Multimedia Tools and Applications (2021) 80:34297–34309
https://doi.org/10.1007/s11042-021-10809-z

1135T: SOCIAL MULTIMEDIA PROCESSING

Design of an intelligent video surveillance system for crime prevention: applying deep learning technology

Chang-Soo Sung[1] · Joo Yeon Park[2]

Received: 17 January 2020 / Revised: 16 November 2020 / Accepted: 10 March 2021 /
Published online: 24 March 2021

Abstract

As the security threat and crime rate have been increased all over the globe, the video surveillance system using closed-circuit television (CCTV) has become an essential tool for many security-related applications and is widely used in many areas as a monitoring system. However, most of the data collected by the video surveillance system is used as evidence of objective data after crime and disaster have occurred. And, often time, video surveillance systems tend to be used in a passive manner due to the high cost and human resources. The video surveillance system should actively respond to detect crime and accidents in advance through real-time monitoring and immediately transmit data in case of an accident. This study proposes developing an intelligent video surveillance system that can actively monitor in real-time without human input. In solving the problems of the existing video surveillance system, deep learning technology will be carried through the data processing model design to visualize data for crime detection after building an artificial intelligence server and video surveillance camera. In addition, this design proposes an intelligent surveillance system to quickly and effectively detect crimes by sending a video image and notification message to the web through real-time processing.

Keywords Video surveillance system · Deep learning · Artificial intelligence · Crime prevention

✉ Joo Yeon Park
jooyeon.park@murdoch.edu.au

Chang-Soo Sung
redsun44@dongguk.edu

[1] Department of Technology Entrepreneurship, Dongguk University-Seoul, Seoul 04620, Korea

[2] Discipline of Information Technology, Media and Communication, Murdoch University, Murdoch, Western Australia 6150, Australia

1 Introduction

Today, the video surveillance system using closed-circuit television (CCTV) has become an essential tool for many security-related fields and is widely used in public not only for monitoring traffics, human behaviors but also for preventing crime and disaster. With increasing security threats and crime rates worldwide, the video surveillance market continues to grow in various sectors [2]. In addition, as video surveillance systems have become widespread, the role of big data and its processing power have become equally important. Recent new technologies such as artificial intelligence and big data infrastructure play a big role in improving the scalability and accuracy of video surveillance solutions. Therefore, research into intelligent-based surveillance systems is becoming an important area for disaster and crime prevention as well as industrial and public security protection. There are several proposed models for intelligent traffic video surveillance and accident identification [43] and enhanced algorithms for the resolution of intelligent video system images [20]. However, few studies have been explored video surveillance systems with cutting-edge technologies such as deep learning [12, 23].

Moreover, video surveillance systems are widely used in public and various areas, especially for people and traffic observation. But still, they are not yet widely used for crime and disaster surveillance and prevention [44]. Criminal investigations and disaster observations are still analyzed based on objective data collected from video surveillance systems. In other words, video surveillance systems are used passively and are mostly used only as evidence after a crime has occurred. In addition, the existing video surveillance system requires a lot of labor and cost to manage because the video must be continuously monitored in real-time to activate it [16, 17]. However, for better crime detection and prevention, video surveillance systems should actively respond to incidents and process images in real-time in a cost-effective manner. This can be managed through big data and artificial intelligence models.

Therefore, this study proposes developing an intelligent video surveillance system to actively monitor itself in real-time without direct detection from human labor. In solving the problems of the existing video surveillance system, deep learning technology will be carried through the data processing model design to visualize data for crime detection after building an artificial intelligence server and video surveillance camera. In addition, this design proposes an intelligent surveillance system to quickly and effectively detect crimes by sending a video image and notification message to the web through real-time processing. The related work regarding video surveillance systems and deep learning will be discussed in the following section before addressing the system configuration and development process. The insights and contributions to this design of an intelligent surveillance system will also be addressed.

2 Intelligent video surveillance system

The closed-circuit television (CCTV), often known as video surveillance, is a system that automatically identifies particular objects and behaviors through the programs [13]. In other terms, advanced video surveillance is referred to as “an intelligent video processing technique designed to assist security personnel by providing reliable real-time alerts and to support efficient video analysis for forensic investigations” [31]. It transmits a signal to a specific place from a limited set of monitors using a video camera [14]. It captures light to convert it into an electrical signal to process recording and display on the screen with a video signal [5].

Recently, video surveillance system usage has been largely increased in many areas with various sectors for public surveillance, and its market has been growing continuously. For example, it could be used in crime prevention, traffic monitoring, employee monitoring, and even a sporting event. As such, an image processing-based intelligent robotic system has been used in various sectors. It can be used to detect crop diseases in agricultural sectors and measure the number of sperms for sperm health in the medical field [29, 45]. Some research has classified such intelligent video surveillance systems into their desirable characteristics and features; object detection, tracking and movement analysis, detect abnormals, vehicle detection and traffic analysis, object counting systems, privacy-preserving systems, etc. [46].

Through full human intervention, the traditional monitoring systems have several disadvantages, such as high labor costs, long-term capture, and limited ability to monitor simultaneous screens. Thus, the conventional surveillance systems have been replaced with advanced intelligent systems to capture better abnormal behavior and unidentified patterns with developing technologies such as artificial intelligence technology and pattern recognition system. Moreover, as technologies have been advanced, video surveillance is evolving with the new technology trends such as big data infrastructure, artificial intelligence, and neural network systems [35]. The video surveillance market has shifted from traditional analog video to an IP video surveillance system incorporating new technologies that improve processing power. The evolution of artificial intelligence by deep learning is especially directly used in video surveillance systems, enabling proactive prediction of security accidents and predictive analysis that can be prepared in advance. Intelligent video surveillance systems provide flexible control over the speed of video data collection and increase data collection speed whenever a security accident indicator is detected, providing more information, enabling accurate and reliable analysis. Big Data also provides a new way to store and access video data seamlessly and cost-effectively to video surveillance systems. These advanced technology-applied surveillance systems can be automatically monitored objects and actions with accuracy, even with fewer observers.

Several different types of an architectural model for image captioning are used in the surveillance system, including image feature-based and RCNN object-based. The semantic-based architectural model uses many techniques such as Artificial vector machines (ANN), Support Vector Machines (SVM), and Bayesian networks (BN) [18, 30, 38]. The convolution neural networks (CNN), the most popular type used in deep learning technology and models such as VGG, AlexNet, Inception, are most predominant. Table 1 shows the types of architecture for image captioning with different features [40].

Even though the increasing demands of intelligent video surveillance systems, there are contradictive issues to deal with when using in public spaces such as large traffic density, heterogeneity, and privacy. Privacy has especially been an ongoing debate issue with the

Table 1 Architecture type of image captioning

Architecture type	Feature
Image Feature-based	Inception, AlexNet, VGG, RGB feature
Object aggregation based	Detected objects, Objects sequential combination
Semantic-based	ANN, SVM, BN
RCNN object feature-based	CNN, Fast CNN, faster CNN.
others	nearest neighbor (NN), ranking, query, reinforcement learning

surveillant systems embedding with all the high-end intelligent technology [3]. People like to feel safe, but at the same time, they generally dislike to be watched their activities everywhere. Therefore, when developing intelligent video surveillance, concern needs to be taken to satisfy the capturing surveillance object while ensuring privacy to people [47]. This development aims to provide real-time and objective data through active monitoring using artificial intelligence to prevent crimes and disasters.

3 Deep leaning technology

Recent activities in various networks, including social media, streaming video and images, video surveillance systems produce a tremendous amount of data every day and second, so-called big data. Considering big data requires the computation of complex functions that need to develop complex hierarchies of concepts using knowledge of deep learning and sophisticated algorithms [1]. Moreover, the big data from the video surveillance systems placed in almost every corner of public places require substantial data warehouses. However, it would reduce storage space if the only analysis result needs to be stored. Deep learning technology is required to effectively implement various objectives in different fields due to big data [40]. Thus, deep learning techniques have been drawn significant attention in recent years due to its ability to solve complicated issues with the highest accuracy in big data evolved in the increasing number of public surveillance video systems.

Deep learning is "a machine learning technique with artificial neural networks and representation learning" that uses multilayers in the network to optimize implementation [4, 37]. Deep learning techniques involve components of both training and learning [39]. It uses neural learning networks and algorithms with big data and powerful computational resources regarding learning. That is, deep learning techniques use multiple layers of neural network algorithms based on raw input data to depict higher information levels in different layers. The more layers, the finer the model and the higher the performance [22].

There are several model architectures used for object detection; R-CNN, Fast R-CNN, Faster R-CNN. Convolutional neural networks (CNN) are among the most popular types of neural networks used in deep learning [19]. The convolutional neural network, CNN, in short, is the innovative deep learning model that divides the image into multiple regions and classifies each area into various classes. However, due to its requirements of multiple regions for an accurate prediction, CNN has drawbacks of inefficient process and high computation time with a higher volume of data. RCNN model has been proposed to reduce the massive volume of data and the processing time to a millisecond level. It also accelerates accuracy and efficiency [34]. RCNN is an algorithm with regions with CNN features combining regional proposals with convolutional neural networks. It performs high computation time as extracting around 2000 regions from each image [9]. It is widely used to extract visual information from multiple visual data sources. However, training is expensive in space and time and still needs a high computation time for object detection.

Fast RCNN is a Fast Region-based Convolutional Network model using deep convolutional networks for object detection with higher training and testing speed and detection accuracy. It is implemented with Python and fetches images with an input. The image is then processed with convolutional layers to create a convolutional feature map [11].

Lastly, the faster RCNN model is a canonical model for deep learning-based object detection that replaces the selective search method with region proposal network (RPN). It is fast and the best performing in detecting objects [24].

Deep learning techniques have been largely used in various fields in fraud detection, bioinformatics, speech and image recognition, and 3D point clouds [10] and in organizational strategy and customer relationship management. Smartphones and video cameras are the essences of connected networks. The relevance of images, video and audio in social media, streaming analytics and web browsing has created a necessity of producing and processing massive amounts of data. The computation of such complex features requires knowledge of deep learning networks and the ability to develop complex hierarchies of concepts using sophisticated algorithms. Excellent working knowledge of deep learning techniques, deep learning types, and deep learning applications can help users execute it for various purposes. In unsupervised data, machine learning may not always be feasible because manual labeling of data is expensive and time-consuming. Deep learning networks are designed to help overcome these issues. In short, deep learning provides better performance on many problems with complete automation [7]. As seen in Table 2, various fields such as healthcare, human behavior, and accident and disaster management are paying a lot of attention to deep learning models and applying their application.

4 Design for an intelligent video surveillance system configurations

4.1 System configurations

Figure 1 depicts the system configuration. Raspberry Pi camera is used to configure with GPU server for the study instead of the existing video surveillance system.

The system environment includes python flask for WEB, python TensorFlow in deep learning, and python socket for Raspberry Pi. GPU server consists of four vCPU 30G memories and one 24G Tesla p40 GPU.

Table 2 Deep learning application in various fields

Author	Year	Deep learning application
Nie et al. [25]	2015	Disease inference
Gilbert et al. [8]	2015	Railway track inspection
Ravi et al. [33]	2016	Human activity recognition
Cheng et al. [6]	2016	Recommendation systems
Hung et al. [11]	2017	Biological image data.
Ogle et al. [28]	2018	Disaster risk communication
Turchini et al. [41]	2018	Surveillance system
Mittal and Hasija [21]	2019	Healthcare and biomedicine
Saravi et al. [36]	2019	Preparedness against adverse flood events
Lei et al. [15]	2019	Face representation
Sur [40]	2019	Visual captioning
Qi et al. [30]	2019	Ship target detection

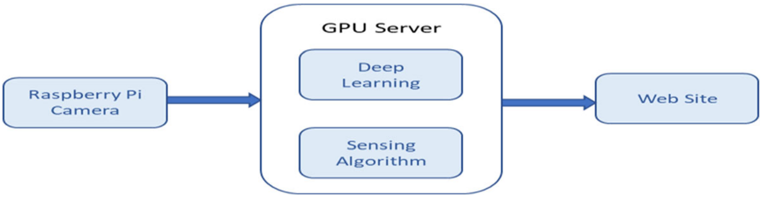

Fig. 1 System configuration

A. Raspberry Pi (Camera)

Raspberry pie plays the same role as the existing video surveillance system. It functions as a transmission function of a camera image frame and captures images.

B. GPU servers

GPU server performs in three parts: socket communication with Raspberry Pi, automatic recognition deep learning and notification algorithm, and website opening. Each function consists of three threads within the GPU server and operates simultaneously, as shown in Fig. 2, thread configuration of the GPU server.

4.2 Deep learning algorithm

The socket communicates with Raspberry Pi; if there is data received in 5001 ports when opening the server socket, it transfers corresponding data to an image. The translated image data is utilized as INPUT data for deep learning algorithms (Fig. 3). Raspberry Pi performs TCP socket communication by opening 5001 port in the public IP of the GPU. The data transmitted in the communication between the Raspberry Pi and the GPU server is the real-time image frame converted into a byte form.

The second model of the deep learning algorithm is the inception V3 model based on faster R-CNN. Faster R-CNN is the most advanced solution to detect the object in images with high accuracy and reliability [27]. Thus we have used it as a base to design the intelligent video surveillance system.

The model forms a hierarchy, as shown in Fig. 4 and has a very high recognition rate for an image. It enables us to recognize the targets of people and various objects when using the

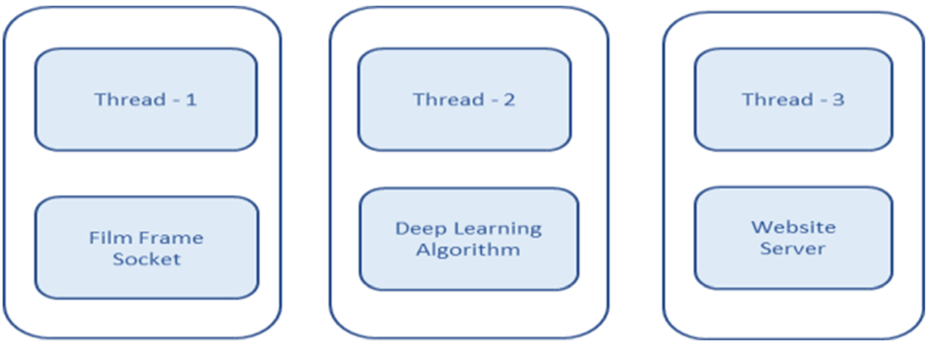

Fig. 2 GPU server thread configuration

```
server_socket = socket.socket()
server_socket.bind(('0.0.0.0', 5001))
server_socket.listen(0)
connection = server_socket.accept()[0].makefile('rb')
```

Fig. 3 Python socket communications receipt

dataset from the COCO net, a large-scale dataset for object detection, segmentation, and captioning. However, specific objects such as deadly weapons and fires cannot be figured out by the dataset of the COCO net. It has to be trained directly by labeling additional images for recognition.

The processed data, the input data of the model, is generated, as shown in Fig. 5. Regarding the data training with a deep learning algorithm, the corresponding label in the image was found and named as shown in Fig. 5, and then converted to XML file. The changed XML file is converted to CSV file, then the existing ckpt was called and converted to the new tf-record. Having used the transferred tf-record, the repeated training took place multiple times to recognize various weapons and fires that are essential functions to the system. The training is implemented based on the following model to improve the detection rate.

batch size = 1;

Repeat 900,000 times with a learning rate of 0.00003;

Repeat 1,200,000 times with a learning rate of 0.00003.

Developed artificial intelligence deep learning algorithm is constructed so that real-time image frame data received through socket communication can be instantly identified and detected by working simultaneously with socket communication and using a function thread. As a result, the detection rate has been increased up to 99 % with deep learning. Figure 6 shows 99 % of fire detection with the deep learning model.

4.3 Web site application

The third thread is a web server and the back-end technology is Python's flask. The Python's flask model opens a Web server to configure socket communication and an artificial intelligence deep-learning algorithm. As shown in Fig. 7, the Python flask Web server was opened to transmit image frames where artificial intelligent deep learning has been performed. It was also configured to allow the streaming of videos while repeatedly sending image frames that appear in real-time.

Finally, image data streaming to the server via raspberry pie is processed, as shown in Fig. 8. The processed data results in a final result through the Faster R-CNN model. The

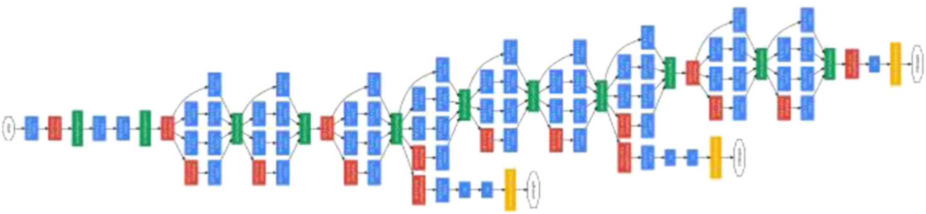

Fig. 4 Inception V3 model

Fig. 5 Data image processing diagram

system sends push notifications to the application via fcm BROKER if there is anything unusual.

The web application is developed using the javascript and react framework. Using an application, you can watch the real-time image from a video surveillance system and receive a notice in case of an emergency. Figure 9 shows the main screen of the application that allows users to view images on the video surveillance system in real-time or receive notifications in a particular situation. It also allows users to check a specific zone. It is configured to display the map form by region or inside the building so that users enable to check the video when they click a particular image icon. As shown in Fig. 10, if an intelligent video surveillance system recognizes a crime or disaster situation, the app is configured to save such cases and push notifications as images and send them to the text message. The Fig. 10 screen shows the detection of the crime scene and the notification message sending to an individual message with an image.

There is an increasing number of proposed systems using deep learning model to enhance detection speed and accuracy. As shown in Table 3, CNN-based models have been proposed in various surveillance systems but show some limitations of real-time detection speed, accuracy and inefficiency. Previous research of a similar approach also concerns a trade-off between speed and accuracy [12, 23, 27]. However, this proposed model reveals a right balance with speed and accuracy in detecting crime and disaster. This proposed model shows up to 99 % accuracy and transmits images and notification to a user's application in real-time. Unlike other proposed systems, this proposed system enables users to identify and detect immediately through real-time image data from socket communication. Video streaming is also capable while continuously transmitting real-time image frames performed by artificial intelligence deep learning. By providing real-time notifications to the user's application, crimes can be more proactively prevented. This system may not be the simplest or fastest

Fig. 6 Deep learning with 99 % chance of fire detection

```
if __name__ == '__main__':
    frame = None
    do_something_only_once()
    threading.Thread(target=app.run(host='0.0.0.0', port=5000, debug=True, threaded=True, use_reloader=False
    #app.run(host='0.0.0.0', port=5000, debug=True, threaded=True)
```

Fig. 7 Web server creation using Python flask

way to detect an object, but it may be one of the best performing and accurate model for crime detection.

5 Discussion and conclusion

Deep learning has been drawn attention in dealing with big data with an accurate analysis. However, the research is yet to be explored, and there is still a lack of understanding of the underlying theory working behind the video surveillance system. By suggesting a deep learning model into a video surveillance system, it would solve a wide variety of problems with manual systems and provide significant assistance in preventing disaster and crime. This research proposed an intelligent video surveillance system that actively detects and protects from crimes. The key to the intelligent video surveillance system proposed through this research is a deep learning algorithm that is assigned to the distributed servers. This proposed model suggests that if the deep learning technology is applied to the servers linked to the notification system, crime and disaster notifications can be made faster with a high accuracy rate through enriched information analysis. With a deep learning model into servers in the video surveillance system, this system provides a higher image processing speed and accuracy. This intelligent video surveillance system also sends images to web applications simultaneously to notify users of any accidents or crimes. In this way, crime and disaster could be detected faster and taken further action for prevention.

Even though deep learning has been utilized in various research areas, the video surveillance system design with deep learning still needs to be explored. This study opens other ways of developing surveillance systems in better performance with crime prevention. This study also provides insights into deep learning technology application and utilization by indicating that it is important to create and implement an architecture suitable for the video surveillance infrastructure. It depends on the decision of developers and managers who fully understand the purpose, use, and cost of designing video surveillance systems. Moreover, objective data from surveillance systems such as CCTV are widely used to analyze crime and disaster. However, it

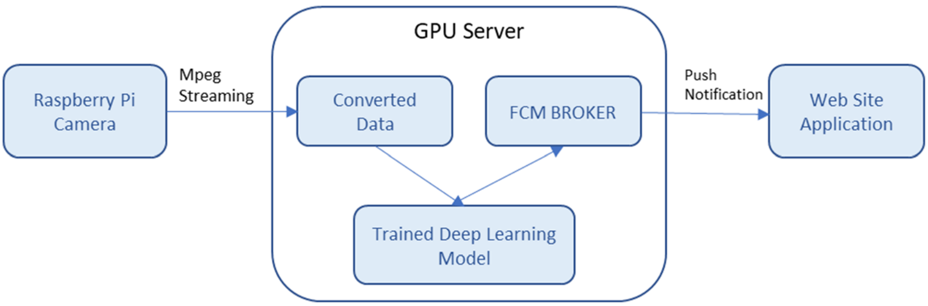

Fig. 8 Data processing process

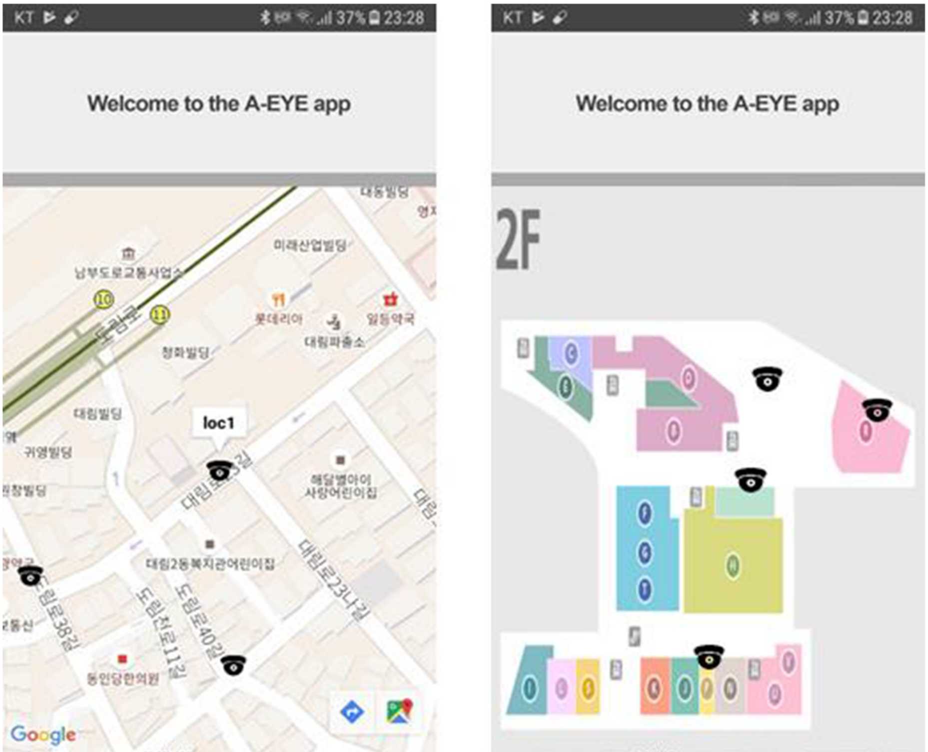

Fig. 9 Main screen

is still insufficient to prevent crime and disaster in advance through real-time analysis data. The intelligence video surveillance system enables us to capture and predict serious incidents by analyzing real-time images without human intervention. This could provide cost-effective and efficient ways to ensure safety. However, there have been contradictive issues with privacy and security, in that the more robust surveillance for security could cause threatening privacy issue [26]. Several techniques have been developed and studied to identify the security and privacy issues in the video surveillance system [32]. And thus, developers also need to keep track of negative responses raised by the intelligence surveillance system like privacy and

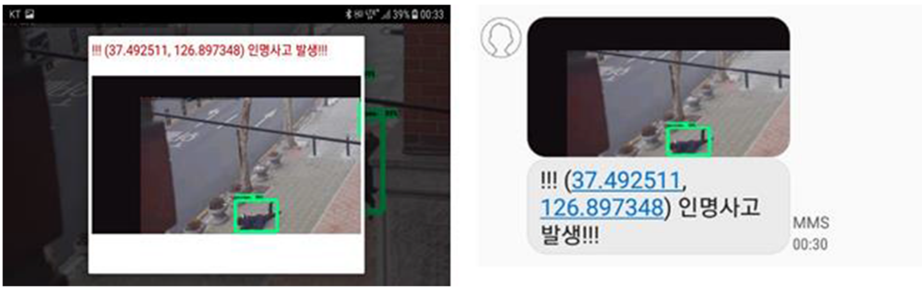

Fig. 10 Crime and disaster situations

Table 3 Comparison with other systems using a deep learning model

Author	System	Deep learning model	Limitation
Niu and Song [27]	Indoor Navigation graph from building designs	Faster R-CNN	A large amount of manual work in preparing training data and lack of accuracy of extracted results
Vishnu et al. [43]	Intelligent traffic video surveillance and accident detection system	Multinominal logistic regression (MLR)	Lack of real-time detection of criminal activity and other unnatural activities
Jain et al. [12]	Weapon detection	Faster R-CNN	Higher accuracy (84.6%) but relatively lower real-time speed.
Muhammad et al. [23]	Disaster management system	Deep CNN	A balance between speed and accuracy but heavy model size, no authentication

social justice risks [42]. Also, as all advanced technologies have negative aspects and their advantages, ethical issues such as who will manage and control the algorithms of the artificial intelligence surveillance system when the use of them are common will need to bring attention in future research.

Acknowledgements This work is based on the design constructed by Mr. Minsu Kim, the CEO of CLOMOUNT Co., Ltd. The authors would like to thank Mr. Kim for providing data sources to develop this research.

References

1. Ahn H, Kim D, Kim YS (2015) Generating new ground truth data by editing previous data from integrated video annotation database. BigDAS '15: Proceedings of the 2015 International Conference on Big Data Applications and Services, pp 208–212
2. Allied, Telesis, Intelligent video surveillance: Recent trends and what lies ahead. https://www.alliedtelesis.com/blog/intelligent-video-surveillance-recent-trends-and-what-lies-ahead. Accessed 10 Jan 2020
3. Barhm M, Qwasmi N, Qureshi F, el Khatib K (2011) Negotiating privacy preferences in video surveillance systems. In: Mehrotra WK, Mohan WK, Oh J, Varshney P, Ali M (eds) Modern approaches in applied intelligence, vol 6704. Springer, Berlin Heidelberg, pp 511–521
4. Bengio Y, LeCun Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
5. CCTV Cameras explained, tech cube: security sales success. https://www.techcube.co.uk/blog/cctv-cameras-explained/. Accessed 10 Jan 2020
6. Cheng HT, Koc L, Harmsen J, Shaked T, Chandra T, Aradhye H, Anderson G, Corrado GS, Chai W, Ispir M, Anil R, Haque Z, Hong L, Jain V, Liu X, Shah H (2016) Wide & deep learning for recommender systems. Cornell University, arXiv:1606.07792v1, June 24
7. Chollet F (2018) Deep learning with python. Mannng Publications Co., Shelter Island
8. Gibert X, Patel VM, Chellappa R (2015) Deep multitask learning for railway track inspection. IEEE Trans Intell Transp Syst 18:153–164
9. Girshick R, Donahue J, Darrell T, Malik J (2013) Rich feature hierarchies for accurate object detection and semantic segmentation, CVPR 2017: computer vision for microscopy imae analysis (CVMI) workshop, Cornell University, arXiv:1311.2524v5
10. Guo Y, Wang H, Hu Q, Liu H, Liu L, Bennamoun M (2019) Deep learning for 3D point clouds: a survey. arXiv: 1912, 12033v1[cs.CV]
11. Hung J, Ravel D, Lopes SCP, Rangel G, Nery OA, Malleret B, Nosten F, Lacerda MG, Ferreira MU, Renia L, Duraisingh MT, Costa FTM, Marti M, Carpenter AE (2017) Applying faster R-CNN for oject detection

on malaria images, CVPR 2017: computer vision for microscopy imae analysis (CVMI) workshop, Cornell University, arXiv:1804.09548v2
12. Jain H, Vikram A, Mohana, Kashyap A, Jain A (2020) Weapon detection using artificial intelligence and deep learning for security applications, Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020), IEEE Xplore Part Number: CFP20V66-ART, ISBN: 978-1-7281-4108-4, pp 193–198
13. Ju YW, Yi SJ (2013) Implementing database methods for increasing the performance of intelligent CCTV. Int J Secur Appl 7:113–120
14. Kumar V, Svensson J (2015) Promoting social change and democracy through information technology, IGI Global, Hershey, pp 75. ISBN 9781466685031
15. Lei J, Zhang B, Ling H (2019) Deep learning face representation by fixed erasing in facial landmarks. Multimed Tools Appl 78:27703–27718
16. Li L, Huang W, Gu IYH, Luo R, Tian Q (2008) An efficient sequential approach to tracking multiple object through crowds for real-time intelligent CCTV systems. IEEE Trans Syst Man Cybern 38:1254–1269
17. Little DD, Ross SE (2013) APMs and airport mobility – Historic trends and future possibilities. Automated people movers and transit systems (ed.): Computer Science Today. Recent Trends and Developments. Sproute, American Society of Civil Engineers, Reston
18. Liu Y, Zhang D, Lu G (2008) Region-based image retrieval with high-level semantics using decision tree learning. Pattern Recogn 41(8):2554–2570
19. Mathworks (2020) Deep Learning, What is deep learning?_3 things you need to know. https://www.mathworks.com/discovery/deep-learning.html. Accessed 28 Apr 2020
20. Mationlansk A, Maksimova A, Dziech A (2016) CCTV object detection with fuzzy classification and image enhancement. Multimed Tools Appl 75:10513–10528
21. Mittal S, Hasija Y (2019) Applications of deep learning in healthcare and biomedicine. In: Dash S, Acharya B, Mittal M, Abraham A, Kelemen A (eds) Deep Learning Techniques for Biomedical and Health Informatics. Studies in Big Data, vol 68. Springer, Cham
22. Morioka K, Kovacs S, Lee JH, Korondi P (2010) A cooperative object tracking system with fuzzy-based adaptive camera selection. Int J Smart Sens Intell Syst 3:338–358
23. Muhammad K, Ahmad J, Baik SW (2018) Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing 288:30–42
24. Mwiti D (2019) A 2019 guide to object detection, Heartbeat, July 18, 2019. https://heartbeat.fritz.ai/a-2019-guide-to-object-detection-9509987954c3. Accessed 3 Mar 2020
25. Nie L, Wang M, Zhang L, Yan S, Zhang B, Chua TS (2015) Disease inference from health-related questions via sparse deep learning. IEEE Trans Knowl Data Eng 27(8):2107–2119
26. Nielsen A (2017) Video surveillance threatens privacy, experts say. The daily universe. June 28. https://universe.byu.edu/2017/06/28/video-surveillance-threatens-privacy. Accessed 28 Apr 2020
27. Niu L. Song YQ (2019) A faster R-CNN approach for extracting indoor navigation graph from building designs. The international archives of the photogrammetry, remote sensing and spatial information sciences, vol XLII-2/W13, ISPRS Geospatial Week 2019, 10–14 June 2019, Enschede, The Netherlands
28. Ogle RI, Rho JC, Clarke RJ (2018) Artificial intelligence in disaster risk communication: A systematic literature review. 2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM). Dec. 4–7
29. Paliwal N, Vanjani P, Liu JW, Saini S, Sharma A (2019) Image processing-based intelligent robotic system for assistance of agricultural crops. Int J Soc Humanist Comput (IJSHC) 3(2):191–204
30. Qi L, Li B, Chen L, Wang W, Dong L, Jia X, Huang J, Ge C, Xue G, Wang D (2019) Ship target detection algorithm based on improved faster R-CNN. Electronics 8(9):959
31. Rai M, Asim A, Husain TM, Yadav RK (2018) Advanced intelligent video surveillance system (AIVSS): a future aspect. In: Neves AJR (ed) Intelligent video surveillance. IntechOpen, November 5. https://doi.org/10.5772/intechopen.76444
32. Rajpoot QM, Jensen CD (2014) Security and privacy in video surveillance: Requirements and challenges. In: Cuppens-Boulahia N, Cuppens F, Jajodia S, El Kalam A, Sans AT (eds) ICT Systems Security and Privacy Protection. Sections 2014, 428. IFIP Advances in Information and Communication Technology, Springer, Berlin, pp 169–184
33. Ravi D, Wong C, Lo B, Yang GZ (2016) Deep learning for human activity recognition: a resource efficient implementation on low-power devices. In: BSN 2016–13th Annual Body Sensor Networks Conference
34. Ren S, He K, Girshick R, Sun J (2016) Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems. Jan 6:1–14, arXiv:1506.01497v3
35. Salahat E, Saleh H, Mohammad B, Al-Qutayri M, Sluzek A, Ismail M (2013) Automated real-time video surveillance algorithms for SoC implementation: A survey. IEEE International Conference on Electronics Circuits and Systems. December 2013

36. Saravi S, Kalawsky R, Joannou D, Casado MR. Fu G, Meng F (2019) Use of artificial intelligence to improve resilience and preparedness against adverse flood events. Water 11(973):1–16
37. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
38. Singh V, Gupta R (2019) Novel framework of semantic based image reterival by convoluted features with nonlinear mapping in cyberspace international. J Recent Technol Eng (IJRTE) 8(1C2):939–942
39. Sreenu G, Saleem Durai MA (2019) Intelligent video surveillance: a review through deep learning techniques for crowd analysis. Journal of Big Data 6:1–27
40. Sur C (2019) Survey of deep learning and architectures for visual captioning – transitioning between media and natural languages. Multimed Tools Appl 78:32187–32237
41. Turchini F, Seidenari L, Uricchio T, Bimbo AD (2018) Deep learning based surveillance system for open critical areas. Inventions 3(69):1–13
42. Vincent J (2018) Artificial intelligence is going to supercharge surveillance, What happens when digital eyes get the brains to match? Jan. 23, 2018, The Verge. https://www.theverge.com/2018/1/23/16907238/artificial-intelligence-surveillance-cameras-security. Accessed 13 Jan 2020
43. Vishnu VCM, Rajalakshmi M, Nedunchezhian R (2018) Intelligent traffic video surveillance and accident detection system with dynamic traffic signal control. Cluster Comput 21:135–147. https://doi.org/10.1007/s10586-017-0974-5
44. Williem A, Madasu V, Boles W, Yarlagadda P (2012) A suspicious behavior detection using a context space model for smart surveillance systems. Comput Vis Image Underst 116:194–209
45. Yang HM, Lim DW, Choi YS, Kang JG, Kim IH, Lin A, Jung JW (2019) Image-based human sperm counting method. Int J Soc Humanist Comput (IJSHC) 3(2):148–157
46. Zablocki M, Gosciewska K, Frejlichowski D, Hofman R (2014) Intelligent video surveillance systmes for public spaces – a survey. J Theor Appl Comput Sci 8:13–27
47. Zhang P, Thomas T, Emmanuel S (2012) Privacy enabled video surveillance using a two state Markov tracking algorithm. Multimed Syst 18(2):175–199