International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Apr-24)

http://dx.doi.org/10.12785/ijcds/1501109

Deep Learning Model For Autism Diagnosing:

Mazin R. Swadi 1[1] and Muayad S. Croock 2[2]

[1]*Department Control and System Engineering, University of Technology-Iraq, Baghdad, Iraq [Email: cse.22.08@grad.uotechnology.edu.iq]*
[2]*Department Computer Engineering, University of Technology-Iraq, Baghdad, Iraq*

Received 18 Jan. 2024, Revised 15 Mar. 2024, Accepted 18 Mar. 2024, Published 1 Apr. 2024

Abstract: : The brain development and physical appearance of the face are both impacted by the neurologic disorder known as Autism Spectrum Disorder (ASD). Children with ASD exhibit different facial landmarks from children who normally grow with Typical Developing (TD), despite the fact that the disorder is thought to be inherited. When a child's behavioral traits and facial features are examined, the likelihood of an accurate diagnosis is highest. In recent years, deep convolutional neural networks widely used in image processing because they can effectively extract features from images and learn to recognize patterns, making them well-suited for tasks such as object detection, image segmentation, and classification. In this paper, we provide a deep learning CNN-based model for diagnosing autism that divides children into two categories depending on their facial features: possibly healthy and potentially unhealthy. The Tesorflow and Keras libraries are used by the suggested deep learning model to carry out feature extraction and picture classification. The dataset obtained from the Kaggle repository is used to train and evaluate the model. The dataset that was used to test this model consisted of 2,122 that were excluded from the original 2,940 images due to the quality and race. The testing of the proposed model results in an Area Under Curve (AUC) of 99.8% and an accuracy of 97.3%. This model proves its high diagnosis accuracy, ease of use, and fast decision.

Keywords: : Autism Spectrum Disorder, Convolutional neural network, Deep learning, Keras, Tensorflow.

1. Introduction

Asperger's syndrome, childhood disintegrative disorders, and autism are under the group of severe neurodevelopmental brain illnesses known as Autism Spectrum Disorders (ASD). These illnesses can vary greatly in terms of severity and symptomatology, as the term "spectrum" indicates [1]. The International Statistical Classification of Diseases (ISCD) and Related Health Problems (RHP) now include these diseases as Omnipresent Developmental diseases (PDD) under Mental and Behavioral Disorders (MBD) [2]. Early childhood symptoms of ASD often appear in the first few years of life (age of 3-6 years) and may include a lack of eye contact, apathy toward caregivers, and a failure to react when called names [3]. A person's perception of and interactions with others are also impacted by these diseases, they have trouble socializing and communicating with society throughout their first years of life, and they abruptly become withdrawn or hostile. ASD often persists throughout adolescence and maturity, even if it first manifests in infancy [4].

In order to give computer systems visual perception abilities similar to those of humans, computer vision is used. It is an interdisciplinary field that makes it possible for computer systems to correctly process, examine, and understand our visual environment. For instance, computer vision enables machines to recognize important data from photos and video files in the same manner that people do. The goal is to give computers access to this 'natural' visual quality so they can comprehend and evaluate complicated digital systems just like people can and possibly even better. Modern computer vision makes use of machine learning, an aspect of artificial intelligence that is concerned with "teaching" machines to pick up new skills on their own over time. A machine learning system, however, will consider prior experiences and judgments to determine the most appropriate response, in contrast to a system that constantly follows a set of predefined rules or instructions. Furthermore, all of this can be accomplished with little to no human involvement [5].

A simple part or subset of machine learning is deep Learning (DL). It enables machines to examine and grasp digital data without human intervention. To create models that can generate judgments from input data, deep learning often makes use of both statistical principles and algorithms. Deep learning is therefore used in a variety of industries, from supercomputers to sophisticated software engineering [6] ,[7]. In order to recognize the facial expressions of children with autism, Haque and Valles [8] updated the

Facial Expression Recognition 2013 dataset in 2018. They did this by using deep learning techniques. Parikh et al. [9] formulated a technique to extract the characteristics associated with autism using machine learning techniques. DL techniques, in particular Convolutional Neural Networks (CNNs) and recurrent neural networks (RNNs), have been utilized or suggested for the purpose of identifying autism in kids [10] , [11]. Therefore, it is essential to encourage early intervention for ASD children in every household. This means developing an objective, affordable, and simple-to-understand diagnosis or screening solution.

2. Related Work

Throughout recent years, researchers focused on the diagnosing of autism in children using modern techniques of Artificial Intelligent (AI). For example, the author of [12] created the ASDTest app, which was a mobile-based ASD screening tool, to do away with the time-consuming and expensive ASD diagnosis methods. The app logged over 1400 cases, including those involving infants, children, teenagers, and adults. Due to the uneven nature of the toddler data set, which makes it challenging to identify diseases in that category. Additionally, the condition may be identified in children based merely on an image analysis since individuals have a shared pattern of different facial malformations [13]. Depending on this the author of [14] and [15] introduced a DL model can accurately categorize kids as possibly autistic or as healthy with a 94.6% classification rate. 3,014 photos of kids with and without autism were used to train and test the model, with 90% of the data going toward training and 10% going toward testing. The researcher 0f [16] proposed an enhanced framework for autism face recognition based on transfer learning to detect ahead-of-time signs of ASD in kids. Using a variety of machine learning and deep learning classifiers as well as other transfer-learning-based pre-trained models, they gathered facial photos of kids with autism spectrum from the Kaggle data pool. Noted that the highest accuracy was 90.67%. the researcher used the same dataset that was used by [13].

At the other hand, the authors of [17] introduced a study aimed to help families and psychiatrists in diagnosing autism using a simple method, a deep learning-based web application for identifying autism established on face characteristics evaluated in experiments utilizing a CNN with transfer-learning and a flask framework. The classification was done using the pre-trained models MobileNet, Xception, and InceptionV3. The face photos were extracted from a Kaggle dataset pool that is openly accessible and includes over 3,000 photos of kids in a diversity of facial expressions. the accuracy of the classification findings for the proof data: InceptionV3 attained 89%, Xception attained 94%, and MobileNet gained 95% accuracy. The East Asian dataset contains 1122 photos equally split between children with ASD and typical developing children (TD children). Using VGG16, children of the same race were able to generate classification accuracy of 95% and an F1-score of 0.95 [18]. By using the pre-learned models, NASNET-Mobile, Xception, and Visual Geometry Group Network (VGG19) the writer presented a categorization model [19]. The author of [19] used the Kaggle dataset. The results showed that Xception scored 91%, VGG19 scored 80%, and NASNETMobile scored 78%. The Kaggle dataset is used for the purpose of validating and teaching autism spectrum diagnosis models that use transfer learning from face photo data [20]. This study examined the following models: VGG16, EfficientNetB0, EfficientNetB7, Xception, InceptionV3, and MobileNet. The results showed that the accuracy is 88%, 87.7%, 86.1%, 85.6%, 82.6%, and 86.3%, in that order. Using the CNN model, an empirical study is carried out to determine the optimal optimizer and hyperparameter set for improved prediction accuracy [21]. The reformed models, Xception, VGG19, ResNet50V2, MobileNetV2, and EfficientNetB0, showed accuracies of 95%, 86.5%, 94%, 92%, and 85.8%, respectively, after training and validating using the optimal configuration. Classification models employing OpenCV and the VGG16 method of SVM classifier, CNN, and Haar Cascade are the main emphasis of [22]. The accuracy findings using these models were: Convolution Neural Network (VGG16)- 90%, Support Vector Machine- 65%, and Haar Cascade Classifier- 72%. In order to distinguish between a normal kid's face and a kid with ASD, face characteristics are retrieved utilizing a DL-model, which was the author's target in upgrading a new child face image-based autism diagnosis in children [23]. The suggested deep learning architecture based on dense nets successfully classifies the input face photos. The dense block is essential to raising the system's overall accuracy. The classification accuracy of the suggested model was 91.50%.

Underdeveloped countries have proportionally fewer people diagnosed with ASD. Thus, the accessibility to mechanisms used to identification of ASD is important in these countries. The authors used two-dimensional face photographs as input to assess CNN's utility in aiding in the detection of ASD. The outcome shows good achievement for the studied approach. MobileNet and DenseNet201 achieved the best outcome with an average of 90.7% accuracy and standard deviations of 0.68% and 1.64%, respectively. Where DenseNet201 obtained a good accuracy of 93.5% in the best cases [24]. The authors tried to look into the effectiveness of various modalities of facial pictures for diagnosing ASD under DL-based neural networks. The focus of the author's study is on how deep learning models perform when applied to various face image modalities, as well as the difficulties and potential fixes that may arise for each modality. Using the appropriate datasets, the models are trained, and tested, and their assessment metrics are analyzed as part of the technique. Based on face image analysis, deep learning techniques based on neural networks may be able to reliably diagnose ASD. To enhance accuracy on 3D photos, more training on bigger 3D datasets is necessary, since the models perform better on 2D data [25]. Xception, Visual Geometry Group Network (VGG16) the classification job was carried out using the previously trained models. The authors used Dataset obtained from Kaggle for the testing of those models. Outputs of the 3 models of deep learning have been evaluated with the utilization of frequent measures and performance, namely accuracy, sensitivity, and specificity. With a 91% accuracy rate, the Xception model had the highest consequences [26].

In this study, we introduce a DL model that can, with 97.3% accuracy, categorize children as either healthy or perhaps autistic. A dataset from the Kaggle repository is used to train and test the model. Photos, equally split after augmentation between children with and without autism, made up the dataset utilized to evaluate these models. It is divided into 85% for training, while the remaining 15% is used for validation and testing. This classification approach is presented to help with early intervention and diagnosis for children with ASD by bridging high diagnosis accuracy, usability, and quick results.

3. Convolutional Neural Network

Convolutional neural networks (CNN) were developed from multilayer perceptrons (MLPs) , a CNN offers a great level of fabric dimensionality for deep learning. It is hence effective for categorizing visual data [27],[28]. Features were extracted from input training data using it. A collection of filters is included in each convolution layer to aid in the extraction of characteristics. Typically, the intricacy of the characteristics that convolution layers learn grows with CNN model depth. The first convolution layer, on average, catches simple aspects of training data; the final convolution layer, on the other hand, collects complicated features. The convolution filter window's shape and the input's form dictate the convolutional layer's output shape. We employ several methods, such as stride convolutions and padding, that modify the output's size in various instances. Effective methods for modifying the dimensionality of the data include stride and padding. The height and width of the output can be increased via padding. Giving the output the same height and width as the input is a common purpose for this. The stride can lower the output's resolution; for instance, it can limit the output's height and width to only 1/n of the input's (where n is an integer larger than 1). However, Extracting features by taking the convolution of a section of the data specimen under consideration. Both the stride length and the padding value determine how much of the data section the filter visits on each iteration.

Convolutional, pooling and fully linked layers are the three layers of neurons that make up the standard CNN design. It consists of a tiny weighted filter that convolves the one-dimensional input into a matrix-like picture [29]. Figure 1, depicts the architecture of the CNN.

$$h(t) = \sum \mathrm{I(a).K(t\text{-}a)} \quad \infty \quad a = -\infty \qquad (1)$$

where $h(t)$ refer to convolution result, $I(a)$ refer to the input image, and $K(a)$ denote the kernel.

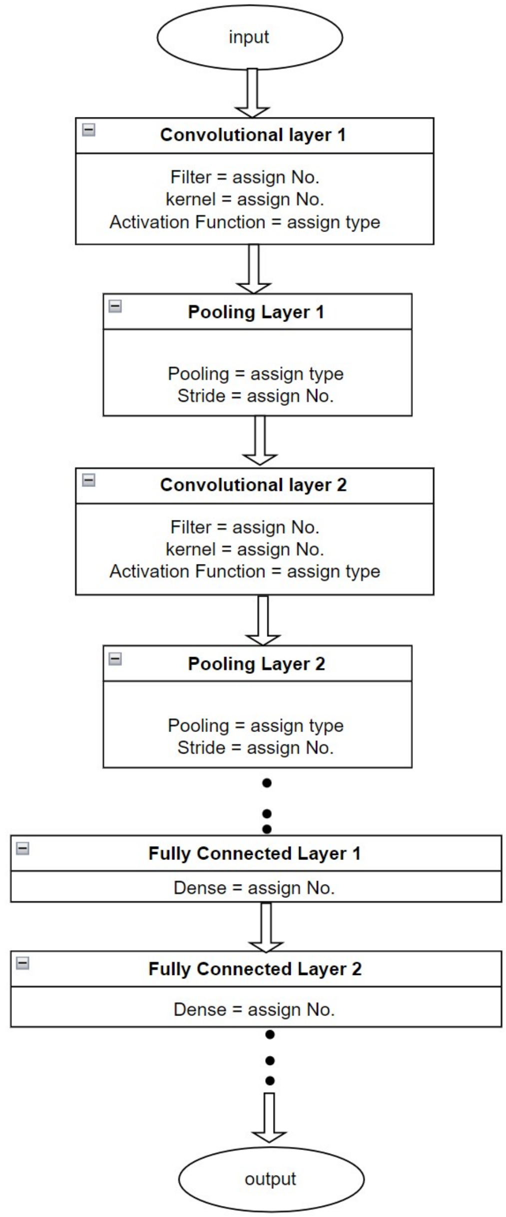

Figure 1. Architecture of Convolutional Neural Network.

A. Convolutional Layer

1) Convolution

Convolution is the name given to an integral transformation process that applies a certain operator to a function [30]. A new portrayal of the resemblance between two characteristics inside a particular window is created by reformatting the original function. It is repeatedly employed in the processing of images to identify features with absence of erasing the spatial correlations among distinct pixels. Using a matching operator, it searches a much bigger pixel set for a certain characteristic. Over the whole original picture, the convolution kernel is displaced by a predetermined

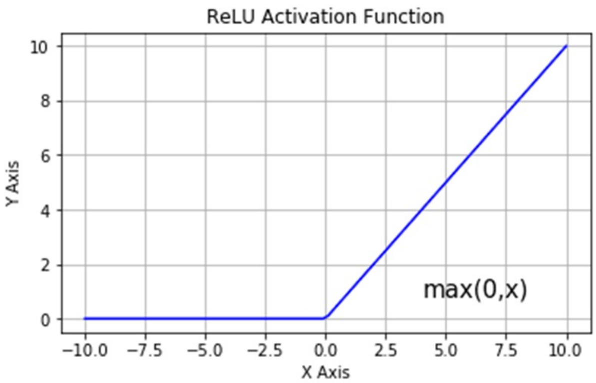

Figure 2. Shows the ReLU function.

stride (the kernel should not exceed the image range).

Convolution operation is a detailed method of feature extraction in image processing that is competent at reducing the size of the data and producing a less unnecessary data collection, often referred as a feature map. As a characteristic recognition [31], each kernel filters out the areas of the original picture where the characteristic is found. In the end, a map with an altitude that shows the distribution of these characteristics is produced images must be processed using some generic feature-describing techniques like convolution since directly studied original data require a lot of preprocessing, such as picture partitioning, and are with difficulty able to dismantle the emblematic content of the pictures. It seeks to automatically extract desired traits and increase their computer-readable visibility. Not only can operating convolutions accurately depict the input's properties, but they also reduce the need for human procedures.

2) Activation Function

There needs to be an activation following the weighted addition and a bias. Enhances CNN achievement by adding Rectified Linear Units (ReLUs) [32]. Figure 2, illustrates the ReLU function. cites shortly after as the reasons for ReLU's exceptional achievement: its hard non-linearity, non-differentiability at zero, and sparse features [33]. ReLUs are after all frequently used to turn on convolutional outputs.

ReLU provided by the "(2)" below.

$$f(x) = max(0, x) \quad (2)$$

B. Pooling Layer

A common intermediary between many convolutional layers is pooling, also known as subsampling. The two pooling techniques types that CNNs utilize most commonly are max-pooling and average-pooling [34].

The benefits of pooling for CNNs are numerous. Pooling explicitly aims to avoid overfitting by reducing the data's dimensionality and concentrating local data into a pooling window [35]. Reducing data dimensionality also facilitates computation efficiency. The stride and pooling size are parameters of the pooling layer [36]. However, cut any inessential features that are captured during convolution layer decreasing the size of the data sample [34]. Aside from that, pooling works on the assumption that neighboring picture pixel values are almost identical [37]. To do pooling, the average, minimum, and maximum of four nearby pixel values are employed. Generally speaking, a 2*2 filter reduces the input's size picture by 50 percent. Figure 3, explains max and average pooling.

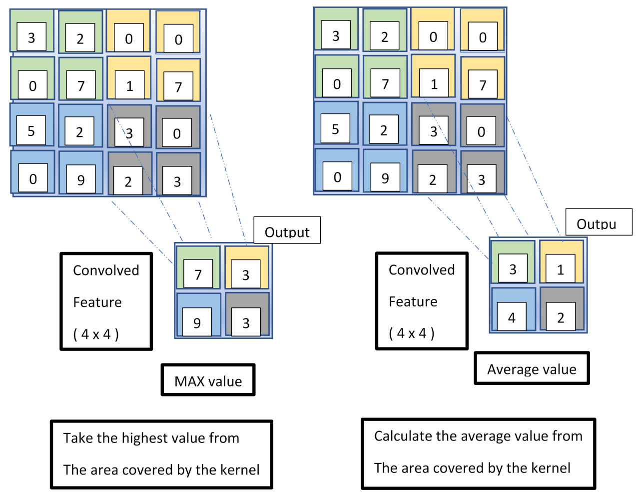

Figure 3. Explains max and average pooling.

C. Flattened Layer

A network's top layer, known as the classification layer, gathers the last complex information and produces a column vector where each row points towards a class. To be more exact, every element in the output vector represents the probability estimate for every class [38], and the total of the elements equals one. Every node in an adjacent layer is connected to every other node in a fully connected layer [39]. The fully connected layer classifies an item using the sigmoid function and a probabilistic value that is either "0" or "1" for binary classification [40]. The sigmoid function can be explained in Figure 4.

4. Proposed Methodology

The proposed autism diagnosing approach for children based on facial images is explained in this section and is divided into subsections to ease the reading flow.

A. Dataset

This dataset comprises about three thousand facial pictures of children with and without ASD, distributed equally. The original dataset consisted of 3014 images total [30], whose evident problems were displayed in [41]. The Kaggle dataset's whole collection of photos was obtained from the internet, as the benefactor expressed difficulty obtaining any

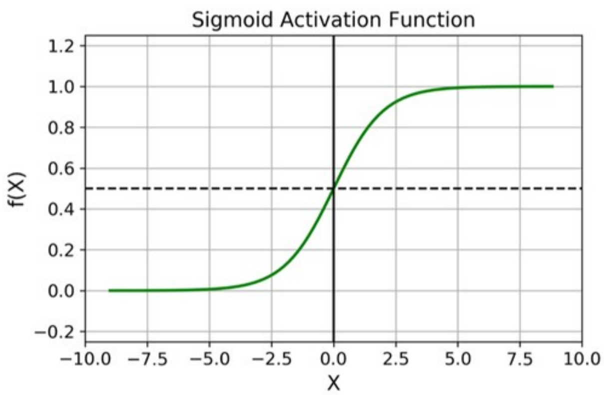

Figure 4. The sigmoid function.

autism spectrum disorder images from reputable organizations or trustworthy resources [41]. We made use of the approved dataset of [42], that comprises 11% of children of race and 89% of white children. This dataset was only used to demonstrate the effects of deep learning development based on facial images. Table I, shows the details of the races in the ASD dataset.

As previously mentioned, the dataset we used in this study was located on Kaggle and contains roughly 3,000 photos of children with and without autism. An effort has been made to exclude the different races with small ratios and some other images with bad quality. Thus, 2000 images are adopted to do argumentation operations with a balanced dataset (an equal number of images for each type). The data separation is done into training, validation, and testing groups. The validation and testing set each contains 15% of the total facial photographs (a combined 10% and 5%, respectively), whereas the training set contains 85% of them.

B. Proposed Deep Learning Model

DL-based CNN models have been widely used in computer vision in recent years because they exhibit strong discriminative ability while maintaining high-performance levels [14]. When it comes to processing and picture recognition, the CNN algorithm excels. This structure is composed of layers that include convolutional, pooling, and totally linked layers. In our proposed model, used three Conv2d layers to uncovers hidden patterns and extracting essential features. These features are like building blocks that help the network make sense of the image's content. The Conv2d layer uses filters, to analyze the image. These filters slide over the image, examining small regions at a time. As they move, they extract relevant information and transform the raw pixels into meaningful representations. Using this process, it can detect edges, shapes, and other important characteristics of the image. We added with each Conv2d one MaxPooling layer. The three MaxPooling layers like taking a step back to observe the view, downsize the spatial dimensions of the feature maps. The function of this layer is to retain the most important information

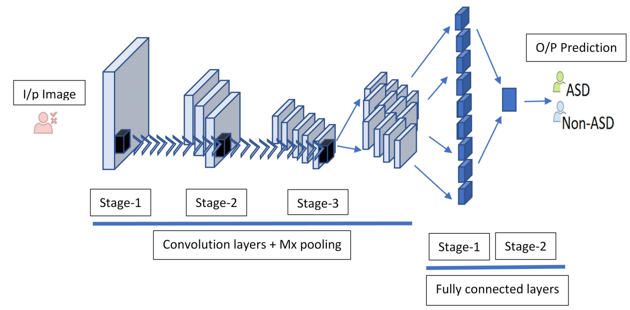

Figure 5. The proposed structure of the autism DL-classifier model.

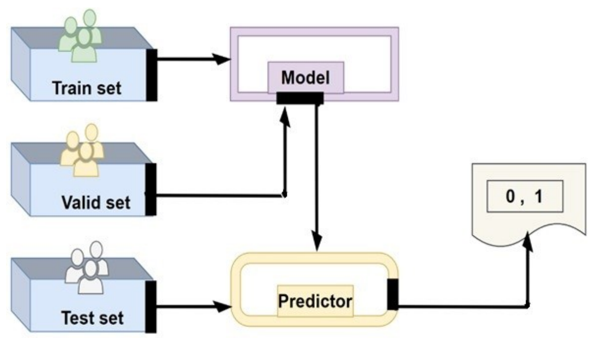

Figure 6. DL model training, validation, and test.

while reducing computational load. The activation function that we used is ReLU( ) , which introduces non-linearity to the model, It ensures that the network can learn complex patterns by allowing it to approximate any function.The adopted model is explaining in Figure 5.

The blank model has to be trained and verified using a certain collection of train and validation photos in order to generate any DL model. The model is next tested using a collection of never-before-seen photos (test photos) to see how well the model performs, as illustrated in Figure 6. Pictures in test datasets are label-free, whereas pictures in training and validation datasets include unique labels indicating which class (autistic or non-autistic) them belong to. With the information it has gained, the trained machine learning model can correctly predict the class of any unknown data (new data) while retaining essential traits that were taken from the training pictures.

C. Proposed Algorithm

Figure 7, shows the workflow of the proposed autism diagnosing approach that uses the DL model for classifying the facial images of patients. The collected data set is cleaned from any noisy ones and preprocessing steps of normalization and resizing are applied. The preprocessed dataset is divided into 85% training set, 10% for validation, and 5% for testing. Different functions are implemented for performing the training of the proposed DL model to produce a well-trained one that is also validated. The trained

TABLE I. Race percentage in the ASD dataset

Dataset	White children	Black children	East_asian children	Other children
Kaggle ASD dataset	89%	4.20%	1.10%	5.70%

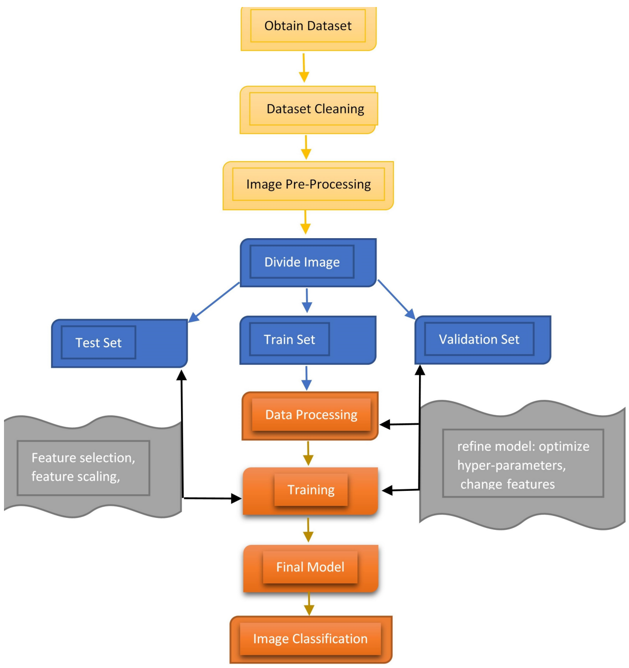

Figure 7. Proposed autism diagnosing approach workflow.

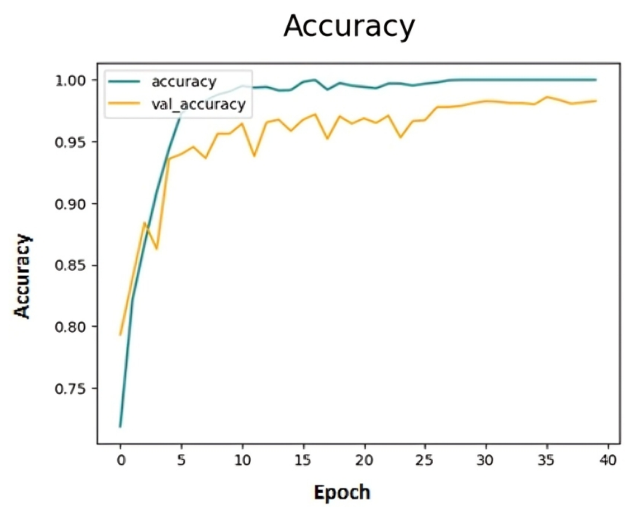

Figure 8. Accuracy and validation accuracy of the proposed model.

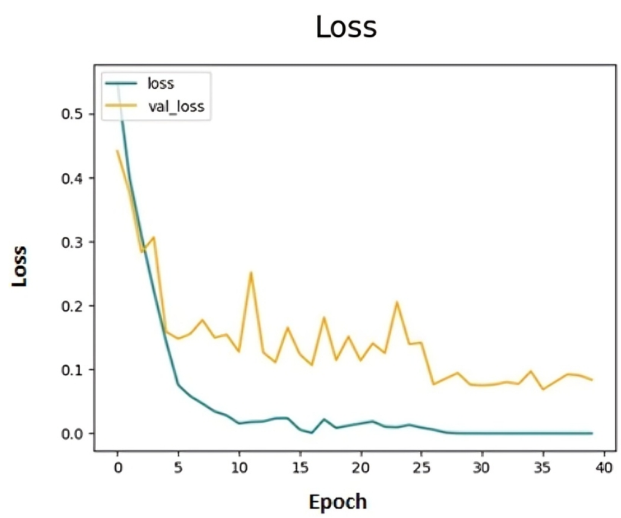

Figure 9. Loss and validation loss of the proposed model.

model is tested over the 5% ratio of the dataset images to ensure the efficiency and accuracy of the designed model. The obtained results of the trained model represent the classes of ASD that show the level of the patient's image.

5. Results And Discussion

As mentioned earlier, the trained model is tested over 5% of the dataset images, thus the computed accuracy and loss curves for our model validation and training are displayed in Figure 8, and Figure 9. These curves illustrate how well our model fits the training and validation information sets.

Ahead of time diagnosis, referral, and treatment face several challenges, especially for children from low-income parents or who take care of them. One of the major obstacles to early diagnosis of ASD is the shortage of sufficient numbers of qualified experts. The heterogeneous symptoms of ASD are another significant challenge. The proposed DL-CNN model achieves a testing accuracy of 97.3%. To guarantee the accuracy of our model's diagnosis, it has been evaluated on a trustworthy dataset comprising pictures of kids who have been diagnosed with autism and those who have not, all of which were determined by an expert using a conventional method for diagnosing ASD that takes DSM-5 criteria into account.

The graph that illustrates a classification model's performance across all classification thresholds is called the Receiver Operating Characteristic (ROC) curve. Two metrics, the False Positive Rate (FPR) and True Positive Rate (TPR), are shown on this curve. Figure 10, shows the Receiver Operating Characteristic curve plots including the TPR vs. FPR of the proposed DL model.

To compute the points in an ROC curve, there's an efficient, sorting-based algorithm that can provide this information for us, called AUC. Think of it as integral calculus:

TABLE II. shows a summary of the accuracy and AUC scores of the proposed approach in comparison with the literature works.

Models	Accuracy %	AUC %
[18] - VGG16-T-L	95	_
[19] - Xception model	91	_
[20] - MobileNet model	88	_
[21] - modified Xception	95	98
[22] - VGG16	90	_
[23] - densenet model	91.5	_
[25] - MobileNet +2 dense layers	94.6	94.48
[30] - CNN-RunPoolmodel	94.6	_
[31] - EVAM-CNN	90.6	90.67
[Proposed model]	97.3	99.8

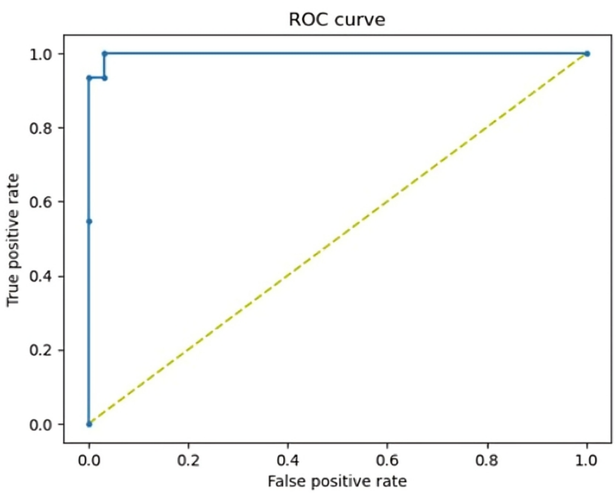

Figure 10. ROC curves of the model.

the Area Under Curve (AUC) estimates the complete two-dimensional area beneath the entire ROC curve from (0,0) to (1,1). 99.8% is the recorded AUC for our proposed model . Table II, shows a summary of accuracy and AUC scores of the proposed approach in comparison with the literature works.

Identifying autism using face images is a sensitive topic, different from traditional image categorization algorithms. The performance of this model is greatly influenced by the uniformity and general caliber of the training set of photos. A wider upper face, a shorter middle face that includes the nose and cheeks, wider eyes, a larger mouth, and the philtrum which was stated in the introduction are some of the common visual indicators of autism. A picture of face show impartial expression is the only way to precisely provide these details since the above characteristics are closely associated with facial emotions. The trial runs revealed that photos with a impartial impression obtained more elevated accuracy and came out better when we excluded different races and performed augmentation. In order to guarantee proper head alignment, the photos need to be shot on a basic, light-colored backdrop with both of the child's ears clearly visible and their eyes open and noticeable that is, without hair hiding the eyes.

The recommended method's simplicity, speed, and accuracy make it advantageous for diagnosing ASD. The obtained results with an AUC of 99.8% and an accuracy of 97.3% outperform the results of the research work in the literature.

6. Conclusion

An autism diagnosing approach for children was proposed based on the DL model classifier that has been trained using facial images. The DL model was trained, validated, and tested using the adopted dataset that includes an efficient number of images for children with TD and ASD. The DL was considered due to recorded sufficient accuracy in diagnosing children with autism. The proposed approach was tested and the obtained results of 99.8% for AUC and 97.3% for accuracy proved the ability to diagnose the patients in accurate rates. These findings demonstrated that a child's static face photos effectively provide the distinctive characteristics of ASD, enabling a rapid and precise ASD screening process. It also could be valid for everyone with a mobile or PC-GUI application that supports the child's parents (or who take care of the child) as the first diagnostic step .

References

[1] C. S. Paula, S. H. Ribeiro, E. Fombonne, and M. T. Mercadante, "Brief report: prevalence of pervasive developmental disorder in brazil: a pilot study," *Journal of autism and developmental disorders*, vol. 41, pp. 1738–1742, 2011.

[2] L. C. Nunes, P. Pinheiro, M. C. D. Pinheiro, M. Pompeu, M. S. Filho, R. Comin-Nunes, and P. G. C. D. Pinheiro, "A hybrid model to guide the consultation of children with autism spectrum disorder," in *Research & Innovation Forum 2019: Technology, Innovation, Education, and their Social Impact 1*. Springer, 2019, pp. 419–431.

[3] L. Kanner *et al.*, "Autistic disturbances of affective contact," *Nervous child*, vol. 2, no. 3, pp. 217–250, 1943.

[4] P. C-DAC, "International Statistical Classification of Diseases and Related Health Problems (ICD)."

[5] M. Du, “Machine vs. human, who makes a better judgment on innovation? take gpt-4 for example,” *Frontiers in Artificial Intelligence*, vol. 6, 2023.

[6] G. Yolcu, I. Oztel, S. Kazan, C. Oz, K. Palaniappan, T. E. Lever, and F. Bunyak, “Deep learning-based facial expression recognition for monitoring neurological disorders,” in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2017, pp. 1652–1657.

[7] ——, “Facial expression recognition for monitoring neurological disorders based on convolutional neural network,” *Multimedia Tools and Applications*, vol. 78, pp. 31 581–31 603, 2019.

[8] M. I. U. Haque and D. Valles, “A facial expression recognition approach using dcnn for autistic children to identify emotions,” in *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2018, pp. 546–551.

[9] 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*, vol. 13, p. 9, 2019.

[10] M. Li, D. Tang, J. Zeng, T. Zhou, H. Zhu, B. Chen, and X. Zou, “An automated assessment framework for atypical prosody and stereotyped idiosyncratic phrases related to autism spectrum disorder,” *Computer Speech & Language*, vol. 56, pp. 80–94, 2019.

[11] F. B. Pokorny, B. Schuller, P. B. Marschik, R. Brueckner, P. Nyström, N. Cummins, S. Bölte, C. Einspieler, and T. Falck-Ytter, “Earlier identification of children with autism spectrum disorder: An automatic vocalisation-based approach,” 2017.

[12] A. S. Mohanty, P. Parida, and K. Patra, “Identification of autism spectrum disorder using deep neural network,” in *Journal of Physics: Conference Series*, vol. 1921, no. 1. IOP Publishing, 2021, p. 012006.

[13] M.-P. Hosseini, M. Beary, A. Hadsell, R. Messersmith, and H. Soltanian-Zadeh, “Deep learning for autism diagnosis and facial analysis in children,” *Frontiers in Computational Neuroscience*, vol. 15, p. 789998, 2022.

[14] T. Vo, T. Nguyen, and C. Le, “Race recognition using deep convolutional neural networks,” *Symmetry*, vol. 10, no. 11, p. 564, 2018.

[15] A. Khalil, S. G. Ahmed, A. M. Khattak, and N. Al-Qirim, “Investigating bias in facial analysis systems: A systematic review,” *IEEE Access*, vol. 8, pp. 130 751–130 761, 2020.

[16] T. Akter, M. H. Ali, M. I. Khan, M. S. Satu, M. J. Uddin, S. A. Alyami, S. Ali, A. Azad, and M. A. Moni, “Improved transfer-learning-based facial recognition framework to detect autistic children at an early stage,” *Brain Sciences*, vol. 11, no. 6, p. 734, 2021.

[17] Z. A. Ahmed, T. H. Aldhyani, M. E. Jadhav, M. Y. Alzahrani, M. E. Alzahrani, M. M. Althobaiti, F. Alassery, A. Alshaflut, N. M. Alzahrani, A. M. Al-Madani *et al.*, “Facial features detection system to identify children with autism spectrum disorder: deep learning models,” *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022.

[18] A. Lu and M. Perkowski, “Deep learning approach for screening autism spectrum disorder in children with facial images and analysis of ethnoracial factors in model development and application,” *Brain Sciences*, vol. 11, no. 11, p. 1446, 2021.

[19] F. W. Alsaade, M. S. Alzahrani *et al.*, “Classification and detection of autism spectrum disorder based on deep learning algorithms,” *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[20] A. Singh, M. Laroia, A. Rawat, and K. Seeja, “Facial feature analysis for autism detection using deep learning,” in *International Conference On Innovative Computing And Communication*. Springer, 2023, pp. 539–551.

[21] M. S. Alam, M. M. Rashid, R. Roy, A. R. Faizabadi, K. D. Gupta, and M. M. Ahsan, “Empirical study of autism spectrum disorder diagnosis using facial images by improved transfer learning approach,” *Bioengineering*, vol. 9, no. 11, p. 710, 2022.

[22] S. Ganesan, J. Senthil *et al.*, “Prediction of autism spectrum disorder by facial recognition using machine learning,” *Webology*, vol. 18, no. Special Issue on Information Retrieval and Web Search, pp. 406–417, 2021.

[23] P. Sandeep and N. S. Kumar, “Autism detection in children with facial cues using dense net deep learning architecture,” *Design Engineering*, vol. 2021, no. 9, pp. 7506–7520, 2021.

[24] J. N. A. Rodrigues, K. R. T. Aires, A. C. B. Soares, V. P. Machado, and R. d. M. S. Veras, “Classification of facial images using deep learning models to support asd identification,” in *2023 XLIX Latin American Computer Conference (CLEI)*. IEEE, 2023, pp. 1–9.

[25] M. S. Alam, Z. Tasneem, S. A. Khan, and M. M. Rashid, “Effect of different modalities of facial images on asd diagnosis using deep learning-based neural network,” *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 32, no. 3, pp. 59–74, 2023.

[26] A. Rashid and S. Shaker, “Autism spectrum disorder detection using face features based on deep neural network,” *Wasit Journal of Computer and Mathematics Sciences*, vol. 2, no. 1, 2023.

[27] J. Naskath, G. Sivakamasundari, and A. A. S. Begum, “A study on different deep learning algorithms used in deep neural nets: Mlp som and dbn,” *Wireless personal communications*, vol. 128, no. 4, pp. 2913–2936, 2023.

[28] Z. Xiang, R. Zhang, and P. Seeling, “Machine learning for object detection,” in *Computing in Communication Networks*. Elsevier, 2020, pp. 325–338.

[29] L. Cuimei, Q. Zhiliang, J. Nan, and W. Jianhua, “Human face detection algorithm via Haar cascade classifier combined with three additional classifiers,” in *2017 13th IEEE International Conference on Electronic Measurement Instruments (ICEMI)*. IEEE, 2017, pp. 483–487.

[30] J.-P. Fanton, “Convolution and deconvolution: two mathematical tools to help performing tests in research and industry,” *International Journal of Metrology and Quality Engineering*, vol. 12, p. 6, 2021.

[31] K. Yakovlev, I. E. I. Bekkouch, A. M. Khan, and A. M. Khattak, “Abstraction-based outlier detection for image data,” in *Intelligent Systems and Applications: Proceedings of the 2020 Intelligent Systems Conference (IntelliSys) Volume 1*. Springer, 2021, pp. 540–552.

[32] R. O. Ogundokun, R. Maskeliunas, S. Misra, and R. Damaševičius, “Improved cnn based on batch normalization and adam optimizer,” in *International Conference on Computational Science and Its Applications*. Springer, 2022, pp. 593–604.

[33] R. Maalej and M. Kherallah, "Improving the dblstm for on-line arabic handwriting recognition," *Multimedia Tools and Applications*, vol. 79, pp. 17 969–17 990, 2020.

[34] H. J. Jie and P. Wanda, "Runpool: A dynamic pooling layer for convolution neural network." *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, pp. 66–76, 2020.

[35] H. Gholamalinezhad and H. Khosravi, "Pooling methods in deep neural networks, a review. arxiv 2020," *arXiv preprint arXiv:2009.07485*.

[36] T. Zebin, P. J. Scully, N. Peek, A. J. Casson, and K. B. Ozanyan, "Design and implementation of a convolutional neural network on an edge computing smartphone for human activity recognition," *IEEE Access*, vol. 7, pp. 133 509–133 520, 2019.

[37] J. I. Forcen, M. Pagola, E. Barrenechea, and H. Bustince, "Learning ordered pooling weights in image classification," *Neurocomputing*, vol. 411, pp. 45–53, 2020.

[38] A. K. Bhadani and A. Sinha, "A facemask detector using machine learning and image processing techniques," *Engineering Science and Technology, an International Journal*, pp. 1–8, 2020.

[39] C.-H. Yeh, M.-H. Lin, P.-C. Chang, and L.-W. Kang, "Enhanced visual attention-guided deep neural networks for image classification," *IEEE Access*, vol. 8, pp. 163 447–163 457, 2020.

[40] S. Bhise, S. Gadekar, A. S. Gaur, S. Bepari, and D. Deepmala Kale, "Breast cancer detection using machine learning techniques," *Int. J. Eng. Res. Technol*, vol. 10, no. 7, pp. 2278–0181, 2021.

[41] B. R. G. Elshoky, E. M. Younis, A. A. Ali, and O. A. S. Ibrahim, "Comparing automated and non-automated machine learning for autism spectrum disorders classification using facial images," *ETRI Journal*, vol. 44, no. 4, pp. 613–623, 2022.

[42] L. K. Gaddala, K. R. Kodepogu, Y. Surekha, M. Tejaswi, K. Ameesha, L. S. Kollapalli, S. K. Kotha, and V. B. Manjeti, "Autism spectrum disorder detection using facial images and deep convolutional neural networks." *Revue d'Intelligence Artificielle*, vol. 37, no. 3, 2023.

Eng. Mazin Radhi Swadi holds a Bachelor of Engineering (B.Eng.) in Control and System Engineering Department , He was an Atomic Energy Commission employee between 2001 till 2004 , He is Ministry of Science and Technology employee between 2004 till now . currently he is student in Master of Computer Engineer email: cse.22.08@grad.uotechnology.edu.iq.

Muayad S. Croock earned his Ph.D. (2012) in Computer Engineering at Newcastle University, School of Electrical, Electronics, and Computer Engineering. He finished his M.Sc. degree in Computer Engineering from the University of Technology-Iraq (2003), and a B.Sc. degree in Computer Engineering from the University of Technology (1998). Dr. Muayad is a staff member at the University of Technology since 2004, where he currently holds a professor in Software Engineering. His research interest includes Software Engineering, Security, and WSN. muayad.s.croock@uotechnology.edu.iq.