Autism Spectrum Disorder Detection: Video Games based Facial Expression Diagnosis using Deep Learning

3 authors:

Morched Derbali
King Abdulaziz University
21 PUBLICATIONS   198 CITATIONS

SEE PROFILE

Mu'tasem Jarrah
UA
16 PUBLICATIONS   277 CITATIONS

SEE PROFILE

Princy Randhawa
Manipal University Jaipur
60 PUBLICATIONS   498 CITATIONS

SEE PROFILE

Autism Spectrum Disorder Detection: Video Games based Facial Expression Diagnosis using Deep Learning

Morched Derbali[1], Mutasem Jarrah[2], and Princy Randhawa[3]

Department of IT, Faculty of Computing and IT, King Abdulaziz University, Jeddah, Saudi[1, 2]

Department of Mechatronics Engineering, Manipal University Jaipur, Jaipur, India[3]

Abstract—**In this study, a novel method is proposed for determining whether a child between the ages of 3 and 10 has autism spectrum disorder. Video games have the ability to immerse a child in an intense and immersive environment. With the expansion of the gaming industry over the past decade, the availability and customization of games for children has increased dramatically. When children play video games, they may display a variety of facial expressions and emotions. These facial expressions can aid in the diagnosis of autism. Footage of children playing a game may yield a wealth of information regarding behavioral patterns, especially autistic behavior. You can submit any video of a child playing a game to the interface, which is powered by the algorithm presented in this work. We utilized a dataset of 2,536 facial images of autistic and typically developing children for this purpose. The accuracy and loss function are presented to examine the 92.3% accurate prediction outcomes generated by the CNN model and deep learning.**

Keywords—Autism in children; machine learning; deep learning; convolution neural network (CNN); video games; prediction

I. INTRODUCTION

Autism is a complicated, behaviorally defined, static condition of an immature brain which is of significant concern to practicing pediatricians due to a staggering 55.6 per-cent rise in pediatric incidence from 1991 and 1997, surpassing spina bifida, cancer, and Down syndrome [1]. Rather than new environmental effects, this increase is due to increased awareness and evolving diagnostic criteria. Autism is a condition with numerous nongenetic and genetic origins, rather than a disease. Autism (autism spectrum disorders) is defined as a group of developmental disorders characterized by deficiencies in three behavioral domains: [2]

1) interpersonal interaction.
2) a diverse set of areas of interest and hobbies; and
3) speech, communication, and creative play.

A. Autism and its Characteristics:

Early childhood autism is a pervasive developmental disorder. Autism affects communication, relationships, and self-control. Infants often get autism. Autism is a "spectrum disorder" that affects people differently. It has many traits. [3].

Early diagnosis can help a person with autism live a full life. According to the DSM-5, autism is characterized by persistent differences in communication, interpersonal relationships, and social engagement [4]. Example: Being nonverbal or having abnormal speech patterns, having trouble understanding nonverbal communication, developing and maintaining relationships, and having trouble maintaining a traditional back-and-forth conversational manner [5]. Repetitive habits, interests, and behaviors. Excessive awareness to or significantly reduced sensitivity to many sensory stimuli, repetitive sounds or phrases (echolalia), preference for homogeneity and complexity with transition or regimen, rigid or heavily restricted and strenuous interests, hyper sensitivity to or dramatically reduced sensitivity to many sensory stimuli, rigid or severely regulated and intense interests, hyper sensitive. According to the American Psychological Association's Diagnostic statistical, autistic traits must be present in early childhood, but they may not fully express until social pressure builds the person's strength to deal with them, and difficulties may be covered up by learned coping skills [6].

B. The Role of Video Games in Autism

According to research, enabling youngsters to play games using smart phones might help detect autism. Dr. Jonathan Delafield-Butt, a senior professor in childhood development, said it was important to detect autism early so parents and children could receive a variety of support services. [7]. Autism is a neurodevelopmental disorder with many shared traits, challenges, and abilities. Many autistic people have visual-spatial thinking, pattern identification, and a visual preference. Games that require visual clues and spatial skills are rewarding to such people. Games are creative but structured. RPGs and scrolling shooters satisfy research participants' desire for imagination without requiring self-generated creativity, which many autistic people lack. Video games have many audio and visual cues. Autistic people value rules and objectivity more than neurotypical people. To avoid anxiety and sensory meltdowns, follow clear guidelines. Video games reinforce clear expectations. Autistics need routine and repetition. Unknown circumstances cause anxiety, discomfort, and a desire to escape. Video games allow for controlled practice and mastery [8,9]. Games are more controlled than real life. Autism makes unpredictable human behavior difficult. Understanding social signs, idioms, humor, sarcasm, and satire can also cause anxiety. Playing a game that becomes more familiar each time helps autistic gamers overcome these challenges in a safe, controlled environment.

Parents and educators worry that autistic students spend too much time gaming instead of socializing. Playing has many benefits if encouraged and controlled [10].

II. RELATED WORK

Modern diagnostic tools for mental diseases were developed in the late 1800s, although their origins may be dated back to the 4th century B.C (Before Christ) Era [11]. The gold standard for diagnosing often these mental-disorders relies heavily on information gathered from various respondents (e.g., parents, teachers) about the onset, direction, and duration of various behavioral descriptors, which is then considered by providers when making a diagnosis predicated on DSM-5 (Diagnostic and Statistical Manual) Categorization of Diseases-10th Installment (ICD-10) requirements [12]. Providers employ a variety of strategies to collect this data, ranging from subjective (e.g., assessment scale) and unstructured (e.g., semi - structured or unstructured interviews) to much more objective (e.g., actual observations) and organized (e.g., structured diagnostic interviews) [13].
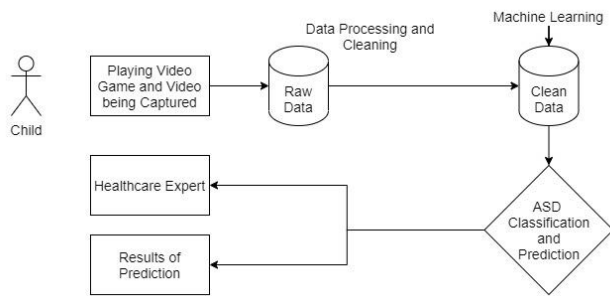

Fig. 1. Machine learning model for prediction and classification

Fig. 1 shows the architecture of the machine learning model that is used commonly for prediction of machine learning along with video games. The video of the child playing the game is captured and then the data is preprocessed, and machine learning is applied on it to predict the results.

Autism spectrum disorder (also referred to as ASD) and attention deficit hyperactivity disorder (also referred to as ADHD) are both conditions that are relatively common in children and can continue into adulthood. Autism spectrum disorder (ASD) is a developmental condition that causes patients to have difficulties with speech, behavior, and social interaction. Patients with ASD also tend to engage in repetitive behavior, have impatience issues, and attention problems. Since the publication of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), the term autism spectrum disorder (ASD) has been used to refer to a more comprehensive diagnostic entity that formerly referred to a number of distinct disorders, including Autistic Disorder, Asperger's, and other Pervasive Chromosomal Anomalies [14]. According to findings from recent studies, the prevalence of autism spectrum disorder (ASD) in children and adolescents has increased from 1 in 100 to 1 in 59 in the past 14 years (from the year 2000 to 2014). Inattention, hyperactivity, and impulsivity are all symptoms of attention deficit hyperactivity disorder (ADHD), which is a common

brain condition found in children and adolescents. Both autism spectrum disorder and attention deficit hyperactivity disorder (ADHD) tend to continue into adulthood. A diagnosis of the condition is made in approximately 5.9–9.4% of all infants. Because autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD) are so common in children, accurate and timely diagnosis of these conditions is critical. [15,16,17].

The field of artificial intelligence known as machine learning has the potential to significantly improve the use of computer methods in the field of neuroscience. A significant amount of research has been done to establish machine-learning models and deep-learning approaches in order to interpret high-dimensional MRI (Magnetic Resonance Imaging) data in order to simulate neural networks that regulate the brains of people who suffer from a variety of mental illnesses. [18,19]. These studies resulted in the development of machine-learning methods for the classification of Alzheimer's disease, mild cognitive impairment, right temporal epilepsy, schizophrenia, Parkinson's disease, dementia, attention deficit hyperactivity disorder, autism spectrum disorder, and major depressive disorder [20]. These statistical algorithm-based machine-learning models are well-suited to complex issues that require a combinatorial explosion of options or non-linear processes. This is because typical computer models fail in terms of quality or scalability in these situations, but these statistical algorithm-based machine-learning models succeed [21].

Senju et al, discussed the approaches to early detection of autism in infants. Early detection here refers to before 18 months of age. It gives an overview of the known processes of early social development in children that can fall under the category in which "core deficits" are manifested in young children and summarizes a criterion for the same. The paper concludes by discussing how the preferences for social stimuli changes in the infants over time [22]. They develop a preference for familiar people in first couple of months of life. They begin to have one to one interaction with their caregivers between three and six months of age. This is the period where they understand the interactive styles of their usual social partners. During four to seven months, infants can differentiate between emotional expressions of the face. Soon they develop preferences for caregivers and after 12 months, they display a pattern of response to separation and reunion that demonstrates what they have learned to expect from their experiences with their caregivers [23].

So, in the case of the infants with autism, one would speculate that the dyadic interactions would be less intense and lower capacity to recognize the emotional expressions of others. However, by three to five years of age, autistic children would be capable of forming secure attachments to their caregivers. The author concluded that similar patterns of attachment security and insecurity are found in typically developing children, it is very unlikely to be considered as a core deficit in autistic infants. In infants, failure to discriminate emotions can be considered as an important feature to characterize autism. But however, the rate of false positives could be very high. Children with autistic tendencies make use of different methods and skills to develop secure

attachments and to obtain the capacity to differentiate between the facial expressions and emotions than the typically developing children.

Whalen et al. [23,24] discovered that using a computer game to teach children with autism improved motivation and engagement when compared to the traditional methods of trying to teach children with autism. They established the computer game Teach Town to assist autistic children improve social, emotional, academic, and adaptive skills [25].Whalen's findings are in line with what others in the area have found: video-game-like tools are beneficial with students with ASD because they are constant and predictable, entail few social variables, and enable children to control and set the speed of the activity. Many children on the autism spectrum have visual-spatial abilities that create videos actual gameplay an area of expertise. While it's normal to be concerned about inattention, behavioral concerns, and even addictions, there really are simple ways that may help autistic children get the most out of computer games and computers without causing additional problems at home or at school. The digital world provides a wealth of resources for teaching issue, social skills, adaptability in new contexts, and even motor skill development. Raising autistic children may be difficult, but apps, games, and technologies can make contacting and educating autistic children simpler. Tbatah et al [26] laid out a principle called anger superiority hypothesis as its foundation. This hypothesis states that "angry faces capture attention faster than happy faces" in common individuals. The authors aim to test and compare the threat detection abilities in autistic people with individuals with a history of typical development using a facial visual search paradigm.

Thabtah et al [26] theory is based on building classification systems using machine learning, specifically a new method called Rules-machine Learning. This approach helps in detecting autistic traits and offers user knowledge bases (rules) which enables the professionals to make better analysis of the reasons behind the classification.

The primary objective of this technique is rule discovery by search method which can be done using covering classification. Then, evaluation is performed on the discovered rules to discard any redundancies and to optimize it further by reducing the number of discovered rules. In order to improve the overall efficiency of the training process, this phase contributes to the narrowing of the search space for individual data items. The classifier, which is utilized to make predictions regarding the value of the class, is more comparable to an outcome of the rule evaluation phase described above. For the purpose of generating the necessary data from the participants, a mobile application known as ASD Tests is utilized [27]. It implements four screening methods for toddlers, children, adolescents, and adults based on the Q-CHAT-10 (Quantitative Checklist for Autism in Toddlers), AQ-10-child, AQ-10-Adolescent, and AQ-10 adult, respectively. The Q-CHAT-10 is a quantitative checklist for autism in toddlers. In addition to this, the author made use of the datasets that were previously deposited in the University of California Irvine Data Repository by the authors.

Wu et al [28] analyzed the performances of various ML techniques such as Bagging, Boosting, rule induction, and decision tree classifiers on child, adolescent and adult ASD screening datasets. The error rates of adult dataset for the above techniques were between 5.68 and 8.23 per cent whereas the Rules Machine Learning (RML) model outperformed them with an error rate less than 5.6 per cent. Hence the paper was concluded by clearly revealing how ML approaches like covering can be used for obtaining promising results [29].

Jacob et al [29] obtained a high standard clinical data of children at risk for ASD to implement machine learning algorithms. The aim is to build a low-cost and easy to use ASD screening tool. To implement this, the author chose to proceed with a combination of two approaches. Two different algorithms are trained to combine their outputs as a final screening assessment. One is based on the short, structured parent- reported questionnaires and the second is based on tagging key behaviors from casual home videos of the test subjects. The first classifier was trained using data from ADI-R (Autism Diagnostic Interview) score sheets with labels corresponding to established clinical diagnoses. The training of second classifier i.e., the video classifier was done using ADOS (Autism Diagnostic Observation Schedule) instrument score sheets and diagnostic labels. To ensure sufficient training volume, progressive sampling was used in both the cases. After evaluating multiple machine learning algorithms, the author chose Random Forests for its robustness against overfitting [30].

In the clinical sample, the results showed that the parent questionnaire classification approach performed better than some of the more established screening tools, such as the M-CHAT (Modified Checklist for Autism in Toddlers) and the CBCL (Child Behavior Checklist). By combining the two different methods of classification into a single evaluation, performance was improved. The author concluded by stating how ML can play a crucial role in enhancing the performance of the behavioral health screeners and how this research demonstrated a significant improvement over established screening tools for autism. The author also mentioned how the research demonstrated how this research demonstrated significant improvement over established screening tools for autism. Chorianopoulou et al. [31] presented an ML-based approach to early diagnosis of ASD from videos of infants by identifying specific behaviours from them. This approach was based on using videos of the infants. They used a dataset that contained 2000 short videos with various behaviours of interest, such as directed gaze towards faces or objects of interest, positive affect vocalization, and other similar behaviours, all of which were manually coded by expert raters [32]. This dataset was used to conduct their research.

The authors addressed the issue by employing a deep learning model that was image-based and that was based on facial behavior features. Gorriz et al [32] has applied the various feature transformation techniques such as Log, Z-score, sine functions to the collected datasets of toddlers, children, adolescents, and adults [33]. In the next stage, various classification techniques were implemented with these

transformed ASD datasets, to evaluate and assess their performance [33].

For toddler dataset, the median highest result was calculated by Adaboost for Log transformation, Adaboost and SVM (Support Vector Machine) for Scale transformation respectively as 99.06%. The mean highest result which was reported to be 98.77% was calculated by SVM for Log and Sine transformations. The maximum highest result was recorded to be 100% was calculated by Adaboost, GLMboost and SVM for all feature transformation methods and C5.0 for Scale transformation [27, 34].

For child dataset, the median highest accuracy of 100% was achieved by LDA (Linear Discriminant analysis) and PCA (Principal Component Analysis) for Log and Scale feature transformations. The mean highest accuracy of 97.2% was achieved by Adaboost for Log and Scale, respectively. Finally, the maximum highest accuracy of 100% was achieved by all classifiers and feature transformation methods [34,35].

For adolescent dataset, the median highest accuracy of 95% was obtained by C5.0, LDA, PCA where LDA and PCA for both Log and Scale as well as C5.0 for Scale. The mean highest result of 93.89% was obtained by PCA for Log and GLMboost for Scale, respectively. The maximum highest result of 100% were achieved by all classifiers and feature transformation methods [35,36].

Table I summarizes previous studies and its findings and limitations on usage of algorithms and the different features for the autism spectrum disorder

TABLE I. SUMMARIZED PREVIOUS STUDIES AND ITS FINDINGS FOR AUTISM SPECTRUM DISORDER

Reference	Findings	Limitations
Thabtah et al, 2018 [6]	Inside the ASD screening tool, SVM was used to integrate the ML algorithm. 97.6 % accuracy	Datasets were not balanced. The sample is small comprising 612 autism patients and 11 non-autism instances.
Thabtah et al, 2020 [26]	Rules-Machine Learning is a machine learning approach based on rule induction (RML). Covering learning was used to generate non-redundant rules in a simple method. RML classifies with greater prediction accuracy than typical algorithms such as boosting, bagging, and decision trees, thanks to the use of ten times cross-validation to split the dataset into ten subsets.	In terms of class labels, RML appeared to be ineffective when dealing with unbalanced data sets. There were no examples of toddlers in this article.
Vaishali et al, 2018 [27]	Optimal feature selection was automated using the Binary Firefly algorithm (ten out of twenty-one features were chosen as the best). There was no concern with class imbalance (there are 151 occurrences with class 'yes' and 141 instances with class 'no' in the ASD youngster's dataset). Models such as NB, J48, SVM, and KNN were used. SVM obtained the highest accuracy of 97.95 percent.	In the ASD kid dataset, there were some occurrences that are missing. There was a risk of model overfitting on the dataset because to the smaller number of occurrences in the dataset. Swarm intelligence wrappers had certain drawbacks (Binary Firefly algorithm)
Al banna et al, 2020 [33]	Analyzed the patient's condition using facial expressions and emotions, employing an AI system and sensor data. Sent out frequent messages to parents, assisting the patient in coping with ASD during COVID-19. A smart wristband with an integrated monitor and camera is linked to a smartphone app in this system. Used real-time grayscale photos from one Kaggle dataset of 35,887 images to detect ASD. The Inception-ResNetV2 architecture had the greatest accuracy of all the models, at 78.56 percent.	When compared to other methods, the accuracy is poor. The research is still in its early phases.
Sen et al, 2018 [37]	The authors devised a new algorithm that combined structural and functional characteristics. Drew many different depictions of the brain's functional connections. The results showed that incorporating multimodal characteristics improves case discrimination accuracy the most.	In contrast to earlier studies, the ML models utilized demonstrate a 4.2 percent improvement in the accuracy of the predictions for Autism. Datasets suffer significantly from fluctuations.
Van den et al, 2017 [39]	SVM, Naive Bayes, and Random Forest classification algorithms are used. There were 95,577 kid records with 367 variables, of which 256 were deemed to be adequate. Different qualities were well delineated. Created a dataset with four classifications (ASD: None, Mild, Moderate, Severe). The J48 algorithm attained the highest accuracy of 87.1 percent (2 class) and 54.1 percent (4 class) (decision tree)	Doesn't predict the severity of ASD. A cursory collection of traits (criteria) used to identify ASD, which may or may not always correspond to an instance of ASD.

III. METHODOLOGY

This study demonstrated the use of Deep Learning and Image processing techniques for the detection of Autism using facial expressions. The initial approach was to build and train a neural network based on the available data on Autism. Following this, any video of the patient which clearly shows their facial expressions could be taken as the input through an interface created for the users. This input was used for the detection of autistic characteristics using the previously trained model [38].

The methodology followed can be divided into five steps as shown in Fig. 2:

1) Capturing the facial expressions while playing a Video Game.
2) Data Preprocessing.
3) Model Building and Training.
4) Prediction and Optimization.
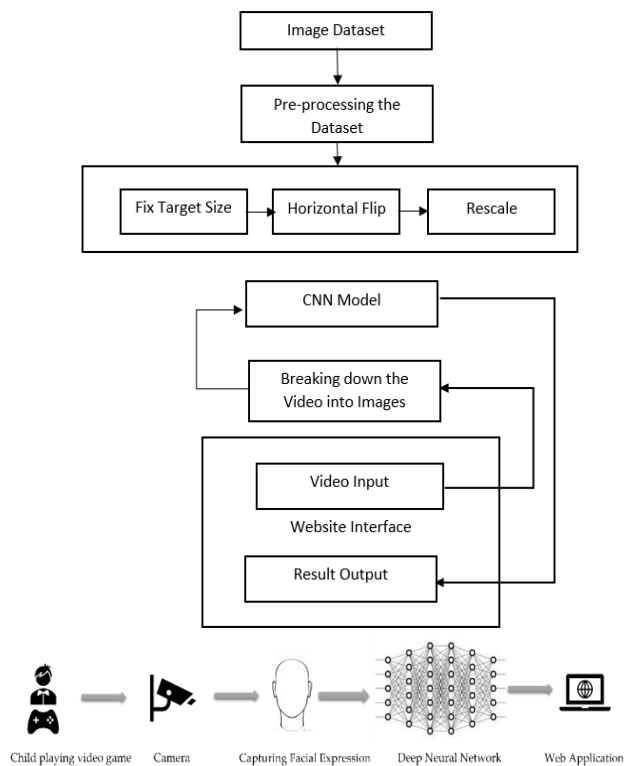*5)* Uploading the video to the Web Interface.

Fig. 2. Work flow of the method for the detection of facial expression while playing video games

A. Capturing the Facial Expressions while Playing the Video Game:

The facial expression of the child is captured while she/he is playing the video game using the web camera attached to the personal computer. The video is captured and saved locally and then the same video is uploaded to the website user interface and then the trained Convolutional Neural Network (CNN) model predicts if the child in the image broken down from the video is autistic or not [40].

B. Data Preprocessing

The collected image dataset was preprocessed by 3 processes so that the CNN model can train the dataset to predict if the child is autistic or not. They were:

1) Dividing the video into frames: OpenCV was used to fragment the video captured into images or frames. The video is captured using a webcam while the child is playing the game to monitor and analyze the video captured which would be broken down into images and then train the CNN model. The dataset consists of 2536 images belonging to the autistic and non-autistic classes in the training set and 300 images in the test set. The required image data is ob-tained from Kaggle. The ImageDataGenerator class of keras library enables us to read the images from the folders. Rescaling of images is done by dividing each pixel value of an image by 255. The images in the dataset are of various sizes which calls for resizing into one final size of 64x64 pixel and the same can be implemented using a function called "flow_from_directory" of the above class [40,41].

2) Fix target size: The collected image dataset had to be resized for uniformity in processing and to do so, all the images were resized to 64x64 pixel so that the convolution neural network model can be trained using those images [42]. This was done by using the function "flow_from_directory" and the argument to the function changed to 64x64 pixels. Whereas the neural network was training on the training data, the flow from directory () function was used to read photos straight from the directory and enhance them. The technique assumed that photos from various classes stored in separate directories but are all included within the same parent directory.

3) Horizontal flip: The images needed to be flipped horizontally to maintain uniformity, so this is achieved by using the function "ImageDataGenerator" and adjusting the arguments.

4) Rescaling: The ImageDataGenerator class may rescale pixel values from 0-255 to the recommended 0-1 range for neural network models. Normalization is the process of reducing data to a number between 0 and 1. Setting the rescale parameter to a ratio that may be multiplied by each pixel to produce the required range will do this [41,42].

C. Model Building and Training:

Deep learning is a technique for automatic learning that implements the use of examples to teach machines how to learn in the same way that people do. A self-driving car's ability to recognize a stop sign or differentiate between a pedestrian and a lamppost is dependent on its use of deep learning, which is an essential component of the technology. It makes it possible to control consumer electronics with one's voice, such as mobile phones and tablets, televisions, and hands-free speakers. The concept of "deep learning" has been receiving a lot of attention as of late, and with good reason. It's about achieving things that weren't possible before you started working on them.

During the process of deep learning, a computer model will learn to perform categorization tasks directly from either pictures, text, or sound. Models that use deep learning have the potential to achieve an accuracy that is on par with or even exceeds that of humans in certain circumstances. In order to train models, a significant amount of labelled data as well as various topologies of multilayer neural networks are utilized. The acronym "CNN" stands for "Convolution Neural Network". Image recognition and processing are two applications that make use of a type of artificial neural network known as a convolutional neural network (also abbreviated as CNN). CNNs are designed to focus specifically on analyzing pixel input. CNN was used to train the model based on the image data sets that were taken by the webcam of the computer that the child was using to play the video game on. CNNs are image processing, artificially intelligent (AI) systems that utilize deep learning to perform both generating and informative tasks. These tasks frequently include machine vision, which includes image and video identification, as well as recommendation systems for natural language processing (NLP) [40, -42]. A neural network is a piece of computer hardware and/or software that mimics the way neurons in the

human brain communicate with one another. Traditional neural networks were not intended to be used for image analysis, so in order for them to do so, they require the images to be broken up into smaller chunks. The "neurons" that make up CNN are organized more similarly to those in the prefrontal cortex, which is the part of the brain in humans and animals that is responsible for processing visual input. The difficulty of processing images in pieces that is inherent to traditional neural networks can be circumvented by arranging the layers of neurons in such a way that they cover the entirety of the visual field [48]. A CNN makes use of a technology similar to a perceptron that is designed to have minimal requirements for processing [43]. The layers of a CNN are comprised of an input layer, a layer, and a hidden layer. These layers are followed by several convolutional layers, average pooling, fully connected layers, and normalizing layers. A system that is significantly more effective and easier to train for image analysis and natural language [43] has been made possible as a result of the elimination of constraints and improvements in the efficiency of image processing.

D. Prediction and Optimization

Optimization plays a crucial role for any machine learning problem. Gradient descent is an optimization algorithm that finds the lowest possible value or the minimum value of a function through iterations. While the loss function which is also known as the cost function is all about calculating the loss/errors for every prediction that the neural network makes, gradient descent can be used to find the minimum of this loss function. The goal is to estimate the values of coefficients of a function that can minimize the cost function. In other words, the new coefficients will have a significantly lower cost.

This technique is initialized by taking small random values as coefficients for the function. The cost is then evaluated by inserting them into the function. The next step is to change the values of the coefficients in a direction that can lead to a lower cost in the next iteration. This direction can be estimated with the help of derivatives. The derivative gives a slope (gradient) at the desired point on a curve, or a function and the sign of that slope can be used to determine the direction in which the coefficients can be moved in further iterations [44]. Now that the algorithm is aware of the direction of progression of coefficients with the help of the gradient at current position, the next move is to make a step by scaling it and subtracting the obtained value from the current position. Subtracting is done as the aim is to minimize the function [40]. Another parameter called learning rate is used to scale the gradient and control the step size. Learning rate can affect the performance in a significant way. Smaller learning rate can lead the algorithm to reach the final iteration before even reaching the optimum point [41]. The Adam algorithm is implemented by

this optimizer. Adam optimization is a gradient descent approach based on adaptive first- and second-order moment estimation. The approach is "computationally more efficient, has small memory demand, is robust to diagonally resizing of gradient, and is well suited for situations with huge data/parameters," as according to Thabtah- et al., 2020.

E. Uploading the video to the Web Interface

Streamlit is one of the recent and fastest python-based model deployment tools. This open-sourced python based framework simplified the whole model deployment cycle along with providing an easy way to structure the functionalities of the interface [45,46].
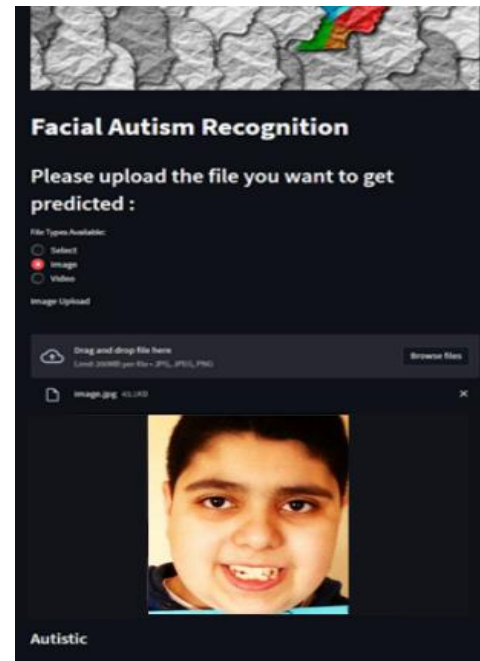

Fig. 3. Web interface for prediction of the disease

Once the neural network was trained, the model of an epoch with the best validation accuracy was saved as a '.model' file. The interface enables the users to select the type of media file that is to be uploaded. If the selected media type is an image, the saved model is used to classify the uploaded image to Autistic/Non-Autistic for which the result can be displayed on the interface [45]. But if the selected media type is a video, then the uploaded video is divided into frames using OpenCV and each image is classified into Autistic/Non-autistic with the help of the saved model. The mean value of classification of all images is considered as the final classification for the entire video and displayed on the interface.
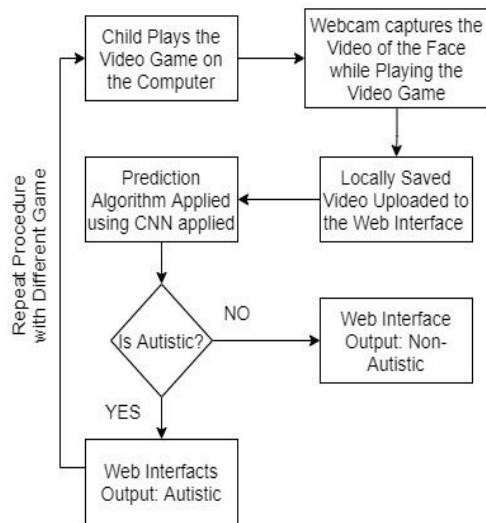
Fig. 4. Prediction mechanism if autistic or not

Fig. 3 describes the entire working of the prediction mechanism implemented in this project. The video of the child while playing the game is captured by a webcam and then the video is saved locally. The same video is then uploaded to the website interface for the CNN deep learning algorithm to work on it. After processing the file and after the prediction model is applied on the video, the outcome or result of the prediction is shown on the interface as "Autistic" or "Non-Autistic" as shown in Fig. 4.

IV. RESULTS AND DISCUSSION

This study explores the use of video games to discriminate between children with and without ASD. Compared to previous qualitative techniques, the activities and learning metrics in the evaluation game give a quantifiable depiction of children's abilities, making the identification of ASD more accurate and practical. The use of these games as a supplementary tool in educational interventions for kids with ASD is also possible. In previous studies, the model has been trained using different algorithms such as SVM, Neural Network, RML Classifier, Random Forest, Naïve Bayes with a maximum accuracy of 97.5% [46]. They have used images of a facial expression and behavior of a patient while in our study we trained the model using CNN when the child is playing video game and capture the image of his behavior and facial expression. The Convolutional Neural Network (CNN) model is used to extract and produce the proper patterns of the face features when the child is playing video game and capture the image of his behavior and facial expression such as inappropriate snickering and laughing, Lack of pain sensitivity, Inability to maintain proper eye contact, unable to communicate with gestures, Inadequate reaction to sound etc. The model was trained in the cloud using Google Colab with python, which supports TensorFlow and Keras. The epoch number declared was 159 and it had a batch size of 20. The VGG model was used for implementing the convolutional neural network. Karen Simonyan & Andrew Zisserman of Oxford University's Visual Geometry Group (VGG) proposed VGG models, which performed well in the ImageNet

Challenge [47 -48]. This model gets 92.7% top-5 accuracy on ImageNet's 14 million photos from 1000 classes. 300 images were used for the testing and 100 images were used for validation of the CNN model [48].

The model achieved 92.3% accuracy for the testing dataset and 87.3% accuracy for the validation dataset. For this study, many evaluation metrics were calculated, and the results are summarized in the following sections. The results of the comparative prediction analysis of our model and existing model of the same dataset as shown in Table II.

TABLE II. COMPARISON OF THE RESULTS USING THE SAME DATASET

Author	Model	Accuracy
Haque et al [49]	ResNet50	89.2%
Shaik et al [50]	VGG19	84.0%
Our model	VGGFace	92.3%

A. Evaluation Metrics

Sensitivity, Specificity, Precision, Recall, F1 Score and Accuracy were calculated for evaluating the performance of the CNN model. The convention True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) were used.

1) Sensitivity: Sensitivity measures a model's ability to recognize positive cases. TPR is sometimes called recall. Sensitivity measures how many positive cases a model accurately identified. High-sensitivity models have minimal false negatives, meaning they overlook some positive examples. Sensitivity is a model's capacity to recognize good cases. We need our models to locate all good examples to generate accurate forecasts. True positive rate plus false negative rate equals 1. Higher true positive rate means model correctly identifies positive situations. The sensitivity was calculated using the following formula as shown in eq. (1):

$$\text{Sensitivity} = \frac{TP}{TP+FN} \qquad (1)$$

The Sensitivity of the model was 0.9560 or 95.60%.

2) Specificity: Specificity assesses the model's ability to identify real negatives. This means there will be a percentage of true negatives forecasted as positives, or false positives. True Nega-tive Rate (TNR). True negative rate plus false positive rate equals 1 always. Low specificity suggests the model is mislabeling a lot of negative data as positive. The specificity was calculated using the following formula as shown in eq. (2):

$$\text{Specificity} = \frac{TN}{FP+TN} \qquad (2)$$

The specificity of the model was 0.8865 or 88.65%.

3) Precision: Precision is the ratio of True Positives to total positive samples (either correctly or incorrectly). It is calculated using the formula as shown in eq. (3):

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (3)$$

The precision of the model was 0.9048 or 90.48%.

4) Accuracy: Model accuracy measures which model is better at finding correlations and patterns in a dataset based on training data. Accuracy is calculated using the formula as shown in eq. (4):

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

The accuracy was found out to be 92.3% for the testing dataset consisting of 300 images and 87.3% for the validation dataset.

5) F1 score: The harmonic mean of accuracy and recall is used to get the F1 score. It is calculated using the formula as shown in eq. (5):

$$F1\ Score = 2 * \frac{Precision.Recall}{Precision+Recall} \quad (5)$$

The F1 Score was calculated to be 0.9297 or 92.97%.

6) Confusion matrix: A matrix called the confusion matrix is used to assess how well classification models perform given

a particular set of test data. As the model cannot afford to predict non-autistic when the patient is autistic, the confusion matrix was calculated as shown in Fig. 5, which indicates that the false positive condition case value should be lower. 155 of the 300 images were correctly classified.

	Autistic	**Non-Autistic**
Predicted Autistic	152	16
Predicted Non-Autistic	7	125

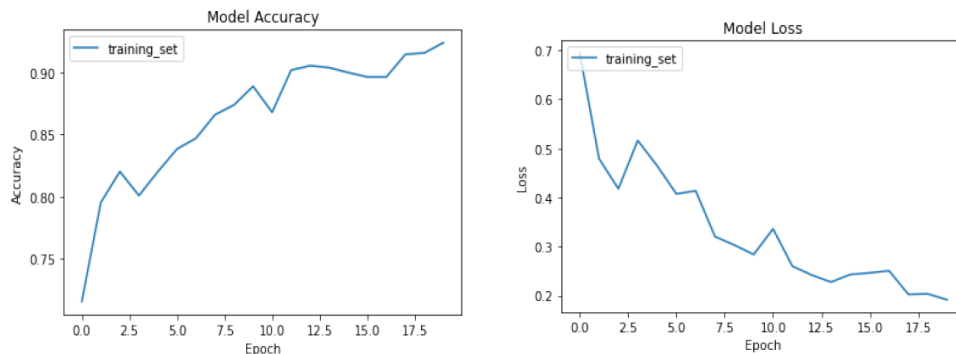Fig. 5.   Confusion matrix using CNN for autism spectrum disorder

Fig. 6.   a) and b) . Model Training Accuracy and Training Loss

Fig. 6(a) and (b) depicts the plot of the model's training accuracy, training loss, respectively for 10 epochs. The highest accuracy achieved was 92.3 % with the VGG model by applying two dense layers with specific parameters. A loss function optimizes ML algorithms. The loss is determined during training and validation, and its interpretation depends on how well the model performs. It's the total of training or validation set mistakes per example. Loss value indicates model's performance after each optimization cycle. A performance metric measures the algorithm's accuracy. Model accuracy is generally estimated as a percentage based on model parameters. It measures how well your model predicts actual data.

V.   CONCLUSION

The study built a deep learning web app to diagnose autism using a convolutional neural network and camera footage of a youngster playing a video game. CNN's architecture can extract facial attributes by generating facial feature patterns and assessing facial landmark distances, classifying faces as autistic or not. VGG CNN Model produced accurate results. Testing accuracy was 92.3%, validation accuracy was 87.3%, and precision was 90.4%. Future research will improve this model by broadening psychologists' autistic kid diagnoses. This programme helps

identify ASD. A precise autism diagnosis can help pick a treatment plan for autistic children. More precision might improve autism diagnosis. The platform might reveal this neurological disease, bringing treatment closer. This research is part of how humans utilize technology to tackle the world's healthcare issues. Future studies may use machine learning and deep learning algorithms to help individuals recognize a range of ailments using the same platform. Although young, digital technologies offer unlimited potential. In the case of autism, a mix of digital tools and in-person therapy visits is expected. Autism therapy isn't one-size-fits-all. These tools will help researchers produce medicines for ASD patients more swiftly. Language is being studied for potential therapies. SFARI-funded researchers are employing cellphones and automatic transcription software to record speech from autistic youngsters. Based on the data collected by recording footage of children playing video games, a video game can be designed for children with autism or autistic symptoms. This game will help parents or guardians determine whether their kid has autism such as Kinect game.

ACKNOWLEDGMENT

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://www.kaggle.com/code/basmarg/vgg-16-autism-image-data/data

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

[1] Aljero, M.K.A.; Dimililer, N. A Novel Stacked Ensemble for Hate Speech Recognition. Applied Sciences 2021, 11, doi:10.3390/app112411684.

[2] Suman Raj, Sarfaraz Masood, Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques, Procedia Computer Science, Volume 167, 2020, Pages 994-1004, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.399.

[3] Chae, S.; Choi, A.; Jung, H.; Kim, T.H.; Kim, K.; Mun, J.H. Machine Learning Model to Estimate Net Joint Moments during Lifting Task Using Wearable Sensors: A Preliminary Study for Design of Exoskeleton Control System. Applied Sciences 2021, 11, doi:10.3390/app112411735.

[4] Nishio, M. Special Issue on "Machine Learning/Deep Learning in Medical Image Processing". Applied Sciences 2021, 11, doi:10.3390/app112311483.

[5] Pană, M.-A.; Busnatu, Ș.-S.; Serbanoiu, L.-I.; Vasilescu, E.; Popescu, N.; Andrei, C.; Sinescu, C.-J. Reducing the Heart Fail-ure Burden in Romania by Predicting Congestive Heart Failure Using Artificial Intelligence: Proof of Concept. Applied Sci-ences 2021, 11, doi:10.3390/app112411728.

[6] Thabtah, Fadi, Firuz Kamalov, and Khairan Rajab. (2018) "A new computational intelligence approach to detect autistic features for autism screening." International journal of medical informatics 117: 112-124.

[7] Tamang, L.D.; Kim, B.W. Deep Learning Approaches to Colorectal Cancer Diagnosis: A Review. Applied Sciences 2021, 11, doi:10.3390/app112210982.

[8] Abbas, H.; Garberson, F.; Glover, E.; Wall, D.P. Machine learning approach for early detection of autism by combining questionnaire and home video screening. Journal of the American Medical Informatics Association : JAMIA 2018, 25, 1000-1007, doi:10.1093/jamia/ocy039.

[9] Akter, T.; Shahriare Satu, M.; Khan, M.I.; Ali, M.H.; Uddin, S.; Lio, P.; Quinn, J.M.W.; Moni, M.A. Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders. IEEE Access 2019, 7, 166509-166527, doi:10.1109/access.2019.2952609.

[10] M. N. Parikh, H. Li, and L. He, "Enhancing diagnosis of autism with optimized machine learning models and personal characteristic data," Frontiers in Computational Neuroscience, 2019.

[11] Chu, H.-C.; Tsai, W.W.-J.; Liao, M.-J.; Chen, Y.-M. Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning. Soft Computing 2017, 22, 2973-2999, doi:10.1007/s00500-017-2549-z.

[12] F. Thabtah and D. Peebles, "A new machine learning model based on induction of rules for autism detection," Health Informatics Journal, vol. 26, no. 1, pp. 264–286, 2020.

[13] M. H. Al Banna, T. Ghosh, K. A. Taher, M. S. Kaiser, and M. Mahmud, "A monitoring system for patients of autism spectrum disorder using

[14] artificial intelligence," in Proceedings of the International Conference on Brain Informatics, pp. 251–262, Springer, Padua, Italy, July 2020.

[14] Ganesh, K.; Umapathy, S.; Thanaraj Krishnan, P. Deep learning techniques for automated detection of autism spectrum disorder based on thermal imaging. Proc Inst Mech Eng H 2021, 235, 1113-1127, doi:10.1177/09544119211024778.

[15] Deshpande, G., Wang, P., Rangaprakash, D., and Wilamowski, B. (2015). Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data. IEEE Trans. Cybernet. 45, 2668–2679. doi: 10.1109/TCYB.2014.2379621

[16] A. V. Dahiya, C. McDonnell, E. DeLucia, and A. Scarpa, "A systematic review of remote telehealth assessments for early signs of autism spectrum disorder: video and mobile applications," Practice Innovations, vol. 5, no. 2, pp. 150–164, 2020.

[17] Hassouneh, A.; Mutawa, A.M.; Murugappan, M. Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods. Informatics in Medicine Unlocked 2020, 20, doi:10.1016/j.imu.2020.100372.

[18] Gwyn, T.; Roy, K.; Atay, M. Face Recognition Using Popular Deep Net Architectures: A Brief Comparative Study. Future Internet ,2021, 13, 164

[19] N. Zaman, J. Ferdus and A. Sattar, "Autism Spectrum Disorder Detection Using Machine Learning Approach," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-6, doi: 10.1109/ICCCNT51525.2021.9579522.

[20] Elshoky, B. R. G., Younis, E. M. G., Ali, A. A., and Ibrahim, O. A. S., Comparing automated and non-automated machine learning for autism spectrum disorders classification using facial images, ETRI Journal 44 (2022), 613– 623. https://doi.org/10.4218/etrij.2021-009721.

[21] Sato, W.; Sawada, R.; Uono, S.; Yoshimura, S.; Kochiyama, T.; Kubota, Y.; Sa-kihama, M.; Toichi, M. Impaired detection of happy facial expressions in autism. Sci Rep 2017, 7, 13340, doi:10.1038/s41598-017-11900-y.

[22] Lee JH, Lee GW, Bong G, Yoo HJ, Kim HK. Deep-Learning-Based Detection of Infants with Autism Spectrum Disorder Using Auto-Encoder Feature Representation. Sensors. 2020; 20(23):6762. https://doi.org/10.3390/s20236762

[23] Levy, S.; Duda, M.; Haber, N.; Wall, D.P.J.M.a. Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. 2017, 8, 1-17.

[24] Król, M.E.; Król, M.J.N. A novel machine learning analysis of eye-tracking data reveals suboptimal visual information extraction from facial stimuli in individuals with autism. 2019, 129, 397-406.

[25] Center for Autism Research. Social Responsiveness Scale, 2nd Edition (SRS-2). Available online: https://www.carautismroadmap.org/social-responsiveness-scale-/?print=pdf (accessed on 26 October 2020).

[26] Thabtah, F.; Peebles, D. A new machine learning model based on induction of rules for autism detection. Health Informatics J 2020, 26, 264-286, doi:10.1177/1460458218824711.

[27] Vaishali, R., & Sasikala, R. (2018). A machine learning based approach to classify autism with optimum behaviour sets. International Journal of Engineering & Technology, 7(4), 18.

[28] Wu, C.; Liaqat, S.; Helvaci, H.; Cheung, S.S.; Chuah, C.N.; Ozonoff, S.; Young, G. Machine Learning Based Autism Spec-trum Disorder Detection from Videos. Healthcom 2021, 2020, doi:10.1109/healthcom49281.2021.9398924.

[29] Jacob, S.; Wolff, J.J.; Steinbach, M.S.; Doyle, C.B.; Kumar, V.; Elison, J.T.J.T.p. Neurodevelopmental heterogeneity and computational approaches for understanding autism. 2019, 9, 1-12.

[30] Chorianopoulou, A.; Tzinis, E.; Iosif, E.; Papoulidi, A.; Papailiou, C.; Potamianos, A. Engagement detection for children with Autism Spectrum Disorder. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Sig-nal Processing (ICASSP), 5-9 March 2017, 2017; pp. 5055-5059.

[31] Achenie, L.E.; Scarpa, A.; Factor, R.S.; Wang, T.; Robins, D.L.; McCrickard, D.S.J.J.o.d.; JDBP, b.p. A machine learning strategy for autism screening in toddlers. 2019, 40, 369.

[32] Górriz, J.M.; Ramírez, J.; Segovia, F.; Martínez, F.J.; Lai, M.-C.; Lombardo, M.V.; Baron-Cohen, S.; Consortium, M.A.; Suckling, J.J.I.j.o.n.s. A machine learning approach to reveal the neurophenotypes of autisms. 2019, 29, 1850058.

[33] Al Banna, M.H.; Ghosh, T.; Taher, K.A.; Kaiser, M.S.; Mahmud, M. A monitoring system for patients of autism spectrum disorder using artificial intelligence. In Proceedings of the International Conference on Brain Informatics, 2020; pp. 251-262.

[34] Duda, M.; Ma, R.; Haber, N.; Wall, D.P.J.T.P. Use of machine learning for behavioral distinction of autism and ADHD. 2016, 6.

[35] Sen, B.; Borle, N.C.; Greiner, R.; Brown, M.R.J.P.o. A general prediction model for the detection of ADHD and Autism using structural and functional MRI. 2018, 13, e0194856.

[36] Thabtah, F.; Peebles, D.J.H.i.j. A new machine learning model based on induction of rules for autism detection. 2020, 26, 264-286.

[37] Sen, B., Borle, N. C., Greiner, R., & Brown, M. R. (2018). A general prediction model for the detection of ADHD and Au-tism using structural and functional MRI. PloS one, 13(4), e0194856..

[38] F. C. Tamilarasi and J. Shanmugam, "Convolutional Neural Network based Autism Classification," in 2020 5th Interna-tional Conference on Communication and Electronics Systems (ICCES), pp. 1208–1212, IEEE, (2020, June).

[39] van den Bekerom, B. Using machine learning for detection of autism spectrum disorder. In Proceedings of the Proc. 20th Student Conf. IT, 2017; pp. 1-7.

[40] Carette, R.; Elbattah, M.; Cilia, F.; Dequen, G.; Guérin, J.-L.; Bosche, J. Learning to Predict Autism Spectrum Disorder based on the Visual Patterns of Eye-tracking Scanpaths. In Proceedings of the HEALTHINF, 2019; pp. 103-112.

[41] Guimarães, A.J.; Araujo, V.J.S.; Araujo, V.S.; Batista, L.O.; de Campos Souza, P.V. A hybrid model based on fuzzy rules to act on the diagnosed of autism in adults. In Proceedings of the IFIP International Conference on Artificial Intelligence Applications and Innovations, 2019; pp. 401-412.

[42] Sumi, A.I.; Zohora, M.F.; Mahjabeen, M.; Faria, T.J.; Mahmud, M.; Kaiser, M.S. fassert: A fuzzy assistive system for chil-dren with autism using internet of things. In Proceedings of the International Conference on Brain Informatics, 2018; pp. 403-41

[43] Hassan, T.-u.; Abbassi, R.; Jerbi, H.; Mehmood, K.; Tahir, M.F.; Cheema, K.M.; Elavarasan, R.M.; Ali, F.; Khan, I.A. A Novel Algorithm for MPPT of an Isolated PV System Using Push Pull Converter with Fuzzy Logic Controller. Energies 2020, 13, doi:10.3390/en13154007.

[44] Tariq, Q.; Fleming, S.L.; Schwartz, J.N.; Dunlap, K.; Corbin, C.; Washington, P.; Kalantarian, H.; Khan, N.Z.; Darmstadt, G.L.; Wall, D.P.J.J.o.m.I.r. Detecting developmental delay and autism through machine learning models using home vide-os of Bangladeshi children: Development and validation study. 2019, 21, e13822.

[45] Pavithra, D., Jayanthi, A.N., Nidhya, R. and Balamurugan, S. (2022). Autism Screening Tools With Machine Learning and Deep Learning Methods: A Review. In Tele-Healthcare (eds R. Nidhya, M. Kumar and S. Balamurugan). https://doi.org/10.1002/9781119841937.

[46] P. Mazumdar, G. Arru, and F. Battisti,Early detection ofchildren with autism spectrum disorder based on visual explora-tionof images, Signal Process. Image Commun.94(2021), 116184.

[47] Fawaz Waselallah Alsaade, Mohammed Saeed Alzahrani, "Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms", Computational Intelligence and Neuroscience, vol. 2022, Article ID 8709145, 10 pages, 2022. https://doi.org/10.1155/2022/8709145.

[48] Miao, Y., Dong, H., Jaam, J. M. A., & Saddik, A. E. (2019). A deep learning system for recognizing facial expression in real-time. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 15(2), 1-20.

[49] M. I. U. Haque and D. Valles, "A facial expression recognition approach using DCNN for autistic children to identify emotions," in Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 546–551, Vancouver, Canada, 1–3 November 2018.

[50] Jahanara, S., & Padmanabhan, S. (2021). Detecting autism from facial image, International Journal of Advance Research, Ideas and Innovations in Technology, Volume 7, Issue 2 - V7I2-1181.