



Real-time facial emotion recognition system among children with autism based on deep learning and IoT

Fatma M. Talaat¹

Received: 29 October 2022 / Accepted: 13 February 2023 / Published online: 7 March 2023
© The Author(s) 2023

Abstract

Diagnosis of autism considers a challenging task for medical experts since the medical diagnosis mainly depends on the abnormalities in the brain functions that may not appear in the early stages of early onset of autism disorder. Facial expression can be an alternative and efficient solution for the early diagnosis of Autism. This is due to Autistic children usually having distinctive patterns which facilitate distinguishing them from normal children. Assistive technology has proven to be one of the most important innovations in helping people with autism improve their quality of life. A real-time emotion identification system for autistic youngsters was developed in this study. Face identification, facial feature extraction, and feature categorization are the three stages of emotion recognition. A total of six facial emotions are detected by the propound system: anger, fear, joy, natural, sadness, and surprise. This section proposes an enhanced deep learning (EDL) technique to classify the emotions using convolutional neural network. The proposed emotion detection framework takes the benefit from using fog and IoT to reduce the latency for real-time detection with fast response and to be a location awareness. From the results, EDL outperforms other techniques as it achieved 99.99% accuracy. EDL used GA to select the optimal hyperparameters for the CNN.

Keywords Autism · Emotion recognition · Assistive technology · Deep learning

1 Introduction

This section discusses some important issues such as autism spectrum disorder, emotion recognition problem with people with autism, and assistive technology.

1.1 Autism spectrum disorder

A neurological ailment called autism spectrum disorder (ASD) affects behavior and communication. Although Kanner was the first to recognize it in 1943 [1], our knowledge of ASD has greatly increased in terms of diagnosis and treatment. The initial indicators of this developmental syndrome can be seen in a child's early years, even though it can manifest at any age. The DSM-5 (diagnostic and statistical manual of mental disorders) [2]

states that a person with autism exhibits ineffective behavior, social, and communication abilities. Even though it is a permanent handicap, therapy and support can help a person perform some tasks more successfully. Some autistic individuals have trouble falling asleep and behave rudely. It is simpler to diagnose ASD in children than in adults since some symptoms in adults may be confused with those of other intellectual disorders, such as attention deficit hyperactivity disorder (ADHD).

They maintain a constant demeanor and show no signs of wanting to socialize. Autism is a disorder with a wide range of symptoms, depending on the person's symptoms. The word "spectrum" is included in the disorder's name since it can range from mild to severe [3]. A person with severe autism is typically nonverbal or has speaking difficulties. They struggle to understand their emotions and express them to others. A person with autism therefore struggles to do routine chores. A real-time emotion identification system for autistic children is essential to detecting their feelings and assisting them when they experience pain or rage.

✉ Fatma M. Talaat
fatma.nada@ai.kfs.edu.eg

¹ Faculty of Artificial Intelligence, Kafrelsheikh University, Kafrelsheikh, Egypt

1.2 Autism-related issues with emotion recognition

Empathy is referred to as the capacity for comprehending and reciprocating the feelings of another. Sympathy, on the other hand, is the ability to share comparable feelings with another individual. People with ASD may not be able to empathize or sympathize with others [4]. When someone is harmed, they may express gladness, or they may show no emotion at all. As a result of their inability to respond correctly to others' emotions, autistic people may appear emotionless. Several studies, however, have looked into if someone with autism can genuinely express their emotions to others.

Empathy requires a careful examination of another person's body language, speech, and facial expressions in order to comprehend and interpret their sentiments. People with ASD lack the necessary social skills associated with interpreting body language and reciprocating feelings, whereas youngsters learn to understand the facial expressions needed to demonstrate empathy by seeing and emulating people around them. In people with ASD, the majority of social skills needed to engage with others are significantly hampered.

ASD is linked to a specific social and emotional deficit that is characterized by cognitive, social semiotic, and social understanding deficits. Autism typically prevents a person from understanding another person's emotions and mental state through facial expressions or speech intonation. They may also have trouble anticipating other people's actions by analyzing their emotional conditions. Facial expressions are heavily used in emotion recognition studies. The ability to recognize emotions and distinguish between distinct facial expressions is normally developed from infancy [5].

Children with ASD frequently neglect facial expressions [6]. Additionally, children with autism perceive facial expressions inconsistently [7], suggesting that they lack the ability to recognize emotions. When interpreting emotions, multiple sensory processing is frequently necessary [8]. The measurements of the face, body, and speech can all be used to interpret emotion. The capacity to divide attention and concentrate on pertinent facial information is necessary for the recognition of emotions; this sort of processing is largely subconscious.

1.3 Assistive technology's role in the lives of individuals with ASD

Assistive technology is any piece of equipment or gear that enables people with ASD to do things they could not do before. Such technological aids assist people with

disabilities in carrying out daily tasks. There has been a surge in the development of technology that assists people with autism in recent years. This technology ranges from low-level to powerful and evolving [9]. Assistive technology's main goal is to help people with special needs. The creation of technology-based therapeutic rooms could result from a collaborative effort between such centers, schools, and the government. The majority of academics agree that, depending on the severity of the disease, it is critical to pick appropriate assistive technology for people with autism in a systematic manner [10].

As a result, not all assistive technology is appropriate for everyone. Each person with ASD has their own collection of characteristics. As a result, it is clear that there is no uniform set of assistive technologies. Only professionals are capable of recognizing the distinctions and providing the necessary support. Basic techniques to advanced computing technologies are used in assistive technology [10]. It can be classified into three main categories: (i) Basic assistive technology refers to pictorial cards that are used to communicate between the child and the instructor; (ii) medium assistive technology refers to graphical representation systems; and (iii) advanced technology refers to human-computer interaction applications such as robots and gadgets.

Assistive technology encompasses a wide range of tools that help people with autism overcome their functional limitations. Augmented and Alternative Communication (AAC), which consists of a plan of actions that could enable a nonverbal person to engage with others [12], is another technological application that is used to improve communication for people with ASD. Additionally, the use of computer-based adaptive learning can improve the methods for teaching such exceptional people to learn in order to improve their lives. This medium could be hardware, software, or a hybrid of the two.

1.4 Problem statement

Dynamic assistive technology incorporates control apparatus, touch displays, and augmented and virtual reality applications, among other advanced technological computer gadgets that have evolved. These technologies can be used for both diagnosis and treatment. Pictures and images have piqued the interest of people with ASD [13]. They have proven to be efficient visual learners, and pictorial cards have proven to be an effective tool to teach children how to conduct daily duties.

The aim of this study is to develop real-time emotion recognition system that based on DL CNN model. Emotion detection using face images is a challenge task, the consistency, number and quality of images that used for training model have a significant impact on the model

performance. From the images used in training, model should be able to distinguish between different emotions. It was a challenge due to several reasons including (i) face images have different characteristics, for example some of them have shorter face, boarder face, wider image, small mouth, etc. (ii) some faces may indicate emotions that different from their actual emotions. (iii) Sad emotions may sometimes overlap with anger emotions, and the same for joy and surprise. The results showed that the higher accuracy is for the natural class.

The main contributions can be summarized as:

1. Proposing a real-time emotion identification system for autistic youngsters.
2. Face identification, facial feature extraction, and feature categorization are the three stages of emotion recognition.
3. A total of six facial emotions are detected by the propound system: anger, fear, joy, natural, sadness, and surprise.
4. This research presents a Deep Convolutional Neural Network (DCNN) architecture for facial expression recognition.
5. The proposed emotion detection framework takes the benefit from using fog and IoT to reduce the latency for real-time detection with fast response and to be a location awareness.
6. The suggested architecture outperforms earlier convolutional neural network-based algorithms and does not require any hand-crafted feature extraction

The remaining work is organized as follows. In Sect. 2, some of the recent related work in the emotion recognition techniques is presented. In Sect. 3, the proposed method is presented. Experimental evaluation is provided in Sect. 4. And in Sect. 5, we conclude this work.

2 Literature review

Researchers and technology professionals have long studied a facial detection system that not only recognizes faces in an image but also estimates the type of emotion based on facial traits. Bledsoe [14] documented some of the first research on automatic facial detection for the US Department of Defense in 1960. Since then, software has been developed specifically for the Department of Defense, although little information about the product is given to the general public. Kanade [15] developed the first fully effective autonomous facial recognition system. By discriminating between features retrieved by a machine and those derived by humans, this system was able to measure sixteen distinct facial features.

One of the studies [16] shows how using visual cues can help autistic youngsters learn about emotions. They made a series of movies and activities to teach children about various emotions and track their progress. Another study [17] offered a web application that provides a platform for such special youngsters to engage using a simulated model. In addition, a project called “AURORA” used a robot and allowed interaction between the child and the robot, resulting in a human–computer interaction interface [18]. Another study [19] supports human–computer connection by showing a series of short movies of various types of emotions to teach emotions to youngsters with special needs. The writers of [20] look into the various possibilities for using robots as therapeutic instruments.

Despite several studies on teaching emotions to autistic children, a number of obstacles remain. An autistic individual has a hard time deciphering emotions from facial expressions. In a study [21], the application of giving an emotional hearing aid to special-needs individuals was proven. The facial action coding system (FACS), introduced by Ekman and Friesian [22], is one of the approaches used to recognize facial expressions. Depending on the facial muscular activity, the facial action coding system depicts several sorts of facial expressions.

This method allows for the quantitative measurement and recording of facial expressions. Facial recognition systems have advanced significantly in tandem with advances in real-time machine learning techniques. The study [23] presents a detailed review of current automatic facial recognition technologies and applications. According to the research, the majority of existing systems recognize either the six fundamental facial expressions or different forms of facial expressions.

When it comes to recognizing emotions, emotional intelligence is crucial. Understanding emotions entails biological and physical processes as well as the ability to recognize other people’s feelings [24]. By watching facial expressions and somatic changes and converting these documented changes to their physiological presentation, an individual can reliably predict emotions. According to Darwin’s research, the process of detecting emotions is thought to entail numerous models of behavior, resulting in a thorough classification of 40 emotional states [25]. On the other hand, the majority of studies on facial attribute stratification, on the other hand, refer to Ekman’s classification of six primary emotions [26]. Happiness, sadness, surprise, fear, disgust, and anger are the six primary emotions. However, six other fundamental emotions were later added to the neutral expression.

The ease with which various groupings of emotions may be identified is a big advantage of using this approach. A variety of computer-based technologies have been developed to better read human attitudes and feelings in order to

Table 1 Different techniques for facial emotion recognition

References	Year	Used technique	Description	Advantages	Disadvantages
[25]	2019	One of the most popular SDKs for categorizing emotional states is the chosen program, Affectiva by Affectiva (in partnership with iMotions) The facial action coding system (FACS), an intricate and popular objective taxonomy for coding facial behavior, is employed by Affectiva	The examination of the facial analysis system's recorded reactions is the main topic of this study 50 college students were used as a sample for the experiment Each student saw 100 random images, and their responses to each one were noted The recorded responses were then contrasted with the anticipated response to the visual	The experiment's findings revealed a number of flaws in the face analyzing system	The algorithm has trouble categorizing facial expressions and is unable to recognize and detect any underlying emotions that a viewer may feel when they view the image Face analysis software can only identify emotions that are physically displayed on the face by observing physiological changes in certain facial regions
[28]	2016	The following factors were included in the full factorial design: face sex (female, male); intensity (10%, 55%, 90%); kind of expression (anger, disgust, fear, joyful, sad, surprise); and observer gender (woman, man) as a between subject factor With 10 trials each condition, a total of 360 trials were shown	This study examines the effects of these variables on the dwell duration on the two most important features of the face—the eyes and the mouth—as well as the speed and accuracy of expression recognition Participants were asked to categorize the six facial expressions (angry, disgust, fear, happiness, sorrow, and surprise), each with three levels of intensity, on both male and female faces (low, moderate, and normal)	We were able to investigate the impact of reduced intensity on eye scan routes as well as whether participants were sensitive to very low levels of emotional content thanks to this methodology	The degree to which different expressions serve as a vehicle for social engagement or as a byproduct of an emotional experience may vary Neutral replies that were 90% neutral and 10% expressive for extremely low intensity expressions were recorded as erroneous
[27]	2019	Visualizations	In a programming course, we have used ten different sorts of visualizations that authors have offered to help us understand the learners' current state of effectiveness	The low level data of users' interactions with various technological educational tools were processed to create these visuals, which represent high level information Four categories have been established for the visualizations: time-based, context-based, emotional changes, and cumulative information	Small dataset (limited scale)

improve the user experience [27]. To anticipate meaningful human facial expressions, they mostly use cameras or webcams. With moderate accuracy, one can deduce the emotions of another person who is facing the camera or webcam. Meanwhile, several machine learning and image processing experiments have shown that face traits and eye-glazing behaviors may be used to identify human moods [28]. The Facial Action Coding Method (FACS) is a

classification system for facial impressions based on facial affectation. It was first proposed in 1978 by Ekman and Friesen, and it was modified in 2002 by Hager [29].

Using a convolutional neural network and a modified version of particle swarm optimization, the authors of [30] provide an efficient dynamic load balancing technique (EDLB) that examines the FC architecture for applications in healthcare systems. Authors in [31] proposed a new

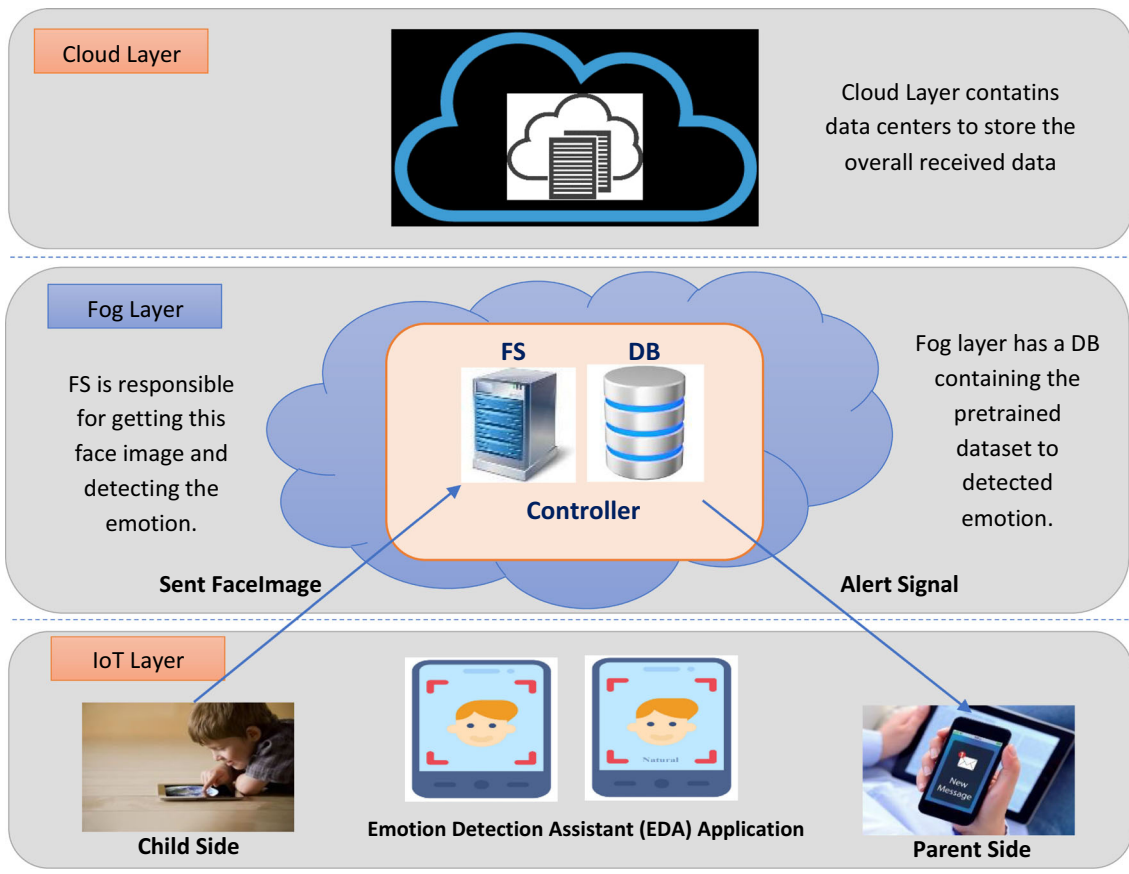


Fig. 1 Proposed emotion detection framework

effective hyperparameters optimization algorithm for CNN.

In order to secure the home, authors in [32] suggest innovative and adaptable ways. The major goal of this essay is to influence the burglar’s mindset and prevent him from carrying out these duties. The Raspberry Pi and Arduino Uno are modified to function as the system’s master controllers.

An online video of people crossing the street is used to test the suggested method, IFRCNN [33]. The number of persons in a location who maintain social distance is tracked in this paper using social distance violation data and live updates of recorded video. Updates will be kept in a cloud-based storage system, where any company or organization can access real-time updates on their digital devices.

Table 1 summarizes different techniques for facial emotion recognition.

3 Problem definition

This section proposes a real-time emotion identification system for autistic youngsters. Face identification, facial feature extraction, and feature categorization are the three stages of emotion recognition. A total of six facial emotions are detected by the propound system: anger, fear, joy, natural, sadness, and surprise. This research presents a Deep Convolutional Neural Network (DCNN) architecture for facial expression recognition. The suggested architecture outperforms earlier convolutional neural network-based algorithms and does not require any hand-crafted feature extraction.

3.1 Proposed emotion detection framework

The proposed emotion detection framework is based on 3 layers which are: (i) Cloud Layer, (ii) Fog Layer, and (iii) IoT Layer. The two main layers are: (i) IoT and (ii) Fog. In IoT, there are the proposed assistant mobile application which is used to capture an image of the child while using the smart device and then sends this face image to the fog layer. In Fog layer, there are a controller which is the fog

server (FS) responsible for getting this face image and detecting the emotion depending on the proposed DL technique. In Fog layer, there is also a database (DB) containing the pretrained dataset. After detecting the emotion, the fog server sends an alert message to the parent device in case of detected emotion is anger, fear, sadness, or surprise.

The main controlling and managing is implemented at the controller at the fog layer to reduce the latency for real-time detection with fast response and to be a location awareness. The overall proposed framework is shown in Fig. 1.

3.2 Cache replacement strategy using fuzzy

There will be a difficulty with the fog server’s cache memory due to the enormous amount of collected images that will be transferred to the fog per second. Therefore, after a predetermined amount of time, we should employ an effective cache replacement mechanism to release the memory from the old collected images.

As shown in Table 2, the fog server has a table that lists information about each image that was taken, including the following: (i) Image ID, which is a sequence number; (ii) Arrival time; (iii) Current Time; and (iv) TTL, or Time To Live, which is determined by combining the arrival time and current time; (v) Priority: based on the place where the photograph was taken. If the picture was shot far away from where the parent was, then it is very important. If the acquired image was taken at the same location as the parent, it has a low priority. (vi) Available: To determine if the photograph should be kept or destroyed.

Fuzzy logic is used to determine if an image should be kept around for a while longer or should be removed. When compared to computationally precise systems, the reasoning process is frequently straightforward, saving processing power. Particularly for real-time systems, this is a really intriguing aspect. Typically, fuzzy approaches require less time to design than traditional ones. The following consecutive steps are used to carry out the fuzzy inference

process: (i) Fuzzified Inputs, (ii) Applying regulations, (iii) Crips values. These steps are demonstrated in Fig. 2.

Each image is ranked based on the following 3 characteristics: (i) Arrival time (AT), (ii) Time To Live (TTL), which are determined by the arrival time and the current time, (iii) Priority (P): This factor is determined by where the photograph was taken. Each photograph receives a Ranking (R) value from the Fuzzy by taking into account its three preset attributes (AT, TTL, and P). The fuzzy process takes into account each of those characteristics. The ranking is quickly and accurately determined by a fuzzy algorithm. Insofar as they are less sensitive to shifting settings and incorrect or forgotten rules, fuzzy algorithms are frequently robust.

- a. Arrival Time (AT): has values: Early, Medium, or Late.
- b. Time To Live (TTL): Small, Medium, or Large.
- c. Priority (P): Low, Medium, or High.

The rating value is R1, meaning that the photograph is significant and will require additional time (15 min). If a photograph has a rating value of R2, it has a low priority and can be deleted to make room for another.

a. Fuzzified inputs

The considered three features for each image are fuzzified. The Fuzzified Arrival Time (FAT), Fuzzified Time to Live (FTTL), Fuzzified Priority (FP) are used to predict the value of R. They are calculated as shown in Eq. (1), (2), and (3) [34], 35.

$$FAT = \left\{ \frac{AT, \mu_{FAT}(AT)}{AT \epsilon T} \right\}, \tag{1}$$

$$\mu_{FAT}(AT) \epsilon [0, 1]$$

where AT = {Early, Medium, Late}, T = [0,100]. FAT: Fuzzified Arrival Time.

$$FTTL = \left\{ \frac{TTL, \mu_{FTTL}(TTL)}{TTL \epsilon TL} \right\}, \tag{2}$$

$$\mu_{FTTL}(TTL) \epsilon [0, 1]$$

where TTL = {Small, Medium, Large}, TL = [0,100]. FTTL: Fuzzified Time To Live.

Table 2 Images information table

Image ID	Arrival time	Current time	TTL	Priority	Available
1	10	11	7	High	Available
2	10	11	7	High	Available
3	10	11	7	Low	Available
4	11	18	7	Low	Deleted
5	11	19	7	Low	Deleted

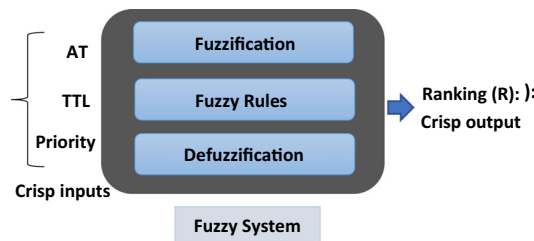


Fig. 2 The steps of the fuzzy process



Fig. 3 Emotion detection assistant (EDA) application

$$FP = \left\{ \frac{P, \mu FP(P)}{P \epsilon p} \right\}, \tag{3}$$

$$\mu FP(P) \in [0, 1]$$

where $P = \{\text{Low, Medium, High}\}$, $p = [0,100]$. FP: Fuzzified Priority.

b. **Applying regulations**

The rules in this stage explain the connection between the output and the specified input variables (AT, TTL, and P) (R). The fuzzy language rules are founded on IF THEN statements like these:

If TTL is small and P is low and AT is early THEN R is R2

If P is high THEN R is R1

iii. **Crips values**

It will be chosen whether to delete the image or keep it for longer than was originally determined based on the fuzzy output value (R).

3.3 Proposed application

The proposed application can be implemented in any smart device such as smartphones and tablets. When the application captures a photograph for the kid, it can detect his feeling as shown in Fig. 3. The Emotion Detection Assistant (EDA) Application can be active in the background while the child uses another application. EDA is used to detect the emotion of the kid. If the detected emotion is

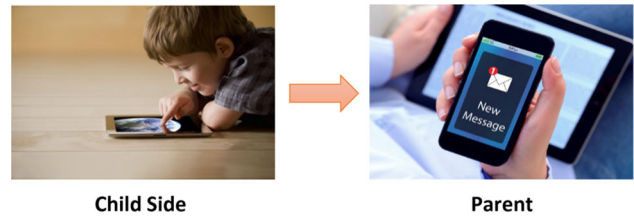


Fig. 4 Emotion detection assistant (EDA) system

natural or joy, then there is no problem. If the detected emotion is anger, fear, sadness, or surprise, it will send an alert signal to a connected application installed on the parent’s device.

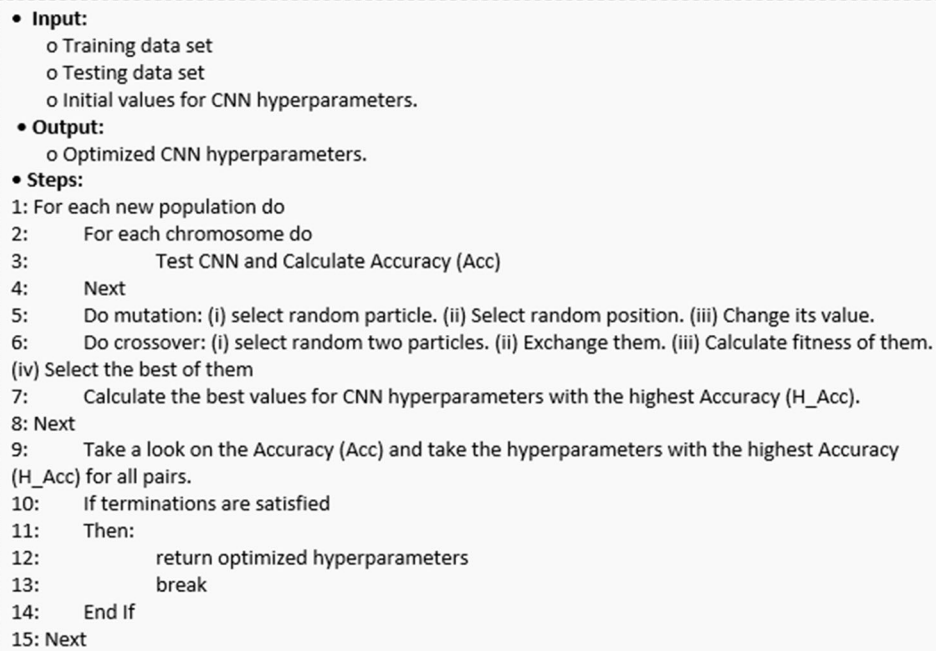
In normal, the child with autism has not enough ability to express his feeling. Hence, EDA is very useful and essential to help his parent to be notified when he is not okay. The child with autism has no ability to ask for help. There are two main sides in this system as shown in Fig. 4: (i) Child Side, and (ii) Parent Side.

3.4 Proposed DL framework

This section proposes an Enhanced Deep Learning (EDL) Technique to classify the emotions using Convolutional Neural Network. The overall architecture of the proposed model is shown in Fig. 5. It consists of two steps: the first step is a hyperparameters optimization followed by an Inception-ResnetV2 training model (trained model). The second step is the segmentation process with the U-net model for a for more fast and more precise detection process. To improve the performance of the algorithm to classify the input image efficiently, the proposed algorithm contains autoencoder for feature extraction and feature selection. The proposed Enhanced Deep Learning (EDL) Technique is based on two sub modules; (i) CNN hyperparameters optimization followed by a trained Model. (ii) Segmentation U-net model.

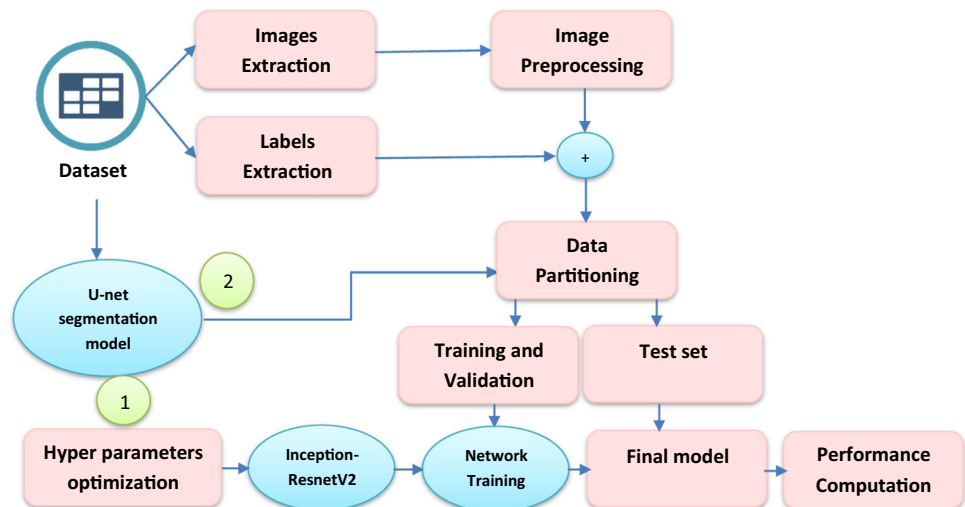
3.4.1 CNN hyperparameters optimization

The optimization of the CNN hyperparameters is based on using the Genetic Algorithm (GA) as shown in Algorithm 1.



Algorithm 1: GA algorithm improvement for CNN hyper parameters

Fig. 5 The proposed EDL model steps



3.4.2 3D U-net architecture segmentation model

U-Net [36] is a network that is used for fast and accurate image segmentation. It comprises an expanded pathway and a contracting pathway. The contracting pathway adheres to the standard convolutional network design. The structure of the 3D U-Net is shown in Fig. 6.

As shown in Fig. 7, for the segmentation process the data set is partitioned into a train, validation, and test datasets.

4 Implementation and evaluation

This section introduced the used dataset and the evaluation of the proposed model.

4.1 Used dataset

This paper uses a set of cleaned images for autistic-children with different emotions [37]. The duplicated images and the stock images have removed. Then, dataset has been categorized into six facial emotions: anger, fear, joy, natural, sadness, and surprise. The six primary used emotions are shown in Fig. 8.

Fig. 6 Structure of the 3D U-Net [34]

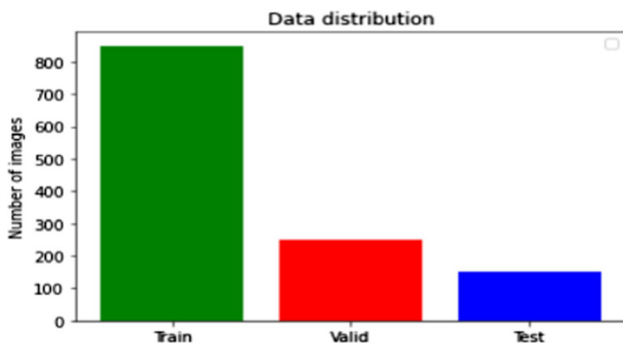
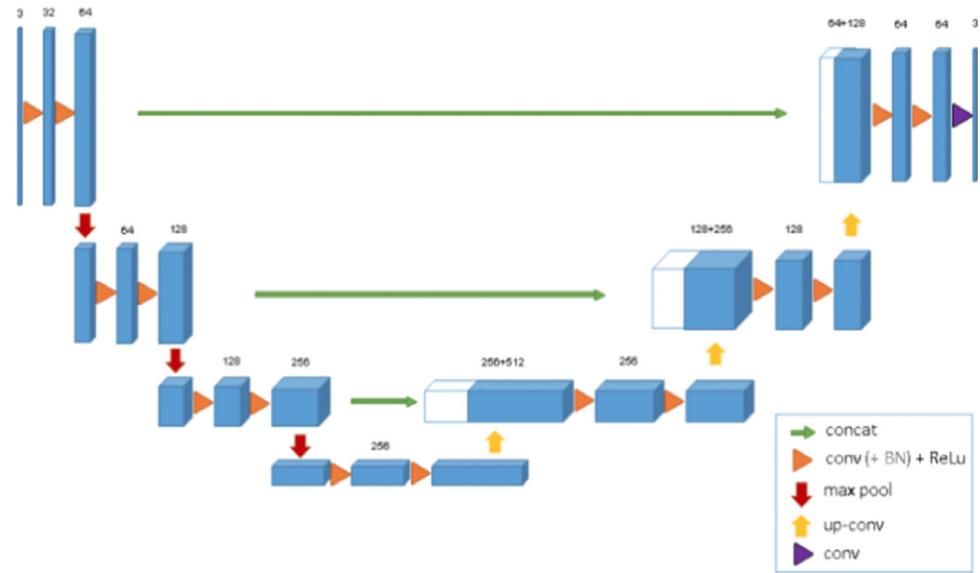


Fig. 7 Dataset distribution for segmentation

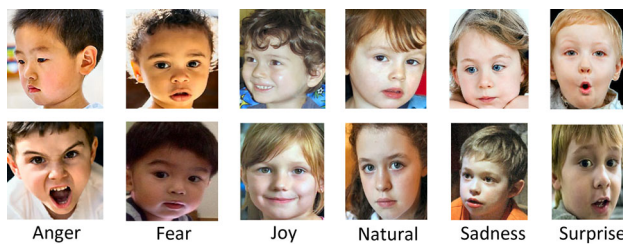


Fig. 8 The six primary used emotions

This paper used 758 images for training (Anger: 67, Fear: 30, Joy: 350, Natural: 48, Sadness: 200, Surprise: 63) and 72 images for testing (Anger: 3, Fear: 3, Joy: 42, Natural: 7, Sadness: 14, Surprise: 6).

4.2 Evaluation of the proposed model

The performance of the CNN with the used data without optimization is shown in Fig. 9.

As shown in Table 3 and in Fig. 10, the effectiveness of the suggested approach (EDL) is evaluated in comparison

with the previously widely used classifiers Decision Tree (DT), Linear Discriminant (LD), Support Vector Machine (SVM), and K-Nearest Neighbor (K-NN).

From Table 3 and Fig. 10, it is shown that the EDL as a classifier achieved the best results due to the enhancement of the performance of the CNN after the optimization of the hyperparameters of the CNN. EDL outperforms other techniques as it achieved 99.99% accuracy. EDL used GA to select the optimal hyperparameters for the CNN.

4.3 Discussion

A data flow generators are used to divide the data into batches and feed it to the proposed models (MobileNet, Xception and ResNet). There were 1200 images in training and 220 image in testing. Model hyperparameters are enhanced and provided in Table 4. Accuracy and loss percentage in training and validation stage are shown in Figs. 11 and 12.

5 Conclusion

This research presented a deep convolutional neural network (DCNN) architecture for facial expression recognition. Assistive technology has proven to be one of the most important innovations in helping people with autism improve their quality of life. A real-time emotion identification system for autistic youngsters is a vital issue to detect their emotions to help them in case of pain or anger. A real-time emotion identification system for autistic youngsters was developed in this study. Face identification, facial feature extraction, and feature categorization are the three stages of emotion recognition. A total of six facial

Fig. 9 The results of the CNN

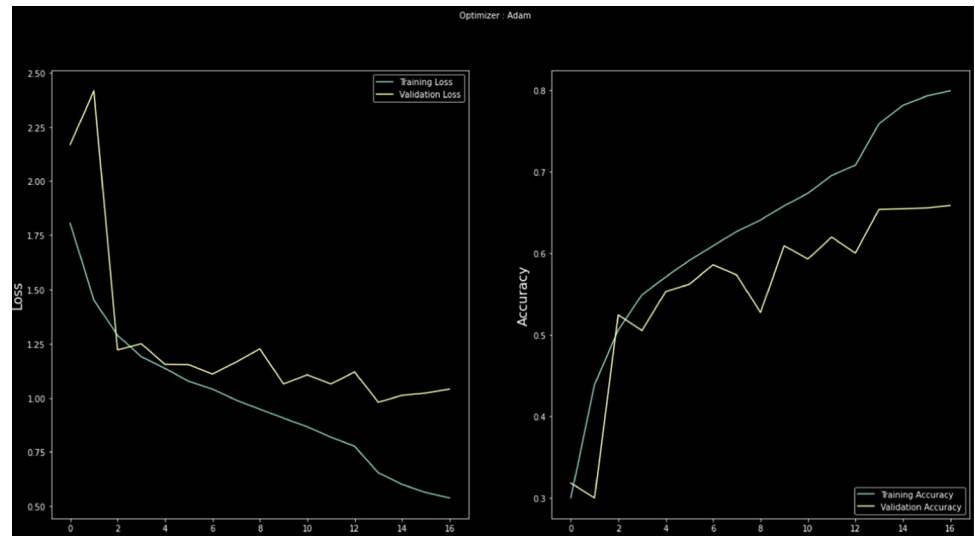


Table 3 The performance of the proposed method (EDL) vs. previous classifiers

Method	Performance metrics			
	Precision (%)	Recall (%)	Accuracy (%)	Specificity (%)
EDL	99.97	100.00	99.99	99.98
CNN	99.65	100.00	99.82	99.65
SVM-Linear	99.93	98.72	99.33	99.93
SVM-Gaussian	99.93	99.43	99.68	99.93
SVM-Cubic	99.93	99.65	99.79	99.93
K-NN	99.64	98.44	99.04	99.65
LD	99.64	97.38	98.51	99.65
DT	95.69	95.96	95.82	95.67

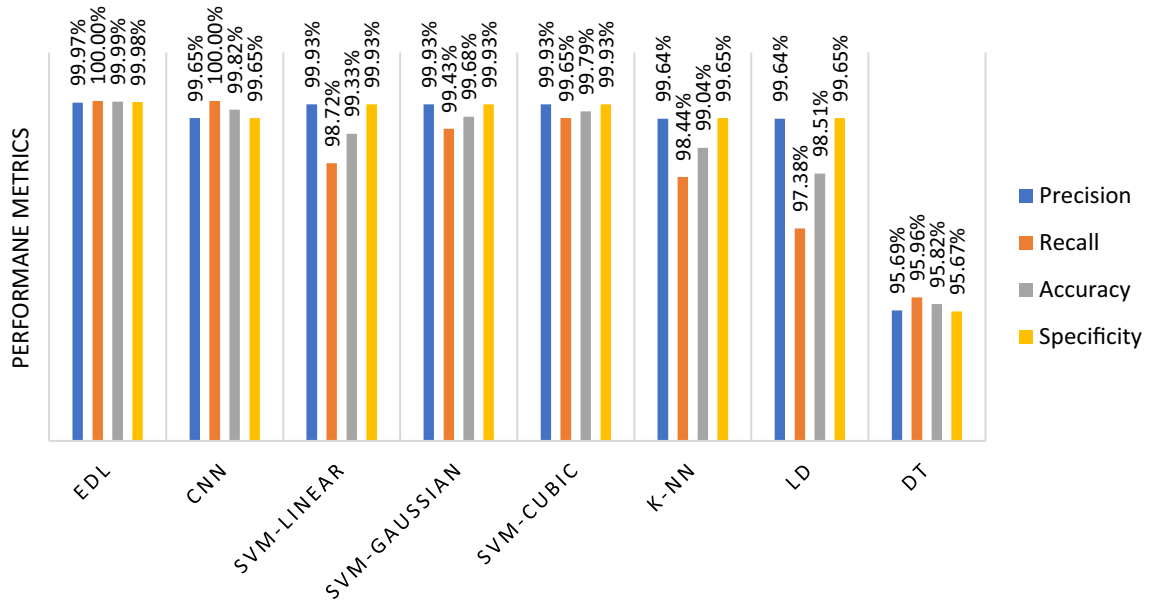


Fig. 10 The results of the CNN

Table 4 Results of the pretrained models

Model	Acc	P	R	F1	K	AUC
MobileNet	0.8812 ± 0.0231	0.9138 ± 0.0281	0.9027 ± 0.0101	0.9002 ± 0.0256	0.813 ± 0.0264	0.892 ± 0.0312
ResNet	0.9143 ± 0.0338	0.9028 ± 0.0335	0.9311 ± 0.005	0.9162 ± 0.0227	0.8021 ± 0.0315	0.8625 ± 0.0103
Xception	0.9523 ± 0.0278	0.932 ± 0.0318	0.9421 ± 0.0134	0.9331 ± 0.0219	0.853 ± 0.0312	0.9134 ± 0.0121

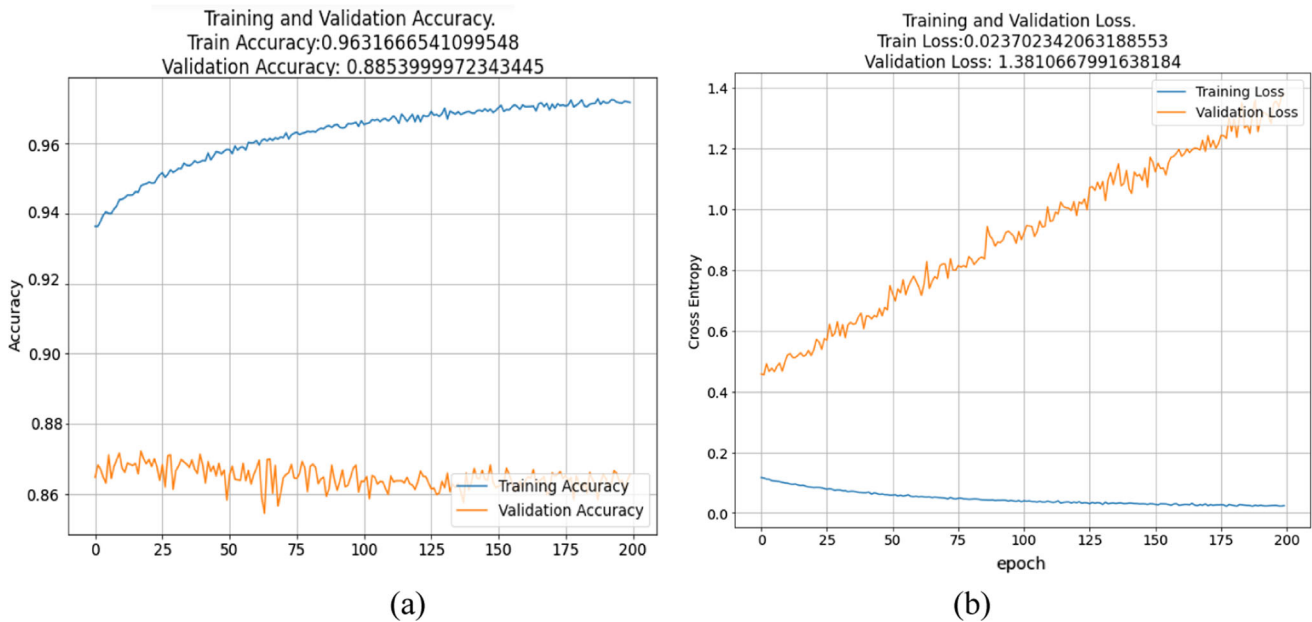


Fig. 11 Deep learning model results, **a** Training and validation accuracy, **b** training and validation loss

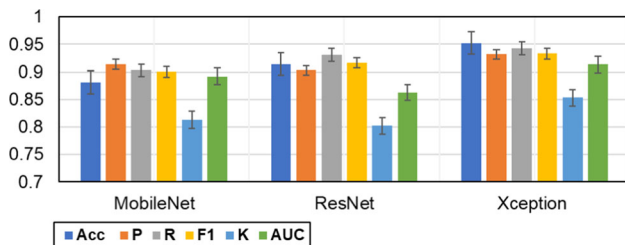


Fig.12 Comparison of the pretrained models

emotions are detected by the propound system: anger, fear, joy, natural, sadness, and surprise. This research presents a deep convolutional neural network (DCNN) architecture for facial expression recognition to help medical experts as well as families in the early diagnosis of autism. The limitation of the proposed technique is that it uses small dataset (limited scale) as the large number of real dataset is not available. In the future work, we intend to use a large real dataset.

Authors’ contributions Single Author.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Data availability <https://www.kaggle.com/datasets/fatmamtaalat/autistic-children-emotions-dr-fatma-m-talaat>.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest to report regarding the present study.

Ethical approval There is no any ethical conflicts.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright

holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kanner L (1943) Autistic disturbances of affective contact. *Nerv Child* 2(3):217–250
- American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Publishing, Washington
- Baio J (2014) Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010
- Baron-Cohen S, Wheelwright S (2004) The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J Autism Dev Disord* 34(2):163–175
- Magdin M, Prikler F (2018) Real time facial expression recognition using webcam and SDK affectiva. *IJIMAI* 5(1):7–15
- Capps L, Kasari C, Yirmiya N, Sigman M (1993) Parental perception of emotional expressiveness in children with autism. *J Consult Clin Psychol* 61(3):475
- McDuff D, Mahmoud A, Mavadati M, Amr M, Turcot J, Kaliouby RE (2016) AFFDEX SDK: a cross-platform realtime multi-face expression recognition toolkit. In: Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems. ACM, pp 3723–3726
- Mazefsky CA, Herrington J, Siegel M, Scarpa A, Maddox BB, Scahill L, White SW (2013) The role of emotion regulation in autism spectrum disorder. *J Am Acad Child Adolesc Psych* 52(7):679–688
- Lopresti EF, Bodine C, Lewis C (2008) Assistive technology for cognition understanding the needs of persons with disabilities. *IEEE Eng Med Biol Mag* 27(2):29–39
- Knight V, McKissick BR, Saunders A (2013) A review of technology-based interventions to teach academic skills to students with autism spectrum disorder. *J Autism Dev Disord* 43(11):2628–2648
- Aresti-Bartolome N, Garcia-Zapirain B (2014) Technologies as support tools for persons with autistic spectrum disorder: a systematic review. *Int J Environ Res Public Health* 11(8):7767–7802
- Blasco S, Cerro P, Elena M, Uceda JD (2009) Autism and technology: an approach to new technology-based therapeutic tools. In: World Congress on medical physics and biomedical engineering, September 7–12, Munich, Germany. Springer, Berlin, pp 340–343
- Miranda P (2003) Toward functional augmentative and alternative communication for students with autism. *Lang Speech Hear Serv Sch*
- Liddle K (2001) Implementing the picture exchange communication system (PECS). *Int J Lang Commun Disord* 36(S1):391–395
- Gates K (2004) The past perfect promise of facial recognition technology. *ACDIS Occasional Paper*.
- Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. *ACM Comput Surv (CSUR)* 35(4):399–458
- Baron-Cohen S, Golan O, Ashwin E (2009) Can emotion recognition be taught to children with autism spectrum conditions? *Philos Trans R Soc B Biol Sci* 364(1535):3567–3574
- Cheng L, Kimberly G, Orlich F (2002) KidTalk: online therapy for Asperger's syndrome. Microsoft Research, Bengaluru
- Goldsmith TR, LeBlanc LA (2004) Use of technology in interventions for children with autism. *J Early Intensive Behav Interv* 1(2):166
- Dautenhahn K, Werry I (2004) Towards interactive robots in autism therapy: background, motivation and challenges. *Pragmat Cognit* 12(1):1–35
- Robins B, Dautenhahn K, Dickerson P (2009) From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot. In: 2009 Second international conferences on advances in computer-human interactions. IEEE, pp 205–211
- El Kaliouby R, Robinson P (2005) The emotional hearing aid: an assistive tool for children with Asperger syndrome. *Univ Access Inf Soc* 4(2):121–134
- Ekman P, Friesen WV (1978) Manual for the facial action coding system. Consulting Psychologists Press, Palo Alto
- Pantic M, Rothkrantz LJ (2000) Automatic analysis of facial expressions: the state of the art. *IEEE Trans Pattern Anal Mach Intell* 12:1424–1445
- Barrett LF, Mesquita B, Ochsner KN, Gross JJ (2007) The experience of emotion. *Annu Rev Psychol* 58:373–403
- Magdin M, Benko E, Koprda Š (2019) A case study of facial emotion classification using Affdex. *Sensors* 19(9):2140
- Batty M, Taylor MJ (2003) Early processing of the six basic facial emotional expressions. *Cognit Brain Res* 17(3):613–620
- Leony D, Muñoz-Merino PJ, Pardo A, Kloos CD (2013) Provision of awareness of learners' emotions through visualizations in a computer interaction-based environment. *Expert Syst Appl* 40(13):5093–5100
- Wells LJ, Gillespie SM, Rotshtein P (2016) Identification of emotional facial expressions: effects of expression, intensity, and sex on eye gaze. *PLoS ONE* 11(12):e0168307
- Bartlett MS, Viola PA, Sejnowski TJ, Golomb BA, Larsen J, Hager JC, Ekman P (1996) Classifying facial action. In: Advances in neural information processing systems, pp 823–829
- Talaat FM, Gamel SA (2022) Predicting the impact of no. of authors on no. of citations of research publications based on neural networks. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-022-03882-1>
- Talaat FM (2022) Effective deep Q-networks (EDQN) strategy for resource allocation based on optimized reinforcement learning algorithm. *Multimed Tools Appl* 81:39945–39961
- Reddy GT, Kaluri R, Reddy PK, Lakshmana K, Koppu S, Rajput DS (2019) A Novel Approach for Home Surveillance System Using IoT Adaptive Security (February 23, 2019). Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur—India, February 26–28. <https://ssrn.com/abstract=3356525> or <https://doi.org/10.2139/ssrn.3356525>
- Babulal KS, Das AK, Kumar P, Rajput DS, Alam A, Obaid AJ (2022) Real-time surveillance system for detection of social distancing. *Int J E-Health Med Commun (IJEHMC)* 13(4):1–13. <https://doi.org/10.4018/IJEHMC.309930>
- https://www.tutorialspoint.com/fuzzy_logic/fuzzy_logic_membership_function.htm
- <https://codecrucks.com/what-is-fuzzy-membership-function-complete-guide/>
- Talaat FM, Alshathri S, Nasr AA (2021) A new reliable system for managing virtual cloud network. *Comput Mater Cont* 73(3):5863–5885. <https://doi.org/10.32604/cmc.2022.026547>
- <https://www.kaggle.com/datasets/fatmamtaalat/autistic-children-emotions-dr-fatma-m-talaat>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.