ISSN: 2752-3829 Vol. 3 No.2, (December, 2023)

Stochastic Modelling and Computational Sciences

AUTISM DETECTION WITH FACIAL EXPRESSION USING A MODIFIED CONVOLUTIONAL NEURAL NETWORK

C. Karpagam[1] and C. Deepa[2]

[1]Assistant Professor, Department of Computer Science with Data Analytics, Dr.N.G.P.Arts and Science College Research Scholar, Sri Ramakrishna College of Arts and Science Coimbatore, India.

[2]Associate Professor & Head, Department of Computer Science (AI & DS)Sri Ramakrishna College of Arts and Science Coimbatore, India

[1]karpagachandra@gmail.com and [2]deepapkd@gmail.com

ABSTRACT

A complex neurodevelopmental disease, autism spectrum disorder (ASD) is typified by repetitive behaviors and difficulty with social communication. The precise origins of ASD are still unknown, although they most likely stem from a confluence of neurological, environmental, and genetic factors. Considering how complicated and varied ASD may show, doctors frequently struggle to correctly diagnose the disorder at an early age. For those with ASD, early identification and intervention are essential to improve outcomes. There have always been significant obstacles for the medical and health fields in identifying autism spectrum disorder early on. For ASD to receive proper treatment and early intervention, an early diagnosis is essential. Numerous types of literature have demonstrated that children with ASD have varied degrees of difficulties with handwriting tasks. Accordingly, this research advises the establishment facial image dataset of participants with and without ASD for deep learning categorization. Early diagnosis enables prompt access to suitable therapies, improving the development of social and communicative skills. The accuracy of the suggested pre-tuned CNN procedures is 82% higher than that of the current CNN.

Keywords—autism spectrum disorder, autism detection, convolutional neural network, facial images

INTRODUCTION

"Autism spectrum disorders" (ASD) are a group of complicated neurodevelopmental brain diseases that include Asperger's syndrome, autism, and childhood disintegrative disorders. The term "spectrum" denotes the wide range of severity and symptoms associated with various illnesses. These conditions are currently listed under the title Pervasive Developmental Disorders, Mental and Behavioral Disorders in the International Statistical Classification of Diseases and Related Health Problems. ASD symptoms include early indicators such as lack of interest, minimal eye contact, and not responding. These symptoms often show up in the first year of life. Only a few children seem to develop autistic traits during this time period. However, these kids start to exhibit symptoms of autism between the ages of 18 and 24 months. These symptoms include restricted and repetitive behavioral patterns, a narrow range of interests and activities, and weak language skills. Due to these disorders, which also affect how a person perceives and interacts with others, youngsters who fail to engage and communicate with society during their first five years of life may abruptly become violent or introverted.

Although a great deal of research has been done on individuals with autism spectrum disorder (ASD), less is known about how other people perceive facial expressivity in ASD. The circumstances of the study involved generating and reinforcing emotions in a realistic manner. In contrast to expectations, ASD emotions were identified more accurately even though they were perceived as being more intense and less authentic. ASD is an immensely heterogeneous condition. Because of this heterogeneity, determining the exact cause of ASD has proven to be a very challenging task, as individuals with the disorder present with a wide range of symptoms and no single genetic or environmental factor that unites them all.

LITERATURE SURVEY

Takahashi, Y. et al. [1] used a hierarchical predictive processing paradigm to show that facial feature predictive learning is adequate for the self-organization of psychological categories at the upper level of network hierarchy, even in the absence of explicit emotional labels. These findings suggest that a hierarchical RNN with PB34 can be used to discriminate between real and artificial emotions through the use of self-organized higher-level neural representation. Combined with their findings, it appears that the predictive computing paradigm might be applied to understand the collection of abstract data from dynamic facial expressions. Only perceptual input was required for the predictive processing framework to achieve both adaptability and emotion-based grouping; nevertheless, the higher-level representations of neurons showed overlap with the Parametric Bias clusters.

A kind of mental illness called autism spectrum disorder (ASD) can be identified using biomedical images and social media data. Autism spectrum disease (ASD) is a neurological ailment that affects how the brain develops and, as a result, the external appearance of the face. The facial landmarks of children diagnosed with ASD vary greatly from those of children who are developing normally (TD). What distinguishes the proposed research as innovative is the development of a facial recognition system for social media autism spectrum disorder diagnosis. Deep learning algorithms were used by Alsaade, F.W. et al. [2] to identify these landmarks; nevertheless, these methods require precise technology in order to extract and create the exact patterns of the face traits. Their study aids communities and psychiatrists in the empirical diagnosis of autism based on facial attributes through the use of a straightforward web application developed on a deep learning system, i.e., a convolutional neural network employing transfer learning and the flask framework.

Zhang, K., et al. [3] investigated the variations in eye tracking characteristics and facial expression processing between kids with ASD and kids without developmental disabilities. They demonstrated that, in comparison to TD children, children with ASD had a considerably worse rate of facial expression recognition across a variety of experimental settings (spatial frequency, orientation). It may be deduced that kids with ASD have less capacity to comprehend facial expressions than kids with TD. The participants in this study were two groups of children with similar verbal IQ scores since developmental delays are common in children diagnosed with ASD. Even though the ASD children were older than the TD children, they nonetheless performed noticeably worse on the experimental tasks, further supporting the theory that the ASD children had difficulties perceiving facial expressions.

Garcia-Garcia et al. [4] presented an innovative approach that utilized visible user interfaces carried out with NFC technology as well as face-based emotion detection applications to help kids with ASD recognize and express emotions with the aim to promote their suggestion based on the findings of research on specialized therapies for children with ASD. The majority of their use is dependent on cutting-edge interaction methods, such as automatic emotion identification using the device's built-in camera and tangible user interfaces (TUIs). In addition, NFC (near field communication) technology was additionally employed to give kids easy-to-use interfaces for using the many game-related products. Using TUIs, which provide a recognizable and simple approach, allows kids to interact with the game in a pleasant and natural way. Professionals claim that the creator created a serious game that removes distracting elements so that children's focus may be directed toward learning how to identify emotions in various contexts and how to communicate those emotions. The software itself has been tested in a real-world context with children diagnosed with ASD and their mental health providers. The results were excellent, and the experts approved of the system as a helpful teaching aid for topics pertaining to emotions.

To provide an overview of the current state of knowledge about multichannel emotion perception in the ASD population and to suggest goals for future research, Zhang, M., et al. [5] reviewed 24 pertinent literatures. They highlighted the variety of methodologies used in this understudied area and suggested that variables related to participants and tasks might account for variations in results. An integrated viewpoint on emotion, language, and cognition is necessary to understand language's role in emotion perception and development. People with ASD who are native speakers of tonal languages need special treatment since pitch variation is the primary emotional signal in spoken language. This gives more information and understanding the domain-specific cross-language components of the identified deficiencies in linguistic and emotional pitch processing.

A novel method for determining if a child, aged 3 to 10, has autism spectrum disorder was developed by Derbali, M. et al. [6]. A child can be drawn into a deep and engaging experience by playing video games. The last ten years have seen a huge expansion in the availability and personalization of games for kids due to the rise of the gaming industry. Youngsters can display a wide range of emotions and facial expressions while they play video games. The identification of autism may be aided by these facial expressions. Videos of kids having fun can reveal a lot about behavioral tendencies, especially those associated with autism. For this purpose, the author employed a set of 2,536 facial pictures of children who were autistic and typically developing. To examine the 92.3% correct prediction results generated by the CNN model using deep learning, the accuracy and loss functions are given.

The issues with homeschooling during the COVID-19 pandemic were noted by Zhang, R., et al. [7], who also stressed the need of developing useful visual teaching materials for kids with ASD. They found that when children with ASD watched video lectures, their motor learning increased when the faces were pleasant instead of neutral. Moreover, better accuracy of motor execution was linked to enhanced visual attention. These results show that facial expressions made by the teacher can improve learning in kids with ASD. The design of educational materials for kids with ASD as well as educators and therapists looking to enhance learning outcomes will be greatly impacted by these findings.

ViTASD, the first ViT-based baseline for diagnosing juvenile autism, was introduced by Cao X et al. [8]. The author showed that a ViT, which offers state-of-the-art performance in terms of accuracy and AUROC while generating attention maps consistent with identifiable ASD features, may be used to organize juvenile ASD as a face image classification problem. The study inspire more investigation into the potential applications of explainable ViTs in other face analysis applications, as well as provide valuable insights to the signal processing and biomedical imaging community for ASD research.

According to Rashid, A. et al. [9], people with autism have difficulty understanding who they are, their needs, feelings, and ideas, as well as their surroundings. An autistic person experiences the world as a horror movie, finding certain sights, sounds, and even food tastes and smells to be frightful and even painful. Consequently, when something unexpected occurs in their environment, they fear that no one else will understand. A precise diagnosis of autism is necessary to save the lives of numerous youngsters. The development of artificial intelligence (AI)-based intelligence systems can help in autism early diagnosis. This study used three state-of-the-art deep learning models—Xception, VGG16, and others—to try and diagnose autism. Upon presenting empirical data for various models, the Xception model had the greatest accuracy rate of 91%.

Awaji, B. et al. [10] suggests artificial intelligence strategies, in particular face feature extraction utilizing machine learning algorithms, show potential in assisting the early detection of ASD. Artificial intelligence algorithms recognize the symptoms of ASD by analyzing small cues and facial expressions. This study created a variety of hybrid systems by utilizing convolutional neural network (CNN) properties to assess face feature photos for an ASD dataset. The first method used pre-trained MobileNet, ResNet101, and VGG16 models. The second method used a hybrid of XGBoost and RF techniques with CNN models (VGG16, ResNet101, and MobileNet). The third method involves diagnosing ASD using XGBoost and an RF based on characteristics from the VGG-16-ResNet101, ResNet101-MobileNet, and VGG16-MobileNet models. With features from the VGG16-MobileNet models, the hybrid RF approach performed exceptionally well, obtaining an AUC of 99.25% as well as 98.8%, 98.9%, 99.1%, and 98.8% for accuracy, precision, sensitivity, and specificity, respectively

This research found a significant gap that hinders the best possible results when diagnosing ASD. This disparity emerged from the inability to properly extract CNN models' characteristics, reduce their dimensions, and integrate them in a way that maximizes their potential. This work carefully extracted high-efficiency feature vectors from many CNN models in order to close this scientific gap in a novel way. The data from two CNN models is combined in these feature vectors.The produced feature vectors are then used to a classification framework that makes use of advanced algorithms, such as a CNN that has been updated. Identifying and monitoring facial features, as well as differentiating between individuals with ADS and TD, are among the categorization tasks.

DATASET DECRIPTION

There are 1470 files in the Autism Spectrum Disorder in Children with Facial Images Expression collection, and another 1470 files are non-autistic. There are 2940 face photos in this dataset, split equally between kids with TD and kids with ASD. There were 3014 images in the original dataset. All of the photos in the Kaggle dataset were found online There are around 89% white children and 11% children of color in this sample.

Usage Notes - The data set is shared publicly on the kaggle repository. The files can be downloaded directly from the URL [11].

METHODOLOGY

Detecting autism using facial expressions with a modified Convolutional Neural Network (CNN) is lamentedted. It entails training a CNN to identify patterns in facial expressions that might be suggestive of autistic spectrum disease. The different stages of proposed framework are explained below and the same is illustrated in Fig. 1.

Data Collection

Data Collection collects a varied variety of facial expressions from people with and without ASD. It is assured that the dataset is properly tagged to identify the presence or absence of ASD. .

Data Preprocessing

Cropped and scaled photos of faces are used to ensure consistency. Standardize image pixel values to a range that is accepted, such [0, 1] or [-1, 1]. The size and diversity of the dataset are enhanced by transformations including rotation, scaling, and flipping.

Data Splitting

Make training, validation, and test sets out of the dataset. 70% teaching, 15% validation, and 15% testing is a common split.

Model Architecture

CNN architecture is tailored for the identification of facial expressions. Start with a CNN that has already been trained (such as VGG, ResNet, or Inception) and modify it as needed. Modify or add layers to the network in order to adapt it to identify facial expressions associated with autism.

Training

The training dataset should be used to train the modified CNN. Use an optimization method (like Adam) and the proper loss function (like cross-entropy) for training. Use strategies like batch normalization and dropout to enhance network performance and prevent overfitting.

Validations

Use the validation set to track the model's performance during training. To optimize the model, adjust hyperparameters (for example, the learning rate).

Testing and Evaluation

Use the testing dataset to evaluate the trained model's accuracy, precision, recall, F1-score, and other pertinent metrics.

Pseudocode for Modified CNN

```
# Import python packages

import numpy as np

import tensorflow as tf
```

```
from tensorflow.keras import layers, models
# Data preprocessing, Loading, Splitting and model building
#Sequential model is chosen to manipulate one input and one output at each layer.
model = models.Sequential()
# Convolutional layer  with pixel size 32
model.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=(img_height, img_width, img_channels)))
# MaxPooling is applied to retaing only the essential information in the image
model.add(layers.MaxPooling2D((2, 2)))
# steps repeated with pixel size 64 and 128
model.add(layers.Flatten())
# activation relu to resolve the vanishing gradient issue
model.add(layers.Dense(256, activation='relu'))
# Compile the model
model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
# Training and Evaluation of  the model
model.fit(train_images, train_labels, epochs=num_epochs, validation_data=(val_images, val_labels))
test_loss, test_acc = model.evaluate(test_images, test_labels)
print(f'Test accuracy: {test_acc}')
```

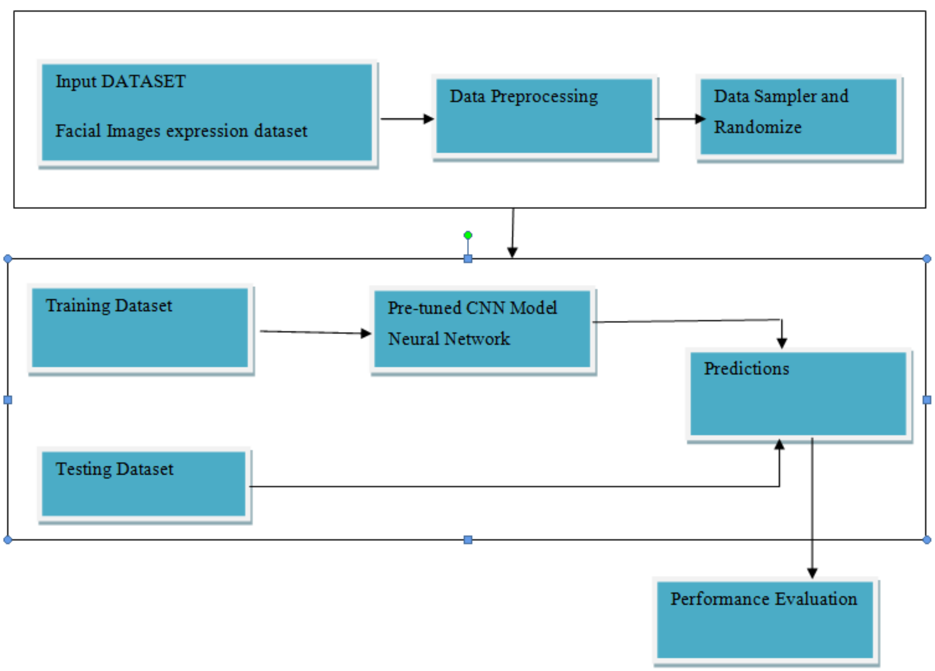

Fig. 1. Proposed Framework

RESULTS AND DISCUSSIONS

ASD is a developmental disease that has an impact on behaviour and communication. It is typified by issues with communication, social contact, and repeated habits. The proposed framework depicted in Fig. 2 has proved an improved accuracy of 82%. The graphic compares the accuracy of present CNN with the proposed pre-tuned CNN approach. It demonstrates that the suggested pre-tuned CNN achieves higher accuracy than existing approach. Further investigation on training and validation accuracy is performed.

While assessing a machine learning model's performance and spotting any issues like underfitting and overfitting, it is essential to interpret the accuracy and loss of training and validation. Fig.3. and Fig.4. depicts the model's ability to learn is indicated by the training graph (shown in green), and its ability to generalize successfully on new data is indicated by the validation graph (shown in red). Accuracy rises over time, and a steady decrease in error loss is observed. This implies that the model has not yet overfitted and is still learning from the training set.

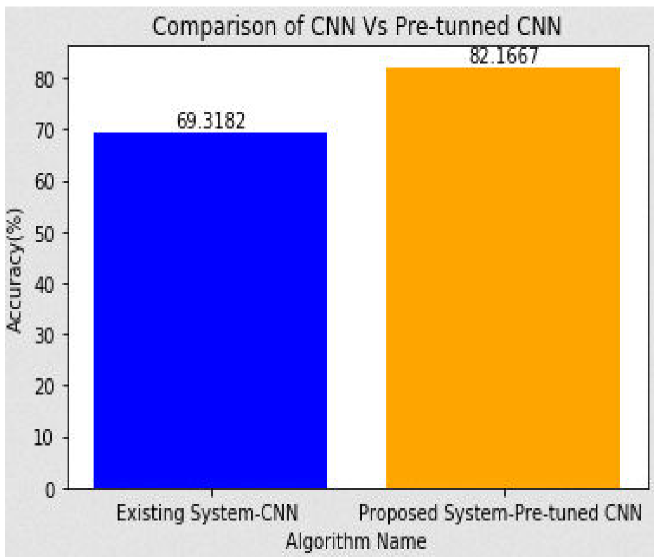

Fig. 2. Accuracy Comparison of CNN and Pre-tunned CNN

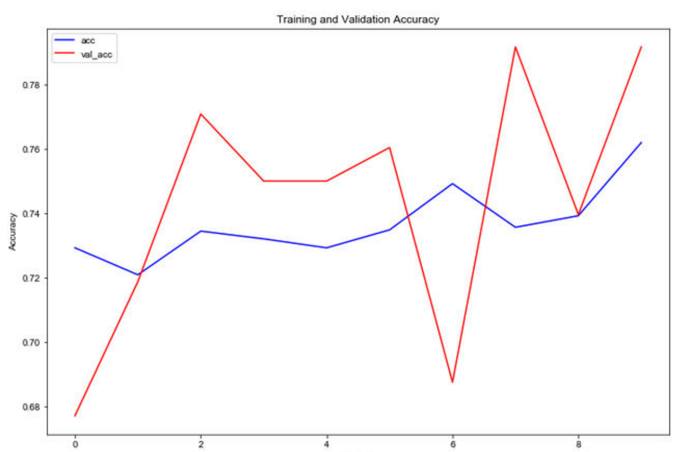

Fig. 3. Plot against Training and Validation image Accuracy

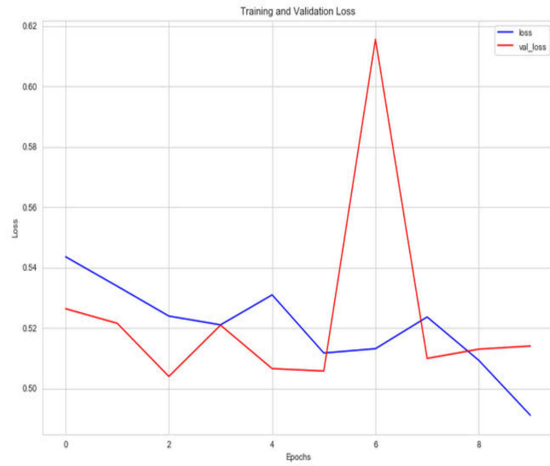

Fig. 4. Training and Validation Loss

The receiver operating characteristic curve, or ROC curve, is a graph that displays a classification model's performance over all categorization at different threshold levels. Fig. 5. displays two parameters as a curve, such as the rates of false positives and true positives.

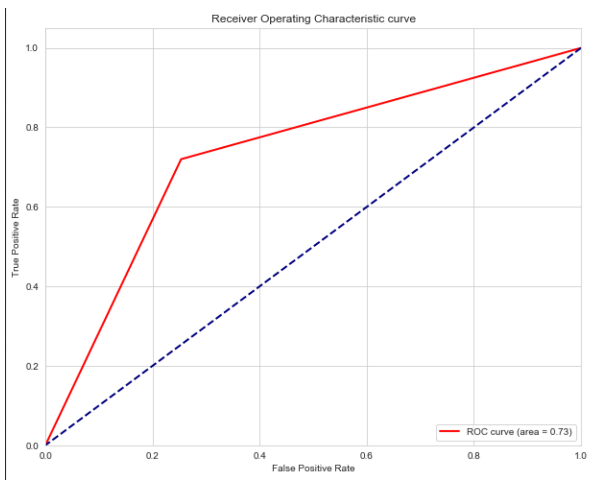

Fig. 5. ROC of Proposed System

Even though it is more difficult to detect ASD in younger children, it is commonly identified in early childhood. The manual early diagnosis of autism spectrum disorder has several limitations. The degree of symptoms associated with autism spectrum illness varies widely. This makes it difficult to diagnose ASD since no single test or collection of criteria can provide a firm diagnosis. Some of the limitations related to manual early diagnosis of ASD can be addressed using advanced learning algorithms. This paper suggests a method for evaluating facial feature pictures in order to identify ASD in its early stages.

CONCLUSION

The knowledge of autism has evolved throughout time. When autism was originally discovered sixty years ago, it was just an undetected developmental delay that was occasionally mistaken with mental retardation. It is currently recognized as a severe neurological disorder in its own right, a major public health concern, and a widely researched topic. According to the study's concept, a dataset suitable for conducting its experiments was obtained; moreover, based on an examination of cutting-edge research on ASD diagnosis, this dataset is the first of its kind. The dataset, which included both patients with ASD and those without, was exposed to image processing techniques. A transfer learning network was then used to train the photos that had been processed. The suggested model was individually tested and trained before the estimated predictions were added together to get the final prediction. The recommended model's favourable findings, including an accuracy of 82% compared to the present approach, urge more research and testing in this field.

REFERENCES

[1] Takahashi, Y., Murata, S., Idei, H., Tomita, H. and Yamashita, Y., 2021. Neural network modeling of altered facial expression recognition in autism spectrum disorders based on predictive processing framework. Scientific reports, 11(1), p.14684.

[2] Alsaade, F.W. and Alzahrani, M.S., 2022. Classification and detection of autism spectrum disorder based on deep learning algorithms. Computational Intelligence and Neuroscience, 2022.

[3] Zhang, K., Yuan, Y., Chen, J., Wang, G., Chen, Q. and Luo, M., 2022. Eye tracking research on the influence of spatial frequency and inversion effect on facial expression processing in children with autism spectrum disorder. Brain Sciences, 12(2), p.283.

[4] Garcia-Garcia, J.M., Penichet, V.M., Lozano, M.D. and Fernando, A., 2022. Using emotion recognition technologies to teach children with autism spectrum disorder how to identify and express emotions. Universal Access in the Information Society, 21(4), pp.809-825.

[5] Zhang, M., Chen, Y., Lin, Y., Ding, H. and Zhang, Y., 2022. Multichannel perception of emotion in speech, voice, facial expression, and gesture in individuals with autism: A scoping review. Journal of Speech, Language, and Hearing Research, 65(4), pp.1435-1449.

[6] Derbali, M., Jarrah, M. and Randhawa, P., 2023. Autism spectrum disorder detection: Video games based facial expression diagnosis using deep learning. International Journal of Advanced Computer Science and Applications, 14(1).

[7] Zhang, R., Cheng, G. and Wu, L., 2023. Influence of instructor's facial expressions in video lectures on motor learning in children with autism spectrum disorder. Education and Information Technologies, pp.1-14.

[8] Cao, X., Ye, W., Sizikova, E., Bai, X., Coffee, M., Zeng, H. and Cao, J., 2023, June. Vitasd: Robust Vision Transformer Baselines for Autism Spectrum Disorder Facial Diagnosis. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.

[9] Rashid, A. and Shaker, S., 2023. Autism spectrum Disorder detection Using Face Features based on Deep Neural network. Wasit Journal of Computer and Mathematics Sciences, 2(1).

[10] Awaji, B., Senan, E.M., Olayah, F., Alshari, E.A., Alsulami, M., Abosaq, H.A., Alqahtani, J. and Janrao, P., 2023. Hybrid Techniques of Facial Feature Image Analysis for Early Detection of Autism Spectrum Disorder Based on Combined CNN Features. Diagnostics, 13(18), p.2948.

[11] https://www.kaggle.com/code/iravad/autism-prediction-cnn-vgg16-transfer-learning/input