Educational Administration: Theory and Practice
2024, 30(1), 7494-7501
ISSN: 2148-2403
https://kuey.net/
Research Article

Autism Detection Using Pre-Trained Models Using VGG And CNN

Dr. S. Sahaya Tamil Selvi*

*Principal, St. Joseph's College for Women Kangeyam Road Tirupur Tamilnadu India

Citation: Dr. S. Sahaya Tamil Selvi, (2024). Autism Detection Using Pre-Trained Models Using VGG And CNN, *Educational Administration: Theory and Practice*, 30(1) 7494-7501
Doi: 10.53555/kuey.v30i1.10651

ARTICLE INFO | **ABSTRACT**

GPU technology efficiently trains and tests deep learning architectures through parallelizable mathematical procedures [1]. GPUs consist of multiple cores that accelerate complex computational operations. The number of processing units (cores) that can operate independently is crucial in determining the level of parallelization possible. There is a substantial difference between GPUs with thousands of cores and CPUs with only four or eight cores. More cores available means an increase in the amount of parallelization possible [2]. When GPUs have a lot of cores, CPU cores operate at a higher frequency. For mathematical procedures in Neural Networks, GPUs are vital [4]. Neural Networks have proven to be very successful in solving several real-life problems.
Ensuring high accuracy while working with diseases, such as Autism Spectrum Disorder (ASD), is crucial. ASD brings developmental disability to the brain and is also associated with genetic conditions [6]. Some causes of ASD are known, while others are still unknown. Patients with ASD exhibit different behavior, communication, interaction, and learning styles compared to ordinary people[5]. While some patients can live and work like ordinary people without any support, many patients require assistance from others to live their life. Some patients may have advanced conversation skills, while others may be nonverbal. Usually, ASD starts to develop before the age of three, but some children may show symptoms early in the 12-month period[4]. In the first 12 or 24 months, they gain knowledge and skills, but later stop learning, making it difficult to communicate with peers and adults, make new friends, and understand complex concepts [8]. People with ASD are also more likely to experience serious issues such as anxiety, depression, and attention-deficit/hyperactivity disorder compared to those without ASD. Early detection of ASD is crucial as it significantly decreases symptoms and has a high chance of improving the quality of life. Medical tests such as blood tests and symptom checking's are the most common ways of detecting ASD. The symptoms are usually checked by parents or teachers, and healthcare providers[10].

Keywords: Autism Spectrum Disorder (ASD), Deep Learning, Convolutional Neural Network (CNN), VGG16 / VGG19 Pre-trained Models, Transfer Learning, XGBoost, Neuroimaging (MRI, fMRI), Autism Detection System (ADS)

2. LITERATURE REVIEW

The rapid advancement of deep learning (DL) has revolutionized medical diagnosis, particularly in neurodevelopmental disorders such as Autism Spectrum Disorder (ASD). The computational demands of training deep neural networks (DNNs) necessitate high-performance hardware, with Graphics Processing Units (GPUs) playing a pivotal role due to their parallel processing capabilities[11]. This literature review examines the intersection of GPU-accelerated deep learning and ASD detection, highlighting the importance of early diagnosis and the technological advancements enabling efficient and accurate detection[12].
Deep learning relies on computationally intensive matrix operations, which benefit significantly from parallel processing. GPUs, with thousands of cores, outperform traditional Central Processing Units (CPUs) in handling large-scale neural network training (Shawahna et al., 2019)[6]. Studies have demonstrated that GPU-accelerated frameworks such as TensorFlow and PyTorch drastically reduce training times while improving model accuracy (Abadi et al., 2016; Paszke et al., 2019). The ability to process multiple computations

simultaneously makes GPUs indispensable in medical imaging and behavioral data analysis for ASD detection[7].
ASD is a neurodevelopmental disorder characterized by impaired social interaction, communication difficulties, and repetitive behaviors (Lord et al., 2018). Early diagnosis is crucial, as interventions before age three significantly improve developmental outcomes (Zwaigenbaum et al., 2015)[5]. However, traditional diagnostic methods—relying on behavioral assessments by clinicians, parents, or teachers—are subjective and time-consuming (Volkmar et al., 2014). There is a growing need for automated, data-driven approaches to enhance diagnostic accuracy and efficiency.
Recent studies have explored deep learning models for ASD detection using various data modalities:
Functional and structural Magnetic Resonance Imaging (fMRI, sMRI) have been widely used in ASD diagnosis. GPU-accelerated convolutional neural networks (CNNs) have successfully classified ASD patients from controls with high accuracy (Heinsfeld et al., 2018). Recurrent Neural Networks (RNNs) have also been applied to fMRI time-series data, leveraging GPU parallelism for efficient training (Li et al., 2021).
Eye-tracking studies reveal distinct gaze patterns in ASD children. Deep learning models trained on GPU clusters have achieved high classification accuracy using eye-tracking datasets (Wang et al., 2020). Similarly, automated analysis of video-recorded behaviors using deep learning has shown promise in early ASD screening (Duda et al., 2016).
ASD has strong genetic associations, and deep learning models analyzing genomic data have identified risk markers (Zhou et al., 2019). GPU-based frameworks enable multi-modal data fusion (e.g., combining imaging, genetic, and behavioral data), improving diagnostic precision (Eslami et al., 2019).

Proposed Methodolgy:

XGBOOST-VGG16: At the beginning of our process, we effectively extracted the desired features using the traditional VGG16 model. We then used a popular classifier called XGBoost to classify the features[6]. Extreme Gradient Boosting (XGBoost) is a distributed, scalable gradient-boosted decision tree (GBDT) machine learning framework. It offers parallel tree boosting for regression, classification, and ranking issues. In supervised machine learning, a model is trained using algorithms to discover patterns in a dataset of features and labels, and the model is then used to predict the labels on the features of a new dataset[4].

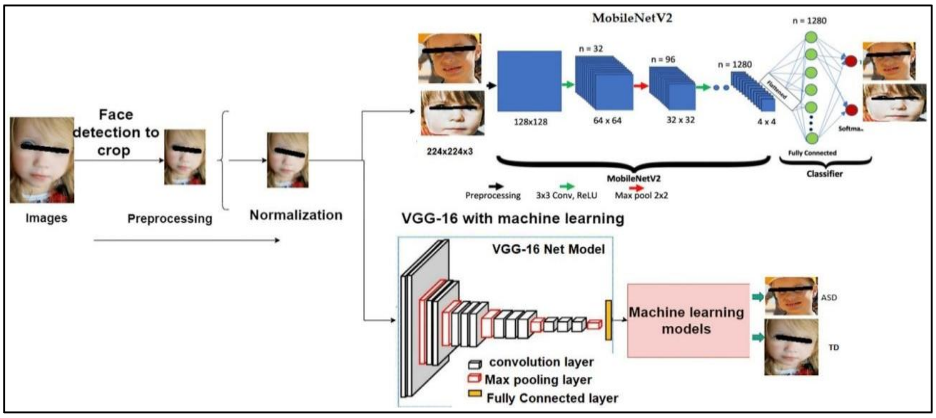

Fig 1.1 Framework of the proposed recognition system.

Convolutional neural networks (CNNs) have established a strong reputation in the field of image data processing, producing superior results compared to traditional methods. However, they require enormous data for the training phase, sometimes socially referred to as data-hungry algorithms especially training from scratch[7]. However, transfer learning (TL) has solved this problemto a very large extent, where a pre-trained model is retrained to perform a specific task with fewer data samples. Pre- trained deep learning models provide substantial benefits in artificial intelligence and machine learning. These models save practitioners time and computing resources by providing powerful starting points for numerous tasks, leveraging knowledge from prolonged training on big and diverse datasets[9]. The ability to adapt pre-trained models to specific tasks using limited labelled data is a major feature of transfer learning. It reduces the requirement for enormous datasets. Pre-trained models are flexible because they can parse many kinds of data for useful hierarchical characteristics[8]. Additionally, these models are useful for practitioners in several areas due totheir resistance to overfitting and rapid convergence during fine-tuning.

Finally, each model was unfrozen for further training for a total of 8 epochs. This approach allows the models to learn from the data more thoroughly and efficiently and allows for the fine-tuning of the models using the selective learning rate[8]. The block architecture of the proposed methodology is illustrated in Figure.

Figure 1.2 Flowchart of the Proposed Methodology

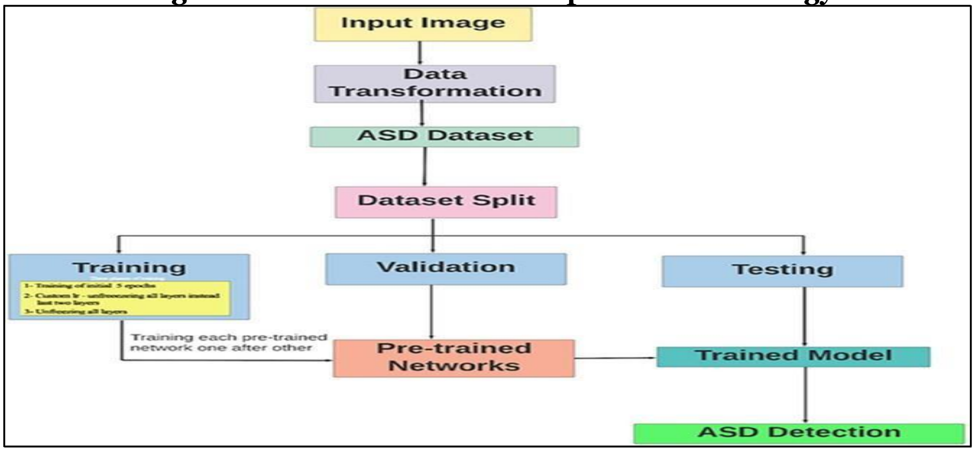

Image analysis can classify autistic people by identifying distinguishing visual traits. Here, facial expressions are important since they reveal emotions and social interactions. Eye gaze patterns are important because autistic people may have unusual eye contact[3]. Body language, including gestures and postural clues, helps explain autism's peculiar behaviours. A holistic view of an individual's relationships comes from contextual awareness, which includes social and environmental aspects. Recognition of repetitive actions, a prevalent autistic feature, and social interaction dynamics aid classification. Machine learning and deep learning can extract and evaluate these properties from photos, revealing autism's complex spectrum. This research must be sensitive to autism heterogeneity and follow ethical norms for visual data classification[4].

The goal of this research is to train and identify the best model based on the given images and subsequently propose the most effective model for the early detection of autism spectrum disorder (ASD)[5].

2.1 DATABASE DESIGN:

The database design for the "Autism Detection using pre-trained models using CNN (VGG19)" project encompasses organizing and structuring data to efficiently store and manage information essential for autism detection[5]. Thedatabase comprises several key entities, including 'Patient', 'Medical History', 'Neuroimaging Data', and 'Diagnosis', each with their respective attributes capturing patient demographics, medical records, neuroimaging information, and diagnostic results[6].

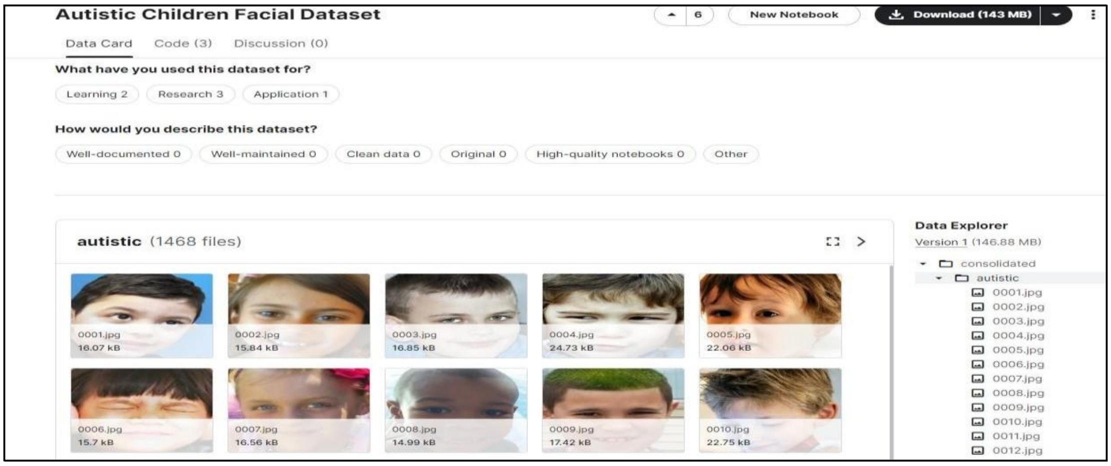

Figure 2.1 The data set of Autism chidren expression

About Dataset:

- Dataset contain two different files of directories each segmented under consolidated, train, train & valid images.
- All the three directories contain 1000+ images segmented under Autistic& Non autistic.
- Presence of the directories helps the project in detecting Autism under pre trained models.

2.2 Input Design

The input design for the "Autism Detection using pre-trained models using CNN (VGG19)" project involves defining the parameters and data inputs required to effectively utilize the pre-trained model for autism detection. This includes gathering neuroimaging data such as MRI or fMRI scans, which serveas the primary input for the VGG19 model[7]. Additionally, demographic information about the patients, including age, gender, and any relevant medical history, may be collected to provide contextual information for the diagnosis. The input design ensures that the data is formatted appropriately and standardized to facilitate seamless integration with the pre-trained CNN model. Preprocessing techniques may also be applied to the input data to handle noise, normalize features, and enhance the model's performance. Overall, a well-defined input design ensures that the necessary data inputs are availableand properly prepared to maximize the accuracy and reliability of autism detection using the VGG19 model.

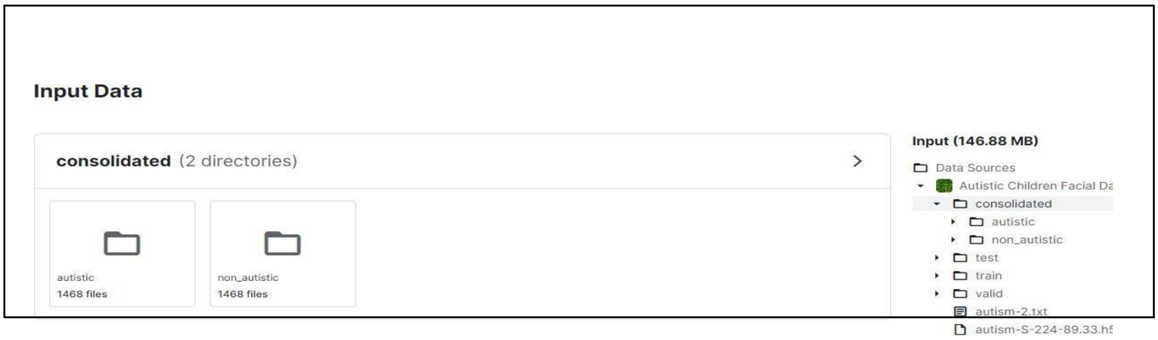

Fig 2.2 Input Dataset

2.3 Output Design:

The output design for the "Autism Detection using pre-trained models using CNN (VGG19)" project focuses on presenting the results of the autism detection process in a clear and interpretable manner[4]. Upon processing the input neuroimaging data through the pre-trained VGG19 model, the output design aims to provide diagnostic information indicating the likelihood or presence of autism spectrum disorder (ASD). This output may include probability scores or confidence levels assigned by the model to indicate the certainty of the diagnosis[5]. Additionally, the output design may involve visual representations such as heatmaps highlighting regions of interest in the neuroimaging data that contribute to the model's decision. The output design ensures that the diagnostic results are presented in a format that is easily understandable and actionable for healthcare professionals, facilitating informed clinical decision-making and patient management strategies.

Fig 2.3 Output Images

3.1 SYSTEM TESTING AND IMPLEMENTATION

System Testing:

In recent years, deep convolutional neural networks (CNNs) have been used extensively in computer vision, showing powerful discriminative capabilities while maintaining high performance levels[2]. As deep learning networks have established themselves as a promising model for facial recognition and CNNs have been used as the deep learning tool in almost all facial recognition systems, our research focused on a deep-learning-based solution[4]. In a recent comparative study of popular deep-learning architectures for facial recognition, Gwyn reported that VGG16/VGG19 showed the highest accuracy levels of image recognition; as such, we further focused our study using VGG16-based deep learning[3].

Transfer learning is a machine learning method where a model developed for one task is reused as the starting point for a model on a second task, and is a popular approach in deep learning. Visual Geometry Group (VGG) is a CNN model proposed by Simonyan and Zisserman that achieved 92.7% accuracy, placing it in the top-five in test accuracy on ImageNet, a dataset of over 14 million images belonging to 1000 classes[2]. VGGFace is a facial image dataset that contains 2.6 million images of different people contributed by the Visual Geometry Group. Tensorflow is an end-to- end open-source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources. Keras is the high-level API of Tensorflow.Keras-VGGFace is an Oxford VGGFace implementation using Keras Functional Framework v2+. A VGG16 model pre-trained with VGGFace is provided in Keras-VGGFace. Thus, VGG16 was adopted forthis research as the pre-trained model for transfer learning[4].
System implementation:

The dataset has been taken from Kaggle which is publicly accessible and we also customized the dataset by adding several images of toddlers. The necessary libraries are included such as Tensor Flow, Keras, NumPy, Pandas, matplotlib, and sklearn. Google Drive is mounted to access the dataset. Then, prepare the data by normalizing the pixel values and resizing the photos to a uniform size of 224x224 pixels[5]. Image labeling is done as 'Autistic' and 'Non Autistic' byassigning 1 and 0 respectively. The ratio of the training set to the validation set is 80:20. Hyperparameters 'epoch' and 'batch size' is set to 23 and 64 respectively. The VGG19 pre-trained model is loaded with an input image and the top layer is set to 'False' means the last layer of fully connected to the original model is not included. The output layer is flattened to 1 dimension via GlobalMaxPooling and the output of the last layer of the pre-trained model is retrieved[6].
A model is trained when it is shown how to categorise or predict something using existing preliminary data. To begin with, the divided original dataset is used to create the train dataset, which will be used to train the model. We will create a validation generator to screen out the data and utilise the verified data to ensure performance quality after preparing and creating the train data[7]. The validation data are the data used to assess any model metrics and the loss at the conclusion of each epoch. These data won't be used to train the model. After using the validation dataset to judge the performance of our model, we will train the data for a particular number of epochs and batch size. We evaluated using a batch size of 12 with 10 epochs to train our model. We trained two models, in which one is based on VGG16 and the other is based on VGG19. Utilising the validation dataset, we will compute the evaluation metrics to assess the performance of our model at the end of each epoch[3].
The results of the deep learning models are presented in this section, and the significant results of developing system are declared.

Experimental Setup: The experiment was executed on different libraries of python and hardware devices for developing an intelligent autism detection system (ADS).

Evaluation metrics

Evaluation metrics are used to assess how effectively a statistical or machine learning model is functioning. The machine learning models or algorithms employed in each project must be evaluated. This study uses different types of performance evaluation metrics for the three pretrained models such as accuracy, sensitivity, and specificity and confusion matrix.
A confusion matrix is a type of measure of classification performance that represents a table of the true and false values of the testing results. A model may be assessed using a variety of different assessment metrics. These consist of measurements such as classification accuracy, logarithmic loss, confusion matrix, and others. When we talk about accuracy, we usually mean classification accuracy, which is the proportion of successful predictions to all input data. Logarithmic loss, commonly referred to as log loss, is used to penalise incorrect classifications. A confusion matrix creates a matrix for us and provides a summary of the model's overall performance. In this work, we have used evaluation metrics like Accuracy, Sensitivity, Specificity, and Precision to evaluate the performance of our model.

```
20/20 [==============================] - 80s 3s/step - loss: 1.7290 - accuracy: 0.4779 - val_loss: 0.6782 - val_accuracy: 0.5400
Epoch 2/100
20/20 [==============================] - 37s 2s/step - loss: 0.6865 - accuracy: 0.5524 - val_loss: 0.6665 - val_accuracy: 0.6100
Epoch 3/100
20/20 [==============================] - 37s 2s/step - loss: 0.6851 - accuracy: 0.5615 - val_loss: 0.6507 - val_accuracy: 0.5900
Epoch 4/100
20/20 [==============================] - 37s 2s/step - loss: 0.6503 - accuracy: 0.6100 - val_loss: 0.6279 - val_accuracy: 0.6600
Epoch 5/100
20/20 [==============================] - 37s 2s/step - loss: 0.6148 - accuracy: 0.6609 - val_loss: 0.6053 - val_accuracy: 0.6200
Epoch 6/100
20/20 [==============================] - 37s 2s/step - loss: 0.5872 - accuracy: 0.6960 - val_loss: 0.5675 - val_accuracy: 0.6800
Epoch 7/100
20/20 [==============================] - 37s 2s/step - loss: 0.5527 - accuracy: 0.7236 - val_loss: 0.5801 - val_accuracy: 0.7300
Epoch 8/100
20/20 [==============================] - 38s 2s/step - loss: 0.5566 - accuracy: 0.7086 - val_loss: 0.5462 - val_accuracy: 0.7200
Epoch 9/100
20/20 [==============================] - 37s 2s/step - loss: 0.5356 - accuracy: 0.7252 - val_loss: 0.6339 - val_accuracy: 0.6800
Epoch 10/100
20/20 [==============================] - 37s 2s/step - loss: 0.5267 - accuracy: 0.7358 - val_loss: 0.5123 - val_accuracy: 0.7400
Epoch 11/100
```

Figure 3.1 Confustion Matrix Results

Accuracy: Accuracy is a common evaluation statistic for classification problems. It displays the percentage of all forecasts that result
in accurate predictions.

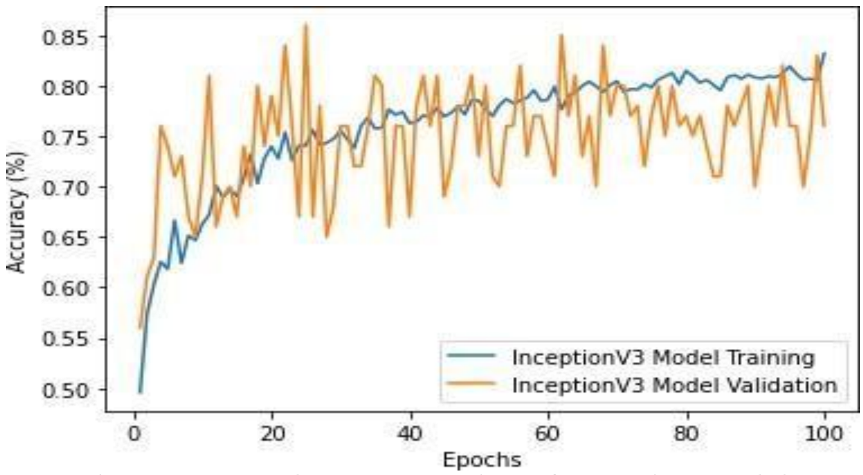

Figure 3.2 Graph representation of Inception model

Sensitivity: Sensitivity provides the true positive rate (TPR), which is the ratio of genuine positives to all positives.
Specificity: Specificity is the proportion of genuine negatives that the model correctly identifies. This means that more real negatives, often known as false positives since they were previously believed to be positive, would be recorded.

$$\text{Accuracy} = \frac{TP + TN}{FP + FN + TP + TN} \times 100\%,$$

$$\text{Specificity} = \frac{TN}{TN + FN} \times 100\%,$$

$$\text{Sensitivity} = \frac{TP}{TP + FP} \times 100\%,$$

where TP is the True Positive, FP is the False Positive, TN is the True Negative, and FN is the False Negative. Specificity is the capacity of the model to correctly identify the normal children, and sensitivity is the capacity of themodel to correctly identify autistic children[4].

Precision: Precision (either rightly or wrongly) is the ratio of correctly classified positive samples (True Positive) to all positively classified samples.

Precision =True Positive/True Positive + False Positive

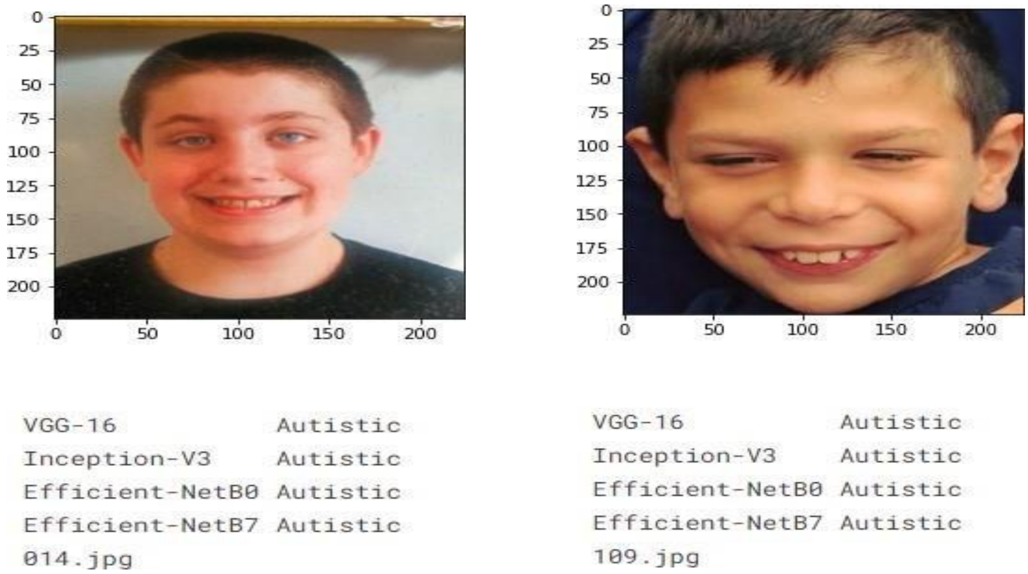

Fig 3.5 Final Output

CONCLUSION AND FUTURE WORK

In this study, we suggested a comprehensive computerised approach for face image-based autism identification. This study used a Convolutional Neural Network with transfer learning to create a deep learning-based online interfacefor identifying autism. The CNN architecture has the right models to extract the facial landmarks, which can classify faces into autistic and non-autistic types[5]. This is done by producing sequences of facial characteristics and calculating the distance 805 between facial features. Parents and physicians will find it easier to recognise ASD in children with the help of this sort of software. Children with autism may gain from having a correct diagnosis of their condition by picking an appropriate therapy route[3]. With VGG16 as its pre-trained model, the model has a precision of 90% and an accuracy of 84%. This study addresses a gap in the literature by screening young children for ASD using facial photos. Clinical observations that children with ASD and normally developing children have distinct facial characteristics are supported by the study's findings[6]

Future work:

Future work may focus on developing the technique into a straightforward mobile application that would let families take a photo with their phone and quickly get a very accurate screening result[4]. More research should be done to combine image- and videobased approaches into one system that enables the detection of both behavioural and facial phenotypic issues in ASD in order to further minimise misclassifications[5].

In the future, we plan to develop a mobile app for identifying children with autism. The users can capture a set of facial images of children and input them into the app. This app determines the probability of autism for each image, and then provides the average probability of autism for all images[4]. In addition, we plan to perform statistical tests and sensitivity analysis to test the robustness and stability of our algorithms in the future work. We also plan to collaborate with kindergartens, rehabilitation centers for children with autism, and child psychology clinics in hospitals. Kindergarten teachers who observe a child with abnormal behavior can use this app to detect whether this child has autism. If the test result is autism, the child's parents can be reminded to send this child to the children's psychological clinic at the hospital for professional examination. By collaborating with the Rehabilitation Center for Children with Autism and the Child Psychology Clinic, we will obtain a lot of facial images of children with autism, and then train the models of the proposed method to improve its performance[1].

BIBLIOGRAPHY

References

[1] Wang, S., Adolphs, R. (2017). Social saliency. In Computational and Cognitive Neuroscience of Vision, pp. 171-193.

[2] Baio, J., et al. (2018). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2014. MMWR Surveillance Summaries, 67(6), 1-23. https://doi.org/10.15585/mmwr.ss6904a1

[3] Zwaigenbaum, L., Penner, M. (2018). Autism spectrum disorder: Advances in diagnosis and evaluation. BMJ, 361: k1674. https://doi.org/10.1136/bmj.k1674

[4] Sidey-Gibbons, J., Sidey-Gibbons, C. (2019). Machine learning in medicine: A practical introduction. BMC Medical Research Methodology, 19(1): 64.

[5] Akter, T., et al. (2021). Improved transfer-learning-based facial recognition framework to detect autistic children at an early stage. Brain Sciences, 11(6): 734. https://doi.org/10.3390/brainsci11060734

[6] Rajaram, M. (n.d.). Concerns with 'Detect Autism' dataset. Kaggle. https://www.kaggle. com/code/melissarajaram/concernswith-detect-autism-datas et, accessed on August 6, 2021.

[7] Musser, M. (2020). Detecting Autism spectrum disorder in children with computer vision. https://towardsdatascience.com/detecting-autismspectrum-disorder-in-children- with- computer-vision8abd7fc9b40a, accessed on August 1, 2021.

[8] Vo, T., Nguyen, T., Le, T. (2018). Race recognition using deep convolutionalneural networks. Symmetry, 10(11): 564. https://doi.org/10.3390/sym10110564

[9] Chaudhuri, A. (2020). Deep learning models for face recognition: A comparative analysis. In Deep Biometrics, pp. 1-17. Springer.

[10] Gwyn, T., Roy, K., Atay, M. (2021). Face recognition using popular deep net architectures: A brief comparative study. Future Internet, 13(6): 144.

[11] Xiao, J., Huang, J., Li, H. (2020). Application of a novel and improved VGG-19 network in the detection of workers wearing masks. Journal of Physics: Conference Series, 1518: 012041. https://doi.org/10.1088/1742- 6596/1518/1/012041

[12] Simonyan, K., Zisserman, A. (2015). Very deep convolutional networks for large- scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

[13] Atluri, S., Bandi, V.S.S., Chunchu, S., Kagitha, P.R.P., Donepudi, S. (2022). Segmentation of potholes from road images. 2022 4th International Conference on Smart Systems\ and Inventive Technology (ICSSIT). http://dx.doi.org/10.1109/ICSSIT53264.2022.9716468

[14] Nitesh, B., Madhuri, A., Sai Manogna, B., Naga Jogendra Babu, K., Ishwarya, N., Mohan Trivendra, G. (2022). Brain tumor segmentation using U-Net based on Inception. 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS). Coimbatore, India.

[15] Alsaade, F.W., Alzahrani, M.S. (2022). Classification and detection of autism spectrum disorder based on deep learning algorithms. ComputationalIntelligence and Neuroscience, 2022: 8709145. https://doi.org/ 10.1155/2022/870914

[16] X. Jiang and Y.-F. Chen, "Facial image processing," in Applied Pattern Recognition, H. Bunke, A. Kandel, and M. Last, Eds., Springer, Berlin/Heidelberg, Germany, pp. 29–48, 2008.

[17] M. Duda, R. Ma, N. Haber, and D. P. Wall, "Use of machine learning for behavioral distinction of autism and ADHD," Translational Psychiatry, vol. 6, no. 2, Article ID e732, 2016.

[18]G. Deshpande, L. E. Libero, K. R. Sreenivasan, H. D. Deshpande, and R. K. Kana, "Identification of neural connectivity signatures of autism using machine learning," Frontiers in Human Neuroscience, vol. 7, p. 670, 2013.

[18] 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.

[19] F. -abtah 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.