea behind this work is to make the early researchers in this specific area to get a comprehensive understanding of ternary reversible logic synthesis. Hence, this work addresses the different ternary reversible logic synthesis approaches and thus allow the researchers to gain more knowledge about the process of ternary reversible logic synthesis in an efficient manner and also gives an insight of how quantum computing can be further explored in the broad area of augmented reality and virtual reality.

Keywords Qutrits · Reversible logic · Synthesis · Ternary reversible gates

P. Mercy Nesa Rani

College of Post-Graduate Studies in Agricultural Sciences, Umiam, Meghalaya 793103, India

P. L. Thangkhiew (✉)

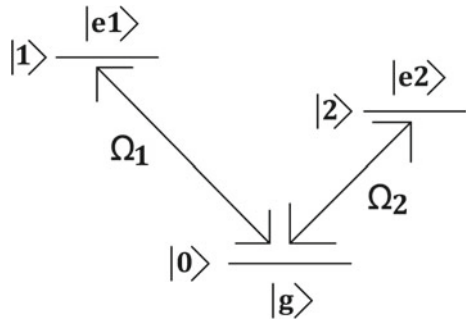
School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Chennai 600127, India

1 Introduction

In a reversible circuit, we have a cascade of reversible gates each of which carry out some bijective mapping between inputs and outputs. They are also structurally different from conventional circuits in the sense that fanout and feedback connections are not permitted. Many researchers have further decomposed the reversible gates into elementary quantum gates, which are considered to be the primitive operations of a quantum computer. In a binary quantum system expressed as a quantum circuit, the number of circuit lines represent the basic unit of information being processed (called *quantum bits* or *qubits*). For a three-valued or ternary quantum system, the unit of information is termed as *qutrit*. For a general multi-valued logic (MVL) quantum system, the unit of information is referred to as *qudit*. As opposed to conventional computing, where the gates represent hardware resources and the total computation time is determined by the gate delays, the concept is somewhat different in reversible or quantum computing. Here, the circuit lines representing quantum bits are the hardware resources, and the gates represent the operations that are carried out in sequence of time. During the process of synthesis, various cost metrics are utilized by researchers to compare the quality of the generated gate netlists, viz. (a) number of reversible gates, (b) number of quantum gates (called *quantum cost*), (c) number of additional circuit lines required (called *ancilla lines*), etc.

Quantum circuits dealing with qubits have two basis states denoted as $|0\rangle$ and $|1\rangle$, [32], and it subsist in a superposition state. For the binary logic circuit line is n , the input combination is 2^n and for the ternary logic circuit line is n , the input combination is 3^n . This has reduced the number of circuit lines to realize some given function. MVL quantum systems have recently picked up the researchers' attention primarily for this reason. In order to store information when compared with binary logic system, 63% reduction in the number of qutrits is present in the ternary logic system [17]. The ternary circuit implementation results obtained with the help of simulation using some of the state-of-the-art technologies reported in [5] have provided a positive push toward ternary logic research. The quantum gate realization of ternary reversible functions are typically carried out in two steps, ternary reversible synthesis followed by decomposition of the ternary reversible gates to elementary quantum gates. Recently, researchers have been attracted to have great interest in incorporating augmented and virtual reality in human-computer interaction (HCI). However, such interaction demands massive computation in which quantum computing techniques may be incorporated in near future. Alternatively, HCI can also be used to understand various aspects of quantum computing. The work in [6] explored the usage of the IBM quantum composer platform through a browser designed based on virtual reality. The work in [37] investigates the usage of virtual reality (VR) to be deployed as a tool to teach the basics of quantum computing and compared it to a traditional desktop environment tool. Ashktorab et al. in [1] identify four applications in which human quantum computer interactions that can have an impact on quantum computing. The fields include the concepts of the state-of-the-art and future quantum users, algorithmic tools, view of the quantum states, and manual for quantum users.

Fig. 1 Ternary quantum system



The rest of this article is organized in the following manner. Section 2 discusses the concrete implementation of three-level quantum system as well as the equivalent ternary reversible logic gates in the quantum representation. Section 4 briefly presents the different synthesis approaches for the ternary logic synthesis followed by the conclusion in Sect. 5.

2 Basic Concepts of Ternary Reversible Logic Circuits

In this section, we briefly look into the concrete implementation of the three-level quantum system as well as the ternary gate representation with elementary gates.

2.1 Physical Implementation of Three-Level Quantum System

Consider a three-level quantum system as shown in Fig. 1. Here, the three qutrit states, viz. $|0\rangle$, $|1\rangle$ and $|2\rangle$, are mapped to three internal states of ions. This is expressed by the *ground state* $|g\rangle = |0\rangle$ and two *excited states* $|e_1\rangle = |1\rangle$ and $|e_2\rangle = |2\rangle$. Here, Ω_1 and Ω_2 represent laser beams, which are applied to the ions to manipulate their transition states such as $|0\rangle \leftrightarrow |1\rangle$ and $|0\rangle \leftrightarrow |2\rangle$.

3 Ternary Gate Representation with Elementary Gates

The cost of a ternary reversible gate is typically estimated as the number of elementary quantum gates required to realize it. This requires the decomposition of higher level reversible gates into fundamental gates which are used in the quantum realization. Two types of ternary elementary gates are explained in the literature,

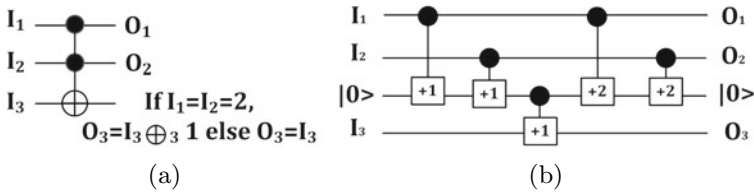


Fig. 2 a Ternary Toffoli gate with both controls $|2\rangle$, b equivalent elementary gate representation

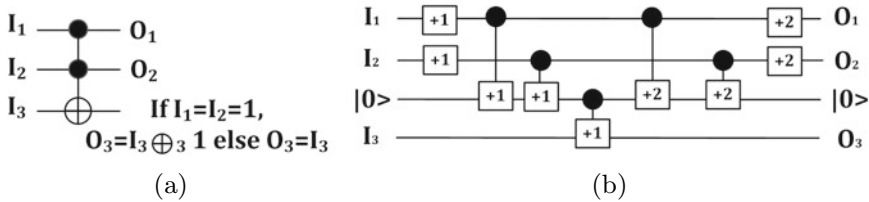


Fig. 3 a Ternary Toffoli gate with both controls $|1\rangle$, b equivalent elementary gate representation

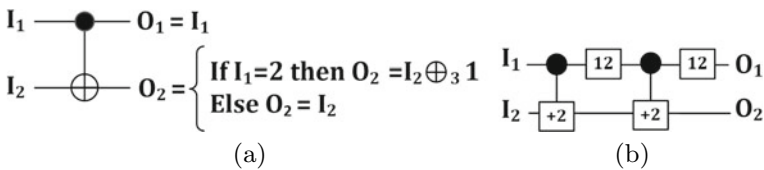


Fig. 4 a Ternary Feynman gate, b equivalent elementary gate representation

viz. Muthukrishnan-Stroud (M-S) gate and ternary shift gates. The elementary gate realizations for ternary reversible gates are discussed below.

The elementary gate realization of ternary Toffoli gate, where the target line C is modified when both control lines A and B are in state $|2\rangle$, requires five M-S gates as depicted in Fig. 2. Similarly, ternary Toffoli gate that modifies the target when both controls A and B are at $|1\rangle$ requires nine M-S gates which is given in Fig. 3.

The realization of ternary Feynman gate requires 4 M-S gates as shown in Fig. 4 [13].

4 Different Synthesis Approaches for Ternary Reversible Logic Circuits

Different kinds of approaches for ternary logic synthesis has been discussed in this section. In general, ternary logic synthesis approaches have been classified into five categories as follows.

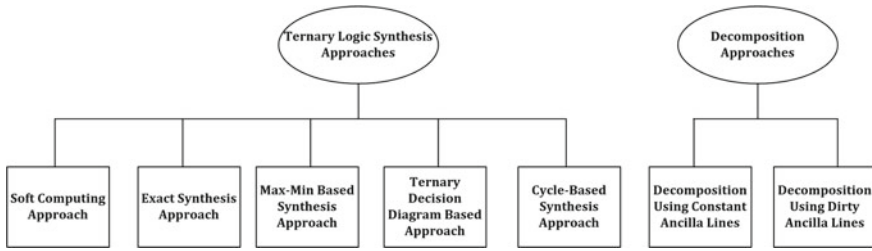


Fig. 5 Summary of works carried out

1. Synthesis Based on Group theory Approach
2. Synthesis Based on Max-Min Approach
3. Synthesis Based on Ternary Galois Field Sum of Products Approach
4. Synthesis Based on Ternary Decision Diagram Approach
5. Synthesis Based on Soft Computing Approach.

The entire work can be summarized in terms of five distinct contributions as depicted in Fig. 5, which are explained below in brief.

4.1 Synthesis Based on Group Theory Approach

This approach discusses about the permutation cycles obtained from the reversible specification. Again these cycles are decomposed into smaller cycles which are then mapped into the ternary reversible gates. Here, we shall discuss the existing literature about the group theory synthesis approach.

In [34], Rani et al. presented the ternary reversible circuits synthesis approach with the help of ternary reversible gates. In this approach, the input given is a Boolean function, and the same is provided as .pla file which has been initially transformed into ternary logic function. Next, this function is expressed as permutation, and it is decomposed into 3-cycles. These cycles are synthesized using the gate library such as Ternary Not (N_T), Ternary Toffoli (T_T) and Ternary Toffoli+ (T_T^+). They could synthesis only 3-cycles from the permutation but unable to synthesis 2-cycles from the permutation.

Yang et al. [36] demonstrated a systematic approach for the realization of ternary quantum in the absence of ancilla bits. All the $p * q$ binary reversible circuits are realized using the gate library, viz. Swap gate, NOT gate, Toffoli gate wherein, p indicates the input, and q denotes the output. In the similar fashion, all $p * q$ ternary reversible circuits in the absence of ancilla bits are synthesized by gate library such as Swap gate (S), ternary NOT (N) gate and ternary Toffoli gates (T). The advantage of this article is that it represents an effective approach thus reduces the space and time resources. The disadvantage is that the approach generates highly suboptimal solutions.

Example 1 Consider the 3-cycle (2, 5, 8). The constructive approach to synthesize the 3-cycle with the help of SNT gate library is illustrated below:

$$\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} 0, 0, 2 \\ 0, 1, 2 \\ 0, 2, 2 \end{pmatrix} \xrightarrow{E_{1,2}} \begin{pmatrix} 0, 0, 2 \\ 1, 0, 2 \\ 2, 0, 2 \end{pmatrix} \xrightarrow{(N_3)^2} \begin{pmatrix} 0, 0, 1 \\ 1, 0, 1 \\ 2, 0, 1 \end{pmatrix} \xrightarrow{(N_2)} \begin{pmatrix} 0, 1, 1 \\ 1, 1, 1 \\ 2, 1, 1 \end{pmatrix} \xrightarrow{T} \\ \begin{pmatrix} 1, 1, 1 \\ 2, 1, 1 \\ 0, 1, 1 \end{pmatrix} \xrightarrow{(N_2)^{-1}} \begin{pmatrix} 1, 0, 1 \\ 2, 0, 1 \\ 0, 0, 1 \end{pmatrix} \xrightarrow{N_3} \begin{pmatrix} 1, 0, 2 \\ 2, 0, 2 \\ 0, 0, 2 \end{pmatrix} \xrightarrow{E_{1,2}^{-1}} \begin{pmatrix} 0, 1, 2 \\ 0, 2, 2 \\ 0, 0, 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \\ 2 \end{pmatrix}$$

$E_{1,2}$ is the Swap gate applied on the first and second column in the matrix. T is the ternary Toffoli gate. N_3 is the ternary NOT gate applied on the third column. N_3^{-1} is inverse of ternary NOT gate applied on the third column. N_2 is the ternary NOT gate applied on the second column. $(E_{1,2})^{-1}$ is the inverse of Swap gate applied on the first and second column.

Miller et al. [27] discussed two different kinds of MVL synthesis with the help of heuristic algorithm. The first approach talks about mapping of a reversible function R to its identity using the MVL gates with the proper sequence of those gates. The next approach discusses about the mapping of inverse function R^{-1} to its identity with the help of MVL gates. After that, the gate counts are compared using different approaches, and the method with less gate count is considered the best approach. The merit of this approach lies faster for the smaller circuits. On the flip side, the synthesis follows a greedy approach that depends on local information and thus backtracking is not used, which in turn generates suboptimal results.

4.2 Synthesis Based on Max-Min Approach

Ternary Max_Min algebra is expressed with the help of the ternary values viz., $T \in \{0, 1, 2\}$ and has three important operations as given below.

1. *Max* operation
2. *Min* operation
3. *Complement* operation.

Ghiye et al. [7] presented a ternary logic function minimization with the help of max–min expression. Specifically, the results for the Boolean function up to input size 3 are reported in this article. However, there exists scalability issues for this approach.

Mandal et al. [26] discussed a realization method wherein ternary function is denoted using minterms. In this article, two algorithms are presented to reduce the minterms. The first algorithm generates the largest prime implicants from a given set of minterms, while the second algorithm outputs the prime cubes from a given list of

prime implicants. The outputs from the two algorithms are expressed in projection operations such as L_i and J_i as defined below, from where the M-S gates netlist is obtained directly.

$$L_i(a) = \begin{cases} 1, & \text{if } a = i \\ 0, & \text{otherwise} \end{cases}$$

$$J_i(a) = \begin{cases} 2, & \text{if } a = i \\ 0, & \text{otherwise} \end{cases}$$

Results for ternary benchmarks are shown wherein the gate cost is shown to be 31% less as compared to earlier works in the literature.

Recently, Khan et al. [10] proposed a synthesis approach using *Ternary Max-Min algebra*. Initially, max–min expressions are generated from a given ternary truth table. After that, a map-based minimization method is used to reduce the expressions. The authors presented synthesis results till the input size 4 only and could not demonstrate the scalability of the approach.

4.3 Synthesis Based on Ternary Galois Field Sum of Products

This synthesis is based on Galois field that contains a finite number of elements. Six basic ternary literals are used in the formulation as follows:

$$t = (012)$$

$$t' = t + 1 = (120)$$

$$t'' = t + 2 = (201)$$

$$t''' = 2t = (021)$$

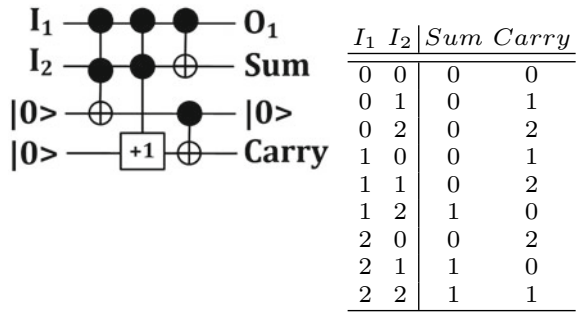
$$t^h = 2t + 1 = (102)$$

$$t^p = 2t + 2 = (210)$$

A *ternary product term* is a ternary Galois field product of a ternary constant, basic ternary literals, and composite ternary literals of ternary variables. A Ternary Galois Field Sum-of-Products (TGFSOP) expression is constructed as a composition of ternary product terms. For example, $t't$ is a product term, and $1 + t^h t'' + t t'$ is a TGFSOP.

Khan et al. [20] proposed the realization of ternary shift gates at quantum level. Two to three qutrits are required for realization of the ternary shift gates. By using one of the reversible ternary shift gates, a ternary signal can be transformed using one of the six gates. An algorithm has been proposed to synthesize multi-output GFSOP. The authors also introduced new ternary gates based on GF(3) addition and multiplication operations and used a graph based data structure called *implementation graph* for

Fig. 6 a Ternary half adder synthesized using [25], **b** truth table



(a)

(b)

synthesis. They showed that the new ternary gates are more efficient for multi-output GFSOP, while the conventional ternary Toffoli gates are efficient for single-output GFSOP.

Khan [14] showed that by grouping pairs of bits together, the conventional binary functions can be expressed as quaternary functions. These quaternary functions are expressed in terms of quaternary Galois field sum of products and implemented using a cascade of quaternary Feynman and Toffoli gates. In [16] Khan presented the Toffoli gate realization without ancilla input, thereby reducing the quantum register width. Mandal et al. [25] proposed a methodology for ternary logic function realization with ancilla bits using generalized ternary gates and showed the results for the synthesis of half-adder and full-adder circuits. Improvement in gate count is achieved using this method as compared to previous methods. The half-adder and full-adder circuits as synthesized are shown in Figs. 6 and 7, respectively. One drawback of this approach is that the number of ancilla lines increases. Khan [12] proposed the realization of 3-qudit Toffoli gate using elementary 2-qudit and 1-qudit gates using eight GTGs. In [13], Khan presented the realization of macro level gates, viz. 2-qudit Feynman gate, 3-qudit Feynman gate and 3-qudit Toffoli gates, using 1-qudit gates and 2-qudit M-S gates. These macro gates are realized using elementary ternary gates. The 1-qudit and 2-qudit gates are physically realizable using quantum technology.

Khan [11] investigated the realization of ternary adder circuit for the first time using GTG in terms of ternary shift operations. The number of gates is reduced using this method and results in less quantum cost. They used four GTGs to synthesize a ternary half adder as shown in Fig. 8. In the first GTG, ancilla input 1 is the controlled input, while the controlling input is I_2 . If $I_2 = 0$, then the value of ancilla input will be the self single shift of ancilla input value. If $I_2 = 1$, then the value of ancilla input will be the self dual shift of ancilla input value. If $I_2 = 2$, then the value of ancilla input will be the dual shift of ancilla input value, and the same procedure has been followed for all the evaluation of GTG. Ten GTGs have been used for the ternary full-adder realization as shown in Fig. 9.

Mondal et al. [28] presented a balanced ternary reversible logic circuit realization, which consists of three states -1 , 0 , and $+1$. They have proposed balanced ternary

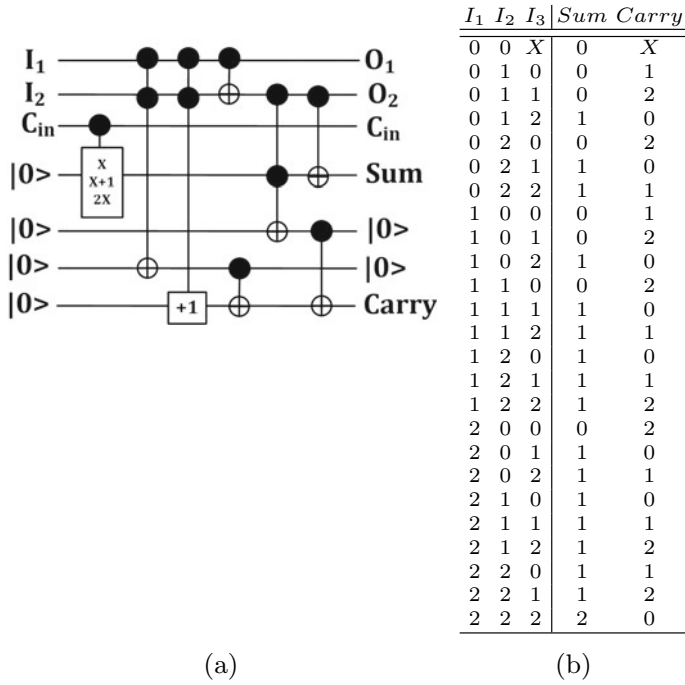


Fig. 7 a Ternary full adder synthesized using [25], b truth table where $X \in \{0, 1, 2\}$

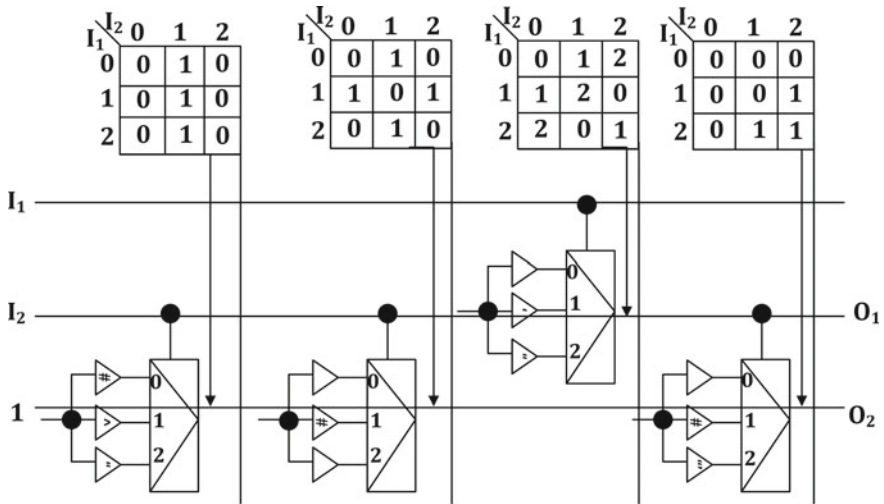


Fig. 8 Ternary half-adder realization using [11]

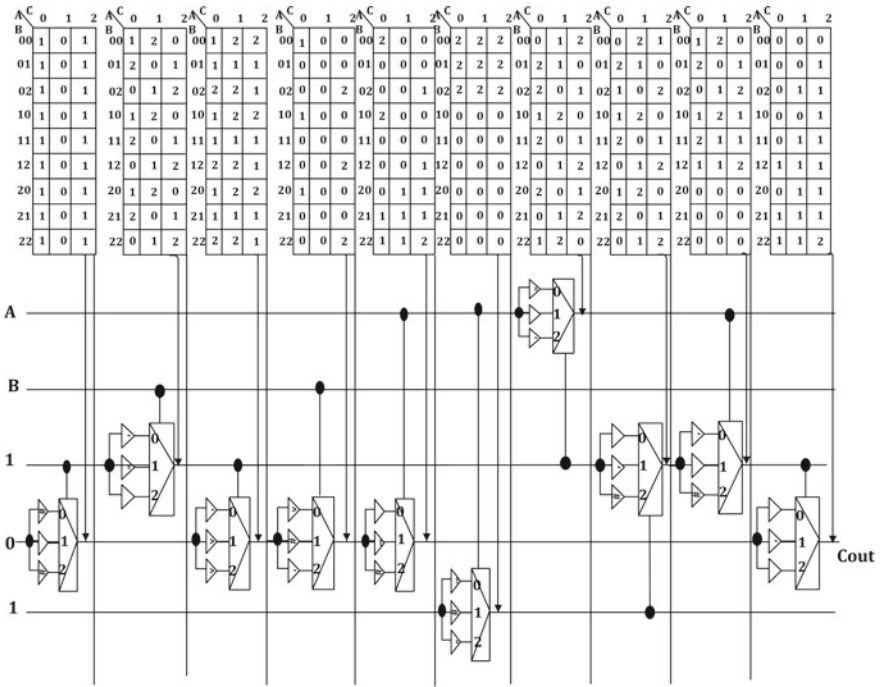
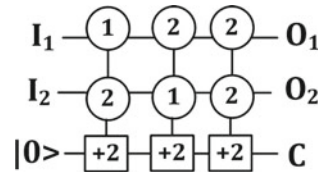


Fig. 9 Ternary full-adder realization using [11]

Fig. 10 Synthesis of carry output of ternary half adder [19]



NOT, CNOT, and C^2 NOT gates. The full adder and half adder, single-trit multiplier, and double-trit multiplier circuits have been realized using balanced ternary gates. In this approach, the hardware complexity has been reduced significantly.

Khan and Perkowski [19] proposed the quantum realization of ternary full adder using ternary Feynman and Toffoli gates. These gates are realized using ternary 1-qutrit and M-S gates. They also realized the carry output of ternary half adder as shown in Fig. 10. Based on the realization of the ternary full adder as shown in Fig. 11, they also proposed the realization of ternary parallel adder with partially-look-ahead carry.

Lisa and Babu [23] presented the optimized design for the ternary adder/subtractor circuit. An algorithm has been proposed to realize the full adder using ternary Peres gate. A 1-digit quantum ternary adder/subtractor circuit is a combinational quantum ternary circuit which adds three digits A , B , and C (where C is the carry-in for the circuit) and generates a sum S and a carryout C_{out} , or subtracts $A - B - C$ result-

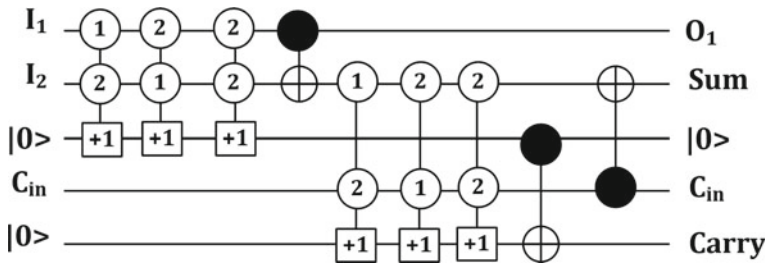


Fig. 11 Ternary full adder synthesis using [19]

ing in a difference D and a borrow B_{out} . Ternary adder/subtractor circuit has been realized using ternary Peres gate and full adder. The comparison of proposed and existing design shows that the proposed circuits perform much better than the existing design [19] in terms of garbage outputs, constant inputs and quantum gate calculation complexity. Moraga [31] presented the basic aspects of quantum computing with respect to ternary systems, and addressed the entanglement issues in quantum computing. Rabadi and Perkowski [33] discussed the Galois field method to synthesize the circuits, using Galois quantum matrix proposed for Swap and Toffoli gates synthesis.

Monfared and Haghparast [29] proposed the realization of subtractor circuits. The authors have presented the realization of ternary half-subtractor and full-subtractor circuits using GTG, ternary Toffoli, and ternary Feynman gates. The number of gates needed for the realization are 3 and 6, respectively. Monfared and Haghparast [30] presented improved designs of ternary half-full-subtractor circuits using ternary M-S and shift gates. They have given the hardware complexity of the circuits as $5\rho + 2\sigma$ and $10\rho + 4\sigma$, respectively, where σ denotes the number of ternary 1-qutrit shift gates and ρ indicates the number of ternary M-S gates used.

Zadeh and Haghparast [38] proposed the design of ternary comparator that it is widely used in tree-based searches. It requires less number of ancilla inputs and less quantum cost. They have used ternary shift gates, ternary Toffoli gates and M-S gates for the realization. Here, the authors have used two circuits for $n = 1$ and $n \geq 2$, where n denotes the number of qutrits used. The reason is that when $n = 1$, they cannot use a 3-qutrit Toffoli gate. To compare two n -qutrits, they have carried out subtraction using addition operations. Better results are obtained when compared with [15]. Haghparast and Dousttalab [8] discussed the synthesis of quaternary flip-flops that are the basic blocks for any storage unit.

4.4 Synthesis Based on Ternary Decision Diagram

In this subsection, we briefly describe the concept of ternary decision diagram, and then the various existing works of ternary function synthesis using TDD. A Boolean

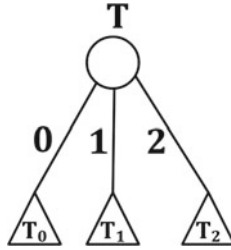


Fig. 12 General structure of TDD

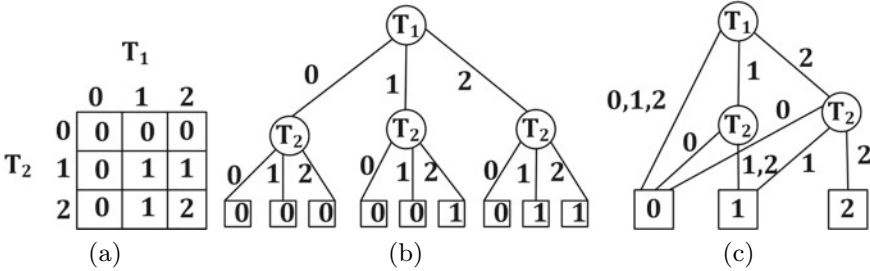


Fig. 13 The Min function: a truth table, b complete TDD, c reduced ordered TDD

function $f : B_n \rightarrow B$ is denoted by a graph-structure known as TDD and is defined as follows.

Definition 4.4.1 TDD is an extension of binary decision diagram (BDD) to three-valued logic. The three-valued Shannon decomposition of a function $T : F^n \rightarrow F$ can be expressed as:

$$T = x^0T_0 \vee x^1T_1 \vee x^2T_2 \tag{1}$$

where x^0, x^1 and x^2 denote the ternary logic states, and $T_0, T_1,$ and T_2 represent Boolean functions.

The representation of an individual node of the TDD is shown in Fig. 12, where there are three child nodes correspond to the three ternary states of a variable.

Example 2 For ternary variables T_1 and T_2 , we define a function Min (T_1, T_2) that gives the minimum value among the two variables T_1 and T_2 . The truth table for the Min function is shown in Fig. 13a, and the corresponding complete TDD in Fig. 13b. For example, Min(0, 2) = 0 and Min(1, 2) = 1.

TDDs can be reduced to get the reduced ordered TDD (ROTDD), which gives a canonical representation of a function for a given ordering of the input variables. The ROTDD for the Min function is shown in Fig. 13c. The complete TDD size is $O(3^n)$. The size of the ROTDD becomes $O(3^n/n)$ after the reduction of TDD using transformations given in [35].

Sasao [35] has presented the generation of Sum-of-Products-TDDs (SOP-TDD) and Exclusive Sum-of-Products-TDDs (ESOP-TDD). The generation of SOP-TDD and ESOP-TDD are illustrated below.

Example 3 We first consider the generation of SOP-TDD for a ternary function $F : T^n \rightarrow B$, where $T = \{0, 1, 2\}$ and $B = \{0, 1\}$. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, $\alpha_i \in T$, denotes a ternary vector. Then, $x_1^{\alpha_1}, x_2^{\alpha_2}, \dots, x_n^{\alpha_n}$ represents a product of an n -variable function, where

$$x^\alpha = \begin{cases} \bar{x}, & \text{when } \alpha = 0 \\ x, & \text{when } \alpha = 1 \\ 1, & \text{when } \alpha = 2. \end{cases} \tag{2}$$

Consider $F = T_1\bar{T}_2 \vee T_2\bar{T}_3 \vee T_3T_2$. By applying Eq. 2 in function F , it can be represented using the array of cubes as shown in Table 1. The complete SOP-TDD is shown in Fig. 14.

In case of ESOP-TDD generation, consider $F = \bar{T}_1T_2\bar{T}_3 \oplus T_1\bar{T}_3 \oplus 1$. By applying Eq. 2 in function F , we get the array of cubes as shown in Table 2 (Fig. 15).

Table 1 Array of cubes for function F

T_1	T_2	T_3
1	0	2
2	1	0
2	1	1

Fig. 14 Complete SOP-TDD

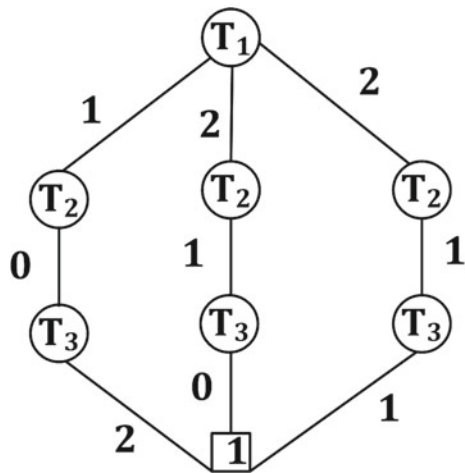


Table 2 Array of cubes for function F

T_1	T_2	T_3
0	1	0
1	2	0
2	2	2

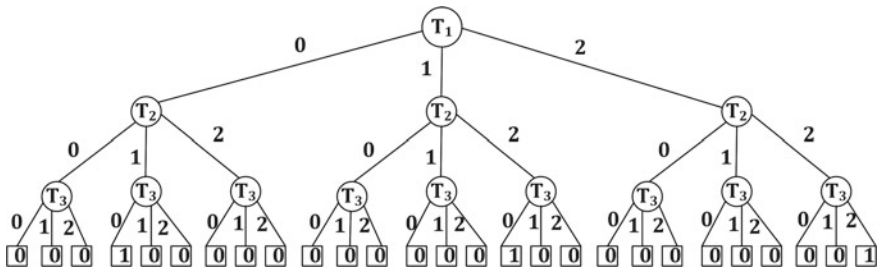
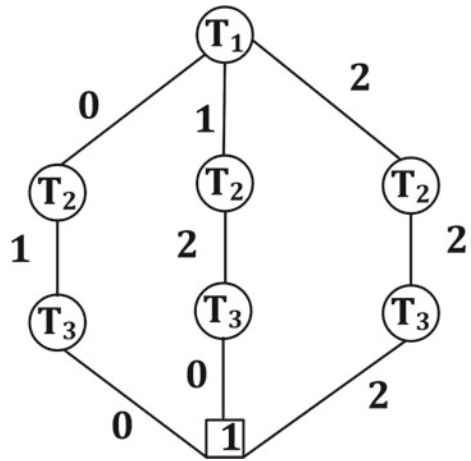


Fig. 15 Initial ESOP-TDD

Fig. 16 Reduced ESOP-TDD



The reduced ESOP TDD is shown in Fig. 16. In this example, we can observe that the size of the generated TDD is smaller for the function F having only 8 nodes as compared to the initially generated TDD with 40 nodes.

Ching and Suaidi [2] presented the different types of TDD used in the boolean function implementation. Kleene DD (KDD) is given more importance because of its effectiveness to determine the unknown inputs in functions. Khan et al. [21] presented sixteen GF expansions and three TDD different types which are useful in the reversible logic design. A heuristic function is incorporated in order to create optimal TDDs.

4.5 *Soft Computing-Based Synthesis*

In this section, soft computing-based synthesis approaches for ternary reversible circuits are discussed. Soft computing is a computational approach to solve the problems which are coming under the category of uncertainty and imprecision. This include methods such as simulated annealing (SA), genetic algorithm (GA), ant-colony optimization (ACO), and particle swarm optimization (PCO).

Khan and Perkowski [18] proposed a GA for the synthesis of both completely and incompletely specified ternary using GTG gates. The authors discussed the realizations of generalized ternary Toffoli gate and swap gate using GTG gates. Lukac et al. [24] presented an evolutionary computation approach to quantum and reversible logic circuits synthesis. They used the standard GA approach to define the cost and fitness functions and defined appropriate functions for mutation and crossover operations.

Chowdhury and Singh [3] presented an evolutionary technique using neural networks for synthesizing multi-valued logic (MVL) functions. The MVL operators are trained using neural networks, which help to reduce the number of gates and interlink connections by 52.94% and 23.38%, respectively.

Deibuk and Biloshytskyi [4] have also developed a GA-based synthesis of ternary reversible circuits using M-S gates. A fitness-function has been proposed which plays an important role in minimizing the amount of logic errors. With proper selection of fitness function and chromosome encoding, they achieved lower-cost implementations in terms of elementary gates. The cost of ternary half adder was 13 that needs less elementary gates as compared to the work in [22], where the cost was 19. Han and Kim [9] proposed a quantum-inspired evolutionary algorithm (QEA) based on the concept and principles of quantum computing. QEA uses a qubit as the smallest unit of information, and an individual is defined by a string of qubits. To evaluate the performance of QEA, experiments have been carried out on the knapsack problem, which shows that QEA performs better without local convergence as compared to the conventional GA methods.

5 Conclusion

In this article, we have explored the different kinds of ternary reversible logic circuit synthesis approaches in a detailed fashion. This shall serve as a motivation for the aspiring researchers in understanding the necessary background of the field in a comprehensive way. Moreover, the discussion about ternary reversible elementary gates and its physical implementation shall allow the researchers to grasp the ternary reversible logic synthesis in quick manner as well as in an efficient way. Further, it helps to know the quantum cost for ternary synthesis since the quantum cost for the ternary elementary gates is considered to be unity. There exists lots of scope to carry out the research in the ternary reversible logic circuits. As we know, there is a need

for creation of benchmarks in ternary domain just like the benchmarks which are present in binary domain. Hence, the future research may be taken up as the creation of a comprehensive set of benchmarks in ternary domain to be used extensively for ternary logic synthesis and also to incorporate the quantum computing functionalities in the virtual reality so that enormous computing power can be utilized to carry out the task quickly.

References

1. Ashktorab, Z., Weisz, J. D., & Ashoori, M. (2019). Thinking too classically: Research topics in human-quantum computer interaction (pp. 1–12). Association for Computing Machinery, New York, NY, USA.
2. Ching, S. P., & Suaidi, M. K. (2002). Ternary decision diagrams. In *Proceedings of the Student Conference on Research and Development* (pp. 76–79), Shah Alam, Malaysia, July 2002.
3. Chowdhury, A. K., & Singh, A. K. (2016). Synthesis and reduced logic gate realization of multi-valued logic functions using neural network deployment algorithm. *Journal of Engineering Science and Technology*, *11*(2), 177–192.
4. Deibuk, V. G., & Biloshytskyi, A. V. (2015). Design of a ternary reversible/quantum adder using genetic algorithm. *International Journal of Information Technology and Computer Science*, *7*(9), 38–45.
5. Dhande, A. P., Ingole, V. T., & Ghiye, V. R. (2014). *Ternary digital system: Concepts and applications*. SM Group.
6. Genç, H. H., Aydn, S., & Erdal, H. (2020). Design of virtual reality browser platform for programming of quantum computers via VR headsets. In *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1–5). <https://doi.org/10.1109/HORA49412.2020.9152931>
7. Ghiye, V., Dhande, A., & Bonde, S. H. (2013). Ternary function minimization by map method. *International Journal of Computer Science and Electronics Engineering (IJCSEE)*, *1*(5), 614–620.
8. Haghparast, M., & Dousttalab, N. (2017). Design of new reversible quaternary flip-flops. *International Journal of Quantum Information*, *15*(04), 1750024:1–11.
9. Han, K. H., & Kim, J. H. (2002). Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Transactions on Evolutionary Computation*, *6*(6), 580–593.
10. Khan, M., & Rice, J. E. (2016). Ternary Max-Min algebra for representation of reversible logic functions. In *Proceedings of the International Symposium on Circuits and Systems (ISCAS)* (pp. 1670–1673), Montreal, Canada, May 2016.
11. Khan, M. H. A. (2004). Quantum realization of ternary adder circuit. In *Proceedings of the 3rd International Conference on Electrical and Computer Engineering (ICECE 2004)* (pp. 249–252), Dhaka, Bangladesh, December 2004.
12. Khan, M. H. A. (2004). Quantum realization of ternary Toffoli gate. In *Proceedings of the 3rd International Conference on Electrical and Computer Engineering* (pp. 264–266), Dhaka, Bangladesh, December 2004.
13. Khan, M. H. A. (2006). Design of reversible/quantum ternary multiplexer and demultiplexer. *Engineering Letters*, *13*(3), 65–69.
14. Khan, M. H. A. (2006). Quantum realization of quaternary Feynman and Toffoli gates. In *Proceedings of the 4th International Conference on Electrical and Computer Engineering ICECE* (pp. 157–160), Dhaka, Bangladesh, December 2006.
15. Khan, M. H. A. (2008). Design of reversible/quantum ternary comparator circuits. *Engineering Letters*, *16*(2), 1–8.

16. Khan, M. H. A. (2009). Quantum realization of multiple-valued Feynman and Toffoli gates without ancilla input. In *Proceedings of the 39th International Symposium on Multiple Valued Logic* (pp. 103–108), Okinawa, Japan, May 2009.
17. Khan, M. H. A. (2010). GFSOP-based ternary quantum logic synthesis. In *Proceedings of the SPIE—The International Society for Optical Engineering* (pp. 1–15), California, United States, August 2010.
18. Khan, M. H. A., & Perkowski, M. (2004). Genetic algorithm based synthesis of multi-output ternary functions using quantum cascade of generalized ternary gates. In *Proceedings of the IEEE Congress on Evolutionary Computation* (pp. 2194–2201), Portland, USA, June 2004.
19. Khan, M. H. A., & Perkowski, M. A. (2007). Quantum ternary parallel adder/subtractor with partially-look-ahead carry. *Journal of Systems Architecture*, 53(7), 453–464.
20. Khan, M. H. A., Perkowski, M. A., & Kerntopf, P. (2003). Multi-output Galois field sum of products synthesis with new quantum cascades. In *Proceedings of the 33rd International Symposium on Multiple-Valued Logic* (pp. 146–153), Tokyo, Japan, May 2003.
21. Khan, M. H. A., Perkowski, M. A., Khan, M. R., & Kerntopf, P. (2005). Ternary GFSOP minimization using Kronecker decision diagrams and their synthesis with quantum cascades. *Journal of Multiple-Valued Logic and Soft Computing*, 11(5), 567–602.
22. Khanom, R., Kamal, T., & Khan, M. H. A. (2008). Genetic algorithm based synthesis of ternary reversible/quantum circuit. In *Proceedings of the 11th International Conference on Computer and Information Technology ICCIT 2008* (pp. 270–275), Khulna, Bangladesh, December 2008.
23. Lisa, N. J., & Babu, H. M. H. (2015). Design of a compact ternary parallel adder/subtractor circuit in quantum computing. In *Proceedings of the 45th International Symposium on Multiple-Valued Logic* (pp. 36–41), Ontario, Canada, May 2015.
24. Lukac, M., Perkowski, M. A., Goi, H., Pivtoraiko, M., Yu, C. H., Chung, K., Jee, H., Kim, B. G., & Kim, Y. D. (2004). Evolutionary approach to quantum and reversible circuits. *Synthesis*, 20, 361–417.
25. Mandal, S., Chakrabarti, A., & Sur-Kolay, S. (2011). Synthesis techniques for ternary quantum logic. In *Proceedings of the 41st International Symposium on Multiple-Valued Logic (ISMVL)* (pp. 218–223), Tuusula, Finland, July 2011.
26. Mandal, S. B., Chakrabarti, A., Kolay, S. S., & Choudhury, A. K. (2016). An efficient synthesis method for ternary reversible logic. In *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)* (pp. 2306–2309), Montreal, Canada, May 2016.
27. Miller, D. M., Dueck, G., & Maslov, D. (2004). A synthesis method for MVL reversible logic. In *Proceedings of the 34th International Symposium on Multiple-Valued Logic (ISMVL)* (pp. 74–80), Ontario, Canada, May 2004.
28. Mondal, B., Sarkar, P., Saha, P. K., & Chakraborty, S. (2013). Synthesis of balanced ternary reversible logic circuit. In *Proceedings of the 43rd International Symposium on Multiple-Valued Logic* (pp. 334–349), Toyama, Japan, June 2013.
29. Monfared, A. T., & Haghparast, M. (2016). Design of new quantum/reversible ternary subtractor circuits. *Journal of Circuits, Systems and Computers*, 25(02), 1650014:1–8.
30. Monfared, A. T., & Haghparast, M. (2017). Designing new ternary reversible subtractor circuits. *Microprocessors and Microsystems—Embedded Hardware Design*, 53(2), 51–56.
31. Moraga, C. (2014). On some basic aspects of ternary reversible and quantum computing. In *Proceedings of the 44th International Symposium on Multiple-Valued Logic* (pp. 178–183), Bremen, Germany, May 2014.
32. Nielsen, M., & Chuang, I. (2000). *Quantum computation and quantum information*. Cambridge University Press.
33. Rabadi, A. A., & Perkowski, M. (2001). Multiple-valued Galois field S/D trees for GFSOP minimization and their complexity. In *Proceedings of the 31st International Symposium on Multiple-Valued Logic* (pp. 159–166), Warsaw, Poland, May 2001.
34. Rani, P. M. N., Kole, A., Datta, K., & Chakrabarty, A. (2016). Realization of ternary reversible circuits using improved gate library. In *Proceedings of the 6th International Conference on Advances in Computing & Communications* (pp. 153–160), Cochin, India, September 2016.

35. Sasao, T. (1997). Ternary decision diagrams survey. In *Proceedings of the 27th International Symposium on Multiple-Valued Logic* (pp. 241–250), Nova Scotia, Canada, May 1997.
36. Yang, G., Song, X., Perkowski, M., & Wu, J. (2005). Realizing ternary quantum switching networks without ancilla bits. *Journal of Physics A: Mathematical and General*, **38**(44), 9689/1–10.
37. Zable, A., Hollenberg, L., Velloso, E., & Goncalves, J. (2020). Investigating immersive virtual reality as an educational tool for quantum computing. In *26th ACM Symposium on Virtual Reality Software and Technology. VRST'20*, Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3385956.3418957>
38. Zadeh, R. P., & Haghparast, M. (2011). A new reversible/quantum ternary comparator. *Australian Journal of Basic and Applied Sciences*, *5*(12), 2348–2355.

Applications in Agriculture and Medical

Tomato's Disease Identification Using Machine Learning Techniques with the Potential of AR and VR Technologies for Inclusiveness



Md. Sadik Tasrif Anubhove , S. M. Masum Ahmed ,
Mohammad Zeyad , Md. Abul Ala Walid, Nawreen Ashrafi,
and Ahmed Mortuza Saleque

Abstract This article uses the machine learning (ML) process of MATLAB to generate an XML file (an image trained file) and uses the XML file and computer vision to illuminate an overview of disease detection in tomatoes. After detecting diseases, the result is analyzed with different camera resolutions, light, and distance conditions. The XML file is designed with 550 images of infected tomatoes and 1100 images of fresh tomatoes using the Viola-Jones algorithm and Haar cascade like feature extraction method. As an initial prototype, a camera-mounted hardware system and a graphical user interface (GUI) can detect and recognize only spot diseases in real-time. For analyzing purposes, the system will detect spots in different lights like sunlight, white light, yellow light, red light, and green light. Then a different resolution camera and different distance of camera position will be used to detect diseases. Results of different light, camera resolution, and distances will analyze the accuracy of the system. Spot detection is effective and can easily be added to many other visible diseases that exist for animals, humans, and crops. A system has been proposed by which it will be possible to the ability to effectively diagnose, plan, visualize, and simulate diseases using virtual platforms and 3D modeling of

S. M. Masum Ahmed and Mohammad Zeyad are equally contributed.

Md. Sadik Tasrif Anubhove · S. M. Masum Ahmed (✉) · M. Zeyad (✉)
Energy and Technology Research Division, Advanced Bioinformatics, Computational Biology
and Data Science Laboratory, Bangladesh, Chattogram 4226, Bangladesh

S. M. Masum Ahmed · M. Zeyad
Heriot-Watt University, Edinburgh, Scotland, UK

Universidad del País Vasco/Euskal Herriko Unibertsitatea, Bilbao, Spain

Md. Abul Ala Walid
Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU),
Gopalganj, Bangladesh

Md. Sadik Tasrif Anubhove · N. Ashrafi
American International University-Bangladesh (AIUB), Dhaka, Bangladesh

A. M. Saleque
The Hong Kong Polytechnic University, 11 Yuk Choi Rd, Hung Hom, Hong Kong

recorded photos could be revolutionary which can be applied to the establishment of augmented reality (AR) and virtual reality (VR). However, the system is a success in terms of reliability, with adoptable false positives.

Keywords Augmented reality (AR) · Disease detection · GUI · Haar cascade · Viola-Jones algorithm · Virtual reality (VR)

1 Introduction

The disease is one of the egregious curbs of any cultivation. Fleeting detection of disease is another oxymoron of lucrative cultivation. Fast detection and prevention of disease are very important. The notion of vertical farms is not in mint condition and with the progression of technology, it is a matter of time before it replaces conventional farms for feeding the world's demography. This research work enables any farmer or technician to inspect crops using an automated system to ensure the disease does not destroy the crops. Conventional farms can have the entirety of crops wiped out without warning so having an inspection system that does not need labor for operation is important. Our system is designed to operate at fixed times and alert users in case of an outbreak of the disease in crops or plants. For irrigation purposes, 50% of fresh water used and agricultural land destroyed, large-scale and intensive agriculture having a detrimental influence on the world's wealth [1–4].

Vertical farms are inevitably the future as the technology, sensors, and software expand and more and more governments embrace the idea of food safety being their number one priority in the face of global crisis. The number of vertical farms being built in the last 30 years was mostly experimental, but now, the technology has proven itself to be sustainable. The biggest challenge to any large agricultural operation is the inspection, detection, and elimination of the disease. Growing crops and rearing animals in a soil-free environment within sterile enclosures sounds foolproof, but the risk of disease and parasites are always looming.

Using a relatively small footprint but building high like a skyscraper can maximize the production of crops in dense environments, but a small outbreak of disease can be devastating for the grower. The effect of different light, camera resolution, and distance also important for the system. The objective of this research work is the build an automated image telemetry system that can routinely inspect and detect any disease and anomaly. The system can detect any disease immediately and alert the grower of which plants are affected and what disease is affecting them.

Building a vertical farm today is expensive, so reducing risks to a minimum is an absolute necessity thus an automated telemetry system is needed to ensure any outbreak of disease can be dealt with quickly without compromising the entire farm. This way, crop losses are at a minimum and fresh produce remains affordable for all [5–9]. The effect of different light, camera resolution, and distance is also important for the system.

Categorizing this Chapter: Sect. 2 describes the historical background, Sect. 3 describes the literature review, Sect. 4 describes the research methodology, Sect. 5 describes the methodology and modeling, Sect. 6 describes the MATLAB training image labeler, Sect. 7 describes hardware implementation, Sect. 8 describes the results analysis of the work, and Sect. 9 describes the cost analysis of the system. Finally, Sect. 10 concludes the paper.

2 Historical Background

The concept of vertical was imagined by Gilbert Ellis Bailey in 1915 in his book "Vertical Farming," where he envisioned plants being grown in soil but stacked vertically in a tower formation. The concept of modern vertical farming is radically different from the 1915 interpretation as the ecologist Dickson Despommier developed the first modern vertical farm in 1999 at Columbia University, where the plants were hermetically sealed off from outside conditions and grown in soil-free environments. Today, the broad consensus of the scientific community is that the definition of vertical farming is the practice of producing crops in a vertical arrangement using a building or warehouse under a controlled environment. These environments will have control of light, nutrients, and the atmosphere. Gilbert Ellis Bailey wrote in 1915 about how important it was for the roots of a plant be exposed to air for it to absorb oxygen for growth but also wrote how mass agriculture was severely destroying the environment by eroding the top layer of soil and exposing many areas to erosion and flooding [10].

He discovered the mass use of plows in modern agriculture was doing damage to the top layer of soil during planting and harvesting times as it dislodged many parts of soil which the wind turned to dust and debris, destroying a fertile layer of soil as farmers were under pressure to plant more and harvest quickly with tractors. The use of herbicide, pesticide, and fungicide only exasperated the problems as both farmers and consumers were consuming dangerous levels of biocidal agents which also leached into the flora and fauna [11, 12].

With the introduction of modern nitrate fertilizer, the commercialization of farming was even easier and with plenty of lands to use, farmers could plant and harvest as they chose fit. For decades, this continued until scientists realized the gravity of the problems in the 1950s. Destroying top layers of soil, deforestation, excessive fertilizer use, dangerous biocidal agents, and no renewable farming method meant crop yields were decreasing year over year in the same plots of land where intensive agriculture was being implemented [13]. With the banning of pesticide DDT, the world took the right step, but something else was going to shake the world as human know it [14].

Hundreds of years of fossil fuel burning and the production of harmful industrial chemicals led to the depletion of the ozone layer and the warming of the earth which was labeled as global warming. Today, it is renamed climate change due to some parts of the world getting warmer while some places face extreme cold winters and storms.

Climate change has a devastating effect on the world today. With natural disasters getting worse and worse as the days go by, the way human feed more than seven billion people is now the most important topic as mass extinction due to crop failure is now a realistic situation. Every year in the developed and developing world, huge amounts of top soil being lost due to agriculture, and much of it never gets replaced. Despite genetically modified crops being able to withstand flooding and drought, crop yields in traditionally fertile places like the Punjab state in India have been declining over the years. With the rise of superbugs (pathogens resistant or immune to antibiotics or other medication) the scenario of mass starvation is more realistic than ever. The depletion of the ozone layer and the increase in UV radiation have led to the extinction of some species of frogs in the Amazon rainforest, and it is only a matter of time when plants are going to be affected.

3 Literature Review

Fabrizio Balducci from crop harvest forecasting the design of focal point of this examination and deployment of practical are errands, ranging, abusing, and contrasting different AI methods with suggestions toward which heading to utilize endeavors and speculations. Three different datasets were exploited to conduct their study, which is the “Istat” (National Institute of Statistics) dataset including 16 attributes, “CNR” (National Research Council) dataset with 4 attributes, and “IoT Sensors” dataset together with 17 attributes. However, dataset “Istat” is built up for estimating future harvest sums on complete time-series analysis and forecasting [15].

Arkeman et al. utilize a greenhouse gas examination to examine the feather palm manor utilized within the creation of biodiesel [16]. Again, Amanda proposes an associate degree skilled process using one thirty-five varieties of tomato in order to determine tomato varieties matching parameters or preferences using fuzzy logic on factors like altitude, resistance to diseases, fruit size, fruit shape, yield potential, maturity, and fruit color. Moreover, their proposed work prototype is developed on the Web using MySQL and PHP [17].

Cassava is the third biggest wellspring of carbohydrates for human food on the planet but a matter of distress that cassava can be infected by virus diseases which can result in large damage. Ramcharan [18] used AI to train cassava leaves dataset where cassava disease images were taken in the field in Tanzania. To detect disease and pest damages, a transfer learning-based deep convolutional neural network was used in their study. It is a matter of fact that their study was conducted to identify two types of pest damage as well as three diseases.

Artificial intelligence can be utilized in many ways like preparing robots, gathering, and keeping up farmland productively that normally requires plenty of human resources, time, and exertion. Self-governing robots can supplant human workers in effectively taking care of basic rural errands, for example, harvesting, planting, and weed control. Advanced mechanics with automation is additionally arising as an answer for tackling the issue of laborers in harvesting [19].

de Luna et al. proposed an observing framework utilizing computer vision techniques where they monitored the development of tomato plants based on the number of flowers and fruits. Both RCNN and SSD models are used as a pre-trained model to detect flowers and fruits. Again, to monitor the maturity of tomato fruits, the machine learning-based classification approaches such as artificial neural network (ANN), K-nearest neighbors (KNN), and support vector machine (SVM) were applied to their study. SSD and SVM both showed eminent berry in their set of data [19].

Using surface modeling and deep network estimation for weight prediction, Franchetti et al. proposed a vision-based plant phenotyping in indoor vertical farming and successfully gauge plant development as well as growth [20].

4 Research Methodology

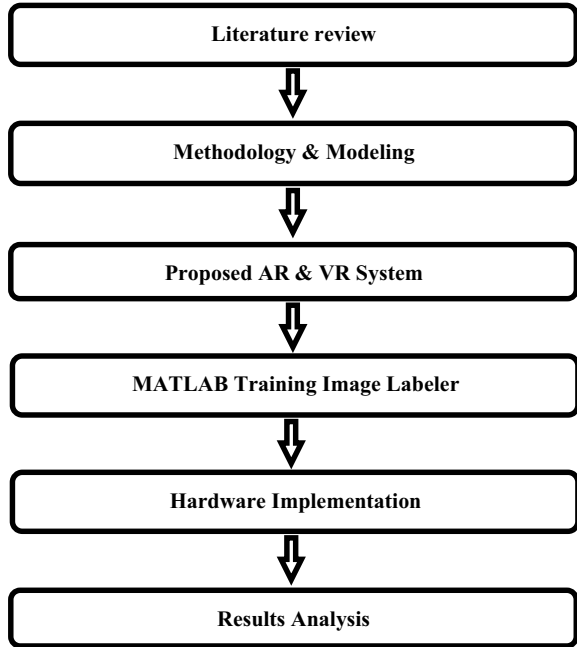
The system was built on the idea that an automated camera system will operate once a day to inspect the crops. The system has to carry the camera, LED array, batteries, Arduino UNO, and any cables attached. A small DIY robot was chosen as these are readily available and can easily carry all the necessary parts. The robot moves in a straight line with the camera detecting all the tomatoes as it goes by. MATLAB software was used to detect the condition of the crops. With the help of training image labeler apps which were previously built in MATLAB software ROI (region of interest) was selected and trained the ROI file with the help of training code to generate the XML file. By using a detection code, a test image was compared with the XML file and detected whether the image was pure or defective. An application was generated for doing this detection for the user's comfort. However, Fig. 1 shows the research methodology of the research work.

5 Methodology and Modeling

The system is mainly a representation of two parts. The first part is hardware; hardware will take power from the power supply. Hardware is made of Arduino and motor driver. Mainly, it is a moveable device that will move the camera. The second part is the Viola-Jones algorithm-based detection process. The camera will send images to MATLAB, then using computer vision, it will detect spots firstly from different distances, secondly in different lighting conditions, and finally using a different resolution camera. A low-resolution camera will send low-quality images and a high-resolution camera will send high-quality images. The system will detect the black spot.

To detect and analyze disorders, image processing in the biomedical engineering domain requires significant development. To make the process more comfortable, a proposed system has been introduced, where a VR device (VR glass) will be connected through a server and the server will collect image or video feed from camera. After

Fig. 1 Research methodology



collecting the image or video stream, it will evaluate and detect disease to deliver the information to the VR device. Finally, the VR device will receive the server’s output and use it to create a virtual platform before sending the data back to the server. Additionally, data from 2D picture or video streams will be evaluated in an AR platform and displayed in 3D utilizing AR. However, Fig. 2 shows the system block diagram.

Figure 3 is the representation of the system flowchart. After switching on the device, the system will start detection. Firstly, the system will scan tomatoes and send the picture to MATLAB. MATLAB will detect the image and show the output. Secondly, the system will detect another tomato as the previous process, and it will show the output. It will continue the process again and again, and finally, if the system reaches the endpoint, then it will stop detection. This is how the system will work.

6 MATLAB Training Image Labeler

For selecting ROI (reason of interest), training image labeler app was used and a training dot mat file was generated by exporting the ROI’s. The dot mat file, a number of positive images, a number of negative images, and a training code were used to train and generate an XML file. The XML file was used to compare and detect diseases as trained.

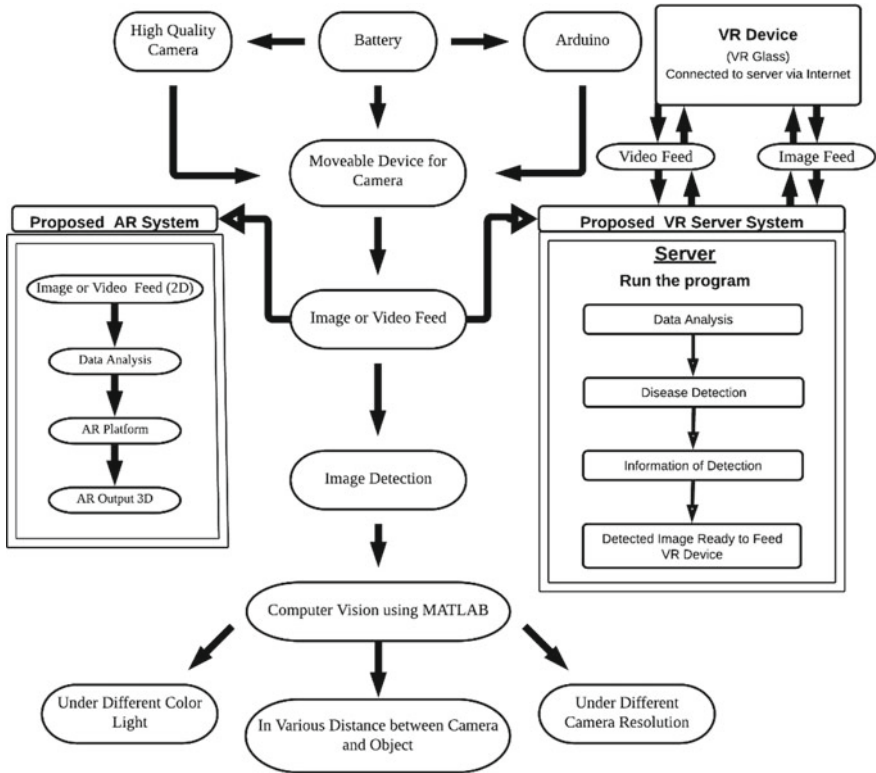


Fig. 2 Block diagram of the system

6.1 ROI

ROI (reason of interest) was selected by a rectangular box. In this work, the black spot was the ROI. The black spot was marked as ROI's. From the positive high-quality images, individually, ROI was selected. (N.B: More positive images for better results). In Fig. 4, ROI was selected. Initially, the analysis focused on determining black spots, and as a result, additional tomato spots were not considered ROI. In this way, ROIs were selected and after selecting all the image's ROI's, it was exported into the workspace and saved as a .mat file.

6.2 Training

A training code was used and the saved .mat file was trained. The directory of both positive and negative folders was clearly mentioned in the training code; otherwise,

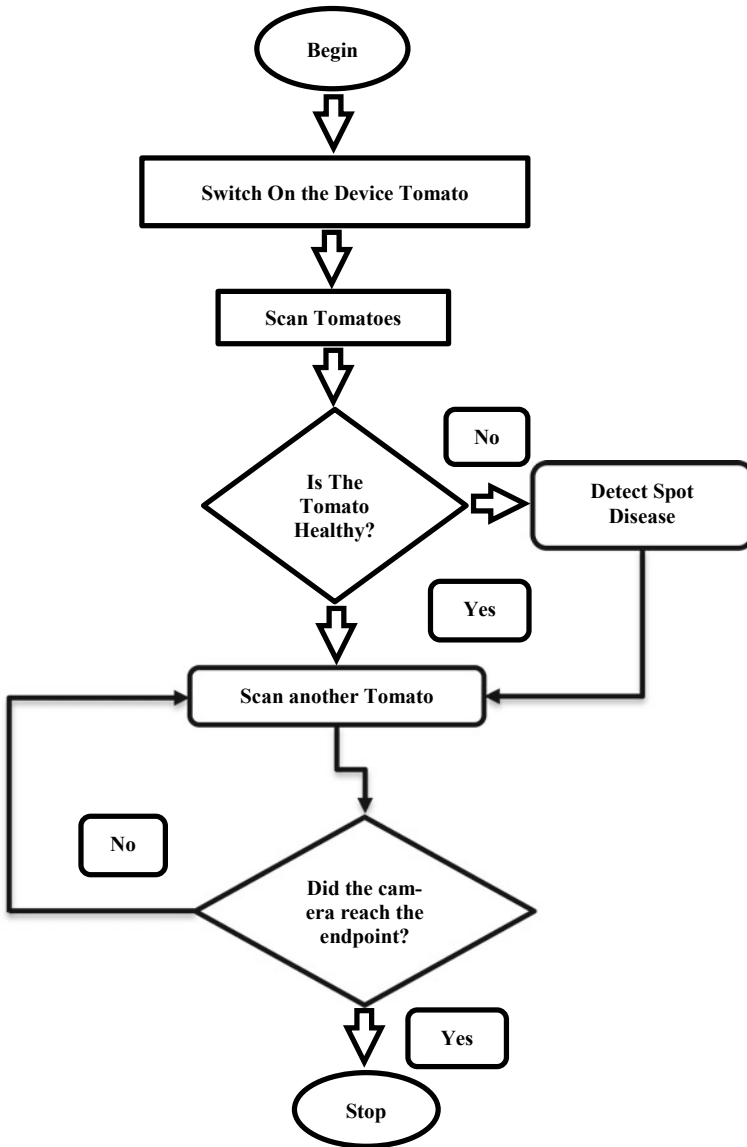


Fig. 3 System flowchart

the code will not be trained. The directory at which the .mat file was saved was also mentioned.

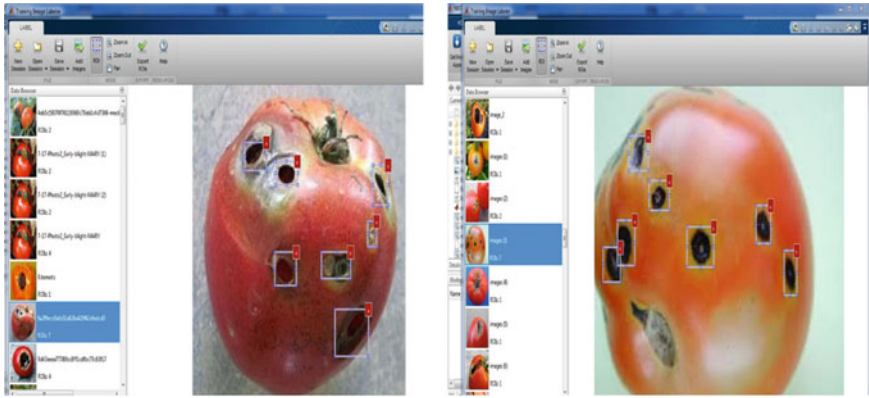


Fig. 4 ROI selection

7 Hardware Implementation

The hardware part was mainly based on a rail track and a moving robot. The robot was used for moving or carrying the camera so that the camera can capture a live image from a different position and detect the images. Robot houses the LED, webcam, batteries, Arduino UNO, and motors which move it in predefined times.

7.1 *Arduino UNO*

The job of the Arduino UNO is to control a robot that has to carry the batteries and camera setup [21, 22]. The camera will be connected to a computer for data analysis as the microcontroller does not have the power to handle such calculations yet but for moving a robot that houses the system, it is more than enough.

7.2 *Rail Track*

The robot was moving through the rail track. The rail track was made with plastic wood. For the moving simplicity of the robot, some parallel rail line was made so that the wheels could get proper friction and could easily move through the line.



Fig. 5 a Robot, b robot in the track

7.3 Robot

Robot houses the webcam, batteries, Arduino UNO, and motors which move it in predefined times. The robot was mainly made of Arduino, motor driver, gear motor, and battery. A camera was attached to the robot. Some code was used to move the robot forward and backward direction. The timing was set in such a way that it will remain in between the track. However, Fig. 5 shows the implemented (a) Robot and (b) Robot in the track of the system.

7.4 User-Friendly App

An app was designed to perform the task easier for users using MATLAB app designers. Moreover, training code, test camera code, and detection code were used to design this app. However, Fig. 6 represents the developed app for this scheme design view.

7.5 Complete Hardware Setup

The camera was controlled by the app and the hardware was automated. For this research work, the delay was set 4 s both for the robot and app detection. The app was generating a report according to the detection, if there were some diseases, it shows “Black spot has been detected”; if not, then it shows “No spot has been detected”.

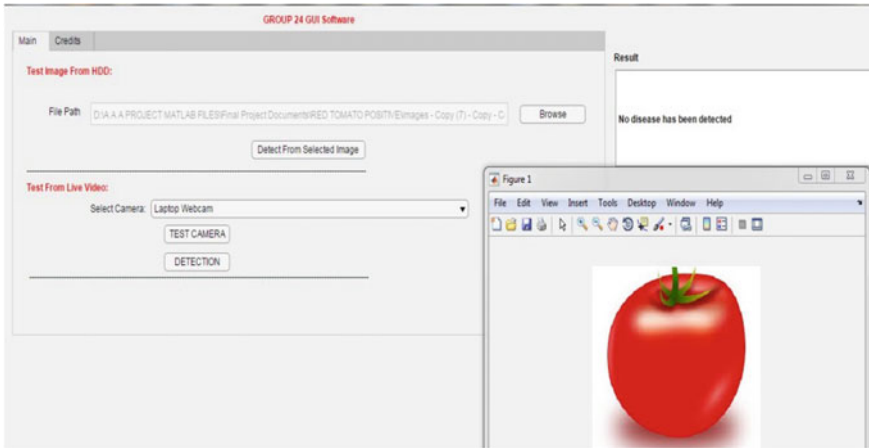


Fig. 6 App developed for this scheme design view

7.6 Algorithms Description

The process of making something visible or understandable by a computer is computer vision. In other words, the vision of the computer is ultimately the computer's eye. Using computer vision, a computer understands or can see images or videos. Image recognition is a practical application of computer vision.

The Haar classifier or Viola-Jones algorithm was used to detect diseases using MATLAB image training labeler application. This algorithm is a feasible target detection obtainment. This feature was first used by Paul Viola and Michael Jones in their research work in 2001. The Haar classifier or Viola-Jones algorithm is materially a machine learning process where the computer is trained by multiple images. Positive and negative images are used for training purposes. Emerged with the training, other images or objects can be detected. After training, an .xml will be originated. Using the originated file, the positive and negative images algorithm will compare the targeted image and generate a result according to the training. Anyhow, Fig. 7 shows the complete hardware setup of the system.

8 Results Analysis

8.1 Image Detection

In the training session, an XML file was generated. With the help of that XML file and a detection code, captured images were detected. For example, some sample images were taken mentioned below in Fig. 8: sample image (a) and Fig. 9: sample

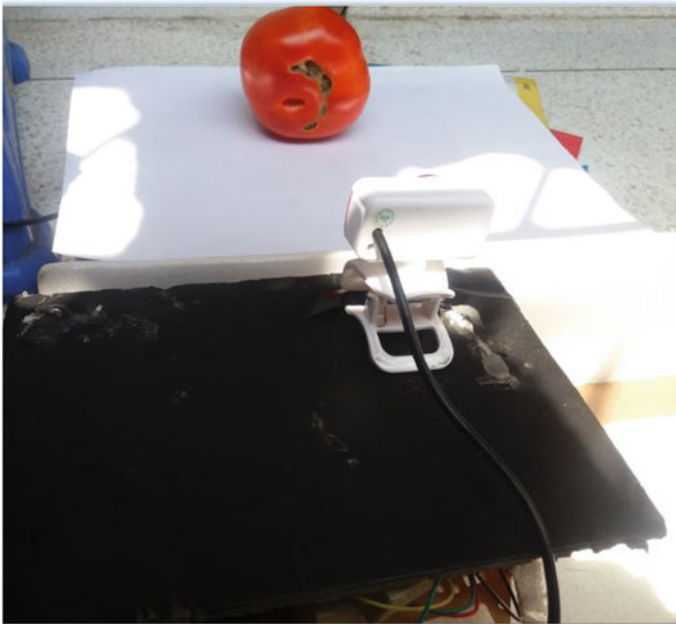


Fig. 7 Complete hardware setup

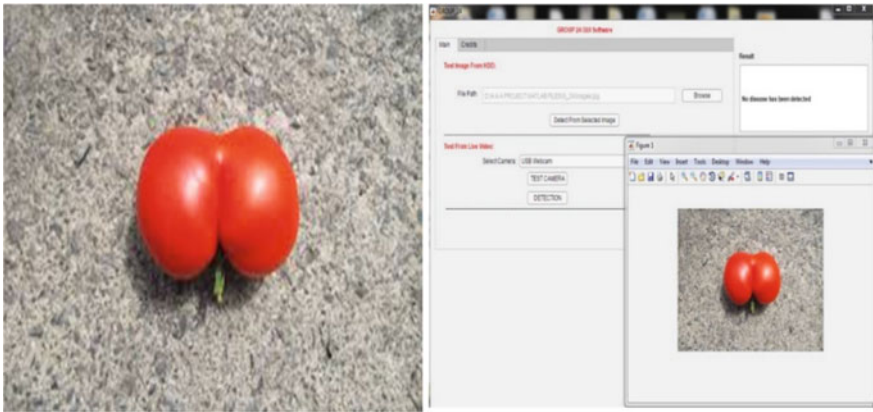


Fig. 8 a Sample image, b detected image

image (a) and after detection, the detected image as shown in Fig. 8: detected image (b) and Fig. 9 detected Image (b).

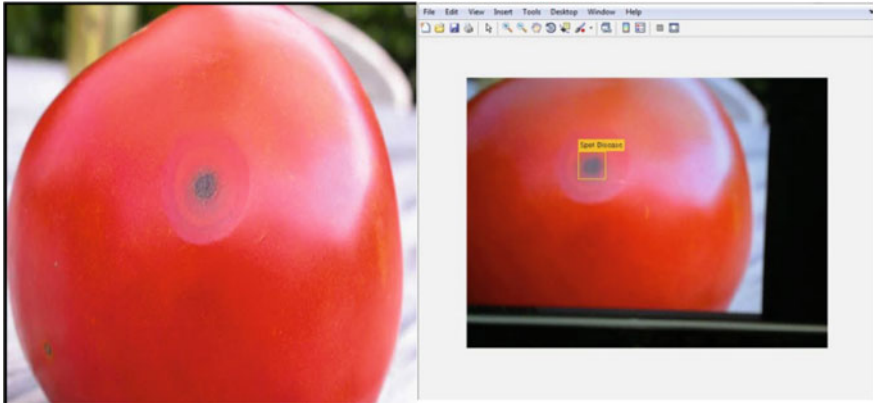


Fig. 9 a Sample image, b detected image

8.2 Camera Resolution

The resolution of the camera is an important factor in image detection. The high-resolution camera takes a better quality picture and a low-resolution camera takes a lower quality snap. First, a low-resolution camera was taken in such a way that the explanation algorithm and training technique operated properly, but it was only for the camera resolution that has been found a false positive or an error in our performance. Since using a high-resolution sensor, an impressive result has been obtained. Figure 10a is the output of a high-resolution camera, and Fig. 10b is the output of a low-resolution camera. The detection of Fig. 10a is fine but the output of Fig. 10b is slightly error. Due to the low image quality, it takes the upper part as a black spot, which is an error.

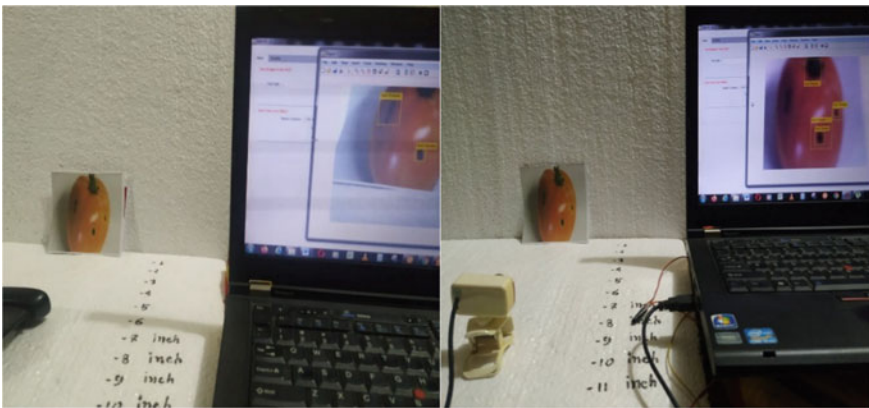


Fig. 10 Output of a high-resolution camera, b low-resolution camera

Table 1 H-R camera versus L-R camera object detection rate

H-R camera	L-R camera
Assumed efficiency of detecting false positive	Efficiency of detecting false positive compared with high resolution camera
100%	70%

For the above pictures, Fig. 10a is the output of a high-resolution camera and Fig. 10b is the output of a low-resolution camera. However, these two images were analyzed for getting a better result. The sample is a picture of an unhealthy tomato. It has a visible spot on its body. In the first image, the system detected a black spot but in the second picture as a low-resolution camera used, the system detects the upper part or core of the tomato as disease or black spot which is counted as a false positive or error. It happens for the low-quality images captured by the camera. Several images had been taken to analyze it.

In comparison with the high-resolution camera, the efficiency of the low-resolution camera is about 70% has been found. Table 1 gives a summary of high-resolution (H-R) camera versus low-resolution (L-R) camera object detection rate.

8.3 Distance

During data collection, it was found that distance between camera and object is an effective term for the system. The camera has been set at different distances like 6–15 in. for getting the desired distance to get high efficient output. Finally, after getting all the output results, it has been found that 6–7 in. are the best distance between cameras and objects as shown in Fig. 11.

Firstly, 6-in. distance data has been collected. In a 6–7 in. distance, the system effectively detects diseases. If the distance increases, the rate of the detection decreases. From 11 to 94-in. distance, the system detects diseases with false positive and above 95-in. system is not able to detect diseases. This distance limit is variable. If a different resolution or a high-quality camera has been used, then the result will be changed slightly. Table 2 gives a summary of the accuracy of detection based on the distance.

Up to 94-in. distance, it can detect with false positive or error as shown in the picture. At a 95-in. distance or above, system cannot detect. It can be varied with camera quality. However, Figs. 12 and 13 show the output of the experiment at 94 in. and 95 in., respectively. Table 3 gives a summary of the highest distance of detection based on the distance.

For data taking purpose of this work, 40 images have been taken and detected it under low light, medium light, and bright light and measured the quantity of perfect detection of black spot. The output results are shown below. In low light, the detection

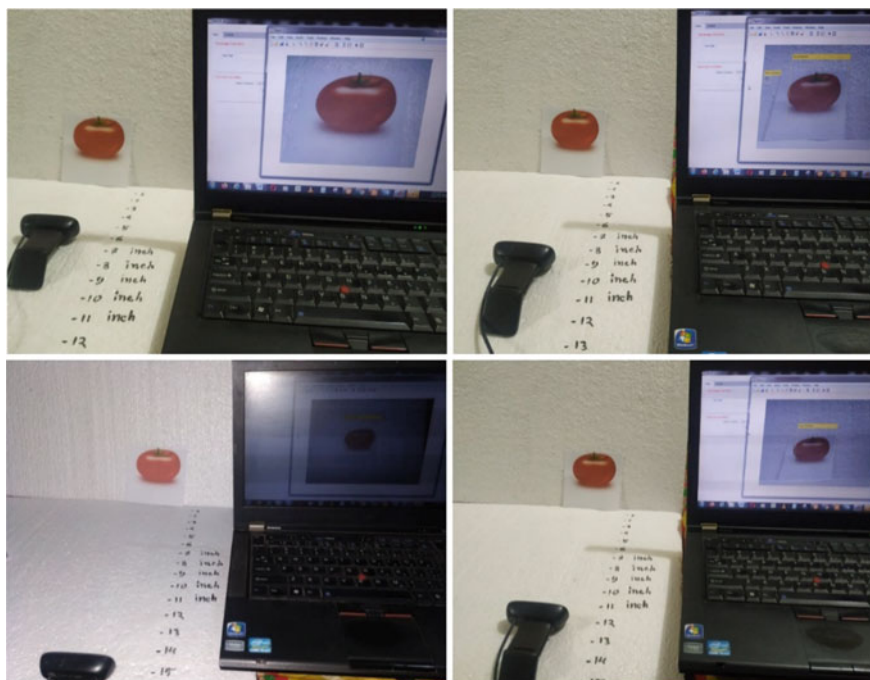


Fig. 11 Output of system in 6, 9, 12, and 15 in. distance

Table 2 Accuracy of detection based on distance

Distance (Inch)	Perfectly detected	Detect with false positive	Not detected
6	Yes	Yes	No
7	Yes	Yes	No
8	No	Yes	No
9	No	Yes	No
10	No	Yes	No
11	No	Yes	No
13	No	Yes	No
14	No	Yes	No
15	No	Yes	No
16	No	Yes	No
17-94	No	Yes	No
>95	No	No	Yes



Fig. 12 Output at 94 in.



Fig. 13 Output at 95 in.

Table 3 Highest distance of detection

Distance in inch	Detected	Not detected
94	Yes	No
95	No	Yes

was quite good but in some cases, under low light, the shades, or some black shades appear. For that reason, the system measures it as a black spot. Table 4 gives a summary of the detection rate according to various lighting condition.

Table 4 Detection rate according to various lighting condition

Method	Different lighting situation (%)		
	Low	Medium	Bright
Viola-Jones algorithm	67.5	87.5	92.5

8.4 Effect of Different Color Light

For analyzing purpose, some research has been done with our system and noted down the outputs. Data has been taken in sunlight, normal white light, red light, yellow light, and green light. It has been found that the system runs fantastically in all the above lighting conditions. In normal light or white light conditions, the first picture shows that the system detected four spots and also it has been show that the sample picture also has four spots. The second picture is in yellow light condition, the third picture is in red light condition, the fourth image is in green light condition, and the output exactly matched with the sample image. Figure 14 shows the output of system in different color condition (normal, yellow, yellow, red, and green light).

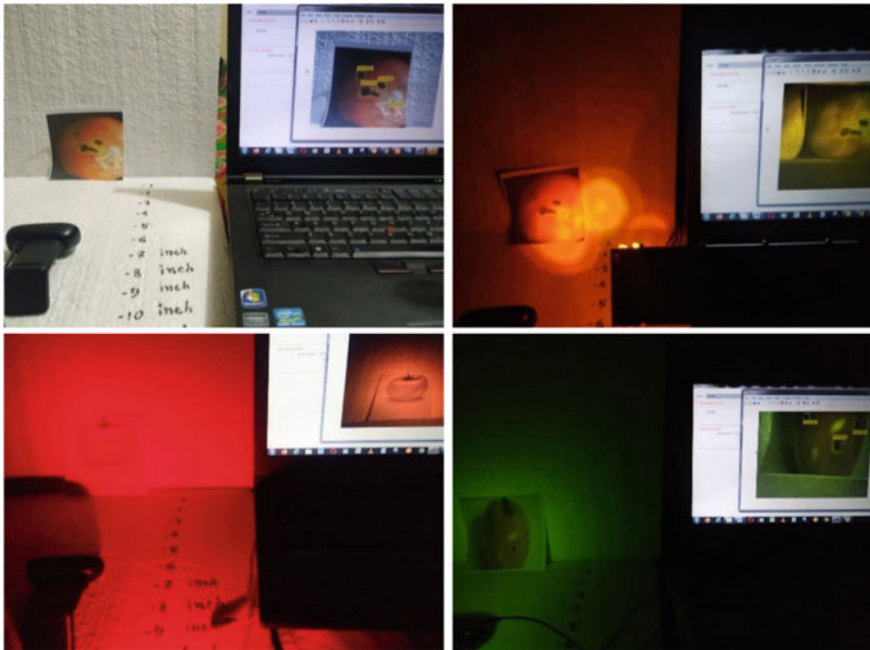


Fig. 14 Output of system in different color condition (normal, yellow, yellow, red, and green light)

Table 5 Cost analysis of the system

Serial number	Name of equipment's	Number of components	Cost (\$)
1	Arduino UNO	01	7.5
2	Motor driver	01	3
3	Camera moveable chassis	01	4
4	USB camera	02	15
5	Cables	04	4
6	Gear motor	02	2
7	Rechargeable batteries	02	11.25
8	2-m rail track	01	10
9	Miscellaneous parts		12.5
10	Red, green, yellow and white light		10.25
Total			79.5

9 Cost Analysis

All the components used in the system are available, and the prices are very low. It is affordable for the majority of users. Table 5 shows the total cost of the system.

10 Conclusion

For vertical farming to grow a large number of crops and inspect them daily is very tough for farmers. If they use this system, their work will be easier and pressure will be reduced. The effect of light does not hamper the efficiency so much that the system can work properly in any lighting condition. The placement of the camera is an important parameter of the system. If the camera resolution and placement of the camera is perfect, then the system will work effectively with efficient output. The number of training images was less because of time constraints, although the system's image recognition is impressive. The system's training and machine learning process are quite easy, and it can be built without too much difficulty. With the help of thousands or millions of images, the system can drastically improve the accuracy and reliability of detection. The system is reliable and works perfectly if it is trained properly with enough resources. It is implementable, improvable and can be utilized in many areas like agriculture, industries, the medical sector, etc. Any visible human diseases, animal diseases, plant diseases, fruits condition, vegetable conditions, clothes color, clothes spot, industrials products quality, etc., can be easily detected through the system. In addition, with the implementation of augmented reality (AR) and virtual reality (VR) technology, this system can be more effective, appropriate, visible, and reliable. As it is an image processing technique, output images are easily connected and converted with augmented reality to improve and analyzing outputs. It may be

presumed that in the near future, there will be much more space for more study and development. The cost of the system is also very low; for that reason, it is affordable for users. With advanced technology and rapid innovation, the system can be more improved. This system can make a revolutionary change in the field of object detection. The field of using this system is unbounded. It can easily identify or detect any visible diseases; for that reason, it can also be used in the medical sector. If it is properly trained and utilized, it will be a very useful device for human beings. Improving it and adding some features, it can be turned into a useful multiple tasker system.

References

1. Debonne, N., van Vliet, J., Metternicht, G., & Verburg, P. (2021). Agency shifts in agricultural land governance and their implications for land degradation neutrality. *Global Environmental Change*, 66, 102221.
2. Angelakts, A. N., Zaccaria, D., Krasilnikoff, J., Salgot, M., Bazza, M., Roccaro, P., Jimenez, B., Kumar, A., Yinghua, W., Baba, A., & Harrison, J. A. (2020). Irrigation of world agricultural lands: Evolution through the Millennia. *Water*, 12(5), 1285.
3. Vlek, P. L., Khamzina, A., Azadi, H., Bhaduri, A., Bharati, L., Braimoh, A., Martius, C., Sunderland, T., & Taheri, F. (2017). Trade-offs in multi-purpose land use under land degradation. *Sustainability*, 9(12), 2196.
4. Anubhove, M. S., Ashrafi, N., Saleque, A. M., Akter, M., & Saif, S. U. (2020). Machine learning algorithm based disease detection in tomato with automated image telemetry for vertical farming. In *2020 International Conference on Computational Performance Evaluation (ComPE), 2020 July 2* (pp. 250–254). IEEE.
5. Lamnabhi-Lagarrigue, F., Annaswamy, A., Engell, S., Isaksson, A., Khargonekar, P., Murray, R. M., Nijmeijer, H., Samad, T., Tilbury, D., & Van den Hof, P. (2017). Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control*, 43, 1–64.
6. Valero Sarmiento, J. M. (2017). *Injectable capsules for physiological monitoring on animals*.
7. Barwick, J., Lamb, D., Trotter, M., & Dobos, R. C. (2017). *On-animal motion sensing using accelerometers as a tool for monitoring sheep behaviour and health status*.
8. Zhang, S., Zhu, C., Sin, J. K., & Mok, P. K. (1999). A novel ultrathin elevated channel low-temperature poly-Si TFT. *IEEE Electron Device Letters*, 20(11), 569–571.
9. Valentin, M. T., Casnor, R. B., Fanwa, J. C., & Dangan, V. S. (2017). Design and fabrication of a pyramid-type plant bed hydroponics of Romaine Lettuce production under lowland condition. *CLSU International Journal of Science & Technology*, 2(2), 1–7.
10. Van Zwieten, M., Stovold, G., & Van Zwieten, L. (2007). *Alternatives to copper for disease control in the Australian organic industry* (A report for the Rural Industries Research and Development Corporation, RIRDC Publication, Vol. 7, p. 110).
11. Cheng, Z. *Ecology of urban lawns: The impact of establishment and management on plant species composition, soil food webs, and ecosystem functioning* (Doctoral dissertation, The Ohio State University).
12. Hicks, S. (2006). *Re-farming the new world order: Imaginations of agriculture in global America*. Vanderbilt University.
13. Ahmed, S. M., Al-Amin, M. R., Ahammed, S., Ahmed, F., Saleque, A. M., & Rahman, M. A. (2020). Design, construction and testing of parabolic solar cooker for rural households and refugee camp. *Solar Energy*, 15(205), 230–240.
14. Metev, S. M., & Veiko, V. P. (2013). *Laser-assisted microtechnology*. Springer Science & Business Media.

15. Balducci, F., Impedovo, D., & Pirlo, G. (2018). Machine learning applications on agricultural datasets for smart farm enhancement. *Machines*, 6(3), 38.
16. Arkeman, Y., Utomo, H. A., & Wibawa, D. S. (2015, August). Design of web-based information system with green house gas analysis for palm oil biodiesel agroindustry. In *2015 3rd International Conference on Adaptive and Intelligent Agroindustry (ICAIA)* (pp. 238–244). IEEE.
17. Amanda, E. C., Seminar, K. B., Syukur, M., & Noguchi, R. (2015). Development of expert system for selecting tomato (*Solanum lycopersicum* L.) varieties. In *2015 3rd International Conference on Adaptive and Intelligent Agroindustry (ICAIA) 2015 Aug 3* (pp. 278–283). IEEE.
18. Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. *Frontiers in Plant Science*, 27(8), 1852.
19. Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A., & Bhansali, S. (2019). Machine learning techniques in wireless sensor network based precision agriculture. *Journal of the Electrochemical Society*, 167(3), 037522.
20. de Luna, R. G., Dadios, E. P., Bandala, A. A., & Vicerra, R. R. (2020). Tomato growth stage monitoring for smart farm using deep transfer learning with machine learning-based maturity grading. *Agrivita*, 42(1), 24.
21. Franchetti, B., Ntouskos, V., Giuliani, P., Herman, T., Barnes, L., & Pirri, F. (2019). Vision based modeling of plants phenotyping in vertical farming under artificial lighting. *Sensors*, 19(20), 4378.
22. Zeyad, M., Biswas, P., Iqbal, M. Z., Ghosh, S., & Biswas, P. (2018). Designing of micro-controller based home appliances governor circuits. *International Journal of Computer and Electrical Engineering (IJCEE)*, 10(2), 94–105.
23. Zeyad, M., Ghosh, S., & Ahmed, S. M. (2019). Design prototype of a smart household touch sensitive locker security system based on GSM technology. *International Journal of Power Electronics and Drive Systems*, 10(4), 1923.

Augmented Reality and Virtual Reality Creating Ripple in Medical and Pharmaceutical World



Siddharth Singh, Pankaj Kumar, Hema Chaudharya, Riya Khurana,
and Ankit Shokeen

Abstract Augmented reality (AR) and virtual reality (VR) emerged as a highly significant affordable and efficient approach in various sectors of health care and medicine such as surgery, diagnostic imaging, medical education, medical training, rehabilitation, nursing, telemedicine, palliative care, therapeutics, and patient care management. In virtual reality (VR), a user interacts and manipulates virtual objects of the virtual world with help of specialized VR devices like display screens. Augmented reality (AR) is a semi-true image, i.e., combination of a real scene viewed by a user and a virtual scene generated by a computer which augments the scene with additional information. The major applications of AR and VR in the healthcare sector are as in surgery, autism, diagnostic imaging, medical education and phantom limb pain, pharmaceuticals production, and so on. In early 1990s, VR technology was first utilized for visualizing medical data during surgery and also for pre-operatively planning surgical procedures. One of the major areas where this approach is highly promising is in plastic surgery and dental surgeries as surgery consequences are directly connected to the external appearance of the patient. Autism affects the nervous system and overall cognitive, emotional, social, and physical health of the person. It impairs the ability of patient to interact and communicate as a normal person. To overcome this issue, AR technology is used by “medical school of Stanford University” under the “Autism glass project.” They use Google Glass AR technology in order to help autistic children interpret others emotions and expecting them to build social relationship as normal people without the requirement of Google glass in future. AR and VR approach are widely used in diagnostic imaging such as in magnetic resonance imaging (MRI) and positron emission tomography (PET) scanning. The chief reasons for using AR and VR technology in

S. Singh (✉)

School of Health Sciences, University of Petroleum and Energy Studies, Dehradun, India

P. Kumar

Department of Applied Sciences, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

H. Chaudharya · R. Khurana · A. Shokeen

Faculty of Pharmaceutical Sciences, PDM University, Bahadurgarh, Haryana, India

diagnostic imaging are enhanced viewing, depth perception, and improvised human-machine interface (HMI). “Anatmage table” is a virtual anatomical table platform for anatomy studying and teaching where detailed structures of each part of human body can be visualized. This can also help in explaining a patient his/her pre-operative plan in order to make them understand their own surgical conditions better and thus improves patient compliance. Phantom pain sensations might include crushing, toe twisting, pins and needles, hot burning feel, etc. It is managed by using AR technology, and this approach lets the patients to see a virtual limb appearing on the screen which moves in the same way as the patient move their amputated limb. This helps in achieving a therapeutic effect by permitting a patient to control their amputated limb with their brain. This article is going to provide overviews on the various applications of AR and VR technology in medical field as well as pharmaceutical sector especially in pharmaceutical production and marketing domain. This is clearly evident that this approach has strong potential of emerging as a fundamental and highly efficient aspect of health care in future.

Keywords Surgery · Pharmaceuticals · Autism · Diagnostic approaches

1 Augmented Reality (AR)

It refers to a field of computer research which combines the real-world and computer-related data. It superimposes the virtual objects with real world and thus augments the physical world environment with virtual (fake or imaginary) world and leads to betterment in skills and senses of human. In simple terms, it acts by adding digital knowledge and details in the physical world [1, 2].

The first AR system named “virtual fixtures” was developed by Louis Rosenberg in 1991, when he was working for US Airforce Armstrong labs [2].

One of the most used examples of AR is the score board displayed during live telecast of sports matches. Also, the replay of any scene of the match during live telecast is also performed using AR as it displays a view of real world with some graphical symbol (e.g., use of a line for depicting path of the ball during cricket match) produced by a computer [3].

The book “Understanding Augmented Reality: Concepts and Applications” by Alan B. Craig defines AR as a mediator between humans and humans, computer and humans, and humans and computer [4, 5].

It should be noted that, it is not necessary that AR will always have some real counterpart.

Augmented reality attributes to three fundamental properties which are as follows:

1. It combines actual and imaginary world.
2. It shows real-time interactions.
3. The virtual (imaginary) and real (actual) are registered in 3D [6, 7].

AR covers the sensorial information such as auditory sensation, virtual sensation, haptic (touch-based) sensation, and olfactory (smell-based) sensation.

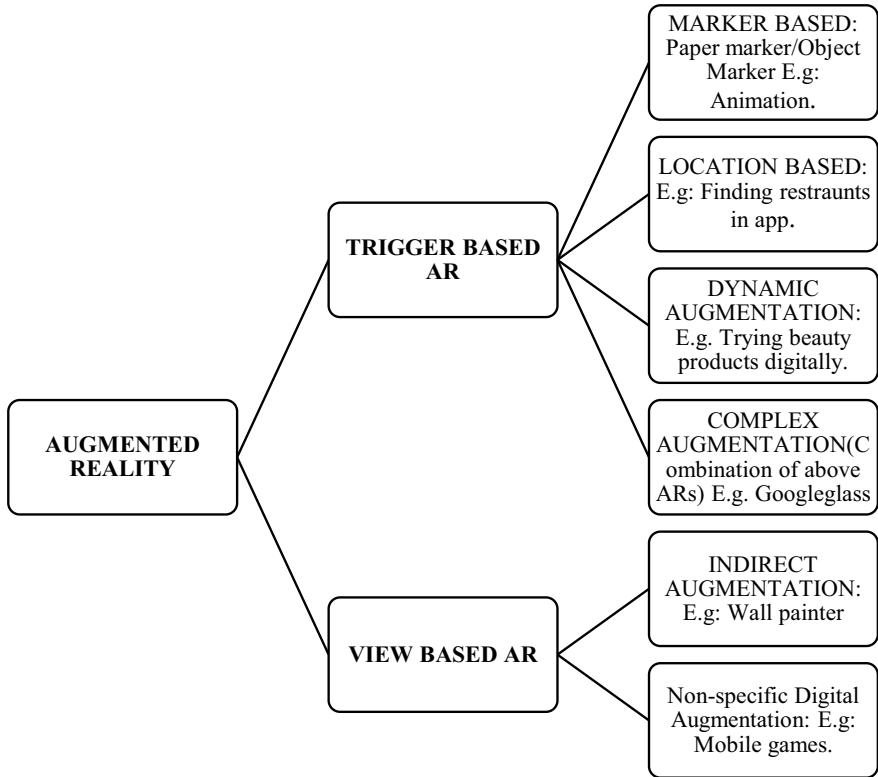


Fig. 1 Types of augmented reality

Augmented reality approach gained huge importance as it allows a person to carry out real-world’s tasks using virtual information.

Augmented reality is broadly classified into two categories [8, 9]:

- A. Triggered augmented reality technology: This involves the use of stimulus or trigger to begin augmentation.
E.g., Object/papermarkers, location, etc.
- B. View-based augmentation: This involves indirect and non-specific digital augmentation that provides a steady perspective of an environment.

These two categories are further subclassified as (Fig. 1).

2 Virtual Reality

Virtual reality came into existence in 1950s, and in 1987s, Jaron Lanier brought the term virtual reality into lime light. Virtual reality refers to a medium consisting of

computers generated simulations that are able to augment or replace the real world and gives the feeling of being immersed in a simulated world/virtual world to the user, i.e., a user can operate the virtual environment in actual time (physical world) [7].

Virtual reality incorporates various technologies, e.g., artificial intelligence (AI), image processing, computer graphics, sound systems, networking, etc. The feedback of VR is felt via sensory perception such as haptic, auditory, and visual perception, similar to AR [10, 11].

The fundamentals of virtual reality technology

The three main fundamentals of virtual reality are:

- (a) **Imagination:** it is also known as presence, i.e., a user thinks of the scenarios that does not exist and feels himself/herself as a part of it which occurs due to manipulation of sensory information by the system.
- (b) **Interaction:** it is a mode by which a user connects with the system. This communication in VR is oftenly aided by feeding (3D) mediums examples: head mounted device (HMD), space ball, etc.
- (c) **Immersion:** According to an article named “Brief Introduction of Virtual Reality & its Challenges” published in the International Journal of Scientific and Engineering Research, Volume 4, Issue 4, April-2013, Immersion is defined as “the extent to which high fidelity physical inputs (e.g., light patterns, sound waves) are provided to the different sensory modalities (vision, audition, touch) in order to create strong illusions of reality in each.” [4–6] (Fig. 2).

3 Types of Virtual Reality Systems

Based on intensity of emersion and the kind of interface being used in the VR system, VR system is classified into three types:

1. Non-immersive VR system:

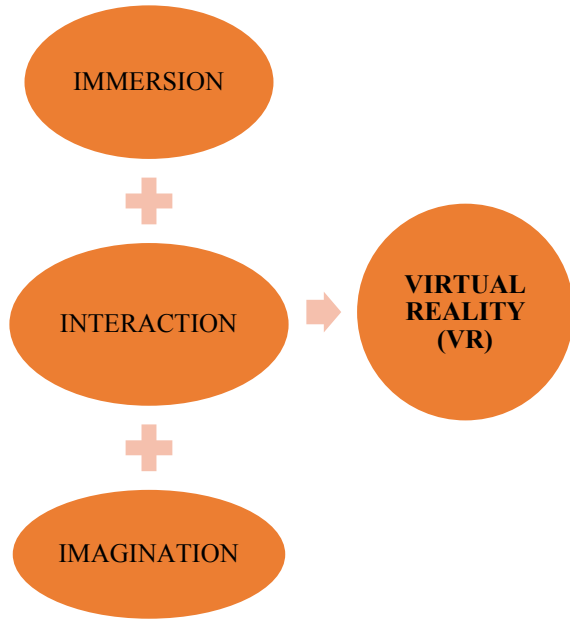
It is also known as “Desktop VR system” or “window on world (WoW)” system. In this system, a user sees the artificial environment either through a single or multiple computer display screens. In this the user connects with the 3D environment with the help of components such as a conventional monitor or a stereo display for displaying images, glasses, space ball, data gloves, keyboard, etc. This system has minimum cost due to the use of unsophisticated elements. This system finds its major role in computer-aided design (CAD) and modeling.

NOTE: In this a user do interact with the 3D environment but he/she is not immersed in it.

2. Semi-immersive VR system:

It is known as “Hybrid System.” This system also uses a conventional monitor system but provides high level of immersion and may also makes use of some model (physical model). Semi-immersive VR system supports head tracking and creates the illusion of being there. It is used in the driving simulator.

Fig. 2 3 Is of virtual reality



3. Immersive VR System:

This system provides the most elevated levels of immersion, i.e., the user feels himself as an integral part of the virtual world. The key elements of this VR systems include HMD (provides a stereoscopic view on the basis of the user’s position and orientation), tracking devices, data gloves, etc. It finds its major role in the walk-through, fly-through, and look-around simulations in the artificial world [12–15].

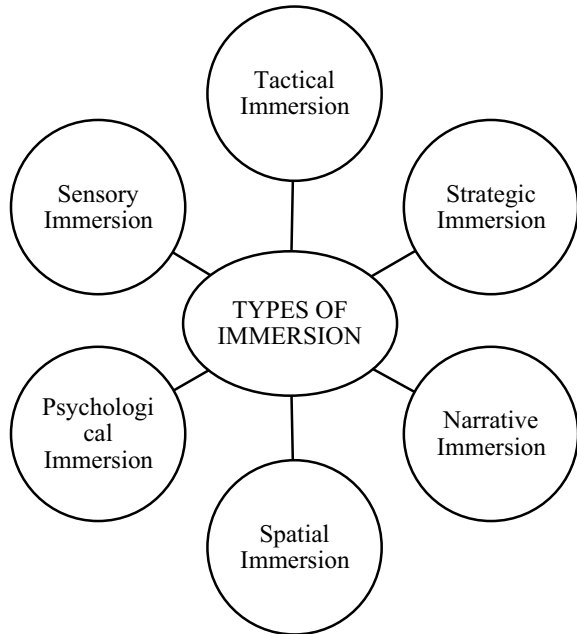
4 Types of Immersion

Immersion in virtual reality are classified as [5] (Fig. 3).

5 Applications of Augmented Reality and Virtual Reality in Medical and Pharmaceutical Field

Augmented reality (AR) and virtual reality (VR) emerged as a highly significant affordable and efficient approach in various sectors of health care and medicine such as surgery, diagnostic imaging, medical education, medical training, rehabilitation, nursing, telemedicine, palliative care, therapeutics, and patient care management. Thus, this approach has strong potential of emerging as a fundamental and highly efficient aspect of health care in future.

Fig. 3 Immersion in virtual reality



1. **AUTISM SPECTRUM DISORDER (ASD)**

Autism spectrum disorder or ASD refers to a set of abnormalities that impairs the development of nervous system (neurodevelopment abnormalities) such as issues in social interaction and communication and behavioral problems like confined interests and recurring behaviors. Generally, these abnormalities start in once early years of childhood, i.e., even before they turn three but this may go unnoticed until their caretaker (guardian, parents, etc.) observes them having abnormal communication and social behavior issues [16].

There are various reasons which might be responsible for autism such as:

- (a) **Malnutrition:** Adequate amount of nutrition is required for proper functioning of both body and brain. The important essential nutrients which are generally deficient in an autism patient includes the following (Table 1):
- (b) **Increased nutritional requirement:** This may occur due to alteration in biochemical pathways and genetic mutation or it can also occur due to poor GI absorption [9, 17].
Example: In some autism patient, methylenetetrahydrofolate reductase enzyme (MTHFR enzyme) gets mutated and leads to increased requirement of folate.
- (c) **Food intolerance:** A considerable subgroup of patients with ASD is intolerant to foods generally casein (major source: dairy products) and gluten (major source: wheat)

Table 1 Types of nutrients deficiency

Nutrients (might be deficient in ASD patient)	Example	Role in body
Vitamins	Vitamin B1, B2, and B3	Helps in optimum microbial function
	Vitamin B6	Enzymatic reactions such as conversion of: (i) Tryptophan to serotonin (ii) Homocysteine to glutathione (iii) Glutamate to gamma-aminobutyric acid (GABA)
	Folic acid and vitamin B12	Proper working of methylation cycle
	Vitamin A	Antioxidant, vision including eye contact function
	Vitamin D	Bone health, cognition, and immune function
Minerals	Zinc	(i) Cofactor for more than 200 enzymatic reactions (ii) Immune functions (iii) Transport of vitamin A (iv) Amino acid metabolism (v) Metallothionein functions (vi) Taste perception
Omega-3 fatty acids	Eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA)	(i) Anti-inflammatory function (ii) Maintenance of intestinal and skin health (iii) Cognitive and vision functions

NOTE: Common allergies are: IgE-mediated (immunoglobulin-E mediated) but individuals with ASD might have IgG-mediated (immunoglobulin-G-mediated allergies) which are able to show up to 72 h after a person has consumed intolerant food for his body.

Mechanism by which food affects normal brain function:

A very common mechanism by which normal brain functioning is affected in individuals in ASD is known as “leaky gut.” Leaky gut refers to changed permeability of the intestine due to which the undigested molecules which were earlier prevented from going into circulation because of tight junctions now gets excess to the circulation. When undigested peptide chains of casein and gluten enters the blood stream, they act as “false neurotransmitters” as both neurotransmitter and undigested molecules of casein and gluten are peptide in nature. Since they act as false neurotransmitters they can cross BBB and impair brain function [16, 17].

NOTE: Digestion of casein and gluten produces opiate like peptides which is digested with the help of enzyme dipeptidyl peptidase-IV (DPP/IV) but autism patients are prone to damage of this enzyme and thus these peptides may enter into circulation (through leaky guts) and impairs normal brain functioning.

Other major factors that can lead to impaired brain functioning includes:

- Intestinal dysbiosis (imbalance in the intestinal flora, i.e., reduction good bacteria excessive growth of harmful bacteria example: yeast and clostridia)
- Mitochondrial disfunction
- Defect in detoxification
- Issue in methylation and transsulfuration cycle
- Inadequate oxytocin hormone (significant for making social connections)
- Immune imbalances (e.g., increase in cytokines, increased serum, IgG, IgG2, IgG4, and autoimmunity, etc.)
- Increased oxidative stress [18]

AR and VR play a vital role in managing autism spectral disorder as it helps in development of communications skills by creating such an environment where patients connect with virtual objects or avatars for communication. This patient who otherwise struggles in managing real-life scenarios can now get control to the environment and social interactions, and this intensifies their self-confidence.

Another advantage of being using such a step is that there is no need for interference for another person for the person of interaction and thus an autistic child gets enough time to ponder over various ways by which he can react to a situations in different condition. Such an approach is very appreciable as it is capable of generating specifically simulated environment according to specific requirements to various autistic patient. The reaction of ASD patient to various stimuli in a controlled simulated environment helps the physician to better understand and analyze the condition (symptoms) of the patient and thus broadens the possibility of improve diagnosis and medication.

Some successful examples where AR and VR technology helped patients with autism includes:

- Michelle Wang and Denise Reid in their research article named “Using the Virtual Reality-Cognitive Rehabilitation Approach to Improve Contextual Processing in Children with Autism” showed that virtual reality when combined with conventional cognitive rehabilitation [VR/CR approach (virtual reality/cognitive rehabilitation approach)] approaches helped in improving contextual processing ability and cognitive functions.
- Medical school of Stanford university used AR technology under the “Autism glass project.” They use Google glass AR technology in order to help autistic children interpret others emotions and expecting them to build social relationship as normal people without the requirement of Google glass in future.

2. **SURGERY**

Surgery is derived from a Greek word, “cheiros” meaning hand and “ergon” meaning work. Surgery can be defined as the branch of medicine where manipulations in body structure are carried out by doctors in order to diagnose, alleviate,

and treat injuries, disease, and other medical conditions. Trepanation or trephination is the oldest known (around 10,000 B.C., i.e., even before the human can write or read) surgical procedure which involves boring and cutting rings or squares in skull bones to expose Dura mater in order to treat conditions related to intracranial pressure (pressure exerted by Cerebrospinal fluid inside the skull and on the brain tissues) [19].

In early 1990s, VR technology was first utilized for visualizing medical data during surgery and also for pre-operatively planning surgical procedures [20].

(a) **AR and VR for future surgeons:**

Very few surgeons are allowed during real surgery but at the same time knowing correct and deep knowledge of the anatomical structure of the surgical target is very critical for the successful surgery. So, to train new and unskilled surgeons, VR and AR technology can be used to provide 360 view in order to experience virtually real surgical operation but in controlled and comfortable environment. One of the major areas where this approach is highly promising is in plastic surgery and dental surgeries as surgery consequences are directly connected to the external appearance of the patient [6].

So, AR and VR not only considerably improves the training exposure of the new surgeons but also it avoids the patient exposure with surgeons which are in early phase of their course. Using AR and VR, it is possible to allow new surgeons to experience the emergency life threatening in virtual world so that they should not panic while dealing with such a scenario in real world.

(b) **AR and VR for real-time surgeries:**

Nowadays, acceptance for minimal invasive surgery (MIS) is increasing over conventional surgical procedures for many different pathologies.

By now, it is well known that AR and VR are capable of simulating real-time conditions and also helps in analyzing complex images by reconstructing them in 3D. This made possible the use of AR and VR in surgical planning and navigations [21].

For example:

- (i) **LAPAROSCOPIC SURGERY:** It is one of the very complex surgical procedures as it requires very skilled hand movements to handle instruments for manipulation, incision, and sewing of tissues. There are several laparoscopic simulative available which use AR and VR technology combined with real-world scenarios, things, surgical tools, and responses which improves performance efficiency, safety of the patient and reduces the time consumed in the operation theater. Some examples of laparoscopic simulative use are:

Pro MIS, computer/enhanced laparoscopic training system (CELTS), LTS3/e, Blue DRAGON [22].

- (ii) **PLASTIC SURGERY:** using AR and VR technology, it is possible to simulate physical world surgical condition and thus leads to the improvement of diagnosis and surgical analysis. So this concept became very significant in case of plastic surgeries as surgical consequences are directly connected to external appearance of the patient.
For example, Kovler et al. demonstrated hip fracture/femur fracture reduction surgery through a haptic device and a 3D eye wear incorporating immersive screen and a user tracking system for efficient elaboration.
- (iii) **DENTAL SURGERIES:** The use of dental implants as a treatment for missing tooth is widely accepted over other methods as implants can efficiently support artificial teeth, i.e., dental prosthetics. For implant placement, a hole is to be made into the jawbone which requires experience and skills as well as careful concern toward the distance between artificial root and nerve canal as failure might cause osteonecrosis, i.e., bone death. The surgical instruments, tracking system, medical images, and computer can be combined using AR and VR technology for carrying out successful dental surgeries by navigating them in real time and ultimately help surgeons with various tasks such as plate fixation/drilling, e.g., a markerless AR-based technology was proposed by Wang et al. in their article named “Video see-through augmented reality for oral and maxillofacial surgery” for supporting oral and maxillofacial surgery. It complemented a 3D patient teeth model with a real patient’s teeth for tracking the position in a real-time video.

3. **PHANTOM LIMB PAIN (PLP) MANAGEMENT/TREATMENT:**

Ambrose Pare, a French military surgeon of sixteenth century, was the first who explained the condition and perception of phantom limb pain (PLP) whereas the term phantom limb pain was coined by a civil war surgeon named Silas Weir Mitchell in 1872. Phantom limb pain refers to the condition in which a patient feels pain in that part of limb which they have lost, i.e., in the limb which no longer exists.

Types of pain associated with amputation. One must understand that phantom sensation, phantom pain, and residual limb or stump pain are the major categories of pain associated with amputation but they all differ from each other [14] (Fig. 4).

1. **Phantom Sensation:**

- It refers to the non-painful sensations felt by an individual in the missing limb. The individuals feel as if their missing limb is still present. It involves paresthesia such as prickling “pins and needles,” hotness or coldness in skin, electric sensations, and fatigue.
- Onset: Immediately after surgery.

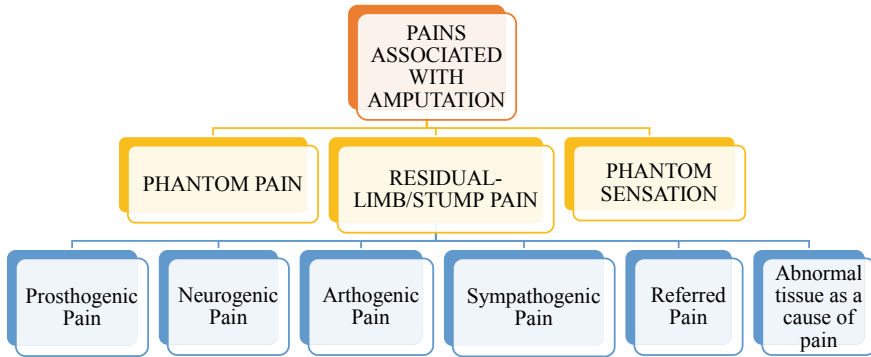


Fig. 4 Type of pain associated with amputations

2. **Residual limb/Stump pain:**

It refers to the pain in residual limb, i.e., the region adjacent to their missing limb and not in the missing limb. Generally, the patient gets relief from this type of pain with the healing of site of incision.

3. **Phantom Pain:**

- It refers to the pain felt by an individual in their amputated limb. It includes burning, shooting, cramping, throbbing, squeezing, and other such sensations.
- Incidence of phantom limb pain:

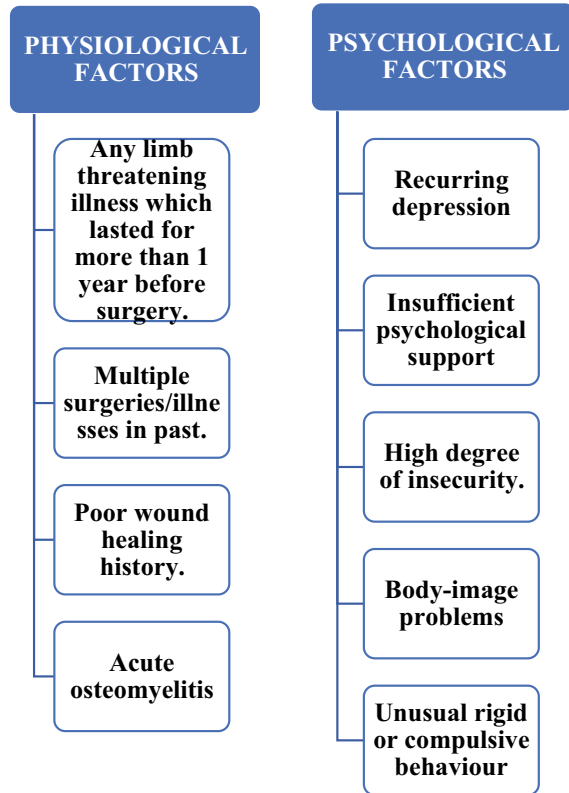
Proximal amputation > Distal amputation

- Onset: Within one week after surgery.
- The conditions which can intensify the phantom pain includes:
 - Emotional stress
 - Cold/local irritants exposure
 - During orgasm (in females)
 - After orgasm (in males)
 - During bladder catheterization
 - Early use of prosthesis
 - Heat (Fig. 5)

Treatment using AR and VR:

It is managed by using AR and VR technology, and this approach lets the patients to see a virtual limb appearing on the screen which moves in the same way as the patients move their amputated limb. This helps in achieving a therapeutic effect by permitting a patient to control their amputated limb with their brain, and thus, treatment using AR and VR technology is more engaging and can improve patients adherence to the treatment.

Fig. 5 Factors which might increase the incidence of phantom limb pain



F. Bach et al. developed a combination of home training system based on AR and fMRI based on VR.

VR-based fMRI: In this mirror training was carried out in a magnetic resonance tomograph; the objective for the development of this system was to analyze the effect on cortical activation during long-term home training. This system consisted of:

- (a) Head mounted display (HMD) build with fiber glass and relevant interfaces
- (b) Data gloves made of plastic optical fibers (POF) sensors
- (c) A PC to operate KISMET (software for simulation and visualization)

The amputee gets immersed into the simulated virtual environment which incorporates a 3D human model created in the MRI scanner specially focusing on the hands. Wearing the HMD and the data glove the amputee masters the command on motion with his/her undamaged hand.

AR-based ARKIT for home training: ARKIT (AR system based at KIT) was developed by F. Bach et al. build a strong system to achieve high level of immersion and subjective reality. This system enabled 3D demonstration of the physical world around the amputee by video see-through incorporating a mirrored hand using image processing methods. The system consists of:

- (a) Stereo camera system
- (b) Binocular stereoscopic HMD
- (c) A desktop computer with high performance to operate the image processing software

The device is quite user-friendly as it allows adjustment of angles and interpupillary distance to appropriately allowing the center of eye wear according to user needs.

The use of ARKIT can be highly efficient as the physician can make the amputee practice-specific exercises according to his/her specific needs, and this can significantly improve the patient condition.

4. **PHOBIAS:**

The term phobia is derived from a Greek word, “Phobos” which means error, panic, fear, etc.

Issac M. marks in his book “Fear and Phobias” defines phobia as a special kind of fear which

- (a) Is out of proportion to demands of the situation
- (b) Cannot be explained or reasoned away
- (c) Is beyond volunteer control
- (d) Leads to avoidance of the feared situation

Phobic patients or phobia is not same for all but its common to all, and the acceptance is rational or by involuntary actions. Simple considered phobias are like from heights, water, etc., but the complex phobias are considered those as of where there are asymmetry in conditions such as cancer or other serious condition disease which are or can be avoided by the patients. Some phobias which arise by themselves cannot be avoided which include phobia of getting diseased, death, being injured or getting its own harmful intentions. The common indications or physiological signs of phobia are: sweating, tremor, nausea, rapid breathing, tightness of chest, diarrhoea, etc.

There are three components of any phobia:

- (a) A subjective in a state sensed by the person
- (b) Outer outcomes visible to observer
- (c) Physiological changes

Treatment of phobias using AR and VR is one of the promising approaches. Whether AR or VR or both will be used for managing a certain type of phobia depends on many factors such as the type of phobia, conditions available and intensity of immersion required, e.g., if the following two criteria are met, then AR can be preferred over VR.

- I Patient can use real elements such as hands and feet for interaction with the application.
- II Possibility of using or reproducing the real environment with minimal investment or time or possibility of using some environmental substitute.

Table 2 Difference between traditional treatment and AR and VR treatment

Traditional treatment	AR and VR treatment
The patient faces actual objects from which he/she fears so there is a possibility that these objects, such as insects, might hurt the patient	In this approach, the objects used for overcoming a patient fear are artificial (virtually made) which assure patient safety
In this approach, patients are taken to the places from which he/she is fears or such scenarios are recreated in some other place and this makes the therapy complex and tedious	In this approach, the physician can create any scenario virtually at any point of time during the therapy. So, the exposure of the patient to the scenarios from which he/she fears is relatively very easy as compared to the other traditional method of treatment
The sequence in which the patient fearing elements occurs cannot be predicted in natural environment element occurs from which the patient fears	The physician not only controls but also generates all the patient fearing elements according to him. So, everything such as the sequence in which these elements appear in front of the patient is all in the hands of physician
The patient safety cannot be fully guaranteed	The patient safety can be fully guaranteed as the elements used are fake
The site where the patient is taken for getting rid of the fears might be a public place which might make him uncomfortable and chances of getting panic attack increases	The site where the treatment is going to be done is selected by the physician, and thus, he can satisfy all the aspects of the therapy according to him and the individual

VIRTUAL REALITY EXPOSURE THERAPY (VRET)

This approach creates an artificial environment where a patient can overcome his fears by interacting with virtual counterparts of real-world element. It offers the advantage of engagement of several senses of the patient, intensifies the fidelity and getting quick response to their fears, and also the sessions are carried out in a safe place everything under complete control of the physician which avoids chances of panic attack in the patient due to embarrassment which is not possible in conventional imaginal exposure therapy and in vivo exposure therapy [23]. This therapy can be used for treatment of various phobias such as acrophobia (fear of heights), glossophobia (fear of public speaking), and aerophobia (fear of flying) (Table 2).

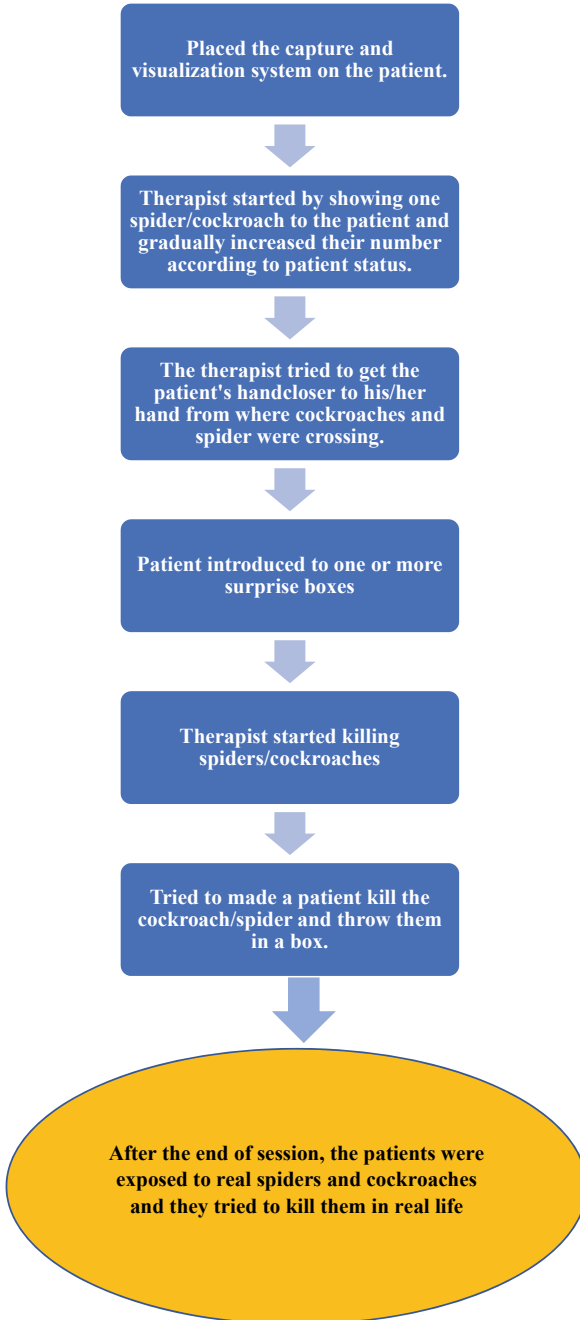
6 Case Studies-Based Applications of AR and VR in Health care

(i) CASE STUDY 1: PHOBIAS USING AR TECHNOLOGY

Let us look at a system made by Juan M.C. et al., using augmented reality to treat phobia of spiders and cockroaches. The system was programmed using C. For development environment, it used visual C++ version and for including AR options, it used AR Toolkit 2.65 with VRML support. They used

one cockroach and three spiders in size small, medium, and large as virtual objects [24–26].

In the exposure session, following steps were employed:



It was found that this exposure session using AR made many patients confident enough that they interacted with their fears in real world and even were able to kill spiders and cockroaches after the session.

(ii) **CASE STUDY 2: PHANTOM LIMB PAIN MANAGEMENT USING AR TECHNOLOGY**

Similarly, Francesco Carrino et al. created a system for management of phantom limb pain incorporating the conventional mirror therapy with augmented reality technology. V. S. Vilayanur Ramachandran invented mirror therapy for phantom limb pain in 1996, which turns the pain into relief in many cases of amputee patients and gives the experience of presence of lost limb as in this patient has to perform an exercise in which he/she keeps his amputated limb behind the mirror, and the mirror gives the delusion of presence of both limbs by reflecting the intact limb. Restricted movements for the patients and availability of only a finite type of exercises were some of the limitations of conventional mirror therapy which Francesco Carrino et al. tried to overcome using “augmented reality mirror therapy.” It intensifies the immersion and allowed the amputees interaction with objects in the virtual world which enabled them to perform various activities such as playing games and complex exercises (movements passing the center of the body) [14, 18].

(iii) **CASE STUDY 3: VIRTUAL REALITY IN NORMALISING EATING PATTERN OF A BULIMIA NERVOSA PATIENT**

Bulimia nervosa is a type of eating disorder characterized by binge eating in alliance with loss of control followed by purging [27]. In a case report described by María Roncero and Conxa Perpiña, a 17 years old girl showed improvement in the management of bulimia nervosa using virtual reality complemented with cognitive-behavioral therapy (CBT). She went through seven sessions of VR simultaneously with Phase 1 and Phase 2 CBT sessions. The VR therapy was comprised of a kitchen environment having two areas including all the required elements such as fridge, stove, cabinets, and table, which were easily excessable or blockable according to the patient desire. The environment also allows other activities such as addressing positive and negative thoughts, with a making call and listening to music, etc., loaded with a feature for the therapist to analyze the displayed behavior or decision of the patient. A noticeable improvement was observed in her thoughts of bingeing and purging which showed a positive impact in normalizing her eating pattern and thus, in managing bulimia nervosa [28].

Other major areas where augmented reality and virtual reality technology are used in pharmaceutical and medical and healthcare field are:

1. **Diagnostic imaging:** AR and VR approaches are widely used in diagnostic imaging such as in magnetic resonance imaging (MRI) and positron emission tomography (PET) scanning. The chief reasons for using AR and VR technology in diagnostic imaging are enhanced viewing, depth perception, and improvised human-machine interface (HMI).

2. **Medical education:** In conventional education system, anatomy, physiology, etc., are taught using pictures, videos, or models but its key limitation was unclear explanations and inability of students to practically understand a patient. But this hindrance can be avoided by using AR and VR technologies, e.g., “Anatomege table” is a virtual anatomical table platform for anatomy studying and teaching where detailed structures of each part of human body can be visualized. This can also help in explaining a patient his/her pre-operative plan in order to make them understand their own surgical conditions better and thus improves patient compliance.
3. **Treating eating disorders:** Technology like virtual reality is emerging as a new platform with a great potential of managing eating disorders like bulimia nervosa and binge eating disorders. Assessment through VR helps us in identification of body image concerns of a patient and also in the identification of particular food or environment responsible of stimulating binge-purging cycle. VR showed betterment in self-esteem and body image issues in such patients [29–32].

7 Conclusion

In this era, technologies are playing massive roles in our lives; so, AR and VR are technologies that are emerging in healthcare field really very fast because they have the potential to improve the healthcare services. This not only helps in the treatment management and improving the surgery efficiently also, education through AR and VR platform will make the platform more interesting for student, and the created real-time experience environment can significantly help in building up their skills. Thus, it can be concluded that AR and VR have a promising future in medical and pharmaceutical world with a wide range of applications such as education, training, diagnosing, palliative care, surgery, and patient care management.

References

1. Ferrari, V., Klinker, G., & Cutolo, F. (2019). Augmented reality in healthcare. *J Healthc Eng., I*(2019), 9321535. <https://doi.org/10.1155/2019/9321535>. PMID:31354932;PMCID:PMC6636465
2. Eckert, M., Volmerg, S. J., & Friedrich, M. C. (2019). Augmented reality in medicine: systematic and bibliographic review. *JMIR Mhealth Uhealth*, 7(4). <https://doi.org/10.2196/10967>
3. Samadbeik, M., Yaaghobi, D., Bastani, P., Abhari, S., Rezaee, R., & Garavand, A. (2018). The applications of virtual reality technology in medical groups teaching. *Journal of Advances in Medical Education & Professionalism* 6(3), 123–129. PMID: 30013996; PMCID: PMC6039818.
4. Criag, B. *Understanding Augmented Reality Concepts and Applications*. Book chapter-1 What is Augmented Reality.
5. Criag, B. *Understanding Augmented Reality Concepts and Applications*. Book chapter-2 Augmented Reality Concepts.

6. Khor, W. S., Baker, B., Amin, K., Chan, A., Patel, K., & Wong, J. (2016). Augmented and virtual reality in surgery-the digital surgical environment: Applications, limitations and legal pitfalls. *4*(23), 454. <https://doi.org/10.21037/atm.2016.12.23>. PMID:28090510;PMCID:PMC5220044
7. <https://www.arvrhealth.com/index.php/category/virtual-reality-in-healthcare/>.
8. Douglas, B. D., Venets, D., Wilke, C., Gibson, D., Loitta, L., Petricoin, E., Beck, B., & Douglas, D. Augmented reality and virtual reality: initial success in Diagnostic Radiology, chapter-2; 2018 edition of State Of The Art Virtual Reality and Augmented Reality Knowhow.
9. Fridhi, A., Benzarti, F., Frihida, A., & Amiri, H. (2018). Application of virtual reality and augmented reality in psychiatry and neuropsychology. *Particular in the Case of Autistic Spectrum Disorder (ASD) Neurophysiology*, *50*(3). <https://doi.org/10.1007/s11062-018-9741-3>.
10. Sherman, R., & Criag, A. (2019). *Understanding Virtual Reality* (2nd ed.) Part-1 Introduction to virtual reality.
11. Burdea, C., & Coiffet, P. (2003). *Virtual Reality Technology* (2nd ed.) 1st Introduction.
12. Douglas, D.B., Venets, D., Wilke, C., Gibson, D., Liotta, L., Petricoin, E., Beck, B., & Douglas, R. Augmented reality and virtual reality: Initial success in diagnostic radiology. <https://doi.org/10.5772/interchopen.74317>. <https://books.google.co.in/books?hl=en&lr=&id=mHqQDwAAQBAJ&oi=fnd&pg=PA5&dq=augmented+reality+and+virtual+reality+in+diagnostic+imaging&ots=JPNQ-pgjFW&sig=t8b8M-TsmaKkHAaJ8QfSfbD2lc#v=onepage&q=augmented%20reality%20and%20virtual%20reality%20in%20diagnostic%20imaging&f=false>
13. <https://www.statnews.com/2019/08/16/virtual-reality-improve-surgeon-training/>.
14. Dunn, J., Yeo, E., Moghaddampour, P., Chau, B., & Humbert, S. (2017). Virtual and augmented reality in the treatment of phantom limb pain: A literature review. *NeuroRehabilitation*, *40*(4), 595–601. <https://doi.org/10.3233/NRE-171447> PMID: 28211829.
15. <https://www.fi.edu/science-of-augmented-reality>.
16. Compart, P. J. (2013). The pathophysiology of autism. *Global Advances in Health and Medicine*, *2*(6), 32–37. <https://doi.org/10.7453/gahmj.2013.092>. PMID:24416706;PMCID:PMC3865380
17. Hsieh, M. C., & Lee, J. J. (2018). Preliminary study of VR and AR application in medical and healthcare education. *Journal of Nursing and Health Studies*, *3*(1:1).
18. Wiggins, L. D., Levy, S. E., Daniels, J., Schieve, L., Croen, L. A., DiGuseppi, C., Blaskey, L., Giarelli, E., Lee, L. C., Pinto-Martin, J., Reynolds, A., Rice, C., Rosenberg, C. R., Thompson, P., Yeargin, M., Young, L., & Schendel, D. (2015). Autism spectrum disorder symptoms among children enrolled in the study to explore early development (SEED). *Journal of Autism and Developmental Disorders*, *45*(10), 3183–3194. <https://doi.org/10.1007/s10803-015-2476-8>. PMID:26048040;PMCID:PMC4573234
19. Kim, Y., Kim, H., & Kim, Y.O. (2017). Virtual reality and augmented reality in plastic surgery: A review. *Archives of Plastic Surgery*, *44*(3), 179–187. <https://doi.org/10.5999/aps.2017.44.3.179>. Epub 2017 May 22. PMID: 28573091; PMCID: PMC5447526.
20. Ellis, H. (2002) *History of Surgery*. Book chapter-1 reprint2002.
21. Monsky, L.W., James, W., & Seslar, S.S. (2019) Virtual and augmented reality applications in medicine and surgery. *The Fantastic VoyageReview Article*, *9*.
22. Jour, B., Sanne, M. B. I., & Jakimowicz, J. J. (2008). What is going on in augmented reality simulation in laparoscopic surgery? 1432–2218. <https://doi.org/10.1007/s00464-008-0144-1>.
23. Francesco, C. F., Rizzotti, D., Gheorghe, C., Bakajika, P. K., Francescotti-Paquier, F., & Mugellini, E. *Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review*. <https://doi.org/10.1007/s11920-017-0788-4>
24. Marks, I.M. *Fears and Phobias*. Book chapter-1 Introduction.
25. Marks, I. M. *Fears and Phobias*. Book chapter-2 section-I and section-II.
26. Juan, M.C., Botella, C., Baños, R., Alcañiz, M., Guerrero, B., & Monserrat, C. (2005). *Using Augmented Reality to Treat Phobias*. <https://doi.org/10.1109/MCG.2005.143>.
27. de Carvalho, M. R., de Santana Dias, T. R., Duchesne, M., Nardi, A. E., & Appolinario, J. C. (2017). Virtual reality as a promising strategy in the assessment and treatment of bulimia nervosa and binge eating disorder: A systematic review.

28. María, R., & Conxa, P. (2015). Normalizing the eating pattern with virtual reality for bulimia nervosa: A case report. *Revista mexicana de trastornos alimentarios*, 6(2). Tlalnepantla jul./dic. <https://doi.org/10.1016/j.rmta.2015.11.001>.
29. Desselle, R., Brown, A., James, R. R., Midwinter, A. J., Powell, M. K., & Woodruff, S. A. Virtual and Augmented Reality Applications in Science and Engineering.
30. Sherman, R., & Criag, A. (2019). *Understanding Virtual Reality* (2nd ed.) Part-2 VR medium.
31. Burdea, C., & Coiffet, P. (2003). *Virtual Reality Technology* (2nd ed.) Input devices: Trackers, Navigation, and Gesture interface.
32. Tao, Z., Kun, Y., Xiong, J., He, Z., & Wu, S.-T. (2020). Augmented Reality and virtual reality displays: perspectives and challenges. *iScience*, 23(8), 101397. <https://doi.org/10.1016/j.isci.2020.101397>. ISSN 2589-0042.
33. Bush, J. Viability of Virtual Reality Exposure Therapy as a Treatment Alternative. <https://doi.org/10.1016/j.chb.2007.03.006>.
34. Riva, G., Wiederhold, B.K., Molinari, E. Virtual environments in clinical psychology and neuroscience book chapter virtual reality: Possibility of virtual reality for mental care.

Autonomous Drones for Medical Assistance Using Reinforcement Learning



Billy Jacob, Abhishek Kaushik, Pankaj Velavan, and Mahak Sharma

Abstract In recent years, usage of the unmanned aerial vehicle in various fields has developed rapidly. Drones have the potential to be reliable medical delivery platforms for microbiological and laboratory samples, pharmaceuticals, vaccines, emergency medical equipment and patient transport. Nowadays, time-sensitive items are most commonly delivered by UAVs, such as medicine, food, or deliveries that would be difficult with traditional services. However, using human resources to operate these UAVs needs a lot of time and investment in training. Mostly, drones use GPS technology to navigate to the location of the customers. A huge number of drones in the airspace create a need for a drone traffic management system to mitigate collision risk. The drone traffic management system again requires human resources and huge investments. To simplify this, we propose to model UAVs' autonomous navigation, which follows the bike lanes in the left end of roadways. This research proposes a framework by using reinforcement learning to allow the UAV to navigate successfully from the start location to the end location by following the bike lanes present on the roads. This concept can be implemented in areas where the UAVs are used to do repetitive transportation of medicines and supplies. As a part of this work, we have used three different reinforcement learning algorithms. Q -Learning, SARSA and Deep Q -Learning were the three algorithms used to train a reinforcement learning agent which in turn was used to operate the drone in a real-world grid environment. The algorithms were evaluated, and the results were tabulated in this paper.

Keywords Drones · Medical purpose · Reinforcement learning

B. Jacob · P. Velavan

Dublin Business School, 13/14 Aungier Street, Saint Peter's, Dublin, Ireland

A. Kaushik (✉)

Adapt Centre, Dublin City University, Dublin 9, Dublin, Ireland

e-mail: abhishek.kaushik2@mail.dcu.ie

M. Sharma

The Maharaja Sayajirao University of Baroda, Baroda, India

1 Introduction

In the COVID times and during natural calamities, unmanned aerial vehicles (UAVs) are employed in various medical transportation fields such as delivery of food, delivery of medicines, emergency materials etc. UAV space is a thriving space. Shakhathreh et al. [1] suggests that drones could potentially allow large-scale deliveries of blood, long-tail medicines, medical samples and even organs. According to Konert et al. [2], the use of drones for medical purposes brings many advantages, such as quick help and shortening the time of travelling to the patient. The research done by Scott and Scott [3] talks about two companies UPS and Zipline who are working on a drone network to deliver vaccines and blood to 20 clinics in remote locations in Rwanda. They also illustrate a scenario from January to March 2016, where DHL's third-generation Parcelcopter tested delivery of over 130 parcels of urgently needed medicines or sporting goods between automated Skyports in two Bavarian Alpine villages. As drones' usage increases in the day-to-day life of the industries mentioned above, there comes another challenge: drone traffic management and drone pilots.

1.1 *Current Drone Traffic Management Systems and Their Limitations*

At present, there are many companies which are working on developing UAV traffic management (UTM) systems. All these organisations are working alongside national space agencies of their respective countries to create one system or a cluster of systems which will coexist. The current legislation of most of the countries all around the world are very stringent with respect to UAV usage in cities and public places. This section gives an overview of the related research works that are carried out in this technological area and their limitations. In the context of Amazon Air, Burzichelli [4] points out the legislation in USA framed by the FAA for small unmanned aircraft systems which does not allow drone operations when it overshoots the visual line of sight of the drone operator. With respect to the Irish laws, flying and operating drones in Ireland is subject to European Union Regulation 2019/947 which too does not allow drone operations when it overshoots the visual line of sight of the drone operator. Apart from legislation, there are also privacy and security issues related to UAV operation in cities. The study conducted by Li et al. [5] outlines the various security and privacy issues and proposes regulating the drone's altitude and its on-board camera capability. When multiple drones fly in the same air space for completing their respective tasks, drone collisions may occur. An article written by Ackerman [6] reported an unfortunate incident in Switzerland where a 10kg delivery drone crashed near a group of kindergarten children and raised questions about the safety of urban delivery drones. The study conducted by Zeng and Lei [7] talks about the UAV collisions when a swarm of UAVs are flying in the same airspace for different tasks and proposes a framework to calculate the minimum distance between drones

using the flight angle and speed by considering the drone anti-collision problem as a geometric method. Another industry which operates drones in populated areas is movie industry, and Mademlis et al. [8] listed the current challenges in autonomous UAV cinematography ranging from legal, ethical and safety factors to operational factors and suggests the future researchers to take this factors into consideration before developing a full autonomous drone for cinematography. Drones are used in medical industry where they are used to deliver medicines to remote locations. The study conducted by Jeyabalan et al. [9] made a survey by taking interviews from 16 individuals from nine different countries working on the front lines of drones for health programs and listed down all the benefits and the concerns surrounding the implementation of the drones in the healthcare sector. Similarly, a study conducted by Alwateer and Loke [10] outlined the challenges and the societal issues accompanying the emerging drone services. The study also reinforces the need for insurance policies so that incidents like physical loss and third-party damage can be covered.

1.2 Solutions for Drone Traffic Management System

The drone traffic management system intends to allow multiple drones to securely and efficiently assign airspace by aggregating the data of relevant controlled airspace and drones flying in that airspace. In 2015, the European commission mandated the Single European Sky Air Traffic Management Research Joint Undertaking (SESAR JU) to produce a blueprint on how to make drone use in low-level airspace safe. The SESAR proposed the U-space framework European Union (2019) to support safe, secure and efficient access to airspace for large number of drones. They have decided to roll out the services in four phases from 2019 to 2050. The current phase is foundation service where the users are provided registration, identification and geofencing facilities. The fourth phase will have automation, connectivity and digitisation for both the drone and the U-space system. Any traffic management system has to solve the multi-agent path finding (MAPF) problem. The paper authored by Stern et al. [11] introduced unified terminology to describe MAPF problems and put out common benchmarks and evaluation procedures for evaluating the MAPF algorithms. A research conducted by Ho et al. [12] proposed an extended formulation of multi-agent path finding using conflict-based search and enhanced conflict-based search to address the movement of heterogeneous drones. The research conducted by Samir Labib et al. [13] also emphasised on the importance of path planning to avoid collision and proposed an UAV Air Traffic Management (UTM) system with the help of IOT. The study conducted by Park et al. [14] described the essential functions of an UTM architecture and demonstrated a working UTM system using four UAVs with the LTE-based wireless communication board and a UTM server and also analysed the data flow between UTM server, UAVs and UAV operators. The architecture proposed by Gharibi et al. [15] to solve the problem of drone traffic management consists of a unified framework which has two important components, zone service providers (ZSP) and drones. Both these components are connected through cloud.

The ZSP provides the navigation information between any two drones in their designated zone, and the drone navigates autonomously with the help of the data provided by the ZSPs. The study conducted by Jiang et al. [16] reviewed the current UTM practice from industry partners like Amazon and Google and proposes a physical architecture for an UTM system which can handle a large number of heterogeneous UAVs with all the required components. Design preferences for several of the services mentioned above will depend on how many drones fly in a specific air space. It, therefore, becomes essential to estimate how many delivery drones would operate in a typical city. The research work performed by Doole et al. [17] proposed a framework for estimating the traffic densities of drone-based package delivery systems from five EU countries and concluded that any city should employ a robust airspace management system to realise commercial drone delivery.

1.3 Reinforcement Learning Solutions for Drone Traffic Management

Reinforcement learning is one of the three machine learning paradigms, alongside supervised learning and unsupervised learning. Reinforcement learning is an area of machine learning concerned with how an agent reacts with actions in an environment based on the current state of the environment with the objective to maximise the cumulative reward for each action. The advantage of this learning technique is that the machine autonomously learns to follow the correct instructions without providing any conditional statements. Numerous reinforcement learning solutions have been attributed for developing an UAV traffic management system. The study conducted by Brittain and Wei [18] proposed a deep multi-agent reinforcement learning framework to enable autonomous air traffic separation in en-route airspace. In this study, each aircraft is represented by an agent, and flight speed selection, altitude selection and flying angle selection are considered as the actions of the agent. The research work done by Munoz et al. [19] applied deep reinforcement learning for autonomous drone delivery. They did the test in the AirSim simulator in a neighbourhood scenario with a plenty of obstacles. The solution uses the double deep Q network, and this algorithm selects or decides the action of the agent which maximises the resultant Q -value. It is vital for the drones with autonomous capability to navigate in new unknown locations as well. Pham et al. [20] proposed a technique to train a quadcopter to learn to navigate to the target point using a Q -learning algorithm in an unknown environment. The study conducted by Wang et al. [21] proposes using only GPS signal and sensory information of the local environment to navigate autonomously and intelligently from arbitrary departure places to arbitrary target positions in a virtual environment by modelling the navigation problem as a control problem and employed deep reinforcement learning to solve it. Researchers at Facebook and the University of California, Berkeley, published a new research paper Belkhale et al. [22] where they showed how a drone can use

AI to learn to handle different delivery payloads under changing conditions. They proposed a meta-learning approach. They used a neural network dynamics model, which takes as input the current state and action and predicts the next state. The study conducted by Wei et al. [23] proposed a deep reinforcement learning model for target searching in an unfamiliar environment. He suggests that research work in such areas can improve the target searching during disaster times where humans can be located in an unfamiliar and adverse environment. The research conducted by Wu et al. [24] proposed a navigation assistance system for quadcopter with Deep Reinforcement Learning. They used a simulating environment with the obstacles on the flight path. They trained the model in flight path with obstacles for 500 flights and used that same model to fly the quadcopter in a new environment with obstacles for 1000 flights and calculated the collision percentage with the obstacle. Yijing et al. [25] solved the UAV obstacle avoidance problem using the Q -learning algorithm where the movement of the UAV is considered as an action and walls, blocks, bricks and traps were taken as the obstacles for training the model.

1.4 Evaluation of Autonomous Drones

Evaluation is a significant part of any research work, and it is important to verify if any developed machine learning model or an algorithm working as per the expectation. The research conducted by Sujit et al. [26] compared five different UAV path following algorithms and compared the algorithms using two parameters, total control effort and total cross track error. Sholes [27] proposed a methodology to categorise the UAVs based on the autonomy capacity of the UAVs. Zhang et al. [28] proposed UAV path evaluation method based on RE-ITOPSIS. Dawnee et al. [29] compared Dijkstra, A^* , Artificial potential field and RRT algorithms and evaluated them using shortness of path, clearance to obstacles, deviation from straight line path and smoothness of path and tabulated the results. The study conducted by Thomas and Brunskill [30] compared various evaluation metrics which is used for off-policy evaluation of reinforcement learning. Jordan et al. [31] says that reproducibility analysis have shown that there is an inconsistency in replicating evaluated performance results. They proposed a comprehensive evaluation methodology for a wide range of reinforcement learning algorithms SarsaParl2, Sarsa (Lambda) and Actor-Critic algorithms.

2 Conceptual Framework

As a part of this research work, we have developed a conceptual framework for drone path planning which can be used by the medical delivery drones to navigate in the cities autonomously utilising the bike lanes at the far ends of the motorways and roadways. The framework which we are proposing can be implemented on all the

drones which are going to be used in the delivery process in cities. We are assuming that the bike lanes will be at both the sides of the roadways and flying height of a drone will be fixed to 20m from the ground. The drones will have GPS technology and an on-board camera. The reinforcement learning model will assist the drones in following the bike lanes in the roadways based on the training provided to the model with the help of the GPS coordinates of the bike lanes in the cities so that the lateral movement of the drone can be reduced to a very great extent. Reduction in the lateral movement of the drones will normalise the drone traffic and reduce the probability of the collision with the other drones, which will be flying in the same air space. The conceptual framework consists of three components: the positioning system of the drone, the lane positioning system and the reinforcement learning component.

2.1 Lane Positioning System

The Global Positioning System (GPS) is a free service that is owned and operated by the U.S. government and is always available. A lot of research works have been done on autonomous navigation of UAVs using GPS. The study conducted by Kwak and Sung [32] proposed a flight path planning technique using GPS way points where the UAV flight paths are planned by setting the way points one by one manually by UAV operators or pilots. We will be using the GPS coordinates to locate the bike lanes. The bike lanes are mostly fixed and the corresponding GPS coordinates are also a constant value which does not change in a short span of time. Let us consider Fig. 1. Let us assume that there is a repetitive task like a medical delivery to a particular

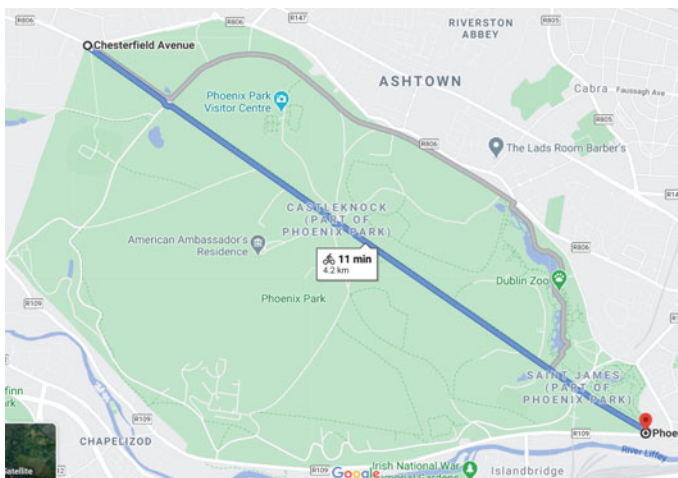


Fig. 1 Positioning of bike lanes using GPS



Fig. 2 GPS points on bike lanes

hospital that needs to be carried out using a drone daily; the position of the lane can be acquired using the GPS coordinates on the bike lanes as shown in Fig. 2.

2.2 Drone Positioning System

As part of this research, we assume that the UAVs which are going to be used for the medical delivery services have GPS on-board. An idea was proposed by Saha et al. [33] to survey remote areas using first-person view technology and GPS in drones. They say that the surveillance drone’s location will be given exactly by the GPS. They also say that GPS is used for stabilising and automatic path planning for the drones. The GPS on-board the UAV will have the ability to determine the location and pinpoint the UAV position during the flight. The geographical position of the UAV is essential to figure out the current and the previous state of the UAV, which is a vital part of reinforcement learning. We will discuss the different states and actions in the upcoming chapter when talking about reinforcement learning (Fig. 3).

2.3 Reinforcement Learning Component

The Reinforcement learning component is the proposed back end module that feeds the drone with the actions to take while flying from the origin to the target location (Fig. 4). It is usually employed by various software systems, machines and robots all around the world to find the best possible behaviour or path it should take in a specific situation. As part of this, we have used three different reinforcement learning

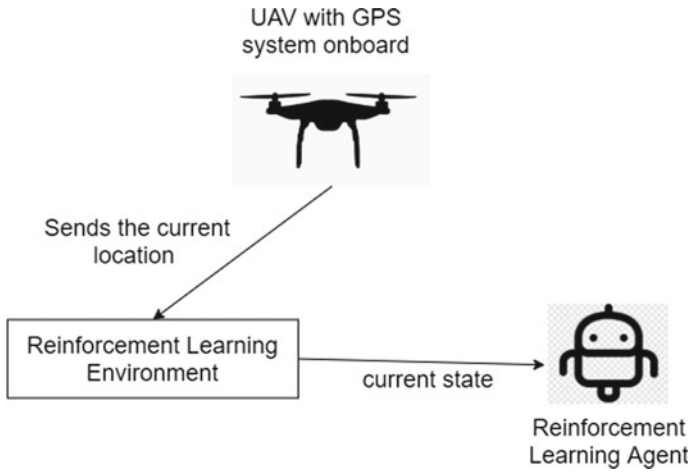


Fig. 3 State detection of drones

algorithms. The three models were trained in gym environment and deployed in a real-world grid setup and the results were compared. The compared results are listed in the Evaluation phase.

1. Model-free reinforcement learning.
2. Model-based reinforcement learning.

2.4 Working Methodology

The working methodology contains the training phase and the deployment phase.

Training phase An open AI gym environment will be created where the different states will be the pre-determined or the fixed GPS coordinates. The actions would be the capabilities of the drone like moving forward, moving backward, moving right, moving left etc. There will be positive rewards if the drone reaches the expected states (prefixed coordinates) and negative rewards if the drone moves out of the expected path. A question might arise how we will determine the current state, i.e. how will we determine whether the drone is following the proper direction or not. This is where the drone positioning system will play its part. A proper reward function will be used to provide rewards to the agent whenever the agent takes an action. The q values will be generated for every action taken by the agent. The Q table will be enriched for the Q learning and SARSA models, whereas in deep Q -learning, the approximated Q values will be calculated using the neural network and the best policy will be determined. The best policy determined from these three algorithms will be used to operate in the UAV in the deployment phase.

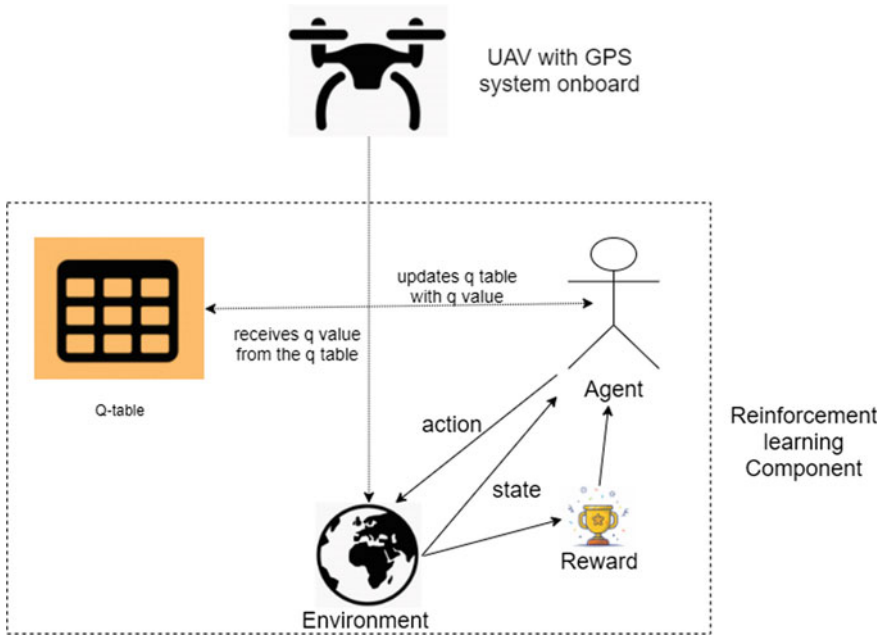


Fig. 4 Conceptual framework diagram

Deployment phase In the deployment phase, there will be no new learning occurring as the best policy is already determined in the training phase and the training environment was a simulation, and now, we have a real environment to deploy to. There are no general rules covering deploying all possible RL agents to production, as there is a huge variety of RL code and approaches. Sometimes, values like distances will be scaled down in the simulation or training environment. This needs to be modelled accurately to get the maximum performance out of the model. In the off-policy Q learning, on-policy SARSA and model-based reinforcement learning and deep Q Learning, the deployed model will be in the form of a Q table where the columns correspond to the states and the rows correspond to the action. First, the agent retrieves the current position of the drone based on the coordinates of the drone. Once the state is retrieved, then the next step is to get the best action to be taken for that state from the deployed Q -table. Once the action is decided, the agent again sends the action to the UAV and the UAV executes that action. This is continuously done until the end is reached or as specified by the controllers or the operators.

3 Model-Free and Model-Based Reinforcement Learning

3.1 *Model-Free Reinforcement Learning*

In reinforcement learning, a model-free algorithm is an algorithm that does not use the probability distribution and the reward function blended with the Markov decision process. A model-free RL algorithm is trial and error algorithm. The agent repeatedly tries all the available actions until it enriches the Q -table with the q values generated from the rewards and the hyperparameters.

Policy A policy is a function which maps the actions and the states which infers that it is the behaviour by an agent in a specific state. A model-free reinforcement learning has two types of policies.

On-policy reinforcement learning In the on-policy reinforcement learning, the agent collects the rewards using the currently learned policy. The agent will use the collected rewards (experience) to optimise the same policy. SARSA is an example of on-policy reinforcement learning.

Off-policy reinforcement learning In the off-policy reinforcement learning, the agent follows no specific policy to collect the rewards. The agent randomly expedites through the various available actions in the environment and collects the rewards provided by the environment and enriches the table during the training phase.

3.2 *Model-Based Reinforcement Learning*

In reinforcement learning, a model-based algorithm is an algorithm that uses a model of the actual environment using the previous experience of the environment. It does not use the reinforcement learning environment as a black box that provides rewards whenever the agent takes action. This model incorporates knowledge of the task's characteristics and the transitional probabilities when moving from one state to another and the different state outcomes. A classic example of model-based reinforcement learning is deep Q -learning.

Deep Q -learning In Q learning, the agent enriches a matrix which it can refer later to maximise the reward in the long run. But one of the disadvantage of this approach is that it is only feasible for very small environment where the random expedition of actions are possible.



Fig. 5 Real-world grid setup model diagram

4 Implementation

4.1 Real-World Grid Setup

Grid setup is a real-world simulation of the actual environment especially for training Q -learning algorithm in an unknown environment. In order to showcase and implement the above-discussed reinforcement learning algorithms in real-world, a grid setup was made on a wooden floor, as shown in (see Fig. 5). There is a 40cm distance between two squares, and the objective of the drone is to traverse through the path shown in Fig. 5 after training. As part of the research, we have not trained the drone in a real environment because of the following disadvantages.

Disadvantages

1. The current legislation in Ireland does not allow the use of autonomous drones in its airspace.
2. The drone, which is used as part of this research, does not have an inbuilt GPS. Drones with inbuilt GPS are expensive and priced three times higher than the drone that we are currently using.
3. High risk of damage to property and to people.
4. Training the drones in real-world bike lanes can be dangerous to the people using the roads and the people who are close to the roads.
5. Uncertainties, such as wind and other dynamics of the environment, present in the system is difficult to take into account.

4.2 *Drone Used*

For completing this research, we have used the DJI Tello EDU drone. Tello EDU is an impressive and programmable drone perfect for education. The drone can be controlled easily using programming languages such as Scratch, Python and Swift. With an upgraded SDK 2.0, Tello EDU comes with more advanced commands and increased data interfaces. The following are the specifications of the Tello EDU Drone. A lot of researchers have used tello drone for completing their studies and applying the machine learning models. The study conducted by Subash et al. [34] for detecting objects using UAV by deploying Mask-RCNN algorithm used the DJI tello drone and successfully completed their objective. Christl [35] solved the autonomous navigation of a MAV towards a visually marked landing platform in an obstructed indoor environment using supervised learning and he used the DJI Tello drone for his research activity (Fig. 6).

Drone Specifications

1. 720P HD transmission .
2. 5 MP photos.
3. 13-min Flight time.
4. Approximate hovering.

4.3 *Custom Open AI Gym Environment*

OpenAI is an AI research and deployment company. Open AI Gym is a great open-source tool for developing and comparing reinforcement learning models. It has various inbuilt environments which simulate the activity of a real-time environment. A reasonable reward mechanism was set for modifying the double deep Q -learning networks which made the algorithm suitable for training a self-driving car. The car was trained in the gym environment for many episodes. After training, the scale car was able to learn a good policy to drive autonomously, and they successfully transferred the learned policy from the gym environment to the real-world environment. For our research, we created a custom gym environment named gym drone. This environment consists of the get-current-state, get-reward, change-state, available-actions and the other required functions which simulate the real-time environment. It has ten distinct states. Each point in the grid has different coordinates and each coordinate points represent different states, and all the points outside the nine coordinate points are considered as state 10. If the agent picks a random action and if the coordinate point of the drone falls outside the nine coordinate points, then the current state of the drone is considered as state ten.

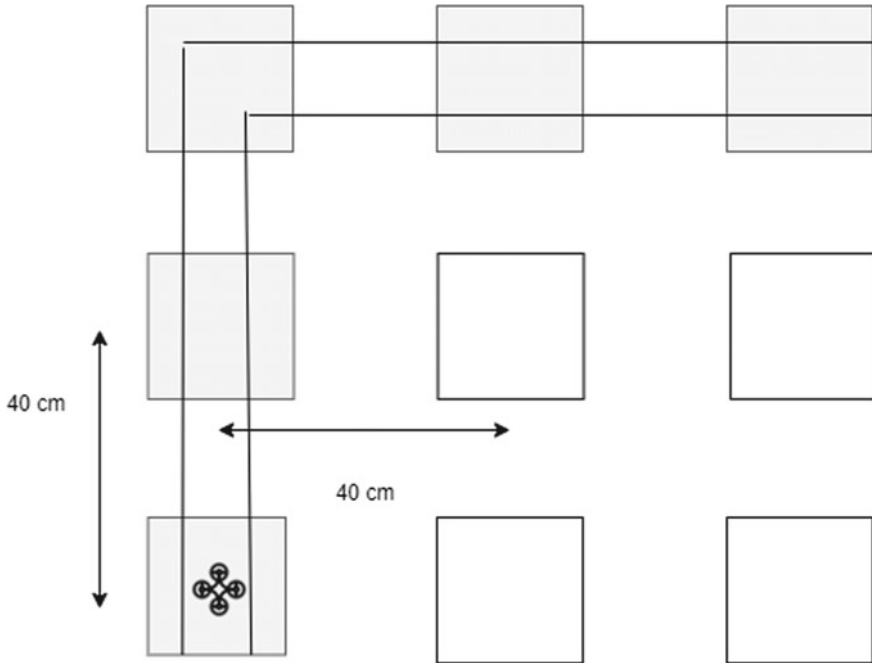


Fig. 6 Grid setup model diagram

4.4 *Q-Learning Implementation*

The reinforcement learning component is the proposed back end module that feeds the drones with the actions. The proposed reinforcement learning model is *Q*-learning, which is a model-free reinforcement learning algorithm to learn the quality of actions determining an agent what action to take under what circumstances [36]. The *Q*-learning model learns in an environment by trial and error using the feedback from actions and states. There are five basic components in *Q*-Learning namely agent, environment, reward, *Q*-table and action.

- **Agent:** The agent is a module that takes decisions based on positive and negative rewards. In this instance, it would be the module which makes the drone to move in four directions: forward, backward, right and left.
- **Environment:** The environment is nothing but a task or simulation. In this instance, the airspace, the drone movement module, the bike lanes in the roadways, the drone origin point and the drone target constitute the environment.
- **Actions:** the movement of the drone by the agent in any direction in the environment.

- States: following the bike lane in the roads, lateral movement of the drones, moving away from the bike lanes, collision with any external objects vertical take-off and vertical landing would be the different states that are available in the environment.
- Q -table: Q -Table is a simple matrix or a lookup table where we calculate the maximum rewards for an action at each state. The Q -table will guide the agent to the best action at each state. For example, a sample Q -table for an environment with 2 states and 3 actions will look like $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ where each row corresponds to a state and each column corresponds to an action.

$$Q(s, a) = Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (1)$$

where

Q = Q -table present in the Q -learning environment.

s = various boxes available for the drone to reach.

a = action, the selection of a certain direction to move by the agent in the environment.

$Q(s, a)$ = new q -value (value present in the q -table in s th row and a th column).

$Q(s', a')$ = old q -value in the q -table in s th row and a th column.

α = learning rate, set between 0 and 1,

γ = discount factor, also set between 0 and 1,

$\max_{a'} Q(s', a')$ = the maximum reward that is attainable in the state following the current state.

4.5 SARSA Implementation

State–action–reward–state–action (SARSA) is an on-policy temporal difference control method. Temporal difference algorithms update the knowledge of the agent on every single step rather than updating it after every single episode. The main difference between the Q -Learning and the SARSA is in the update function. In SARSA update equation, the Q -value is chosen using S' and A' , the next state and the action chosen in the next state, respectively.

$$Q(s, a) = Q(s, a) + \alpha[r + \gamma Q(s', a') - Q(s, a)] \quad (2)$$

where

Q = Q -table present in the Q -learning environment.

s = various boxes available for the drone to reach.

a = action, the selection of a certain direction to move by the agent in the environment.

$Q(s, a)$ = new q -value (value present in the q -table in s th row and a th column).

$Q(s', a')$ = old q -value in the q -table in s th row and a th column.

α = learning rate, set between 0 and 1,

γ = discount factor, also set between 0 and 1,

$\max_{a'} =$ the maximum reward that is attainable in the state following the current state.

4.6 Deep Q-Learning Implementation

The basic working step for deep Q -learning is that the initial state is fed into the neural network, and it returns the Q -value of all possible actions as an output. So in our case, there are ten available states and four available actions. The first step is to build the neural network model. Since the problem is relatively simple, our neural network model is also simple and only contains one hidden layer. The structure and the various components of our neural network model is provided below. The two networks are used to calculate the target and the predict function as shown in Fig. 7. Once in every n iterations, there will be a transfer of weights from one network to the another. This leads to more stable training because it keeps the target function fixed (for a while).

1. Input Layer—10 units since there are 16 states in the environment.
2. Output Layer—4 units since there are 4 actions that Agent can take.
3. Hidden Layer—10 units for demonstration.
4. Loss Function—Mean Squared Error.
5. Method for minimising loss function—Gradient Descent

$$Q(s_{t+1}, a_{t+1}) = Q(s_t, a_t) + \alpha[(r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

where

Q = Q -table present in the Q -learning environment.

s = various boxes available for the drone to reach.

a = action, the selection of a certain direction to move by the agent in the environment.

$Q(s_t, a_t)$ = new q -value (value present in the q -table in s th row and a th column).

$Q(s_t, a_t)$ = old q -value in the q -table in s th row and a th column.

α = learning rate, set between 0 and 1,

γ = discount factor, also set between 0 and 1,

$\max_{a'} =$ the maximum reward that is attainable in the state following the current state.

Fig. 7 Target predict function

$$\left[\left(\underbrace{r + \gamma \max_a Q(s', a'; \theta_i^-)}_{\text{Target}} - \underbrace{Q(s, a; \theta_i)}_{\text{Prediction}} \right)^2 \right]$$

Table 1 Four actions and the corresponding UAV movement

Actions	Column	UAV movement	Distance (cm)
Action 1	Column 1	Move forward	40
Action 2	Column 2	Move right	40
Action 3	Column 3	Move left	40
Action 4	Column 4	Move back	40

4.7 Autonomous Drone Navigation Implementation

This section tells about how the enriched Q -table which is obtained after training any of the three above-given models will be deployed and used for the autonomous navigation of the the drone from the origin point to the target point in real time. A Q -table is generated from SARSA model after just five episodes. In the Q -table, the rows correspond to the ten states and the columns correspond to the four actions. The action taken for every state is based on the maximum value present in the row. For example, if the UAV is at state zero and if we check the Q -table, we can see that the maximum Q -value available in that state is in the first row and hence the action 1 should be taken. The movement done by the drone for each action is provided in Table 1. The video of the actual footage captured while the drone autonomously navigated through the real-world grid setup is uploaded in YouTube and the can be accessed through the link.¹

5 Evaluation and Results

5.1 Q -Learning Evaluation

We can evaluate a reinforcement learning algorithm either by how good a policy it finds or by how much reward it receives while acting and learning. There are three main parameters in Q learning.

α —learning rate.

γ —discount factor.

ϵ —parameter which manages the trade-off between exploration and exploitation.

We will tune these hyperparameters and compare the results.

¹ <https://youtu.be/oaXt7R4YO90>.

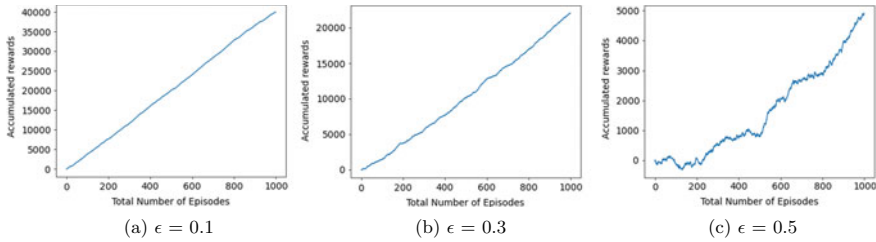


Fig. 8 Rewards for various Epsilon values

We generated graphs by plotting the total number of episodes in the x -axis and the total rewards accumulated in the y -axis. The model was executed for thousand episodes. The graph was generated for three values of Epsilon 0.1, 0.3 and 0.5 (see Fig. 8). From the graphs, we can see that when the epsilon value is less, the agent explores less and exploits the already gained knowledge to maximise the rewards; hence, there are very less amount of negative rewards. When the epsilon value is more, the agent explores more instead of exploiting the already gained knowledge, hence accumulated more negative rewards.

5.2 Tuning the Alpha Value

The learning rate or step size determines to what extent newly acquired information overrides old information. When there are more number of states and actions and the number of episodes are large under certain technical algorithms, the learning rate is decayed eventually starting with high learning rate in the beginning and reducing gradually. In our case as the number of episodes are very less, we have a fixed learning rate which is 0.1. But to explore the impact of learning rate in the model, we compared three alpha values 0.1, 0.4 and 0.8. We tabulated the number of times the agent successfully navigated from the origin point to the target point through the planned path.

5.3 Tuning the Discount Factor

The discount factor is a measure of how far ahead in time the algorithm looks. The discount factor essentially determines how much the reinforcement learning agents cares about rewards in the distant future relative to those in the immediate future. If the discount factor is zero, the agent will be completely myopic and only learn about actions that produce an immediate reward. As we do not have the distant reward incorporated in our reward structure, we will use the optimal discount factor value as 0.6.

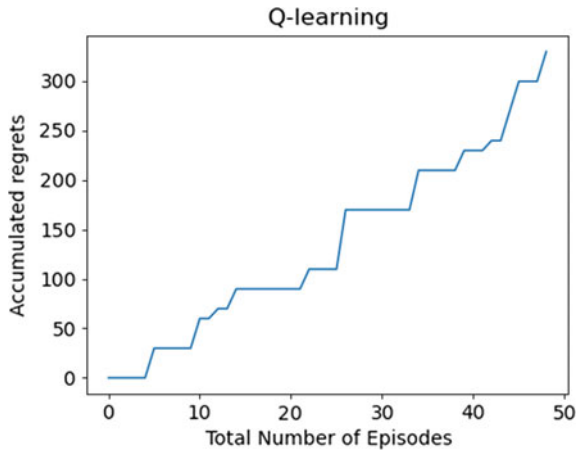


Fig. 9 Regrets accumulated by Q -learning algorithm after fifty episodes

Table 2 Number of successful navigation w.r.t. learning rate

α -value	Episodes	Total rewards	Successful navigation
0.1	1000	80,400	23
0.4	1000	80,500	32
0.8	1000	82,150	39

5.4 Calculating the Regret

The measure which is the expected decrease in reward gained due to the execution of the learning algorithm instead of behaving optimally from the very beginning is known as regret. The regret accumulated can be calculated at every episode and the total cumulative regret can be calculated. The cumulative regret can be plotted against the number of episodes and the quality of the q -learning policy can be evaluated. The total amount of regret accumulated by the Q -learning algorithm in the first fifty episodes is provided in Fig. 9. From the graph, we can understand that the flat part in the graph is produced when the agent is exploiting more and the steep increase in the regrets is when the agent is exploring more (Table 2).

6 SARSA Evaluation

To Evaluate the performance of the SARSA model and to show how it varies from the Q learning model in the reward accumulation front, we executed both the models for hundred episodes as the difference in execution is more visible during the initial

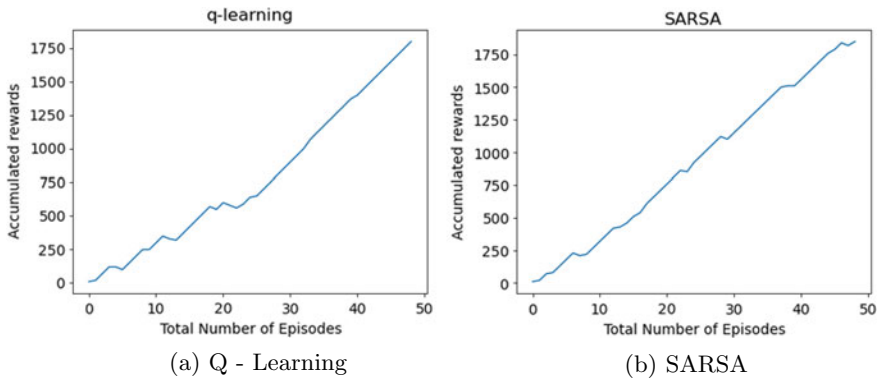


Fig. 10 Comparison of the enriched Q -table between Q -learning and SARSA within the first 5 episodes

episodes of the algorithm. The graph was generated between the number of episodes and the cumulative rewards accumulated (see Fig. 10). From the graphs, we can clearly see that the Q learning accumulated more negative rewards in the first half of the execution and less or nearly zero negative rewards in the second half of the execution, whereas the SARSA algorithm accumulated negative rewards throughout the course of the execution. This difference directly corresponds to the way both the algorithms work. The comparison between both the algorithms for the initial five episodes with respect to their individual Q -tables is provided in Fig. 10. Another evaluation metric is the regret. It is just the difference between optimal and actual reward. Every step that is taken by the algorithm in reinforcement learning has a return or reward associated with it. If you had taken the optimal steps at each trial, your total reward is called the optimal reward. Similarly, the actual reward earned can be computed. The difference is the regret. The total amount of regret accumulated by the SARSA algorithm in the first fifty episodes is provided in Fig. 11. From the graph, we can understand that the flat part in the graph is produced when the agent is exploiting more and the steep increase in the regrets is when the agent is exploring more.

7 Deep Q -Learning Evaluation

The deep Q learning is simply a Q -learning model with two neural networks. The evaluation for this model will be similar to the evaluation of the two other models which was discussed above. First, we will evaluate the reward accumulation capacity of the deep Q -learning model. We trained the model for 400 episodes and calculated the rewards accumulated by the model and plotted them in a graph as shown in Fig. 12. Here is an interesting thing to notice as it is the only algorithm among the

Fig. 11 Regrets accumulated by SARSA algorithm after fifty episodes

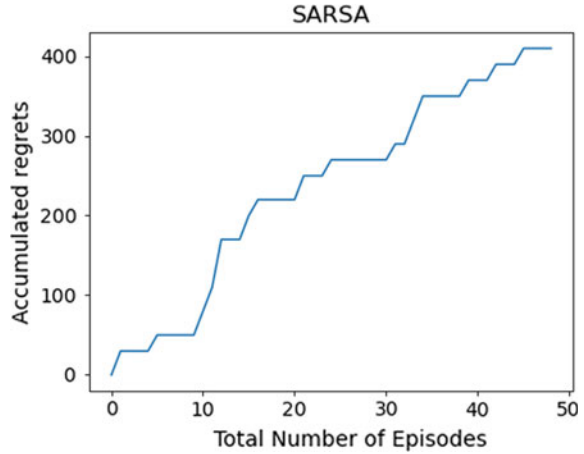
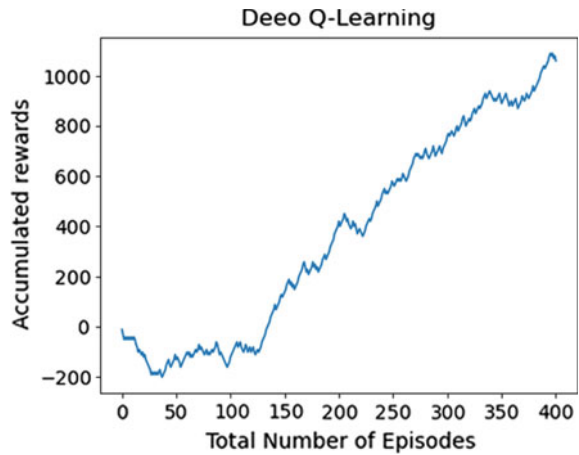


Fig. 12 Rewards accumulated by deep Q -learning algorithm after hundred episodes



three which accumulated cumulative negative rewards in the initial stages of the training and later accumulated cumulative positive rewards. This is due to the fact that it is the only algorithm which will traverse through points which are not traversed by the other algorithms. This can be understood if we go through the Q -table of the deep Q -learning model as shown in Fig. 13. This is the only table which is enriched completely and there are Q -values for all the states and the actions.

8 Shortness of the Path to Reach the Goal

This evaluation metric is usually done to find out which model took the shortest path from the origin to the target. But this metric is irrelevant in our thesis because the

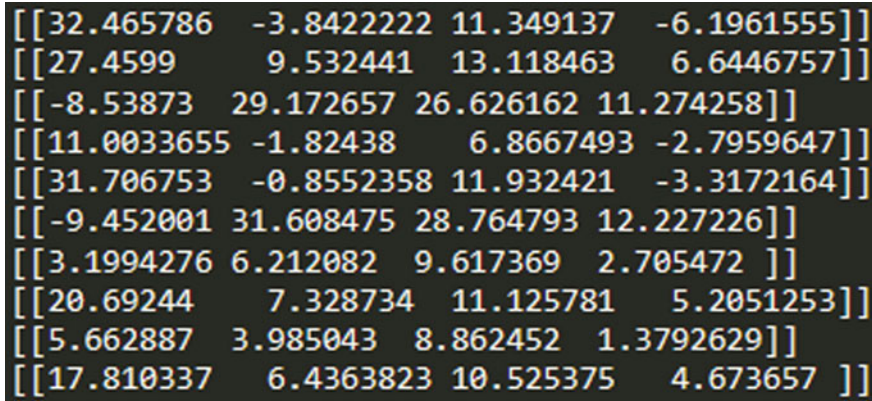


Fig. 13 Q table enriched after training the agent with deep Q learning

Table 3 Distance taken by the UAV to reach the target point

Algorithm	Steps taken	Distance covered (cm)	Successful navigation
Q -learning	4	160	Yes
SARSA	4	160	Yes
Deep Q -learning	4	160	Yes

path taken by all the reinforcement learning models was the same. This is because the agent was trained in such a way that it took only a single path rather than finding different paths. This is because one of the main objectives of this research is to limit the lateral movement of the UAV and thereby reducing the collisions with other drones. In Table 3 we compared the performance of the three algorithms taking into account the distance covered, the steps taken and whether they reached the final target point. The results were similar for all the three algorithms.

9 Smoothness of the Path

This evaluation metric was done to find out the model that took the smoothest path among the three algorithms. This evaluation is also irrelevant in this case because the UAV drone has the capacity to move in straight lines without any issue. But in real time, there were slight variations in the path followed by the UAV, and this was completely due to the fact that the UAV does not have an inbuilt GPS system which in turn affects the stability of the drone and has nothing to do with the efficiency of the algorithms. In Table 4 we compared the performance of the three algorithms taking into account time taken by the UAV to reach the target point, and it yielded similar results.

Table 4 Time taken to complete the task and smoothness of path

Algorithm	Steps taken	Time taken	Variation from target point (cm)
Q -learning	4	20 s	10
SARSA	4	20 s	7
Deep Q -learning	20 s	160 cm	-5

10 Time Taken to Complete the Full Flight

This evaluation metric is usually done to find out the time taken by the agents trained by the three individual algorithms to complete the specified task. This is also irrelevant to the evaluation of the algorithms because the algorithms cannot make any changes to the time factor of the actions. The number of step factor can be altered by the algorithms, whereas the time taken by the UAV to accomplish a certain task or an action is out of control of the algorithm. In Table 4 we compared the performance of the three algorithms taking into account smoothness of the path by finding the variation of landing point of the UAV from the actual target point, and it yielded similar results as well.

From the above results, it is clear that the Q -learning and the SARSA algorithms perform well in environments with less states and actions and the deep Q -learning algorithm performs well in environments with huge number of states and actions.

11 Conclusion and Future Work

In this paper, we have proposed a novel idea for the medical drones to operate autonomously using the bike lanes which are present in the roads by proposing a framework which uses reinforcement learning to enhance the drones path planning functionality through the suggestion of relevant actions from the origin point to the target point.

While only limited forms of these computations were performed on a model environment with less number of states and actions, this research demonstrated a proof of concept that this approach is feasible with further work even if we increase the size of the environment and extend the actions performed by the agent.

An off-policy algorithm Q -learning, an on-policy algorithm SARSA and a model-based reinforcement learning algorithm deep Q -Learning was compared and the results of the thesis is presented and tabulated in the evaluation section.

The models were evaluated using various factors such as shortness of the path to reach the goal, deviation from the straight path, time taken to complete the entire activity from the origin to the target location, smoothness of path etc. and compared with the results taken from a drone completing the same activity while being con-

trolled by a user. Our main aim is to train our reinforcement learning model to follow the cognitive dimension process and help the drone understand, learn, apply, validate and finally achieve the target point autonomously. Further work involves operating the drone in a actual highway environment with a bike lane from a origin point to a target point. In future, we can also consider adding the altitude of the drone as a variable into the model.

References

1. Shakhatareh, H., et al. (2019). Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges. *IEEE Access*, 7, 48572–48634. <https://doi.org/10.1109/ACCESS.2019.2909530>.
2. Konert, A., Smereka, J., & Szarpak, L. (2019). The use of drones in emergency medicine: Practical and legal aspects. *Emergency Medicine International*, 2019. <https://doi.org/10.1155/2019/3589792>
3. Scott, J., & Scott, C. (2017). Drone delivery models for healthcare. In *Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2017.399>
4. Burzichelli, C. D. (2016). Delivery drones: Will Amazon Air see the National Airspace? *Rutgers Computer and Technology Law Journal*, 42, 162.
5. Li, Z., Gao, C., Yue, Q., & Fu, X. (2019). Toward drone privacy via regulating altitude and payload. In *2019 International Conference on Computing, Networking and Communications (ICNC)* (Vol. 42(1), pp. 562–566).
6. Ackerman, E. (2019). Swiss post suspends drone delivery service after second crash. *IEEE Spectrum: Technology, Engineering, and Science News*, 42(1), 10–20.
7. Zeng, G., & Lei, Y. (2019). Research on multiple UAVs cooperative collision prevention. In *2019 IEEE International Conference on Unmanned Systems (ICUS)* (Vol. 1, pp. 756–760).
8. Mademlis, I., Mygdalis, V., Nikolaidis, N., & Pitas, I. (2018). Challenges in autonomous UAV cinematography: An overview. In *2018 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 1–6).
9. Jeyabalan, V., Nouvet, E., Meier, P., & Donelle, L. (2020). Context-specific challenges, opportunities, and ethics of drones for healthcare delivery in the eyes of program managers and field staff: A multi-site qualitative study. *Drones*, 4, 44.
10. Alwateer, M., & Loke, S. W. (2020). Emerging drone services: Challenges and societal issues. *IEEE Technology and Society Magazine*, 39(3), 47–51.
11. Stern, R., Sturtevant, N., Felner, A., Koenig, S., Ma, H., Walker, T., Li, J., Atzmon, D., Cohen, L., Kumar, T. K. S., Boyarski, E., & Bartak, R. (2019). Multi-agent pathfinding: Definitions, variants, and benchmarks. In *Accepted to SoCS 2019: The 12th Annual Symposium on Combinatorial Search* (Vol. 2(1)).
12. Ho, F., Salta, A., Galdes, R., & Goncalves, A. (2019). Multi-agent path finding for UAV traffic management. *Association for Computing Machinery (ACM)*, 54(9), 131–139.
13. Samir Labib, N., Danoy, G., Musial, J., Brust, M. R., & Bouvry, P. (2019). Internet of unmanned aerial vehicles—A multilayer low-altitude airspace model for distributed UAV traffic management. *Sensors*, 19(3), 4779.
14. Park, J., Choi, S., & Ahn, I. (2019). Structure design for unmanned aircraft traffic management system. In *2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)* (Vol. 7(1), pp. 118–120).
15. Gharibi, M., Boutaba, R., & Waslander, S. L. (2016). Internet of drones. *IEEE Access*, 4, 1148–1162.
16. Jiang, T., Geller, J., Ni, D., & Collura, J. (2016). Unmanned aircraft system traffic management: Concept of operation and system architecture. *International Journal of Transportation Science and Technology*, 5(3), 123–135.

17. Doole, M., Ellerbroek, J., & Hoekstra, J. (2016). Estimation of traffic density from drone-based delivery in very low level urban airspace. *Journal of Air Transport Management*, 88, 101862.
18. Brittain, M., & Wei, P. (2019). Autonomous air traffic controller: A deep multi-agent reinforcement learning approach 10. [arXiv:1905.01303](https://arxiv.org/abs/1905.01303)
19. Munoz, G., et al. (2019). Deep reinforcement learning for drone delivery. *Drones*, 3(3), 72.
20. Pham, H. X., La, H. M., Feil-Seifer, D., & Nguyen, L. V. (2018). Autonomous UAV navigation using reinforcement learning 7. [arXiv:1801.05086v1](https://arxiv.org/abs/1801.05086v1) [cs.RO]
21. Wang, C., Wang, J., Zhang, X., & Zhang, X. (2017). Autonomous navigation of UAV in large-scale unknown complex environment with deep reinforcement learning. In *CIIEEE Global Conference on Signal and Information Processing (GlobalSIP)* (Vol. 3(3), pp. 858–862).
22. Belkhale, S., Li, R., Kahn, G., McAllister, R., Calandra, R., & Levine, S. (2020). Model-based meta-reinforcement learning for flight with suspended payloads 858–862. [arXiv:2004.11345](https://arxiv.org/abs/2004.11345) [cs]
23. Wei, X. L., Huang, X. L., Lu, T., & Song, G. G. (2019). An improved method based on deep reinforcement learning for target searching. In *4th International Conference on Robotics and Automation Engineering (ICRAE)* (pp. 130–134).
24. Wu, T. C., Tseng, S. Y., Lai, C. F., Ho, C. Y., & Lai, Y. H. (2018). Navigating assistance system for quadcopter with deep reinforcement learning. In *2018 1st International Cognitive Cities Conference (IC3)* (pp. 16–19).
25. Yijing, Z., Zheng, Z., Xiaoyi, Z., & Yang, L. (2017). Q learning algorithm based UAV path learning and obstacle avoidance approach. In *36th Chinese Control Conference (CCC)* (pp. 3397–3402).
26. Sujit, P. B., Saripalli, S., & Sousa, J. B. (2013). An evaluation of UAV path following algorithms. In *2013 European Control Conference (ECC)* (pp. 3332–3337).
27. Sholes, E. (2007). Evolution of a UAV autonomy classification taxonomy. *IEEE Aerospace Conference*, 39(3), 1–6.
28. Zhang, Q., Feng, J., Yang, J., Ma, Z., & Chen, G. (2018). UAV path evaluation method based on reitopsis. In *Chinese Control and Decision Conference (CCDC)* (pp. 2519–2522).
29. Dawnee, S., Kumar, M. M. S., Jayanth, S., & Singh, V. K. (2019). Experimental performance evaluation of various path planning algorithms for obstacle avoidance in UAVs. In *3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1029–1034).
30. Thomas, P. S., & Brunskill, E. (2016). Data-efficient off-policy policy evaluation for reinforcement learning. In *33rd International Conference on Machine Learning* (Vol. 48, p. 10), New York, NY, USA, 2016.
31. Jordan, S. M., Chandak, Y., Cohen, D., Zhang, M., & Thomas, P. S. (2020). Evaluating the performance of reinforcement learning algorithms. In *Thirty-Seventh International Conference on Machine Learning (ICML 2020)*.
32. Kwak, J., & Sung, Y. (2018) Autonomous UAV flight control for GPS-based navigation. *IEEE Access*, 6, 37947–37955.
33. Saha, A., Kumar, A., & Sahu, A. K. (2017). FPV drone with GPS used for surveillance in remote areas. In *Third International Conference on Research in Computational Intelligence and Communication Networks* (Vol. 6, pp. 62–67).
34. Subash, K. V. V., Srinu, M. V., Siddhartha, M. R. V., Harsha, N. C. S., & Akkala, P. (2020). Object detection using Ryze Tello drone with help of mask-RCNN. In *2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 484–490).
35. Christl, M. (2020). Vision-based autonomous drone control using supervised learning insimulation. *Cornell University Journal Article*.
36. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Integration with Artificial Intelligence

Autonomous Navigation of Drones Using Reinforcement Learning



Billy Jacob, Abhishek Kaushik, and Pankaj Velavan

Abstract An UAV is a small hovering machine that can be remotely guided or flown autonomously through software-controlled flight plans in its embedded systems, operating in combination with onboard sensors and GPS. UAVs have demonstrated their versatility during the COVID-19 pandemic. Medicines and personal protection equipment were airlifted to the remote locations using UAVs. In future, transportation of drugs using UAV will be cost-effective and efficient. However, using human resources to operate these UAVs need a lot of time and investment in training. Most UAVs use GPS technology to travel to their destination from the start point. Many UAVs in the airspace generate a need for a drone traffic management system to mitigate collision risk. The drone traffic management system again demands human experts and massive expenditures. To overlook this challenge, we propose to model UAVs' autonomous navigation using the already available infrastructures present in highways like bike lanes and walking lanes. This research suggests a framework by using reinforcement learning and GPS way-points to allow the UAV to operate successfully from the origin location to the end location by following the bike lanes present on the roads.

Keywords Autonomous navigation · Unmanned aerial vehicle · Reinforcement learning

1 Introduction

Unmanned aerial vehicles are employed in various fields such as photography, building safety inspections, investigation of topography, delivery of food and emergency materials. UAV space is a thriving space. Various studies suggest that drones or UAVs could shortly be offering ordinary tasks like germinating crop fields on an automated

B. Jacob · P. Velavan
Dublin Business School, 13/14 Aungier Street, Saint Peter's, Dublin, Ireland

A. Kaushik (✉)
Adapt Centre, Dublin City University, Dublin 9, Dublin, Ireland
e-mail: abhishek.kaushik2@mail.dcu.ie

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
J. K. Verma et al. (eds.), *Advances in Augmented Reality and Virtual Reality*,
Studies in Computational Intelligence 998,
https://doi.org/10.1007/978-981-16-7220-0_10

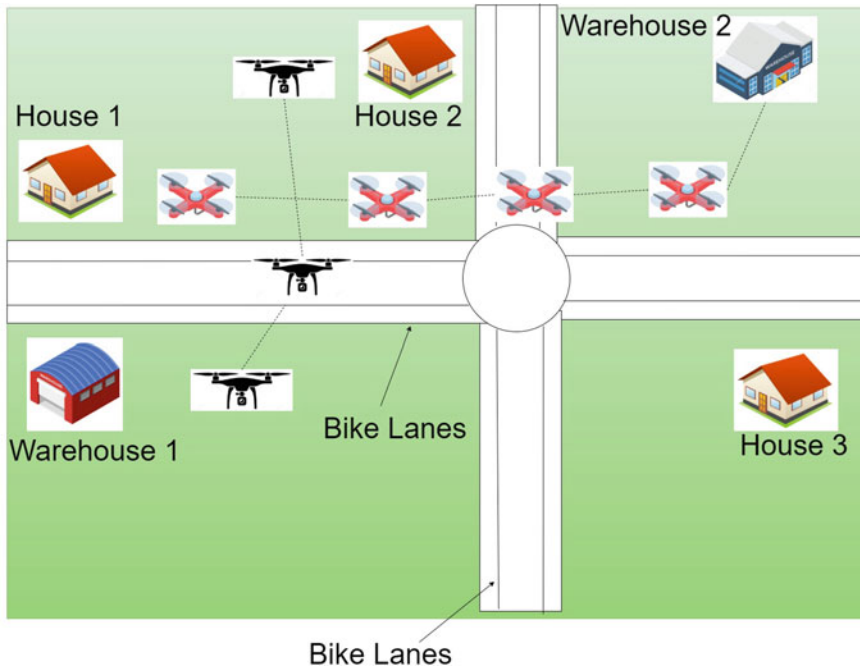


Fig. 1 Representation of the drone delivery using shortest path

basis, controlling traffic incidents, surveying hard-to-reach places, or even delivering food and packages. According to the Association for Unmanned Vehicle Systems International (AUVSI), commercial drones' impact could be \$82 billion and a 100,000-job boost to the U.S. economy by 2025. Irish drone firm Manna Aero is planning to partner with Just Eat, Camile Thai and Ben & Jerrys in Ireland for airborne food deliveries in under three minutes. The company is planning to start its operation on the UCD campus with a target audience of 30,000 students. As drones' usage increases in the day-to-day life of the industries mentioned above, there comes another challenge: drone traffic management and drone pilots. Around the world today, aircraft is guided around the skies by the air traffic controller. Each controller is accountable for a particular area of expertise, having the aircraft safe by communicating directly with pilots utilising wireless radio communications. Similarly, we need a drone traffic management system and pilots to make drone flights safe and successful [1]. Using human resources to operate these UAVs and manage traffic needs a lot of time and investment in training. So, a lot of companies are inclined towards developing autonomous drones. To operate these drones autonomously, we need to make the drones travel in a specific flight path so that mid-air collisions can be reduced to a great extent. To simplify this, we can model autonomous navigation of UAVs which follows the bike lanes present in Dublin in the left end of roadways all around the city. This limits the flight area of drones and diminishes the chance of collision if the speed and the hovering altitude or controlled. This is an exploratory

research proposal which provides a framework for using reinforcement learning to allow the drones to navigate successfully from the start location to the end location by following the bike lanes present on the roads (Fig. 1).

2 Related Works

This section will explain Current Drone Traffic Management system that is prevalent and their Limitations, Solutions for Drone Traffic Management and proposed solution for Managing the drone traffic.

2.1 *Current Drone Traffic Management Systems and Their Limitations*

At present, there are many companies which are working on developing UAV Traffic Management (UTM) systems. All these organisations are working alongside national space agencies of their respective countries to create one system or a cluster of systems which will coexist. The current legislation of most of the countries all around the world is very stringent concerning UAV usage in cities and public places. This section gives an overview of the related research works that are carried out in this specialised area and its limitations. In the context of Amazon Air Burzichelli [2] points out the legislation in USA framed by the FAA for small unmanned aircraft systems which do not allow drone operations when it overshoots the visual line of sight of the drone operator. Concerning the Irish laws, flying and operating drones in Ireland is subject to European Union Regulation 2019/947 [3] which too does not allow drone operations when it overshoots the visual line of sight of the drone operator. Apart from legislation, there are also privacy and security issues related to UAV operation in cities. The study conducted by Belkhale et al. [4] outlines the various security and privacy issues and proposes regulating the drones altitude and its onboard camera capability. When multiple drones fly in the same air space for completing their respective tasks, drone collisions may occur. An article written by Ackerman [5] reported an unfortunate incident in Switzerland where a 10kg delivery drone crashed near a group of kindergarten children and raised questions about the safety of urban delivery drones. The study conducted by Zeng and Lei [6] talks about the UAV collisions when a swarm of UAV's are flying in the same airspace for different tasks. He proposes a framework to calculate the minimum distance between drones using the flight angle and speed by considering the drone anti-collision problem as a geometric method. Another industry which operates drones in populated areas is movie industry. Mademlis et al. [7] listed the current challenges in autonomous UAV cinematography ranging from legal, ethical and safety factors to operational factors and suggest the future researchers consider these factors before

developing a fully autonomous drone for cinematography. Drones are used in medical industry where they are used to deliver medicines to remote locations. The study conducted by Jeyabalan et al. [8] surveyed by taking interviews from 16 individuals from 9 different countries working on the front lines of drones for health programs and listed down all the benefits and the concerns surrounding the implementation of the drones in the health care sector. The study categorised the concerns, viz. privacy and security, safety, scepticism of drone technology, lack of resources, technical challenges, lack of guidelines and regulations and aligning with community priorities. Similarly, a study conducted by Alwateer and Loke [9] outlined the challenges and the societal issues accompanying the emerging drone services. The study also reinforces the need for insurance policies so that incidents like physical loss and third-party damage can be covered.

2.2 Solutions for Drone Traffic Management System

The drone traffic management system intends to allow multiple drones to securely and efficiently assign airspace by aggregating the data of relevant controlled airspace and drones flying in that airspace. In 2015, the European Commission mandated the Single European Sky Air Traffic Management Research Joint Undertaking (SESAR JU) to produce a blueprint on how to make drone use in low-level airspace safe. The SESAR proposed the U-space framework [3] to support safe, secure and efficient access to airspace for a large number of drones. They have decided to roll out the services in four phases from 2019 to 2050. The current stage is foundation service where the users are provided registration, identification and geofencing facilities. The fourth phase will have automation, connectivity and digitisation for both the drone and the U-space system. Any traffic management system has to solve the Multi-Agent Pathfinding (MAPF) problem. The paper authored by Stern et al. [10] introduced unified terminology to describe MAPF problems and put out standard benchmarks and evaluation procedures for evaluating the MAPF algorithms. A research conducted by Ho et al. [11] proposed an extended formulation of Multi-Agent Path Finding using conflict-based search and enhanced conflict-based search to address the movement of heterogeneous drones. The research conducted by Samir Labib et al. [12] also emphasised on the importance of path planning to avoid collision and proposed a UAV Air Traffic Management (UTM) system with the help of IoT. The study conducted by Park et al. [13] described the essential functions of a UTM architecture and demonstrated a working UTM system using four UAVs with the LTE-based wireless communication board and a UTM server and also analysed the data flow between the UTM server, UAVs and the UAV operators. The architecture proposed by Gharibi et al. [14] to solve the problem of drone traffic management consists of a unified framework which has two important components, zone service providers (ZSP) and drones. Both these components are connected through the cloud. The ZSP provides the navigation information between any two drones in their designated zone, and the drone navigates autonomously with the help of the data supplied by the

ZSP's. The study conducted by Jiang et al. [15] reviewed the current UTM practice from industry partners like Amazon and Google and proposes a physical architecture for a UTM system which can handle a large number of Heterogeneous UAVs with all the required components. Design preferences for several of the services mentioned above will depend on how many drones fly in a specific air space. It, therefore, becomes essential to estimate how many delivery drones would operate in a typical city. The research work performed by Doole et al. [16] proposed a framework for evaluating the drone-based package delivery in traffic densities of five EU countries and concluded that any city should employ a robust airspace management system to realise commercial drone delivery.

2.3 Reinforcement Learning Solutions for Drone Traffic Management

Alongside supervised learning and unsupervised learning, reinforcement learning is also another paradigm. Reinforcement learning is a field of machine learning involved with how an agent reacts to actions in an environment based on the different states of the environment to maximise each action's cumulative reward. The advantage of this learning technique is that the machine autonomously learns to follow the correct instructions without providing any conditional statements. Numerous reinforcement learning solutions have been attributed to develop an UAV traffic management system. The study conducted by Brittain and Wei [1] proposed a deep multi-agent reinforcement learning framework to enable autonomous air traffic separation in en-route airspace. In this study, each aircraft is represented by an agent, and flight speed selection, altitude selection and flying angle selection are considered as the actions of the agent. The research work done by Muñoz et al. [17] applied deep reinforcement learning for autonomous drone delivery. They did the test in the AirSim simulator in a neighbourhood scenario with plenty of obstacles. The solution uses the double deep Q network, and this algorithm selects or decides the action of the agent, which maximises the resultant Q -value. It is vital for the drones with autonomous capability to navigate in new unknown locations as well. The work presented by La et al. [18] proposed a technique to train a quadrotor to learn to navigate to the target point using a Q -learning algorithm in an unknown environment. The study conducted by Wang et al. [19] proposes using only GPS signal and sensory information of the local environment to navigate autonomously and intelligently from arbitrary departure places to random target positions in a virtual environment. He modelled the navigation problem as a control problem and employed deep reinforcement learning to solve it. Researchers at Facebook and the University of California, Berkeley, published a new research paper [4] where they showed how a drone could use AI to learn to handle different delivery payloads under changing conditions. They proposed a meta-learning approach. They used a neural network dynamics model, which takes as input the current state and action and predicts the next state. The

study conducted by Huang et al. [20] proposed a deep reinforcement learning model for target searching in an unfamiliar environment. He suggests that research work in such areas can improve the target searching during disaster times where humans can be located in an unknown and adverse environment. The research conducted by Tseng et al. [21] proposed a navigation assistance system for quadcopter with deep reinforcement learning. They used a stimulating environment with the obstacles on the flight path. They trained the model in flight path with obstacles for 500 flights. They used that same model to fly the quadcopter in a new environment with blocks for 1000 flights and calculated the collision percentage with the obstacle. Zheng et al. [22] solved the UAV obstacle avoidance problem using the Q -learning algorithm where the movement of the UAV is considered as action, and walls, blocks, bricks and traps were taken as the obstacles for training the model.

2.4 Evaluation of Autonomous Drones

Evaluation is a significant part of any research work, and it is essential to verify if any developed machine learning model or an algorithm working as per the requirement. The research conducted by Sujit et al. [23] compared five different UAV path

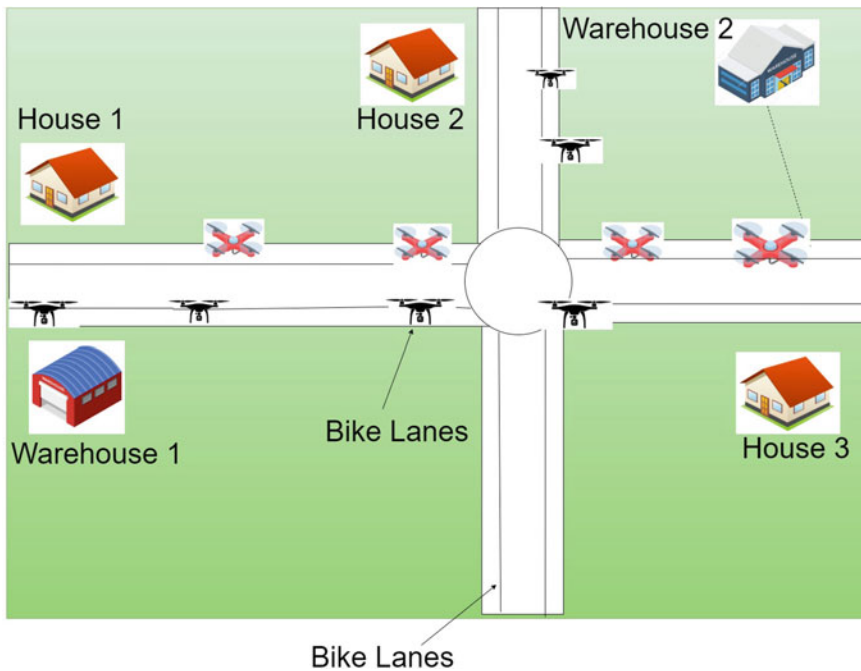


Fig. 2 Representation of the drone delivery using reinforcement learning

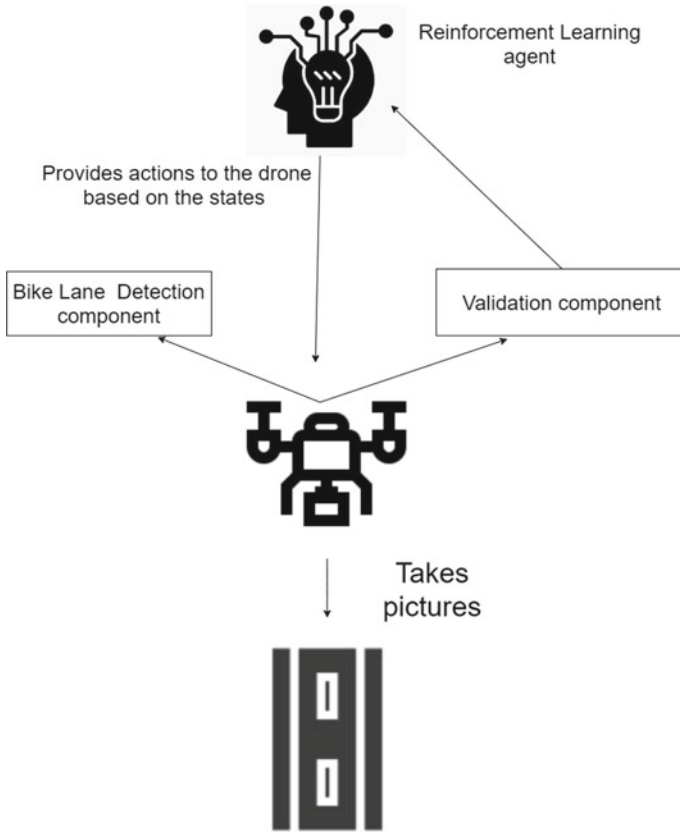


Fig. 3 Conceptual framework

following algorithms and compared the algorithms using two parameters, total control effort and absolute cross-track error. Sholes [24] proposed a methodology to categorise the UAVs based on UAVs’ autonomy capacity. The paper identifies the emerging metrics and taxonomies in the field of autonomy. It presents how the metrics could be applied and how they have evolved and expanded, in theory and practice, as a result of lessons learned during those applications. Feng et al. [25] proposed UAV path evaluation method based on RE-ITOPSIS (Relative Entropy Technique of Order Preference Similarity to the Ideal Solution). Dawnee et al. [26] compared Dijkstra, A^* , Artificial potential field and RRT (Rapidly Exploring Random Tree) algorithms and evaluated them using shortness of path, clearance to obstacles, deviation from straight-line path and smoothness of path and tabulated the results. The study conducted by Thomas and Brunskill [27] compared various evaluation metrics which is used for off-policy evaluation of reinforcement learning.

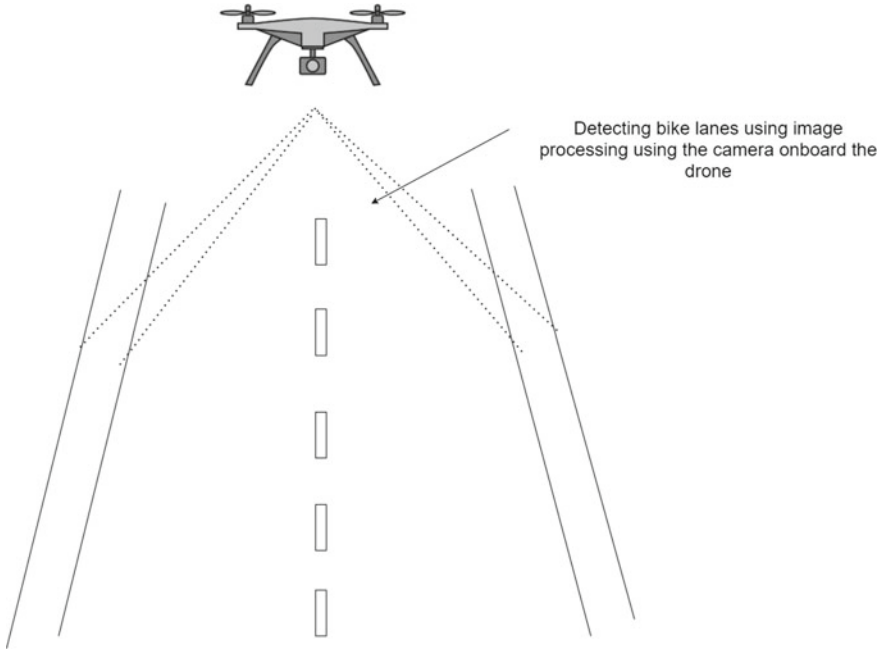


Fig. 4 Representation of the bike lane detection

3 Conceptual Framework for Drone Path Planning Using Bike Lanes

As a part of this research work, we have developed a conceptual framework for drone path planning which can be used by the delivery drones to navigate in the cities autonomously utilising the bike lanes at the far ends of the motorways and roadways. The framework which we are proposing can be implemented on all the drones which are going to be used in the delivery process in cities. We are assuming that the bike lanes will be at both the sides of the roadways, and flying height of a drone will be fixed to 20m from the ground. The drones will have GPS technology and an on-board camera. The reinforcement learning model will assist the drones in following the bike lanes in the roadways so that the lateral movement of the drone can be reduced to a very great extent as shown in Fig.2. Reduction in the lateral movement of the drones will normalise the drone traffic and reduce the probability of the collision with the other drones, which will be flying in the same air space. The framework comprises of the bike lane detection component, validation component and the reinforcement learning component. The representation of the conceptual framework is given in Fig. 3.

3.1 *Bike Lane Detection Component*

The bike lane in the roads can be detected using the camera on-board the drone. Various research works have been carried out in this area where the small autonomous aircraft can be made to follow the bike lanes. The research conducted by Frew et al. [28] described the vision-based control of a small autonomous aircraft following a road. Aufrere et al. [29] proposed a fast and robust vision-based lane following algorithm which will be able to detect the roads based on the previous images of the roads on which the model will be trained. The algorithms and the methodologies proposed in these research works will be incorporated as part of our framework. The bike lane detection component presented in this framework will work based on image processing, and it will be used to detect the lanes at the far end of the roadways. This component will be continuously executed, and the flight path of the drone will be continuously tracked as shown in Fig. 4. As part of this research, we have used the OpenCV library in python. First, the video footage taken by the drone is converted into multiple frames. One of the frames is shown in Fig. 5. After the frame is captured, the following three steps are carried out to detect the bike lane present in the image.

- **Limit the area of interest:** The first step is to limit the area of the picture by frame masking technique. A frame mask is a NumPy array. We change the pixel values of the desired region in that image to 0, or 255, or any other number. When the pixel value is set to zero that area turns black as shown in Fig. 6. In this picture, the pixel values of a certain region in the image has been set to 0. Only a specific area is selected for further analysis.
- **Image thresholding:** In this method, a greyscale image's pixel values are assigned one of the two values representing black and white colours based on a threshold value. So, if the value of a pixel is greater than a threshold value, it is assigned one value, else it is assigned the other value. In this research, we have kept the threshold values as 130. If the pixel values are lesser than 130, then the pixel in that specific area is converted to black colour and if the pixel value is greater than 130, then the pixel is converted to white colour. The image after the thresholding is shown in Fig. 7.
- **Detecting the lane:** In this method, we will be using the Hough transformation technique. Hough transform is a technique to detect any shape that can be represented mathematically. As the 2D view of the bike lane when the drone is flying over the lane is a rectangle, we will use the Hough transform technique to detect the rectangular shapes which are visible in the thresholded image. After this step is completed, the final image will look something as shown in Fig. 8.



Fig. 5 One frame taken from the drone shot video

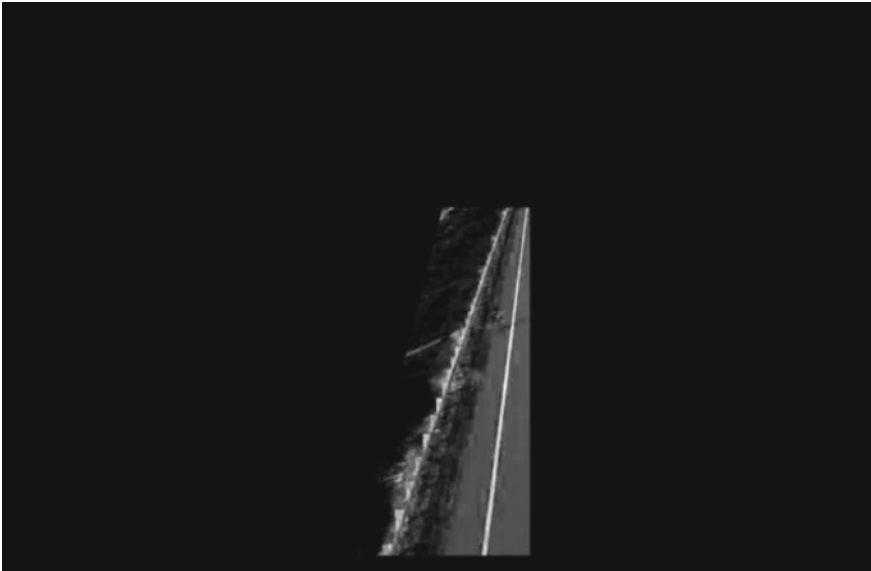


Fig. 6 Picture shot by the drone after frame masking

3.2 Validation Component

The validation component will be a python module which will be responsible for checking the current position of the drone and comparing it with the position of the bike lane detected by the bike lane detection component. The validation component

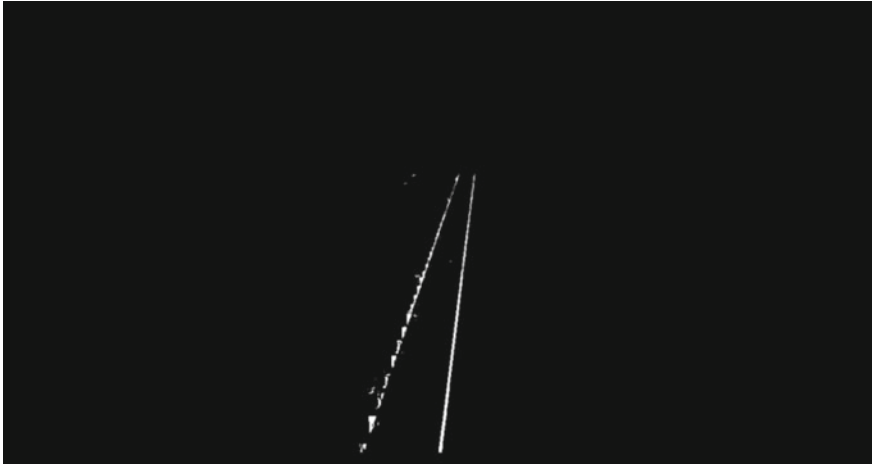


Fig. 7 Picture shot by the drone after image thresholding



Fig. 8 Picture of the detected bike lane

will review based on the altitude at which the drone is flying and the width of the bike lane detected by the bike lane detection component. The environment will use the output of this module to provide the rewards for the reinforcement learning agent. As a part of this research, we are assuming that the altitude of all the delivery drones will be fixed at a height of 20 m from the ground, the sensor width of the camera on board the drone is 6.17 mm, the focal length of the camera is 4.55 mm and the width of the image captured by the drone is 3968 pixels. Now, to validate whether the drone is flying over the bike lane, the image shot by the drone should have the bike lane at the centre of the image. When the drone is flying at a certain height from

the ground, the width of the bike lane can be smaller in perspective as compared with the actual width. To overcome this challenge, we could determine the ground sample distance (GSD) of each pixel. Each pixel on the picture corresponds to a certain distance on the ground. The formula for calculating GSD is given in Eq. (1). When we provide the above-assumed values in the equation, we calculated that $GSD = 1.7224 \text{ cm/pixel}$. If a bike lane with a width of 2 m needs to be validated, then the pixel length between the right and the left edges of the bike lane in the picture taken by the drone would be 116. From the above calculation, it could be concluded that the image shot by the drone should have the bike lane between the pixels 1926 and 2042 as shown in Fig. 10.

$$GSD = (\text{sensor width} * \text{altitude} * 100) / (\text{focal length} * \text{image width}) \quad (1)$$

3.3 Reinforcement Learning Component

The reinforcement learning component is the proposed backend module that feeds the drone with the actions to take while flying from the origin to the target location. The proposed reinforcement learning model is Q -learning, which is a model-free reinforcement learning algorithm to learn the quality of actions determining an agent what action to take under what circumstances [30]. The Q -learning model learns in an environment by trial and error using the feedback from actions and states. There are five basic components in reinforcement learning, namely agent, environment, reward, Q -table and action.

- **Agent:** The agent is a module that takes decisions based on positive and negative rewards. In this instance, it would be the module which makes the drone to move in various directions and altitude.
- **Environment:** The environment is a task or simulation. In this instance, the airspace, the drone movement module, the bike lanes in the roadways, the drone origin point and the drone target constitute the environment.
- **Actions:** the movement of the drone by the agent in any direction in the environment. The various available action would be turn right, turn left, fly straight, fly back, increase the altitude etc. These actions will be provided by the reinforcement learning agent to the drone.
- **States:** Following the bike lane in the roads, lateral movement of the drones, moving away from the bike lanes, collision with any external objects vertical take-off and vertical landing would be the different states that are available in the environment.
- **Q -table:** Q -Table is a simple matrix or a lookup table where we calculate the maximum rewards for an action at each state. The Q -table will guide the agent to the best action at each state. For example, a sample Q -table for an environment with 2 states and 3 actions will look like $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ where each row corresponds to a state and each column corresponds to an action.

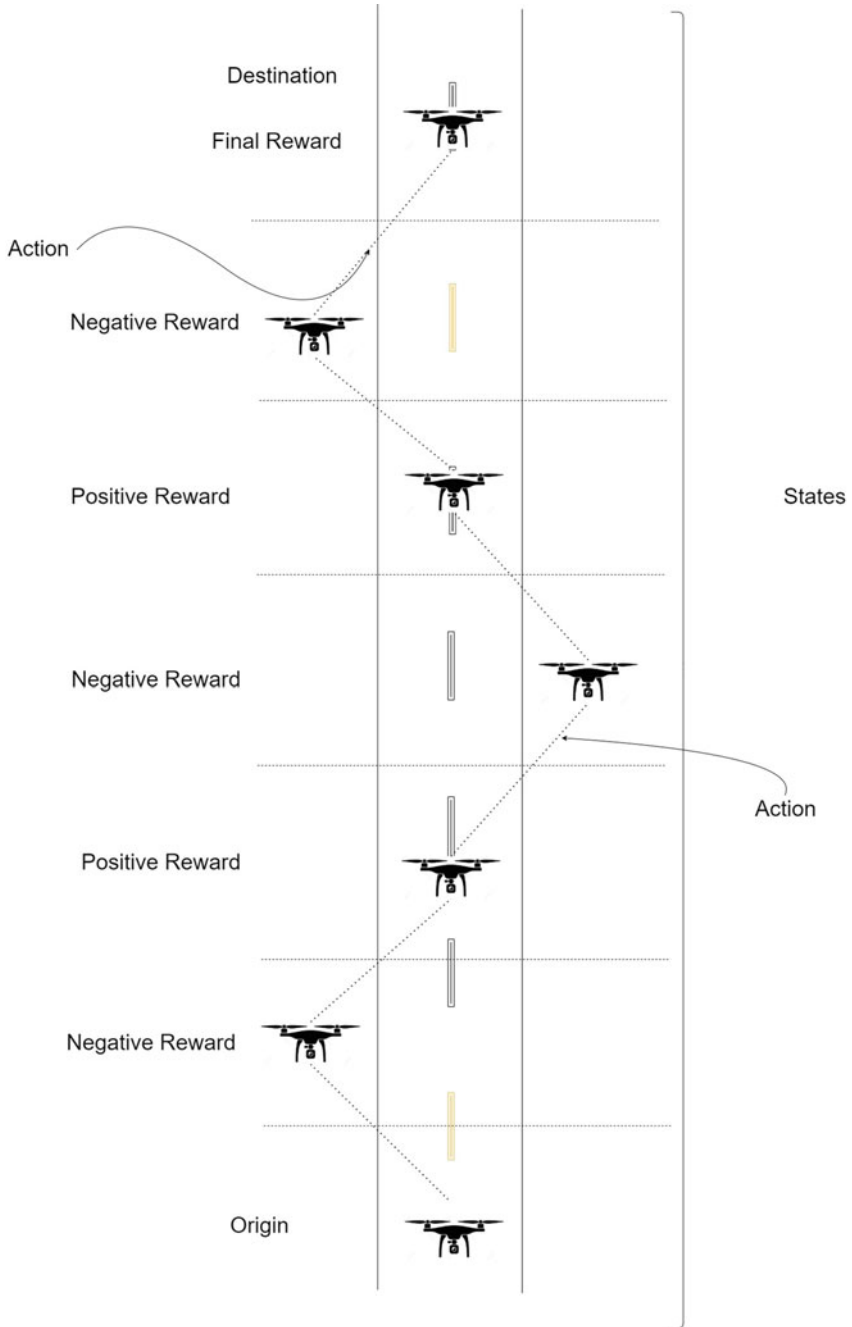
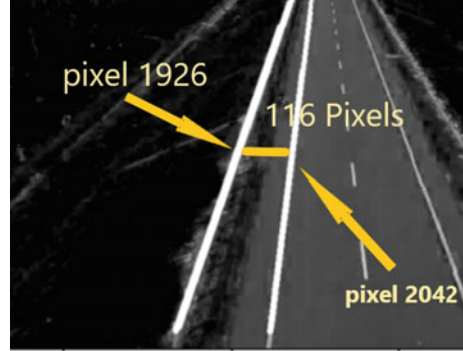


Fig. 9 Representation of the reinforcement learning

Fig. 10 Validation component explanation



Let us consider the sample Q -learning environment with the agent, states and actions (skills,role) as shown in Fig. 9. When a delivery point is given to the delivery drone, that dropping point is updated as the target in the environment as shown in Fig. 9. When the agent feeds the actions to the drone to reach the destination point, if the drone is travelling with the bike lane in the roadways, then the agent receives positive rewards from the environment. The validation component will validate whether the drone is flying in the right path over the bike lane. When the agent advises the drone to move away from the bike lane, then the agent receives negative rewards, and when a drone completes the full flight path by reaching the target point while following the bike lanes in the road and without any collisions, then the agent receives a final big positive reward from the environment. In the training phase, the agent investigates the various actions for the various states and collects the corresponding rewards. In the implementation phase, the agent feeds the proper direction to travel to the drone based on the knowledge gained from the training phase. This way instead of using conditional operators to pre-define a particular flight path or using GPS to reach the destination, the reinforcement learning model will make the agent to learn the correct flight path based on the rewards, states and actions present in the environment (Fig. 10).

$$Q(s, a) = Q(s, a) + \alpha[r + \gamma \max_a Q(s', a') - Q(s, a)] \quad (2)$$

where

Q = Q -table present in the Q -learning environment.

s = states such as UAV take-off, horizontal movement, vertical movement UAV landing etc.

a = flying of the drones in various direction in the environment.

$Q(s, a)$ = new q -value (value present in the q -table in s th row and a th column).

$Q(s', a')$ = old q -value in the q -table in s th row and a th column.

α = the learning rate, set between 0 and 1,

γ = discount factor, also set between 0 and 1,

\max_a = the maximum reward that is attainable in the state following the current state.

Algorithm 1 Reinforcement Learning Training Phase

```

1:
     $\alpha = i, 0 < i < 1$ 
2:
     $\gamma = j, 0 < j < 1$ 
3: for  $episodes = 1, 2, \dots, N$  do
4:   The  $Q$ -learning model explores the various actions (flying path) available and starts to suggest
   flying direction to the drone.
5:   if the drone flies within the bike lane then
6:     the environment provides positive rewards
7:   else
8:     the environment provides negative rewards
9:   end if
10: end for
11:  $Q$  value is calculated using Eq. (2)
12:  $Q$  value will be updated in the  $Q$ -table for that particular state and action as shown in Eq. (2)
13: the model will be trained in many airspaces and many locations with bike lanes and the the
    corresponding  $Q$  values will be calculated

```

4 Training and Evaluation Phase

Two approaches would train the reinforcement learning agent. The agent would be trained by the real users or trained by AirSim built on the unreal engine to buildup the rewards to reach the final goal. The methods which outperform another method would be selected for final training. The training method and evaluation methods are discussed below:

- Training the model using Drone Pilots (User Training).
- Training the model using the Airsim unreal engine environment (Gym Training).

4.1 User Training

The reinforcement learning model can be trained by operating the drone manually and enriching the Q -table. Drone pilots or any person who can operate the drone can be asked to fly the drones from the origin point to the target point through the bike lanes in the roadways. Whenever the pilot operates the drone outside of the bike lanes, then the negative reward will be provided to the reinforcement learning agent from the environment and whenever the drones follow the bike lanes, a positive reward is provided by the environment to the agent. All these rewards will be collected, and the q -values will be calculated and updated in the Q -table. The enriched Q -table which has the Q values for the actions and the states can be deployed in the environment for the autonomous navigation of the drones in the real-time environment (Fig. 11).



Fig. 11 Sample AirSim drone and environment

4.2 Gym Training

The reinforcement learning agent would be trained with the reinforcement learning environment which would be built using Python and the open-AI gym library, which will have the bike lane detection component and the validation component. The gym environment will provide the reinforcement learning agent with pre-click pictures of bike lanes, and the validation component will check whether the drone is following the bike lane; based on this check, the environment will provide the reward to the reinforcement learning agent. The simulation would be executed through multiple episodes. One episode corresponds to one run where the drone starts from the start location and reaches the target location. The optimal number of episodes will be decided based on the performance of the drone in achieving the provided activity.

5 Conclusion and Future Work

In this paper, we have proposed a novel idea for the drones to operate autonomously using the bike lanes which are present in the roads by proposing a framework which uses reinforcement learning to enhance the drones path planning functionality through the suggestion of relevant actions from the origin point to the target point. This platform could be evaluated using various factors such as shortness of the path to reach the goal, deviation from the straight path, time taken to complete the entire activity from the origin to the target location, smoothness of path etc. and compare it with the results taken from a drone completing the same activity while being controlled by a user. Our main aim is to train our reinforcement learning model to follow the cognitive dimension process and help the drone understand, learn, apply, validate and finally achieve the target point autonomously. Further work involves operating the drone in a actual highway environment with a bike lane from a origin point to a target point. This framework can be extended to all kinds of vehicles which operate autonomously without human intervention.



References

1. Brittain, M., & Wei, P. (2019). Autonomous air traffic controller: A deep multi-agent reinforcement learning approach. Available at: [arXiv:1905.01303](https://arxiv.org/abs/1905.01303) [cs, stat]. Accessed: November 9, 2020.
2. Burzichelli, C. D. (2016). Delivery drones: Will Amazon Air see the National Airspace notes and comments. *Rutgers Computer and Technology Law Journal*, 42(1), 1–196. Available at: <https://heinonline.org/HOL/P?h=hein.journals/rutcomt42&i=187>. Accessed: November 7, 2020.
3. Publications Office of the European Union. (2017). U-space: Blueprint. Publications Office of the European Union. Available at: <http://op.europa.eu/en/publication-detail/-/publication/f8613e25-cf38-11e7-a7df-01aa75ed71a1/language-en>. Accessed: November 8, 2020.
4. Belkhal, S., Li, R., Kahn, G., McAllister, R., Calandra, R., & Levine, S. (2020). Model-based meta-reinforcement learning for flight with suspended payloads. Available at: [arXiv:2004.11345](https://arxiv.org/abs/2004.11345) [cs]. Accessed: November 11, 2020.
5. Ackerman, E. (2019). Swiss post suspends drone delivery service after second crash. IEEE Spectrum. IEEE Spectrum: Technology, Engineering, and Science News. Available at: <https://spectrum.ieee.org/automaton/robotics/drones/swiss-post-suspends-drone-delivery-service-after-second-crash>. Accessed: November 8, 2020.
6. Zeng, G., & Lei, Y. (2019). Research on multiple UAVs cooperative collision prevention. In 2019 IEEE International Conference on Unmanned Systems (ICUS) (pp. 756–760), October 2019. <https://doi.org/10.1109/ICUS48101.2019.8995993>
7. Mademlis, I., Mygdalis, V., Nikolaidis, N., & Pitas, I. (2018). Challenges in autonomous UAV cinematography: An overview. In 2018 IEEE International Conference on Multimedia and Expo (ICME) (pp. 1–6). <https://doi.org/10.1109/ICME.2018.8486586>
8. Jayabalan, V., Nouvet, E., Meier, P., & Donelle, L. (2020). Context-specific challenges, opportunities, and ethics of drones for healthcare delivery in the eyes of program managers and field staff: a multi-site qualitative study. *Drones*, 4(3), 44. <https://doi.org/10.3390/drones4030044>
9. Alwateer, M., & Loke, S. W. (2020). Emerging drone services: Challenges and societal issues. *IEEE Technology and Society Magazine*, 39(3), 47–51. <https://doi.org/10.1109/MTS.2020.3012325>
10. Stern, R., Sturtevant, N., Felner, A., Koenig, S., Ma, H., Walker, T., Li, J., Atzmon, D., Cohen, L., Kumar, T. K. S., Boyarski, E., & Bartak, R. (2019). Multi-agent pathfinding: Definitions, variants, and benchmarks. Available at: [arXiv:1906.08291](https://arxiv.org/abs/1906.08291) [cs]. Accessed: November 7, 2020.
11. Ho, F., Salta, A., Geraldès, R., Gonçalves, A., Cavazza, M., & Prendinger, H. (n.d.). Multi-agent path finding for UAV traffic management, p. 9.
12. Samir Labib, N., Danoy, G., Musial, J., Brust, M. R., & Bouvry, P. (2019). Internet of unmanned aerial vehicles—A multilayer low-altitude airspace model for distributed UAV traffic management. *Sensors*, 19(21), 4779. <https://doi.org/10.3390/s19214779>
13. Park, J., Choi, S., & Ahn, I. (2019). Structure design for unmanned aircraft traffic management system. In 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN) (pp. 118–120). <https://doi.org/10.1109/ICUFN.2019.8806075>
14. Gharibi, M., Boutaba, R., & Waslander, S. L. (2016). Internet of drones. *IEEE Access*, 4, 1148–1162. <https://doi.org/10.1109/ACCESS.2016.2537208>.
15. Jiang, T., Geller, J., Ni, D., & Collura, J. (2016). Unmanned aircraft system traffic management: Concept of operation and system architecture. *International Journal of Transportation Science and Technology*, 5(3), 123–135. <https://doi.org/10.1016/j.ijtst.2017.01.004>
16. Doole, M., Ellerbroek, J., & Hoekstra, J. (2020). Estimation of traffic density from drone-based delivery in very low level urban airspace. *Journal of Air Transport Management*, 88, 101862. <https://doi.org/10.1016/j.jairtraman.2020.101862>
17. Muñoz, G., Barrado, C., Cetin, E., & Salami, E. (2019). Deep reinforcement learning for drone delivery. *Drones*, 3(3), 72. <https://doi.org/10.3390/drones3030072>.
18. La, H. M., Feil-Seifer, D., Pham, H. X., & Nguyen, L. V. (2018). Autonomous UAV navigation using reinforcement learning. [arXiv:1801.05086v1](https://arxiv.org/abs/1801.05086v1) [cs.RO]

19. Wang, J., Zhang, X., Wang, C., & Zhang, X. (2017). Autonomous navigation of UAV in large-scale unknown complex environment with deep reinforcement learning. In *CIIEE Global Conference on Signal and Information Processing (GlobalSIP)* (Vol. 3(3), pp. 858–862), November 2017. <https://doi.org/10.1109/GlobalSIP.2017.8309082>
20. Huang, X. L., Lu, T., Wei, X. L., & Song, G. G. (2019). An improved method based on deep reinforcement learning for target searching. In *4th International Conference on Robotics and Automation Engineering(ICRAE)* (pp. 130–134), November 2019. <https://doi.org/10.1109/ICRAE48301.2019.9043821>
21. Tseng S.-Y., Lai, C.-F., Ho C.-Y., Wu, T.-C., & Lai, Y.-H. (2018). Navigating assistance system for quadcopter with deep reinforcement learning. In *2018 1st International Cognitive Cities Conference (IC3)* (pp. 16–19), August 2018. <https://doi.org/10.1109/IC3.2018.00013>
22. Zheng, Z., Xiaoyi, Z., Yijing, Z., & Yang, L. (2017). Q learning algorithm based UAV path learning and obstacle avoidance approach. In *36th Chinese Control Conference (CCC)* (pp. 3397–3402), July 2017. <https://doi.org/10.23919/ChiCC.2017.8027884>
23. Sujit, P. B., Saripalli, S., & Sousa, J. B. (2013). An evaluation of UAV path following algorithms. In *2013 European Control Conference (ECC)* (pp. 3332–3337). <https://doi.org/10.23919/ECC.2013.6669680>
24. Sholes, E. (2007). Evolution of a UAV autonomy classification taxonomy. In *IEEE Aerospace Conference* (Vol. 39(3), pp. 1–6). <https://doi.org/10.1109/AERO.2007.352738>
25. Feng, J., Yang, J., Ma, Z., Zhang, Q., & Chen, G. (2018). UAV path evaluation method based on RE-ITOPSIS. In *Chinese Control and Decision Conference (CCDC)* (pp. 2519–2522), June 2018. <https://doi.org/10.1109/CCDC.2018.8407549>
26. Dawnee, S., Kumar, M. M. S., Jayanth, S., & Singh, V. K. (2019). Experimental performance evaluation of various path planning algorithms for obstacle avoidance in UAVs. In *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1029–1034)
27. Thomas, P. S., & Brunskill, E. (2016). Data-efficient off-policy policy evaluation for reinforcement learning. In *33rd International Conference on Machine Learning* (Vol. 48, p. 10), New York, NY, USA, June 2016. <http://proceedings.mlr.press/v48/thomasa16.pdf>
28. Frew, E., McGee, T., ZuWhan, K., Xiao, X., Jackson, S., Morimoto, M., Rathinam, S., Padiyal, J., & Sengupta, R. (2004). Vision-based road-following using a small autonomous aircraft. In *2004 IEEE Aerospace Conference Proceedings* (IEEE Cat. No.04TH8720) (Vol. 5, pp. 3006–3015). <https://doi.org/10.1109/AERO.2004.1368106>
29. Aufrere, R., Chapuis, R., & Chausse, F. (2000). A fast and robust vision based road following algorithm. In *Proceedings of the IEEE Intelligent Vehicles Symposium 2000* (Cat. No.00TH8511) (pp. 192–197). <https://doi.org/10.1109/IVS.2000.898340>
30. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Human Face Detection and Recognition Protection System Based on Machine Learning Algorithms with Proposed AR Technology



Md. Imran Hossain Chowdhury, Newaz Mahmud Sakib,
S. M. Masum Ahmed , Mohammad Zeyad , Md. Abul Ala Walid,
and Golam Kawcher

Abstract This paper demonstrates the difference between two algorithms, the first one is “Template Matching” and another one is “Local Binary Pattern Histogram (LBPH).” The face identification and recognition security system prototype had implemented using the LBPH algorithm, Python, Raspberry Pi 3 Model B+ , and OpenCV technology. This concept introduces a system for identifying random faces with Haar classifier. This method does not search for individual matches like a biometric device, it compares people to all the same, instead provide matches based on first, second, and third findings in a static collection without having to deal with the databases. Similar to a CCTV, it can identify the persons; however, instead of storing a large amount of material, it contains just a small amount. However, the face detection rate of 90% had achieved when the LBPH method had used in bright lighting conditions. On the other hand, when the template matching method had used in this recognition method in bright light condition then the detection rate was 40%. A method has been proposed that would enable the successful recognition, visualization, and detection of the convicted person utilizing virtual platforms and

S. M. Masum Ahmed and Mohammad Zeyad are equally contributed

Md. I. H. Chowdhury · N. M. Sakib
American International University-Bangladesh (AIUB), Dhaka, Bangladesh

S. M. Masum Ahmed (✉) · M. Zeyad (✉)
Heriot-Watt University, Edinburgh, Scotland, UK

Universidad del País Vasco/Euskal Herriko Unibertsitatea, Bilbao, Spain

Energy and Technology Research Division, Advanced Bioinformatics, Computational Biology and Data Science Laboratory, Bangladesh, Chattogram 4226, Bangladesh

Md. A. A. Walid
Bangabandhu Sheikh Mujibur Rahman Science and Technology University (BSMRSTU),
Gopalganj, Bangladesh

G. Kawcher
University of Siegen, Siegen, Germany

3D modeling of stored pictures, which might be innovative and used in the development of augmented reality (AR). Also, the detection performance of this device had analyzed that could detect the user's face up to 15 m in proper lighting conditions without any hassle. This may be used for home and commercial surveillance, for identification, or in the event of an act of bank robbery, or for counterterrorism. Finally, the device had implemented with the LBPH algorithm and quite economical compare to another biometric security system.

Keywords Augmented reality (AR) · LBPH · Raspberry Pi · Security system · Template matching

1 Introduction

Security systems based on face identification and recognition had implemented for the function of human interaction with computers, security, and protection, surveillance, verification, etc. For example, let's say workers of any organization can be verified using this device. Various techniques of face recognition were represented in literature; they consist of Haris Corner, Viola–Jones, Haar classifier, and principle component analysis [1–7]. The Haar classifier derived from the histogram algorithm of the local binary patterns will be used for facial identification in this work. For position tracking the Eigen features, the project develops this device using Raspberry Pi with two types of the algorithm to justify their accuracy. After detecting the full face, it will be saved into the algorithm. In this way, many more inputs can be saved. Parallel models with Haar-like elements were used with facial recognition technology [8, 9]. Hefenbrock proposed to use GPU to recognize and detect face algorithms [8]. This algorithm has the main drawback of not being unable to process a real-time stream size of 640×480 [10].

Augmented reality (AR) technologies have the potential to improve our view of and interaction with the actual environment. Apart from virtual reality systems, that imitate the actual world, AR systems recognize physical features and overlay computer-generated graphic, auditory, and haptic signals upon real-world input in actual time. AR emphasis has been placed on solving issues with display technology, monitoring, and registration to correctly match virtual and physical objects, user interfaces and human errors, supplementary sensing devices, and the creation of unique AR applications since before the 1990s. The latest generation AR technology, on the other hand, has just lately been available on the market. For instance, Google recently issued a limited amount of Google Glass heads-up glasses for augmented reality (AR) applications. Moreover, the use of fossil fuels nowadays is rapid. Several problems like environmental pollutions, scarcity of energy resources and natural disasters have to be faced in the long-term future generation [11]. So, nowadays, AR system has also been used for updating the energy resources. However, the widespread availability of mobile devices facilitates several more early-stage AR applications. Word Lens, an iPhone app that displays translated text on a view of the camera of a foreign

language, and Layar, a geo-location-based AR interface that enables developers to construct AR layers for the globe, are two examples (for example, for game playing). Certain applications are now available because of the recent introduction of 1 GHz CPUs, location sensors, and high-resolution, autofocus cameras in smartphones [12].

Several new techniques for facial recognition, including principal component analysis (PCA), local binary histogram (LBP), linear displacement analysis (LDA), and support vector machines (SVM), were employed in this application [13, 14]. Some of these analytic approaches consider position, as well as size and shape of facial features like the space between the eyes, the nose and lips [15, 16]. For the reason of the different facial expressions and illumination, some of the methods, mentioned above are not so effective. That is why a real-time face recognition algorithm named Haar classifier had developed by Viola–Jones. Two researchers from India researched face detection and recognition. In this research, the Viola–Jones algorithm had used to perform face detection and recognition. One of the main drawbacks of this work is the long computation time when picture size is large and resolution is high [17]. N. A. Othman and I. Aydin researched face recognition for protection schemes in smart households and cities. In this work, a real-time recognition method was proposed that could handle images quickly. One of the key aspects of this work was to ensure a protection scheme for homes and offices with the help of face recognition. The PIR sensor had used in this work for the detection of an object in the limited region. Moreover, the photos will be taken by a Raspberry Pi. Also, the pictures and messages would be sent to a cellphone through the Telegram service [18].

S. J. Oh and his research teams had presented research on faceless individual recognition of social networks has privacy consequences. Through investigating how often people are identifiable in social media data, this research leads to a better understanding of the security implications of such data exchange. Finally, they proposed a hierarchical individual recognition system that can accommodate vast differences in pose and clothing while only requiring a few training samples to train [19]. M. F. Valstar and his research team researched facial expression recognition and analysis. In this research paper, the primary constraint to evaluating a lack of suitable data. The third challenge in automated face recognition is presented in this paper. The identification of AU occurrence and the calculation of AU intensity are both defined as sub-challenges [20].

Few researchers from Germany worked on a project related to face recognition. Particularly, face recognition systems are vulnerable to attacks focused on the morphed facial image. In this article, we offer a theoretical classification and metrics for evaluating those approaches, as well as a systematic review of relevant research [21]. Characterizing the chapter: Section II describes the analysis of the literature, Section III describes the work description of the HACA-app, Section IV describes the architecture of the system proposed, including hardware and software execution, Section V describes the implemented application design, and Section VI ends the report.

2 Hardware and Software Description

2.1 Raspberry Pi 3 Model B+

The latest model of Raspberry Pi having a 64-bit quad-core Broadcom BCM2837BO processor with 1.4 GHz clock speed and 1 GB SDRAM has been used in this project, practically which gives a better and faster performance. The system is also equipped with dual-band WLAN, Bluetooth 4.2/BLE, swift Ethernet and a dedicated PoE HAT Power-Over-Ethernet service. It consumes less power because of its high configuration chipset and less power rating. It is card size device and can easily plug into a monitor through HDMI.

2.2 Camera Module (I)

NoIR Camera V2 was used to take the picture and create a database. It also can record high definition videos and used during real-time face detection and recognition.

2.3 SD Card

This is one of the important parts of this project and helps to store the operating system and files. Greater storage will give a better opportunity to extend data storage and improve the read–write speed. However, it can be said that the communication or interlink between the processing unit and database hardware will improve and give better performance of the device. But to ensure cost-effectiveness, here we use only 16 GB grade 10 SD cards as it is the experimental prototype design.

2.4 Raspbian


Raspbian is the official operating system of Raspberry Pi. For enabling the camera and Raspberry Pi, the total Raspbian packages were installed.

2.5 SQLite Studio

It is a popular embedded database for the storage of local client data. All the user pictures are stored in the SQLite database management system. During detection,

Fig. 1 Patterns of Haar classifier

a. Edge feature:

b. Line Feature: 

c. Four Rectangle Features: 

code can access the database through this software and matches the real-time data according to the stored data. It is also possible to update and modify the database.

2.6 Camera Module (II)

OpenCV 2.7.13 was installed for the project. It is a real-time application-based programming function library.

3 Methods of Recognition

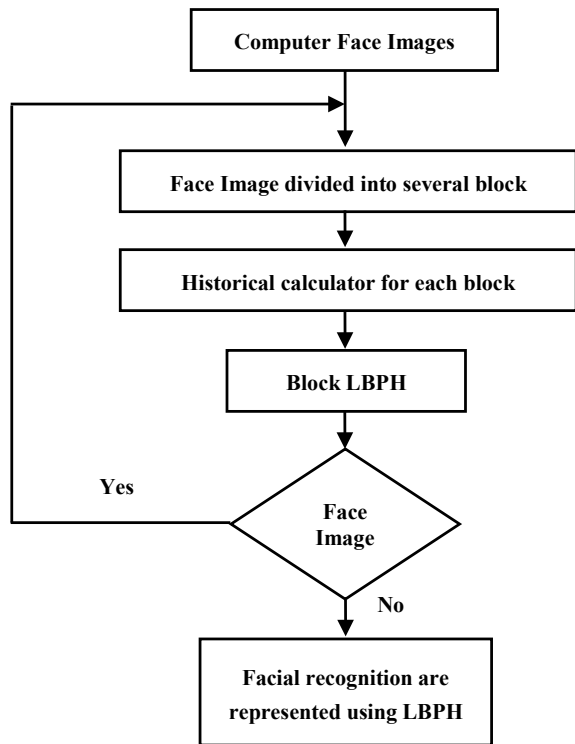
3.1 Haar Classifier

The geometric figures given in Fig. 1 represent five Haar cascade classifier patterns. By preparing the cascade with thousands of negative pictures, this cascade classifier operates with the positive picture superimposed. The facial characteristics may be detected by the sources. Each picture function is the only value that is obtained by removing the number of pixels from the sum of pixels under the black rectangle under the white rectangle.

3.2 Local Binary Pattern Histogram

Here, Fig. 2 shows the LBPH process below. This is one of the efficient methods for image recognition that is applied to images to extract face features. The obtained features measure the similarity between images. The texture operation turns any pixel of a grayscale picture into a binary number which was first came out in 1994 [22]. It is possible to display face images using a sample data vector while LBP and histogram are combined.

Fig. 2 LBPH process flowchart



3.3 Template Matching

A template matching technique is one of the straightforward methods used in facial recognition. When in doubt, do nothing and do something. The first is project-based and the second is project-oriented. We use the root and blueprint to combine features from the input and output. However, on the other hand, instead of trying to locate the actual template inside the picture, the entire template is globally or template-based. Some distinctive features like the nose, the ears, the eyes, and the mouth are essential in harvesting human faces.

4 Block Diagram of Design Prototype

Figure 3 shows the block diagram of the system. The diagram is split into two sections where the first system is based on augment reality (AR) and the final one is a design prototype of a security system based on face detection and recognition. In Fig. 3, a proposed augment reality (AR) system has been proposed. However, the block “Image Acquisition” processed the text data from real-time video streaming.

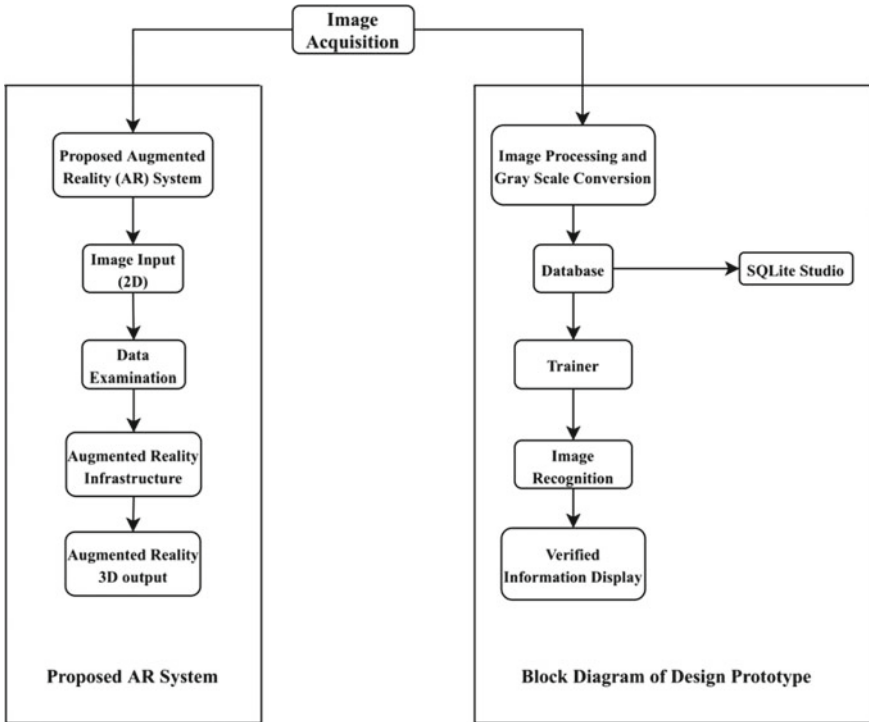


Fig. 3 Block diagram of the system

Then passed it to the next block for image processing where the pixel size will convert through an internal algorithm at the same time, the grayscale process is used to identify the pixel variety of individuals. Afterward, the processed data had transferred to the database and stored in the SQLite as local client data. Then all the data will be trained by the OpenCV platform where those pixels and nodal points of face will convert as a binary machine-readable cryptogram. After that, detector will detect the face and sending those data through the trainer to the database for justification of the user’s data. If the data matched with the database, then it will show on display and the security system will permit access to the user.

5 Methods of Simulations

5.1 Simulation with Template Matching

Figures 4 and 5 represent the face detection data using a template matching algorithm. Angle and light are the biggest problems in this process. It requires the same angle

Fig. 4 Single face detection and recognition

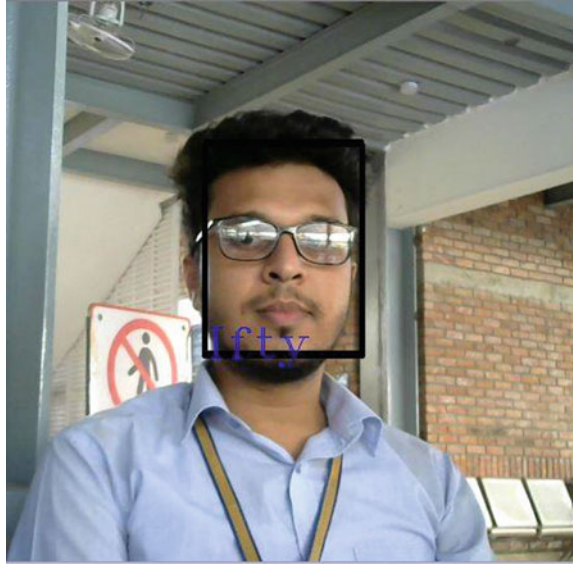
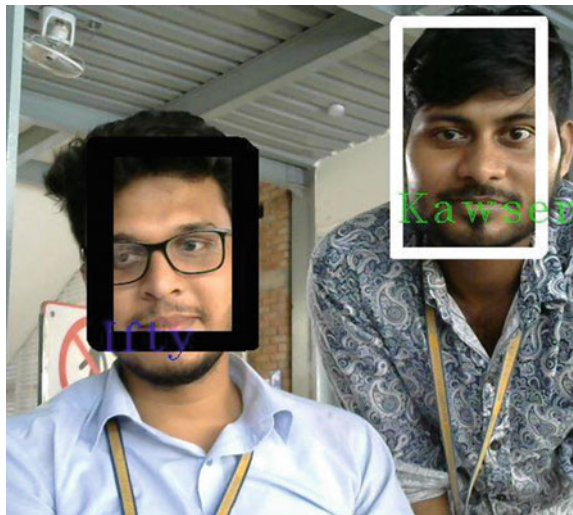


Fig. 5 Multiple face detection and recognition



as the trained picture to detect faces from real-time motion videos. There was a limitation in the low-light situation this algorithm struggles to detect faces. As a result, most of the time, it cannot detect the face. Sometimes, this procedure shows wrong information during face detection. On the other hand, it could be matched with other faces that were already in the database, then this template matching will not match with database data, so it will not proceed with its further system to the

user. So, it could be concluded that this procedure will need the exact light and angle to detect the face of users. Otherwise, it will not work properly.

- This algorithm has low accuracy than the Haar classifier which was used in the final system.
- Most of the time fails to detect the faces accurately according to various user's data.
- In the low-light condition, it cannot have the ability to detect the face as the template of the database is missing because of proper light.

5.2 Simulation with Local Binary Pattern Histogram

An LBPH dataset needs to provide four variables: neighbors, distance, X grid, and Y lattice. With an increase in the number of cells, the structure would be better and boost the vector's complexity. Figure 6 shows the trained dataset of every individual's user. The faces are cropped and resized to 222×222 pixels during the training process.

And Fig. 7 represents the information of every user in SQLite Studio. The application protocol interface was used to capture images. Its main feature was to find the front side of the faces. Every training process takes at least 20 pictures and resized those pictures and then converts them into grayscale images. In Figs. 8, 9, and 10 shows the detection and recognition data of the LBPH algorithm process with different lighting situations and distance.



Fig. 6 Trained database

The screenshot shows the SQLiteStudio interface. The main window displays a table with the following data:

ID	Name	Age	Gender	Department
1	12345 MASHRAFE	34	Male	BD Cricket
2	28751 IMRAN	23	Male	EEE

Fig. 7 Users information

Fig. 8 Face tracking in bright light



6 Final Hardware Implementation

After analyzing the template matching and LBPH algorithm performance, the LBPH has been chosen to design the security system. Here, 5 V and 2 A. A power supply provides power to this device. The display monitor and the Raspberry Pi are connected through an HDMI cable. The device has internet accessibility.

However, Fig. 11 shows the hardware implementation of the proposed system. The mouse and keyboard are connected to the USB port as a controller. The camera module is also connected through USB. After that, the operating system for Raspberry Pi has been installed. All the additional files, software (OpenCV, Python, etc.) and features had implemented gradually in the device. Raspberry Pi becomes ready to

Fig. 9 Face tracking in low light

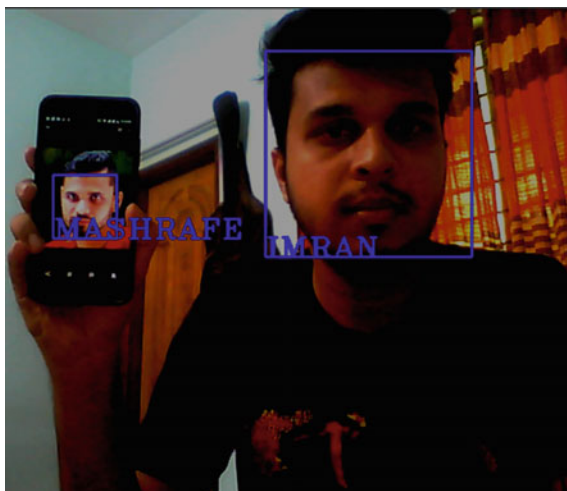


Fig. 10 Real-time single and multiple face recognition



use as a security device. This device works perfectly with one of the popular methods of face recognition known as a local binary histogram.

7 Results and Graphical Analysis

Light is an important element for image processing and face recognition. However, LBPH can work in normal lighting conditions as well as in low light situations. Table 1 shows the performance of template matching and LBPH algorithm with the different light situations.

Fig. 11 Security system implementation in Raspberry Pi



Table 1 Detection rate according to light

C	Detection rate in different lighting situation (%)		
	Low	Mid	Bright
Template matching	7	13	40
LBPH	75	87	90

The graph in Fig. 12 represents the detection rate of algorithms used in this project’s security device. The face detection rate of 90 percent being achieved using the LBPH method while the test was operated under varying lighting conditions. This face recognition and detection method are therefore highly successful in terms

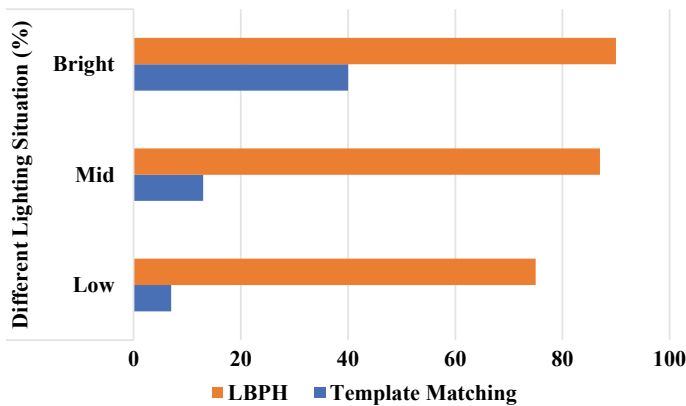


Fig. 12 Algorithm performance graph

Table 2 Detection based on distance

Distance (m)	Detected	Not detected
1	✓	✗
3	✓	✗
5	✓	✗
7	✓	✗
9	✓	✗
11	✓	✗
13	✓	✗
14	✓	✗
14.5	✓	✗
15	✓	✗
15.5	✗	✓
16	✗	✓

of safety. Also, one of the fastest process which takes milliseconds to recognize the user and implemented into this project prototype device.

Here, Table 2 shows how far the device can detect the faces. From 1 to 15 m, this security device detects the faces perfectly. But after 15 m, it struggles to detect faces though sometimes it works.

8 Cost Analysis of Design Prototype

Table 3 shows the total component cost to develop the design prototype of a security system based on face detection and recognition.

Table 3 Cost analysis of design prototype

Equipment name	Cost (USD \$)
Raspberry Pi 3 B+	59.52
Power supply	3.57
HDMI cable	10.15
Camera module	22.62
Total	\$106.51

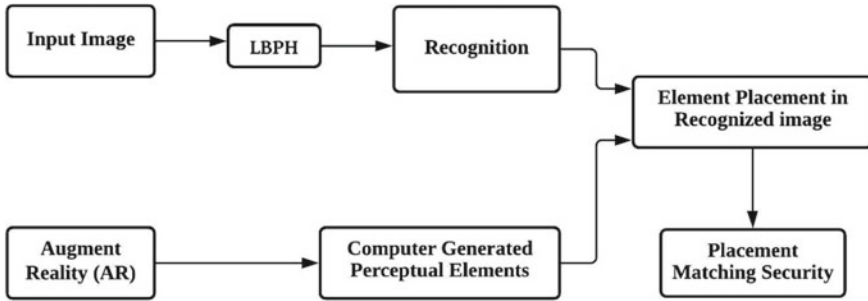


Fig. 13 Algorithm performance graph

9 Future Scope on Security System Based on AR

Augmented reality (AR) is a type of interaction in which a gadget functions as a lens through which a person interacts with their actual world while also being exposed to digital material. It has been researched and implemented in several industries such as gaming and retail and has gained increasing interest in the other sectors in recent years. The proposed LBPH-based face recognition security device prototype can guarantee greater security with the assistance of augmented reality. The detailed procedure to enlarge the security of proposed prototype is illustrated in Fig. 13.

Figure 13 demonstrates the proposed AR-based face recognition security device where security is ensured by placing computer-generated perceptual elements on the defined place of facial image recognized by the LBPH image recognition technique. However, the work has so many future scopes to research based on augmented reality (AR).

10 Conclusion

This face recognition security device prototype had designed especially for third-world countries. The lower cost of this device will help to reach out to every community of society. Although the efficiency rate of the LBPH algorithm is quite higher than that had used in this device. On the other hand, the efficiency rate at lower light also is quite effective and the percentage was 75 percent. By the way, it could detect the human face until 15 m without any issues. After 15 m, it could not detect properly the user's face. Today, the use of facial recognition technologies is economic, efficient, and extremely precise. It can be inferred without a doubt that this device may not have technical or financial obstacles to widespread production from the prototype. Furthermore, the use of augmented reality (AR) technology may make a system more efficient, adequate, transparent, and dependable. Since it is an image recognition approach, the output illustrations may simply be linked and transformed to augmented reality to improve and analyze the results.

References

1. Basha, C. Z., Reddy, M. R., Nikhil, K. H., Venkatesh, P. S., & Asish, A. V. (2020). Enhanced computer aided bone fracture detection employing X-Ray images by Harris corner technique. In: *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* Mar 11, 2020 (pp. 991–995). IEEE.
2. Nehru, M., & Padmavathi, S. (2017). Illumination invariant face detection using viola jones algorithm. In: *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)* Jan 6, 2017 (pp. 1–4). IEEE.
3. Vikram, K., & Padmavathi, S. (2017). Facial parts detection using Viola Jones algorithm. In: *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)* Jan 6, 2017 (pp. 1–4). IEEE.
4. Zeyad, M., Ghosh, S., & Ahmed, S. M. (2019). Design prototype of a smart household touch sensitive locker security system based on GSM technology. *International Journal of Power Electronics and Drive Systems*, 10(4), 1923.
5. Zeyad, M., Biswas, P., Iqbal, M. Z., Ghosh, S., & Biswas, P. (2018). Designing of micro-controller based home appliances governor circuits. *International Journal of Computer and Electrical Engineering (IJCEE)*, 10(2), 94–105.
6. Knežević, K., Mandić, E., Petrović, R., & Stojanović, B. (2018). Blur and motion blur influence on face recognition performance. In: *2018 14th Symposium on Neural Networks and Applications (NEUREL)* Nov 20, 2018 (pp. 1–5). IEEE.
7. Deng, W., Hu, J., & Guo, J. (2017). Face recognition via collaborative representation: Its discriminant nature and superposed representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(10), 2513–2521.
8. Mantoro, T., & Ayu, M.A. (2018). Multi-faces recognition process using Haar cascades and Eigenface methods. In: *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)* May 10, 2018 (pp. 1–5). IEEE.
9. Ai, J., Tian, R., Luo, Q., Jin, J., & Tang, B. (2019). Multi-scale rotation-invariant Haar-like feature integrated CNN-based ship detection algorithm of multiple-target environment in SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12), 10070–10087.
10. Epstein, D., & Feldman, D. (2017). Quadcopter tracks quadcopter via real-time shape fitting. *IEEE Robotics and Automation Letters*, 3(1), 544–550.
11. Ahmed, S. M., Al-Amin, M. R., Ahammed, S., Ahmed, F., Saleque, A. M., & Rahman, M. A. (2020). Design, construction and testing of parabolic solar cooker for rural households and refugee camp. *Solar Energy*, 15(205), 230–240.
12. Roesner, F., Kohno, T., & Molnar, D. (2014). Security and privacy for augmented reality systems. *Communications of the ACM*, 57(4), 88–96.
13. Qi, C., Li, M., Wang, Q., Zhang, H., Xing, J., Gao, Z., & Zhang, H. (2018). Facial expressions recognition based on cognition and mapped binary patterns. *IEEE Access*, 6, 18795–18803 (2018).
14. Uddin, M. Z., Hassan, M. M., Almogren, A., Alamri, A., Alrubaian, M., & Fortino, G. (2017). Facial expression recognition utilizing local direction-based robust features and deep belief network. *IEEE Access*, 5, 4525–4536 (2017).
15. Azzopardi, G., Greco, A., Saggese, A., & Vento, M. (2018). Fusion of domain-specific and trainable features for gender recognition from face images. *IEEE Access*, 6, 24171–24183 (2018).
16. Bouzakraoui, M. S., Sadiq, A., & Alaoui, A. Y. (2019, April). Appreciation of customer satisfaction through analysis facial expressions and emotions recognition. In *2019 4th World Conference on Complex Systems (WCCS)* (pp. 1–5). IEEE (2019).
17. Meena, D., & Sharan, R. (2016, December). An approach to face detection and recognition. In *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)* (pp. 1–6). IEEE (2016).

18. Othman, N. A., & Aydin, I. (2018, April). A face recognition method in the Internet of Things for security applications in smart homes and cities. In *2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)* (pp. 20–24). IEEE (2018).
19. Oh, S. J., Benenson, R., Fritz, M., & Schiele, B. (2016, October). Faceless person recognition: Privacy implications in social media. In *European Conference on Computer Vision* (pp. 19–35). Springer, Cham (2016).
20. Valstar, M. F., Sánchez-Lozano, E., Cohn, J. F., Jeni, L. A., Girard, J. M., Zhang, Z., et al., (2017, May). Fera 2017-addressing head pose in the third facial expression recognition and analysis challenge. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)* (pp. 839–847). IEEE (2017).
21. Scherhag, U., Rathgeb, C., Merkle, J., Breithaupt, R., & Busch, C. (2019). Face recognition systems under morphing attacks: A survey. *IEEE Access*, 7, 23012–23026 (2019).
22. Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, 19(6), 1657–1663.

A Reinforcement Learning Approach for Shortest Path Navigation in Automated Guided Vehicles for Medical Assistance



Pankaj Velavan, Abhishek Kaushik, Billy Jacob, and Mahak Sharma

Abstract With the advent of tremendous innovation in the field of autonomous navigation, the day-to-day chores of humans has been simplified. Autonomous navigation is used in all types of vehicles including aerial, ground and underwater vehicles. They are being used in vast domains like surveillance, delivery, self driving cars, etc. In addition to this, they also find application in the defence sector where they play a crucial role through surveillance in the border. The real challenge for the autonomous vehicles lies in obstacle detection and avoiding it intelligently. Another challenge for experts in the autonomous navigation field is identifying the shortest path. Hence, we propose a novel framework to tackle the obstacle detection problem and to identify the shortest path using a reinforcement learning approach that could be helpful in delivering food and medicines to the old and the special need people. The autonomous vehicle is trained and tested in the grid environment on both Q learning and Double Q learning approach. The performance of the algorithms is evaluated, and the metrics are discussed in the chapter.

Keywords Reinforcement learning · Anki-vector · Internet of things · Medical assistance

1 Introduction

With the outbreak of a pandemic like COVID-19 where close human interaction is at stake a lot of industries along with the population got affected both technically and financially. In such a situation where only essential services are allowed to operate,

P. Velavan · B. Jacob
Dublin Business School, 13/14 Aungier Street, Saint Peter's, Dublin, Ireland

A. Kaushik (✉)
Adapt Centre, Dublin City University, Dublin 9, Dublin, Ireland
e-mail: abhishek.kaushik2@mail.dcu.ie

M. Sharma
The Maharajah Sayajirao University of Baroda, Baroda, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
J. K. Verma et al. (eds.), *Advances in Augmented Reality and Virtual Reality*,
Studies in Computational Intelligence 998,
https://doi.org/10.1007/978-981-16-7220-0_12

one of the inevitable needs for people is food. Of course, many restaurants offer takeaway, and many food delivery companies employ a large workforce to make sure that they reach the customer on time. Even though they follow the safety precautions while delivery, still we are in a state where the experts are not sure about the means through which the virus spreads.

In the earlier days during the development of autonomous ground vehicles, the ultimate purpose was to use it in situations where the human operations are dangerous or highly inconvenient. They were designed with the intention of inspection, smart sensing and mapping in mind. With the development of technology and the increase in the demand of serving more purposes, today the AGVs are employed in various fields like surveillance, delivery, etc. With the colossal innovation in artificial intelligence, they have started to find application in transport systems, control systems, space science and manufacturing as suggested by Das and Mishra [1]. Berecz and Kiss [2] and team in their research have revealed that by the year 2050, around 10 million autonomous vehicles would hit the road and the industry will generate nearly 7 trillion dollars annual revenue. These autonomous vehicles leverage numerous combinations of machine learning algorithms, GPS coordinates and high-level image processing techniques. Even though the applications of AGVs are highly appealing yet they face a lot of shortcomings, and one such notable area as discussed by them is the guidance system for obstacle avoidance.

1.1 Research Question and Hypothesis

In this section, we have discussed about the framed the research question based on the problem statement. Our problem statement is “With the corona virus being highly contagious, the old and the special need people are the most vulnerable. Those type of people who stay at nursing homes are usually given food by the caretakers who in turn are more prone to the virus. So, a contactless viable solution to provide them food safely is the need of the hour”. Based on this, we have framed two research question.

- Exploratory research question: “How can reinforcement learning be used for navigation in shortest path using a robot?”
- Comparative Research Question: “How effective are Q learning and Double Q learning approach in navigation using shortest path?”

1.2 Limitations in Autonomous Navigation

In recent years, numerous research activity is being carried out in areas like autonomous vehicle, driver less car and autonomous food delivery, etc. Many research institutes and tech giants like Google, Waymo and Tesla have already created pro-

totypes and are learning about the real-world scenarios. There is a huge technical challenge in the production of such vehicles since there exists a huge gap in the proper understanding of the dynamic environment in which they are set to operate. The need for a high-level robust and mature product is mandatory for these vehicles to be seen on the road. The major setbacks in the road map of self-driving vehicles are safety, privacy and the law abidance. Recently, many incidents have been reported in news regarding the crashing of self-driving cars. One such fatal incident has been reported in Arizona wherein a 49-year-old woman lost her life [3]. The reason was that the operator was not cautious and failed to monitor the car. Although the driving car company was relieved from all criminal charges, still they faced criticism that the company's inadequate safety risk assessment procedures and ineffective oversight of vehicle operators were contributing factors.

The Irish government on December 2019 claimed that the Road Traffic (Miscellaneous Provisions) Bill will be amended allowing for testing, under strict guidelines, of vehicles in autonomous mode in the public roads [4]. With lots of legislation being passed the challenges for the companies becomes more laborious. The paper authored by Berecz and Kiss [2] sheds some light on the challenges faced by the autonomous vehicle including software related issues, viz. information protection, security breaches, technical issues that may occur due to poor road conditions. Myklebust et al. [5] in their study state that a safety case is required by the international standard for road vehicles, ISO 26262:2018 series. The UK has also issued a code of practice which requires the safety case to be issued. But these safety cases are highly technical for the public to understand. Henceforth, they have decided to come up with a safety case which is simple, precise and easy to understand for the public.

1.3 Solutions for Autonomous Navigation

Numerous groundbreaking ideas have been formulated for the effective and successful navigation of autonomous vehicles both in the real time and simulation environment. Some of the related works have been discussed further. In the healthcare sector, the contribution of autonomous vehicles is invaluable. The study conducted by Sathye et al. [6] was successful in delivering medicines to the patients in their bed without any human intervention. It makes use of the IR sensors which is used by the vehicle for effective manoeuvring in the specified path. In addition to the delivery of medicine, they have also devised a system to track the intake quantity of the patients and update it to the assigned doctors and family. Another major role for autonomous vehicles prevails in the security and surveillance domain. The paper authored by Ayub et al. [7] has developed a hybrid path planning algorithm that converts the path into a matrix. The highlighted path is converted as 1 and the rest of the positions as 0, and then, the UGV is instructed to follow the values 1 in the matrix and then make the decision to chose directions accordingly.

In addition to surveillance and health care, autonomous vehicles are also used widely in the field of agriculture. The research carried out by Gusland et al. [8] in col-

laboration with Norwegian Defence Research Establishment (FFI) clearly explains the difficulties faced by the vehicles for operation in harsh conditions like fog, snow, rain and obstacles present in vegetation. To overcome this, they have come up with the solution of multiple usage radar-S-band (MURA-S) to improve the range of perception of the LIDAR sensor. The usage of time-domain multiplexed multiple-input multiple-output (MIMO) techniques has brought in a remarkable increase in the cross-range resolution of the radar. With everything being able to be purchased online the present-day world is entirely depended on the online stores which in turn rely heavily on the timely delivery of their products. The paper authored by Karthik et al. [9] proposed the usage of donkey and carrot algorithm for the navigation of the autonomous vehicle using Google maps and the direction API.

With a huge number of vehicles on the road arises the problem of pollution which contributes towards global warming. To combat this problem, Liu et al. [10] came up with the optimized solution of two-echelon vehicle routing logic. In this methodology, the conventional delivery vans are used in the first echelon, while the autonomous delivery vehicle (ADVs) serves the delivery in the second echelon. This approach has not only proved to be emission-free but also have made a huge contribution to reducing the delivery cost for the companies. Obstacle awareness is a nightmare for every autonomous vehicle being designed. The success of an autonomous ground vehicle depends on its ability to perceive both static and dynamic obstacle. Henceforth, the team led by Rodrigues et al. [11] came up with a common hybrid control software framework designed to incorporate the autonomous vehicle with the ability of both reactive and deliberative planning. With the autonomous robots being widely used in various areas including the military the demand for a system which operates better in all-terrain is essential. Faisal et al. [12] came up with a fusion model using the fuzzy logic which makes use of various fuzzy rules for effective manoeuvring of the autonomous vehicle. The standard use of three sensors for the detection of obstacle results in the robot being struck in the dead zone which is a major setback. Their experiment exhibits the efficiency of the increased number of sensors combined with a fuzzy module.

1.4 Autonomous Navigation Solutions in Medical Industry

With autonomous vehicles widely being used in multiple fields, they also have not failed to revolutionize the medical industry. The study conducted by Shejwal and Behare [13] created an automated stretcher which is capable of transferring the patient to the allocated room for checking and testing. All the related information about the tests and checkups for the patients are updated in the RFID tag, and it can be used as the data source for effective maneuvering of the stretcher. Another study conducted by Choi and Choi [14] focused on the design and testing of an electric wheelchair platform which has the autonomous navigation capability using the QR code and magnetic band. In order to improve the quality and assistance of several medical services and facilities through new technologies, Fanti et al. [15] had come

up with an innovative idea for effective distribution of drugs to patients with limited number of autonomous vehicles and limited number of trips. The idea is mainly based on collaborative logistics concept. This strategy has also been formalized by an Integer linear programming problem to optimize the delivery tasks.

1.5 Reinforcement Learning Solutions for Autonomous Navigation

The field of autonomous vehicles has been greatly influenced by the advent of various machine learning and artificial intelligence algorithms. Deep reinforcement learning is one such method which has brought in phenomenal innovation and progress in the path planning, obstacle detection and intelligent decision-making skills of the autonomous vehicles. Let us have a look at some of the deep reinforcement learning techniques employed to cater to the purpose of the vehicles. The study conducted by Fu et al. [16] has focused on the problem of autonomous braking in case of emergency simultaneously taken into consideration the accuracy, efficiency and the comfort of the passengers.

1.6 Evaluation of Autonomous Navigation

Any research work without proper evaluation will not be able to predict the performance and credibility of the output model. Similarly, the evaluation of the model used by autonomous vehicles for navigation also is mandatory. The following are some of the remarkable work done in the field of autonomous vehicle navigation evaluation. The research conducted by Huang et al. [17] proposed that the evaluation of autonomous vehicles help in the improvement of driver safety as well would pave for many research problems. They have selected five UGVs from the IVFC 2012 and have analysed the performance of these vehicles on grounds of unstructured zone driving, environment perception, structural on-road driving and dynamic path planning. Gawel et al. [18] have carried out a groundbreaking evaluation on the performance of the autonomous ground vehicles in disasters scenarios using the search and rescue dataset. Different approaches to globally register the UGV generated 3D point-cloud data from LIDAR sensors have been made. Geometrical descriptor matching is the registration algorithm used for the proper exploitation of UAV data on UGVs. Every company in the self-driving industry wants to be the pioneer to introduce complete autonomy for driving. Yet as they stay on track of this goal, they are introducing partial autonomy through every single software update. There arises a scenario when the control transition between human and autonomy is smooth.

2 Conceptual Framework for Autonomous Navigation

As a part of this research work, we have developed a conceptual framework for the effective manoeuvring of the autonomous ground vehicle in the shortest path to reach its destination. This framework has been used in the ANKI vector robot for implementation purpose. This framework can be used in autonomous vehicles in a controlled environment for delivering food to mentally unstable people affected by COVID-19 [19]. The ANKI vector’s inbuilt high-resolution camera has been used for the obstacle detection purpose. Since the vector is a miniature of an autonomous robot, all the objects used as obstacles in the environment are small in size catering to the ease of detection for the robot.

The obstacle detection component helps in the identification of all the obstacles in the environment followed by the operation of the reinforcement learning component, which, with the help of prior knowledge about the shortest path helps the vector traverse the obstacles successfully. One of the main advantages of this framework is that it greatly reduces the time of delivery of food and medication to the affected people. To make this possible, the ANKI vector is trained in the environment to operate in the shortest path using the reinforcement learning component. The core components of this framework are the environment training, obstacle detection component and the reinforcement learning component. Figure 1 represents the flowchart of the entire framework.

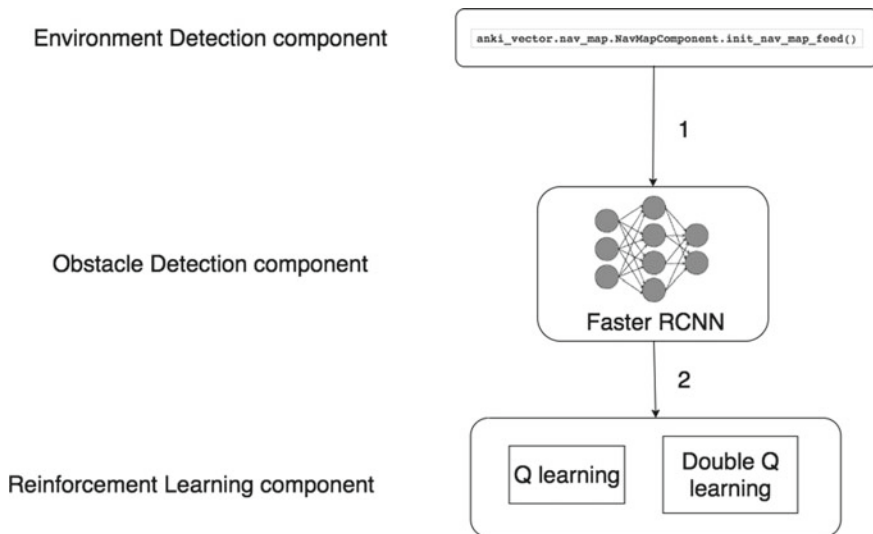


Fig. 1 Flowchart diagram

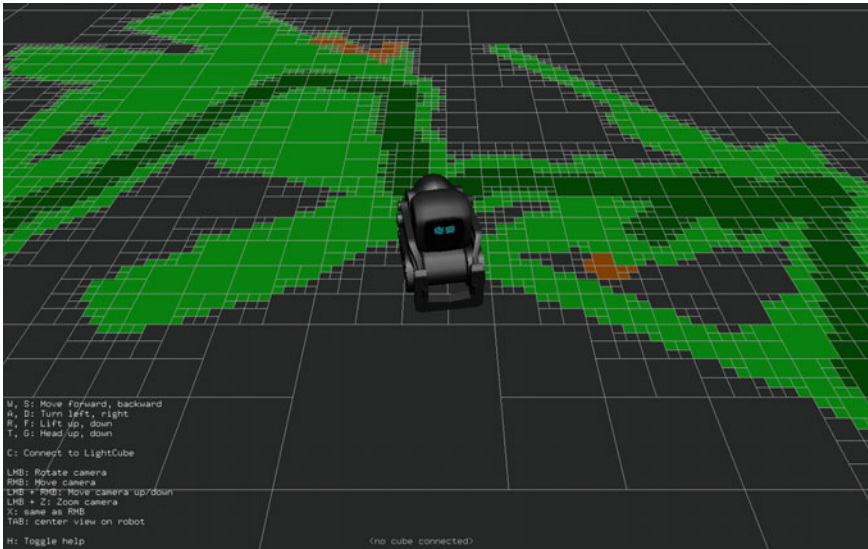


Fig. 2 2D-Mapping diagram

2.1 Environment Learning Component

The capability of the Anki vector to create a 2D memory map of the environment it operates has been leveraged for this experiment. Since all the obstacles that are present in the environment are snapped using the vector during the environment mapping phase, the vector is trained to detect them in the training phase. One of the important advantages of the vector is its ability to automatically detect cliffs/drops and visible edges and navigate around the obstacles. The 2D mapping capability of the vector differs from the standard occupancy map in that it does not take into account the probabilities of occupancy, but instead encodes the type of content present in that node. The study conducted by Chen et al. [20] points out that common problems in the path planning phase are the environment mapping and real-time localization. They have addressed these issues through the formulation of a heuristic Monte Carlo algorithm (HMCA) based on the Monte Carlo localization and discrete Hough transform (DHT). The Hough transform algorithm (HTA) initially captures all the information from the map and greatly reduces the real-time computational complexity. Then, followed by this using the global and local 2D occupancy grid maps (OGMs), the relative rotational angle and the corresponding spatial displacements of the robot are calculated. Thus, the importance of environment mapping can be understood, and we have carefully map our environment by operating the vector robot manually as shown in Fig. 2. Here, the red obstacles are highlighted using the red colour.

3 Obstacle Detection Component

Anki vector's time-of-flight distance sensor has a range of about 30 to 1200 mm with a field of view of 25 °C. The proximity sensor in the robot is located near the bottom between the two front wheels, facing forward. Using the metadata from the time of flight sensor, the vector calculates the distance of the obstacle from it. The distance is resolved to a certain quality value by the core engine based on the time of flight information within a field of view. For implementation purpose, we have trained the vector to detect objects from the Google open image dataset. The study by Leng et al. [21] aims at designing a robust obstacle detection and recognition for driver assistance in autonomous ground vehicles. In the first step, they use U–V disparity algorithm to process the V-disparity map and acquire detailed information about the features on the road. In the second step, the acquired features are fed to a context-aware faster-RCNN. This, in turn, combines the features to improve the recognition accuracy of small and occluded obstacles. The vector that we have used makes use of faster-RCNN for obstacle detection. It consists of a ResNet with an RPN and can detect more than 600 object categories. The faster-RCNN has been chosen for our implementation because it adopts the selective approach for zone segmentation which saves the forecasting time drastically.

3.1 Model-Free Algorithms

Model-free algorithms learn from their experience through trial and error methods. They do not depend on the reward function or the transition probability distribution to make decisions. Consider a traveller (agent) lost in a jungle (environment) with no signal on his mobile phone. He has to reach the nearby village before dusk. He has zero knowledge about the jungle, and the only way for him to reach the village is to try every possible route. Now, the traveller here is our agent. For every correct decision, it takes to move towards the village it is provided with a positive reward while every wrong move incurs a negative reward. These rewards for every action it takes are stored in a matrix. Hence, based on the total experience, the agent will learn to reach the village using the highest reward it had achieved against every state. Since they do not have any prior knowledge about the environment in every state, they take a lot of time to learn. The three different types of model-free reinforcement learning approach, viz. SARSA, Q learning and Actor critic method, have been explained in detail.

Q Learning Q learning is an off-policy reinforcement learning method. Q -learning is a value-based learning algorithm. It updates the value function based on the Bellman equation. The Bellman equation is as follows:

$$V(s) = \max[R(s, a) + \gamma V(s')] \quad (1)$$

where

$V(s)$ = value calculated at a particular point.

$R(s, a)$ = reward at a particular state s by performing an action a .

γ = discount factor.

$V(s')$ = the value at the previous state.

The ‘ Q ’ in Q -learning stands for quality. It represents the usefulness of a given action in gaining some future reward. The research conducted by Xu et al. [22] has focused on the decision-making ability of the autonomous vehicle through constant interaction with the human driver. They propose the adoption of Nash- Q learning-based motion decision algorithm for interaction. The results confirmed that the Nash- Q learning-based algorithm can improve the efficiency and comfort 20.71 % to the no interaction method, while the safety is ensured completely. The major advantage of the Q learning approach is that it does not require any prior knowledge about the environment. Compared to other methods the implementation is less complex because of the openai gym environment framework which enables the easy integration of coding for any use case.

Double Q Learning The double Q learning method is an off-policy-based approach where the policy used for evaluation differs from the one used to select the next action. One of the issues that arise during the calculation of values for the state value pair is the overestimation. This problem can be handled using the double Q learning method which uses double estimation to tackle the overestimation problem. The double Q learning makes use of the approach of double estimators. This means that instead of a single Q table there are two Q tables QA and QB whose values are updated. Here, the expected value of QB for the action ‘a’ is small or equal to the maximum value in the QA table. In layman’s term, with the increase in the number of iteration, the value of QB table will be always smaller than the maximal value in the QA table which assures that the QA values are never overestimated. An interesting research was done by Zhang et al. [23] which focused on the speed control of the autonomous car in a dynamic real-time environment. The usage of an integrated perception approach to creating an environment has helped them to produce an efficient model whose score is 271.73 times better than that of deep Q -learning.

3.2 Model-Based Algorithms

In the model-based reinforcement learning algorithm, the environment is modelled as the Markov decision process with the following elements taken into consideration. The core elements are the set of states, actions, transition probability function and the reward function. The model-based model comprises of a cost function to predict the optimal solution in the long run. One of the crucial traits on the model-based reinforcement algorithms is that one of their greater advantage of having more assumptions and approximations is their disadvantage. Even though they have all the assumptions in every state, still they are restricted to take only those actions as the response.

Deep Q Learning As the name suggests it makes use of the deep learning network to train the agent in the environment. This method comes handy for scenarios where there are more states and actions. For example, in real-time scenarios, if there are more than 30,000 states and 5000 actions then this would result in the creation of a Q table which is huge and requires more memory to store them. Another important issue, in this case, is the amount of energy and time it consumes to traverse and access each cell. The only disadvantage the deep learning initially suffered was the unstable target which is dynamic for each iteration. Hence, to tackle this, there were two separate networks used. One network constitutes the target network with frozen parameter, while the other network acts as the function approximator. Another method that is used to handle this issue is experience replay. This method stores the state, action, reward and the next state in a large table for every frame, and in the future, a subset of these values can be used as a sample for training our model. Since the samples are random, the correlation will be low while simultaneously have higher sampling efficiency. Since our framework has only seven states and 2 actions, the need for a deep neural network is void. Okuyama et al. [24] and his team had successfully implemented the deep Q learning approach to improve the accuracy of the autonomous vehicle to intelligently avoid the obstacles and follow the lane markings.

4 Implementation

4.1 ANKI Vector

The implementation has been done using the Anki vector robot. Vector is a pet robot whose design allows it to intuitively interact with its surroundings. It can recognize people and objects and can also avoid obstacles intelligently. In addition to this, it can be controlled using voice and can also be used to take photographs. In short, the vector acts like a human companion which makes it an ideal choice for our research work. It can be used as a companion for elderly people in care units who are alone, especially during this pandemic situation. As far as our implementation is considered, we are using a grid setup to simulate the environment for our vector to operate in. The 2D mapping capacity of the ANKI vector has been leveraged, and all the obstacles have been recorded. The ANKI vector used in this research work is as shown in Fig. 3.

4.2 ANKI Vector Specification

The tech specification of anki vector is as follows [25]:

1. Quad-core Qualcomm Snapdragon processor, 1.2 GHz
2. 4-microphone array
3. Single point time-of-flight NIR laser, 1 m range

Fig. 3 ANKI vector

4. 720p camera
5. 802.11n Wi-Fi
6. Bluetooth
7. Capacitive top and bottom casing
8. 4 cliff sensors.

4.3 Grid Environment

The implementation has been done in a real-time grid environment, created specially for this research work. The study conducted by Aradi et al. [26] has demonstrated the use of the grid approach for training the autonomous vehicle to plan the path. Similarly, the grid created by us is as shown in Fig. 4. The distance between each node points is marked, and the distance values are 0.1, 0.5, 0.7, 0.8. All the measurements are in metres. To train the model to move in the shortest path, prior knowledge about the shortest path in the grid is essential. Hence, we have used Dijkstra's algorithm to calculate the shortest path and the nodes through which it has to traverse have been identified as shown in Fig. 5. The nodes here are not the obstacles but the junctions at every turn.

4.4 Open AI Framework

The open AI is an artificial intelligence research laboratory founded by Elon Musk and others to develop machine learning open-source simulation environment. In this research work, the gym environment has been used to create a customized vector environment as per our requirement. Meyer et al. [27] in their study have made an

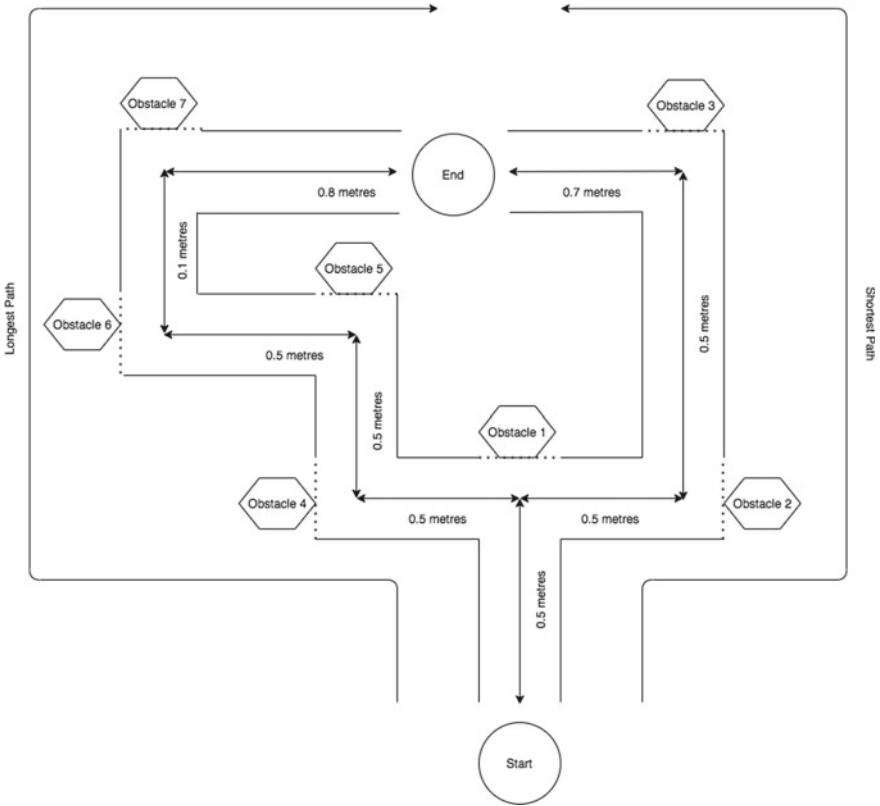


Fig. 4 AGV navigation path

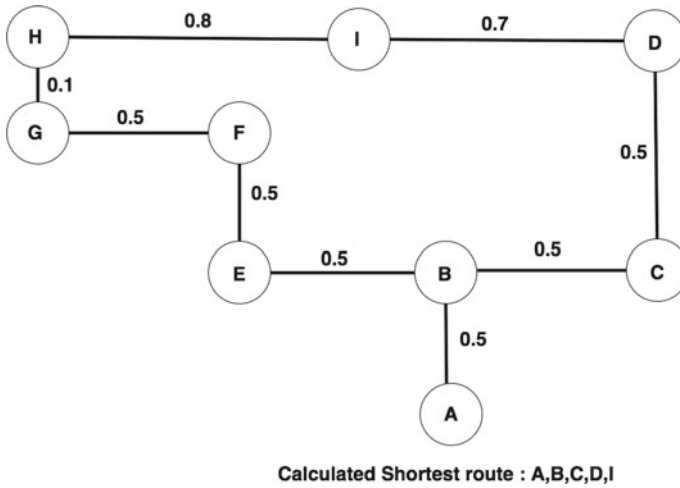


Fig. 5 Dijkstra's algorithm evaluation

autonomous vehicle which is equipped with multiple rangefinder sensors for obstacle detection and have trained it in a simulation environment based on the OpenAI gym Python Toolkit. Similarly, we have also implemented the shortest pathfinding algorithm using the OpenAI gym Python Toolkit. We have 7 states each of which is the obstacle and the available actions are 2.

4.5 *Obstacle Detection Component*

We have used the faster-RCNN algorithm for the detection of objects. The Tensor-Flow model used here is already pretrained on the Google open image dataset with the capability to detect around 600 objects. The ResNet along with regional proposal network (RPN) has been used here. The major advantage of the RPN is the ability to identify multiple objects which is an essential feature when implementing in the real-time scenario. In order to reduce the obstacle detection time of the vector, the library of pretrained objects have been tweaked. This had improved the detection time by three times.

Q Learning Implementation The core functionality of the *Q* learning implementation is to train the agent in the environment and update the state action pair value in the *Q* table. *Q* learning is an off-policy model-free reinforcement learning algorithm. The reinforcement learning algorithm works on the basis of the feedback function to take actions against each state. The five core components in *Q*-Learning are agent, environment, reward, *Q*-table and action.

- **Agent:** The agent is the module which decisions based on positive and negative rewards. In our research work, the agent is the vector, and it takes the decision to move left or right and gets reward during training.
- **Environment:** The environment is the arena in which our vector will operate and the environment comprises of the agent, state and action. The agent vector, along with the obstacles (state) and movement in the right and left direction, (action) constitutes the environment for our work.
- **Actions:** We have two actions for our research purpose. The movement of the vector in the right and left direction.
- **States:** The states in our research work are real-time objects. We have seven states, viz. glasses, watch, bottle, doll, fork, pen, coin, computer mouse.
- ***Q*-table:** *Q*-Table is a simple matrix or a lookup table where we calculate the maximum rewards for an action at each state. The *Q*-table will guide the agent to

the best action at each state. Our Q table will initially look like $\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ where each row corresponds to a state (7) and each column corresponds to an action (2).

$$Q(s, a) = Q(s, a) + \alpha[r + \gamma \max_a Q(s', a) - Q(s, a)] \tag{2}$$

where

- Q = Q -table present in the Q -learning environment.
- s = the various boxes available for the drone to reach.
- a = action, the selection of a certain direction to move by the agent in the environment.
- $Q(s, a)$ = new q -value (value present in the q -table in s th row and a th column).
- $Q(s', a')$ = old q -value in the q -table in s th row and a th column.
- α = the learning rate, set between 0 and 1,
- γ = discount factor, also set between 0 and 1,
- \max_a = the maximum reward that is attainable in the state following the current state.

4.6 Training Phase

The vector was trained for an episode of 200 with the Q table able to learn completely by the 100th episode which was later found from the evaluation. The vector was trained in both the Q learning and the double Q learning method. During training, the rewards have been set accordingly to punish the model if it takes the wrong decision and a positive reward for the correct decision. Based on this, we had trained our vector to traverse in the path, and it was evident that after the 100th episode it had always started to reach its destination using the shortest path. For training purpose, the obstacles used are all real-time obstacles. From our training phase results, it was clear that the Q learning approach had performed better compared to the double Q learning method.

4.7 Deployment Phase

Our main purpose revolves around the problem statement wherein we aim to deliver food and medicine to the aged people at their door as it serves to be contactless during

the pandemic situation. From the training phase, we had identified that the Q learning performs better than the double Q learning method. Hence, we had deployed it to traverse in the shortest path and were able to achieve our goal successfully. The end point in our implementation is to detect a particular obstacle which will indicate the vector that the destination has been reached successfully. Since it had been trained for 200 episodes, it had prelearned the values for all states, and upon deployment, it was able to reach the destination successfully in the first trial itself.

5 Evaluation

The vital part of any research study is the evaluation of the framework developed and to analyse its performance metrics. Moving forward with this approach, we have evaluated our models by broadly comparing the performance of both the Q learning and the double Q learning implementation. The performance of the Q learning and Double Q learning by fine-tuning their hyperparameters have also been compared.

5.1 Comparison of Q Learning and Double Q Learning

Figure 6 depicts the performance of the training for 1000 episodes. From Fig. 6 it is evident that the Q learning approach has learned the environment completely by 100th episode and has started to allocate positive rewards consistently. On the other hand, the double Q learning approach has failed to learn the environment at any given number of episodes and was consistently trying to learn the environment by allocating positive and negative feedback even at the last stage. The Q tables obtained from the training are as shown in Fig. 7. One of the interesting insight observed from the double Q learning is that even though it had failed to learn completely at a particular stage still it had managed to allocate the rewards efficiently with only higher weighted mean.

5.2 Q Learning Hyperparameters

The alpha value was set to 0.2, epsilon value was set to 1, and the gamma value was set to 0.2 for the training purpose with 1000 episodes as criteria. For the evaluation purpose, the number of episodes had been set to 200, and the fine-tuning of hyperparameters like alpha value and epsilon value were done to get a broader picture of the performance of the Q table.

Effect of Learning Rate For better evaluation of the learning rate of our Q table, the total accumulated rewards are calculated and are as shown in Table 1. Here also

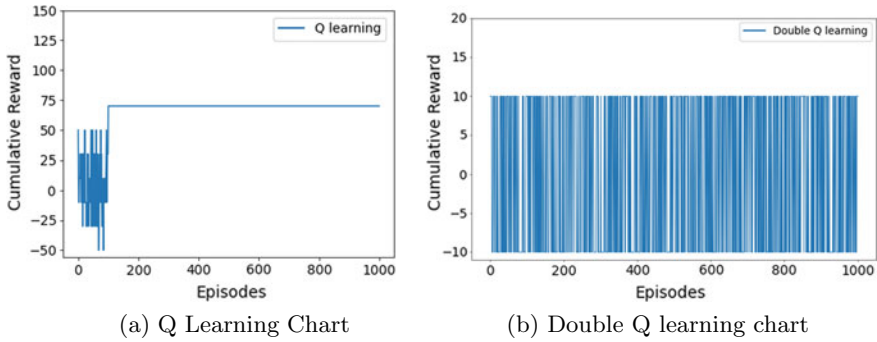


Fig. 6 Q Learning rewards/episodes chart—epsilon values

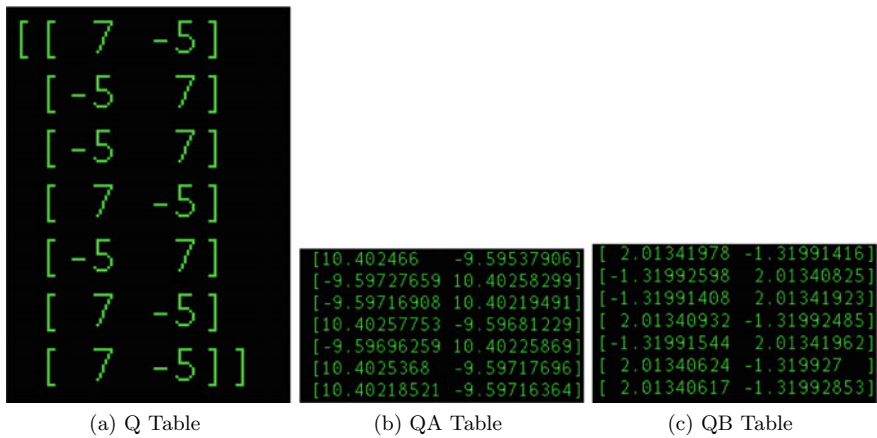


Fig. 7 Comparison of Q table and QA table, QB table

Table 1 Table for successful navigation with respect to learning rate

α -value	Episodes	Total rewards
0.2	200	6860
0.5	200	7080
1	200	7380

the x and y -axis are episodes and rewards, respectively. The plots for the following alpha values were generated and compared. Alpha = 0.2, 0.5, 1. From Table 1, it is evident that with the increase in the alpha value the total reward gain also increases. The learning rate decides on to what extent the overriding of values for the old information can be allowed. With higher the learning rate, the agent will tend to use only the recently learned information and would not depend on the past data. The charts for the varied alpha values are as shown in Fig. 8.

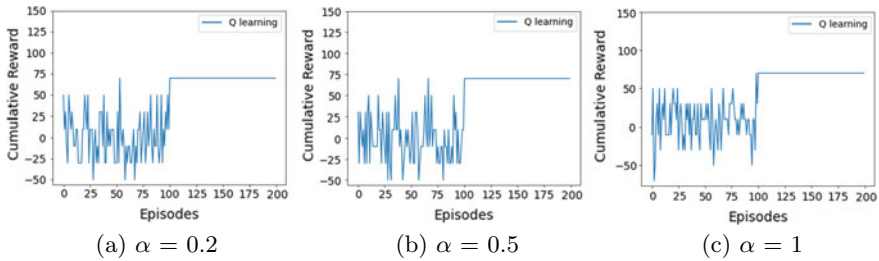


Fig. 8 Q Learning rewards/episodes chart—learning rate

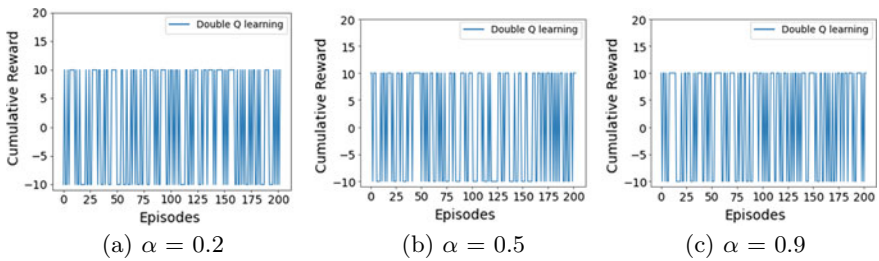


Fig. 9 DQ Learning rewards/episodes chart—learning rate

5.3 Double Q Learning Hyperparameters

Initially, the alpha value was set to 0.2, epsilon value to 1 and the gamma value to 0.2 for training purpose with 1000 episodes same as what was carried out for the Q learning. For the evaluation purpose, the number of episodes had been set to 200 and the fine tuning of hyperparameters like alpha value and epsilon value were done to get a better idea about the individual QA and QB tables as well over the Q table obtained from the Q learning.

Effect of Learning Rate The episodes and rewards are plotted in the x-axis and y-axis respectively. The plots for the following alpha values were generated and compared. Alpha = 0.2, 0.5, 0.9. Here, the learning rate as mentioned previously was not much promising and the learning is constant throughout the training process. But compared to the tuning of the epsilon value. it can be observed that the rewards for the varying learning rates are highly dynamic for every episode. While tuning the epsilon values, the reward allocation was constant for at least a few episodes. The charts for the varied alpha values are as shown in Fig. 9.

6 Conclusion and Future Works

In this research work, a novel framework has been suggested for effective implementation of reinforcement learning for autonomous vehicle navigation in the shortest path. The application of reinforcement learning for the navigation of ANKI vector in the shortest path using the Q learning method answers our exploratory research Question “How can reinforcement learning be used for navigation in the shortest path using a robot?”. This was achieved through the proper analysis of previous research works carried out in the area. Followed by this, the Q learning and double Q learning methods were used to train the Q table. Post-training, the evaluation was done on them, and it was evident that the Q learning table of the Q learning approach was trained within the limited number of episodes. Hence, the final decision to use the Q learning approach for deployment was made. Our comparative research Question “How effective are Q learning and Double Q learning approach in navigation using shortest path?” was answered by comparing their Q tables as well by evaluating their hyperparameter metrics. In this research work, we have showcased the idea of navigating the robot to the destination using the shortest path. This framework was designed with the idea of delivering medicines to the aged people who are affected during the pandemic in care houses, and hospitals where contactless interaction is of prime importance. This method can be extended to realtime by incorporating more storage space for medicines and food. The extension of this idea for multiple destination points would be a more useful use case. As ANKI vector has the capability of interaction, it can be used to interact with the aged people who feel lonely and could act as a better companion for them.

References

1. Das, S., & Mishra, S. K. (2019). A review on vision based control of autonomous vehicles using artificial intelligence techniques. In *Proceedings—2019 International Conference on Information Technology, ICIT 2019* (pp. 500–504), December 2019. <https://doi.org/10.1109/ICIT48102.2019.00094>
2. Berecz, C. E., & Kiss, G. (2018). Dangers in autonomous vehicles. In *18th IEEE International Symposium on Computational Intelligence and Informatics, CINTI 2018—Proceedings* (pp. 263–268), November 2018. <https://doi.org/10.1109/CINTI.2018.8928189>
3. Uber’s self-driving operator charged over fatal crash. BBC News. <https://www.bbc.com/news/technology-54175359>. Accessed January 06, 2021.
4. New rules to allow testing of self-driving vehicles on Irish roads. <https://www.irishtimes.com/business/transport-and-tourism/new-rules-to-allow-testing-of-self-driving-vehicles-on-irish-roads-1.4107190>. Accessed January 06, 2021.
5. Myklebust, T., Stalhane, T., Jenssen, G. D., & Waro, I. (2020). Autonomous cars, trust and safety case for the public. In *Proceedings—Annual Reliability and Maintainability Symposium* (Vol. 2020-January), January 2020. <https://doi.org/10.1109/RAMS48030.2020.9153618>
6. Sathye, R., Surve, M., Deshpabhu, S., Karia, D. D., & Dahatonde, R. (2020). Medicine delivering smart autonomous vehicle. In *Proceedings of 2020 IEEE-HYDCON International Conference on Engineering in the 4th Industrial Revolution, HYDCON 2020*, September 2020. <https://doi.org/10.1109/HYDCON48903.2020.9242651>

7. Ayub, M. F., Ghawash, F., Shabbir, M. A., Kamran, M., & Butt, F. A. (2018). Next generation security and surveillance system using autonomous vehicles. In *Proceedings of 5th IEEE Conference on Ubiquitous Positioning, Indoor Navigation and Location-Based Services, UPINLBS 2018*, December 2018. <https://doi.org/10.1109/UPINLBS.2018.8559744>
8. Gusland, Y., Torvik, B., Finden, E., Gulbrandsen, F., & Smestad, R. (2019). Imaging radar for navigation and surveillance on an autonomous unmanned ground vehicle capable of detecting obstacles obscured by vegetation. In *2019 IEEE Radar Conference, RadarConf 2019*, April 2019.
9. Karthik, S., Bannur, A., Rao, A. T., & Krishna, Y. (2019). Autonomous UGV delivery systems. In *CSITSS 2019—2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution, Proceedings*, December 2019. <https://doi.org/10.1109/CSITSS47250.2019.9031017>
10. Liu, D., et al. (2020). Two-Echelon vehicle-routing problem: Optimization of autonomous delivery vehicle-assisted E-grocery distribution. *IEEE Access*, 8, 108705–108719.
11. Rodrigues, M., McGordon, A., Gest, G., & Marco, J. (2017). Developing and testing of control software framework for autonomous ground vehicle. In *2017 IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2017* (pp. 4–10), June 2017.
12. Faisal, M., Hedjar, R., Al Sulaiman, M., & Al-Mutib, K. (2013). Fuzzy logic navigation and obstacle avoidance by a mobile robot in an unknown dynamic environment. *International Journal of Advanced Robotic Systems*, 10. <https://doi.org/10.5772/54427>
13. Shejwal, Y., & Behare, M. (2020). AGV based stretcher. In *Proceedings of the 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing, ICSIDEMPC 2020* (pp. 200–201), October 2020. <https://doi.org/10.1109/ICSIDEMPC49020.2020.9299646>
14. Choi, J. H., & Choi, B. J. (2018). Design of self-localization based autonomous driving platform for an electric wheelchair. In *2017 Asian Control Conference, ASCC 2017* (Vol. 2018-January, pp. 465–466), February 2018. <https://doi.org/10.1109/ASCC.2017.8287214>
15. Fanti, M. P., Mangini, A. M., Roccotelli, M., & Silvestri, B. (2020). Hospital drugs distribution with autonomous robot vehicles. In *IEEE International Conference on Automation Science and Engineering* (Vol. 2020-August, pp. 1025–1030), August 2020. <https://doi.org/10.1109/CASE48305.2020.9217043>
16. Fu, Y., Li, C., Yu, F. R., Luan, T. H., & Zhang, Y. (2020). A decision-making strategy for vehicle autonomous braking in emergency via deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 69(6), 5876–5888. <https://doi.org/10.1109/TVT.2020.2986005>
17. Huang, W., Wen, D., Geng, J., & Zheng, N. N. (2014). Task-specific performance evaluation of UGVs: Case studies at the IVFC. *IEEE Transactions on Intelligent Transportation Systems*, 15(5), 1969–1979. <https://doi.org/10.1109/TITS.2014.2308540>
18. Gawel, A., Dube, R., Surmann, H., Nieto, J., Siegwart, R., & Cadena, C. (2017). 3D registration of aerial and ground robots for disaster response: An evaluation of features, descriptors, and transformation estimation. In *SSRR 2017—15th IEEE International Symposium on Safety, Security and Rescue Robotics, Conference* (pp. 27–34), October 2017. <https://doi.org/10.1109/SSRR.2017.8088136>
19. COVID-19 has serious effects on people with severe mental illness. <https://www.healthline.com/health-news/covid-19serious/effects/people/with/mental/health/disorders/Strain/on/the/system>. Accessed January 08, 2021.
20. Chen, D., Weng, J., Huang, F., Zhou, J., Mao, Y., & Liu, X. (2020). Heuristic Monte Carlo algorithm for unmanned ground vehicles realtime localization and mapping. *IEEE Transactions on Vehicular Technology*, 69(10), 10642–10655. <https://doi.org/10.1109/TVT.2020.3019581>
21. Leng, J., Liu, Y., Du, D., Zhang, T., & Quan, P. (2020). Robust obstacle detection and recognition for driver assistance systems. *IEEE Transactions on Intelligent Transportation Systems*, 21(4), 1560–1571. <https://doi.org/10.1109/TITS.2019.2909275>
22. Xu, C., Zhao, W., Li, L., Chen, Q., Kuang, D., & Zhou, J. (2020). A nash Q-learning based motion decision algorithm with considering interaction to traffic participants. *IEEE Transactions on Vehicular Technology*, 69(11), 12621–12634. <https://doi.org/10.1109/TVT.2020.3027352>

23. Zhang, Y., Sun, P., Yin, Y., Lin, L., & Wang, X. (2018). Human-like autonomous vehicle speed control by deep reinforcement learning with double Q-learning. In *IEEE Intelligent Vehicles Symposium, Proceedings* (Vol. 2018-June, pp. 1251–1256), October 2018.
24. Okuyama, T., Gonsalves, T., & Upadhyay, J. (2018). Autonomous driving system based on deep Q learning. In *2018 International Conference on Intelligent Autonomous Systems, ICoIAS 2018* (pp. 201–205), October 2018. <https://doi.org/10.1109/ICoIAS.2018.8494053>
25. The new Anki Vector robot is smart enough to just hang out. The Verge. <https://www.theverge.com/2018/8/8/17661902/anki-vector-home-robot-voice-assistant-ai>. Accessed January 11, 2021.
26. Aradi, S. (2020). Survey of deep reinforcement learning for motion planning of autonomous vehicles, January 30, 2020. <https://doi.org/10.1109/tits.2020.3024655>
27. Meyer, E., Robinson, H., Rasheed, A., & San, O. (2020). Taming an autonomous surface vehicle for path following and collision avoidance using deep reinforcement learning. *IEEE Access*, 8, 41466–41481. <https://doi.org/10.1109/ACCESS.2020.2976586>