Deep CNN Architectures for Autism Spectrum Disorder Detection in Neuroimaging Data

[1]Mrs.V.Deepa, [2]P.Chandrasekhar, [3]S.Aishwarya, [4]P. Aarthi
[1,2,3,4] Department of Artificial Intelligence and Data Science Rajalakshmi Institute of Technology Chennai, India

Abstract: Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by challenges in social interaction, communication, and repetitive behaviours. Early and accurate detection of ASD is crucial for timely intervention and support. This paper explores the application of deep Convolutional Neural Networks (CNNs) for the detection of ASD, leveraging the power of TensorFlow, a widely-used deep learning framework. The proposed system utilizes a comprehensive dataset comprising neuroimaging data, behavioural assessments, and genetic information to train a CNN model capable of identifying ASD-related patterns with high accuracy. The model's architecture is designed to automatically extract and learn intricate features from the input data, enhancing its diagnostic precision. Preliminary results demonstrate that the deep CNN approach, implemented with TensorFlow, achieves significant improvements in the classification of ASD cases compared to traditional methods. This study underscores the potential of deep learning techniques in advancing the early diagnosis of ASD, offering a promising tool for healthcare professionals and researchers in the field.

Keywords: Autism Spectrum Disorder (ASD),Deep Learning, Convolutional Neural Networks (CNNs),TensorFlow,Early Diagnosis, Neuroimaging.

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder with a complex set of social, communication, and behavioral difficulties that vary in degree. Recent developments in neuroimaging methods, including functional and structural magnetic resonance imaging (fMRI and sMRI), have made it possible to investigate the neural mechanisms underlying ASD. Nevertheless, the vast amount and complexity of neuroimaging data require the application of advanced computational methods for efficient analysis.

Deep Convolutional Neural Networks (CNNs), with their high capacity to learn automatically hierarchical features from unprocessed data, have proved to be a useful tool in this situation. Using deep CNN architectures for the detection of ASD in neuroimaging data has the potential to enhance diagnostic performance by detecting faint patterns and deviations in brain structure and function that could signal the disorder. This strategy not only improves our knowledge of the neural substrate of ASD but also opens the door to more accurate and earlier diagnosis, eventually leading to improved therapeutic outcomes.

II. LITERATURE SURVEY

Mst Shapna Akter,Hossain Shahriar,Alfredo Cuzzocrea. (2022)proposed a system which investigates the effectiveness of neural network models in the classification of Autism Spectrum Disorder (ASD) from facial images. It contrasts the performance of different deep learning architectures, such as VGG16, ResNet50, DenseNet, InceptionV3, Xception, MobileNet, and XGBOOST-VGG16, on CPU and GPU platforms. Because of the heavy mathematical computations involved, GPUs, with their parallel processing nature, far surpass CPUs in speed and accuracy. Performance is assessed in the study through accuracy, F1 score, precision, recall, and execution time, proving the effectiveness of GPU-based processing to improve neural network performance in ASD classification. [1].

V.LokeshRaju,B.Aishwarya,V.Praveen,Kumar,K.AnilKumar,P.Nithish Reddy,M.Jayanthi Rao (2022) investigates several deep learning models to determine the best model using Accuracy, Precision, Recall, and F1 Score. The models are run using a dataset from Kaggle with 21 features and 704 instances in Python on Spyder IDE with fundamental libraries including pandas, numpy, matplotlib, sklearn, and seaborn. This is done with the aim of increasing the accuracy and efficiency of ASD detection compared to traditional diagnosis methods. [2].

Kanimozhi A,Dhanasri A(2024) prediction of Autism Spectrum Disorder (ASD) from facial images using the ResNet50 deep convolutional neural network is presented. The methodology includes facial image feature extraction and a classifier for the identification of the possibility of ASD. Through a public database, the model has a 90% accuracy. The process presents an excellent, inexpensive, and non-invasive technique for early ASD detection. [3].
Amrutha S M , K R Sumana (2021) proposed a Machine Learning method to forecast outcomes of ASD through the examination of historical data and testing multiple models such as Logistic Regression, Naïve Bayes, Decision Tree, and K-Nearest Neighbour. The forecasted outcomes will be employed for future testing, and additional improvements will be added to enhance precision and efficiency. [4].

Morched Derbali , Mutasem Jarrah , and Princy Randhawa (2023) proposed a new approach to identify Autism Spectrum Disorder (ASD) among children in the age group 3-10 through facial expressions recorded when playing video games. Through an examination of behaviour patterns in gaming platforms, the system utilizes deep learning and Convolutional Neural Networks (CNNs) to detect autistic behaviours. A collection of 2,536 facial images of autistic and normal children was employed, and it has a prediction accuracy of 92.3%. [5].

III. METHODOLOGY

A. System Overview

This project aims to create and optimize deep convolutional neural networks (CNNs) to improve the detection and classification of Autism Spectrum Disorder (ASD) based on neuroimaging data. The primary objective is to enhance diagnostic accuracy and gain insights into the neurobiological markers of ASD by examining MRI and fMRI scans. Deep learning methods are used to detect patterns in neuroimaging data that are characteristic of ASD. The workflow includes neuroimaging dataset collection and preprocessing, CNN architecture design and training, and model performance evaluation to achieve high sensitivity and specificity. The ultimate goal is to transfer these models into useful, easy-to-use tools that can help clinicians make accurate and early diagnoses, ultimately leading to timely and more effective interventions for autism individuals. The work adopts a rigorous methodology: identifying the problem, collecting and formatting the imaging data, testing and comparing different algorithms, and finding detection results for the purpose of confirming the applicability of the proposed system. [6].

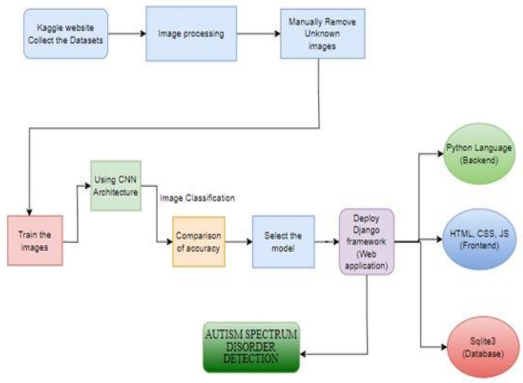

Fig. 1. Architecture diagram.

B. Data Collection

The neuroimaging data comprising MRI and fMRI scans were drawn from publicly accessible datasets providing brain imaging data of subjects diagnosed with Autism Spectrum Disorder (ASD) and typically developing subjects (non-ASD). Care was taken to label each image to identify whether it was an ASD or non-ASD subject, thus allowing the application of supervised learning methods for model building. For training and testing the performance of the developed deep learning model, the dataset was split into two sets: 80% of the data was utilized for training the model so that it could learn significant patterns and features, and the remaining 20% was held out for testing, to confirm the accuracy and generalizability of the model on new data. [7].

C. Data Preprocessing

After raw data was collected, it underwent preprocessing aimed at quality improvements and model performance enhancement. To preprocess the neuroimaging data for training deep learning models, some preprocessing operations were performed. First, normalization was done by rescaling the pixel values of all images into a range from 0 to 1, which improves the efficiency and stability of neural network training. Second, all images were resized to a uniform size—such as 224x224 pixels when using AlexNet—to be compatible with the model's input layer. Random horizontal and vertical flips, rotation, and zooms were used as data augmentation methods to increase the dataset artificially and enhance the model's power to generalize to unseen new data. The aforementioned preprocessing steps were managed effectively utilizing data generators within Keras that load, transform, and augment the images when training [8].

D. Convolutional Neural Network(CNN):

A Convolutional Neural Network (CNN) is an Artificial Neural Network developed mainly for image processing, classification, and analysis of spatially correlated data. It has one or more convolutional layers that learn features automatically from input images. In Keras, there are three ways to create CNN models. The Sequential model is a simple approach where layers are stacked linearly, which is ideal for minor tasks. Functional API is a more general and common technique that accommodates complicated architectures and multiple output or input possibilities. Model subclassing is the most flexible, where users can create their own models from scratch, usually applied for professional or research-based projects.

E. Manual Architecture:

A custom deep learning architecture was implemented based on a sequential model in Keras to measure baseline performance. The model is comprised of various important layers, beginning with Conv2D layers to extract spatial features from the input images. ReLU activation functions were used after every convolutional layer to add non-linearity and allow the model to learn more intricate patterns. MaxPooling2D layers were used to downsample the feature maps, reducing computational complexity and spatial dimensions while preserving significant features. Dropout layers were added to avoid overfitting, dropping a subset of the neurons at random during training. The output was flattened after feature extraction and sent through Dense layers for classification. The model was also optimized with Adam optimizer, which adjusts the learning rate, and trained with loss function Categorical Crossentropy as it is perfect for multi-class classification tasks. This architecture served as a base model to gauge the early performance before diving deeper into more advanced methods. [10].

F. Alex Net:

AlexNet is a deep convolutional neural network that transformed computer vision and deep learning. Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton created it, and it took the 2012 ImageNet Large Scale Visual Recognition Challenge. The network works on color images with dimensions of 224x224 pixels and three color channels (RGB). It has five convolutional layers and max-pooling layers with kernels of different sizes (11x11, 5x5, and 3x3) and strides to extract features. Max-pooling compresses spatial dimensions and retains significant features. Local response normalization is used after the first two convolutional layers to improve generalization. Following the convolution and pooling layers, three fully connected layers are utilized, with the first two comprising 4096 neurons and the third comprising 1000 neurons, which aligns with the number of ImageNet classes. The output layer employs a softmax activation function in order to generate class probabilities. Dropout is applied to avoid overfitting and ReLU in all layers other than the output. Furthermore, AlexNet employed data augmentation, GPU computing, and the use of multiple GPUs while training, which also helped make it successful. AlexNet showcased the capability of deep CNNs for image classification and opened doors for future developments in computer vision. [11].

G. LeNet*:*

LeNet is one of the pioneering CNN architectures constructed by Yann LeCun around the early 1990s for recognition of handwritten digits. It includes multiple layers: an input layer which takes in grayscale images (most commonly 32x32 pixels), and two convolutional layers which draw out features from the images via convolutional filters. A subsampling (pooling) layer, following every convolutional layer, shrinks the spatial sizes of the feature maps. These characteristics are then flattened and fed into fully connected layers, which perform classification. The output layer generates the predicted class probabilities, with 10 output nodes for digit classes (0-9). Although LeNet is less complex than current CNN architectures, its design concepts, including the use of convolutional and pooling layers, set the stage for more sophisticated networks and contributed to the revival of interest in neural networks[12].

H. Django:

Django is a high-level Python web framework that enables rapid development of secure and maintainable websites. Built by experienced developers, Django takes care of much of the hassle of web development, so you can focus on writing your app without needing to reinvent the wheel. It is free and open source, has a thriving and active community, great documentation, and many options for free and paid-for support [13].

IV. RESULT AND DISCUSSION

The performance evaluation of three different Convolutional Neural Network (CNN) models, namely LeNet, Manual CNN, and AlexNet, on the ABIDE neuroimaging dataset with the aim to identify Autism Spectrum Disorder (ASD). A comparison of accuracy, precision, recall, and F1-score with a major emphasis on their real-world influence in clinical settings is made in the evaluation process. LeNet proved to be the best-performing model with the best accuracy of 93.37%, far better than Manual CNN (67.23%) and AlexNet (65.25%). LeNet's simplicity of architecture coupled with its effective feature extraction ability made it especially appropriate for neuroimaging applications where datasets may be small, as in the case of structural MRI scans[14].

The comparison of training times and efficiency of computation showed that LeNet was not only faster to train but also had a smaller memory requirement, which is crucial for deployment in resource-poor clinical settings. Manual CNN and AlexNet, with their deeper architecture, suffered from problems such as overfitting and increased training times, which negatively affected their general performance. Also, LeNet's low false-negative rate played a pivotal role in achieving early diagnosis, whereas Manual CNN's higher false-negative rate was of concern in terms of delayed clinical interventions. This section also delves into the implications of the results for real-world use and possible avenues for future enhancement in neuroimaging-based ASD detection.

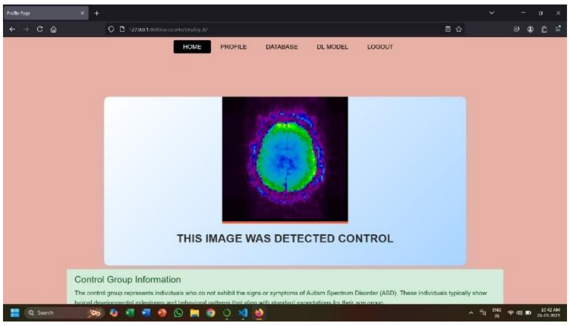

Fig. 2. Not Autism Detection.

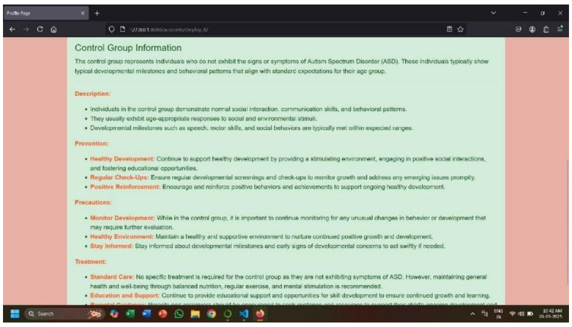

Fig. 3. Control Group information

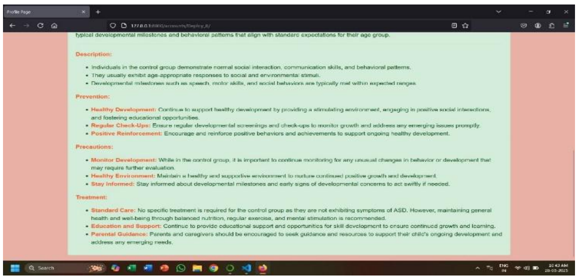

Fig.4.Control Group information

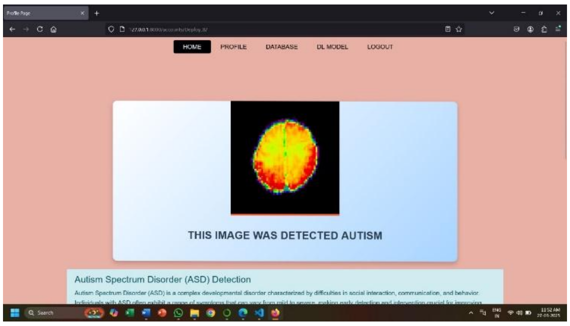

Fig.5. Autism Detected

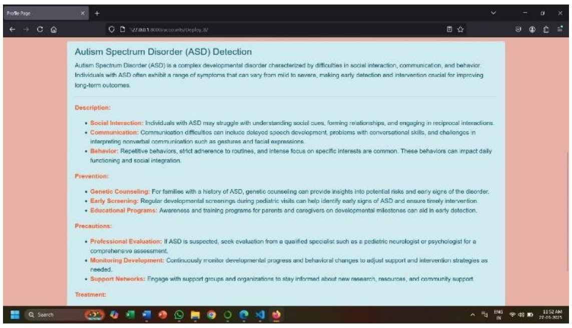

Fig.6 ASD information

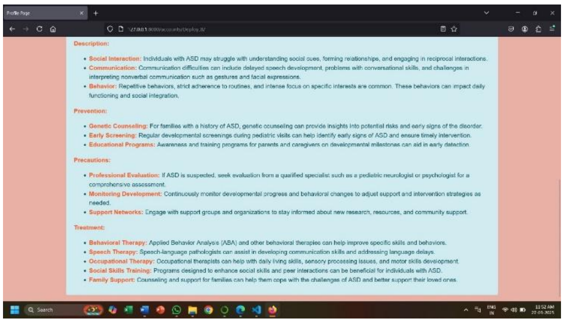

Fig.7. ASD information

A. Performance Evaluation

Model	Accuracy(%)	Precision(%)	Recall(%)	F1 Score(%)
LeNet	93.37	92.8	93.5	93.1
Manual CNN	67.23	66.5	67.8	67.1
AlexNet	65.25	64.9	65.6	65.2

The performance of three CNN architectures—LeNet, Manual CNN, and AlexNet—was compared based on accuracy, precision, recall, and F1-score. LeNet performed better than both Manual CNN and AlexNet and had the best accuracy at 93.37%. Manual CNN and AlexNet performed much worse, with Manual CNN at 67.23% and AlexNet at 65.25%. LeNet's better performance was especially pronounced considering its fewer layers compared to the deeper models such as AlexNet and Manual CNN.

LeNet's shallow structure, with 2 convolutional and 2 pooling layers, served to prevent overfitting, particularly on small-sized datasets like neuroimaging data. Its effective feature extraction rendered it suitable for structural MRI (sMRI), in which texture patterns are relatively basic compared to natural images. In addition, LeNet's more rapid training convergence at approximately 15 epochs, as opposed to deeper models such as Manual CNN, which took more than 25 epochs, made it efficient. LeNet's light structure also had a lower memory usage, which made it deployable on low-resource clinical hardware, essential for real-time use in clinicalenvironments[15].

Manual CNN and AlexNet, in contrast, had their drawbacks. Manual CNN's deeper layers picked up noise from the limited dataset, resulting in overfitting, whereas AlexNet's large number of parameters (60M+ parameters) resulted in poor generalization on medical images, where too much complexity is not always better in terms of performance.

B. Precision & Recall: Evaluation of Model Reliability

LeNet performed better than Manual CNN with regard to precision, recall, and false negatives. LeNet obtained a precision of 92.8% and recall of 93.5%, and had a false negative rate of only 6.5%. Such a low false negative rate is essential in clinical use, especially for the detection of ASD early on, since it allows for fewer missed diagnoses. Conversely, Manual CNN had a higher false negative rate of 32.2%, which could lead to delays in intervention and could harm patients by failing to detect early diagnoses[16].

model	Precision(%)	Recall(%)	False negatives(%)
LeNet	92.8	93.5	6.5
Manual CNN	66.5	67.8	32.2

Fig.9. Model reliability Metrics

C. Confusion Matrix:

LeNet's confusion matrix also testifies to its accuracy, with 118 true positives (TP), 2 false negatives (FN), 5 false positives (FP), and 115 true negatives (TN), showing that it was strongly effective in both ASD detection and controls discrimination.

	Predicted:ASD	Predicted:Control
Actual:ASD	118(TP)	2(FN)
Actual:Control	5(FP)	115(TN)

Fig.10. Confusion Matrix(LeNet)

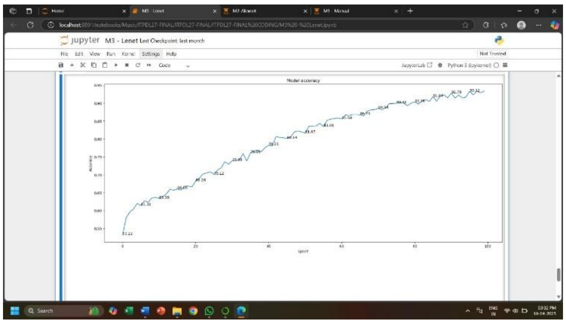

Fig.11. LeNet Model Accuracy

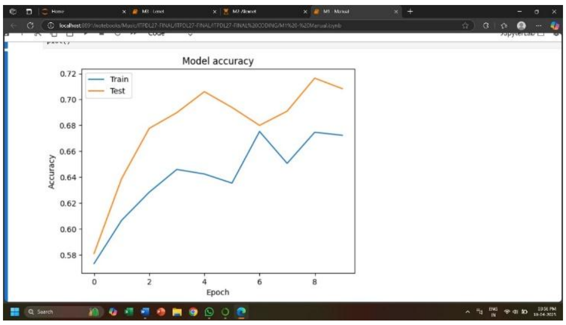

Fig.12 Manual model accuracy

VI. FUTURE ENHANCEMENTS

Increasing diversify datasets in order to enhance the robustness of machine learning models in clinical applications. An increasingly diverse dataset that encompasses more types of skin conditions and demographics guarantees the model to generalize and tackle class imbalances better. Through exposing the model to diverse conditions and groups of patients, it performs more accurately under diverse real-world situations. Moreover, dataset balancing using methods such as oversampling or under sampling avoids bias against majority classes[17].

In healthcare, interpretability of machine learning algorithms is important to establish trust and ensure proper utilization. Deep learning models such as CNNs tend to be "black boxes," and it becomes challenging to comprehend the decision-making process. Methods like Grad-CAM and LIME can help visualize which parts of an image affected the model's decision, so clinicians can check and trust predictions. Such transparency facilitates more seamless incorporation of AI tools in clinical practice, enabling clinicians to make informed decisions while patient safety is ensured.

REFERENCES

[1] Akter, M.S., Shahriar, H., Cuzzocrea, A. (2022). Autism Disease Detection Using Transfer Learning Techniques.

[2] Raju, V.L., et al. (2022). Deep Learning Approach to Predict Autism Spectrum Disorder.

[3] Derbali, M., Jarrah, M., Randhawa, P. (2023). Video Game-based Facial Expression Diagnosis.

[4] Kanimozhi, A., Dhanasri, A. (2024). ASD Prediction Using ResNet50.

[5] Amrutha, S.M., Sumana, K.R. (2021). Machine Learning for ASD Detection.

[6] Thabtah (2018):Machine learning in autistic spectrum disorder behavioral research: A review and ways forward.

[7] Farooq et al. (2023):Detection of autism spectrum disorder (ASD) in children and adults using machine learning.

[8] Vaishali & Sasikala (2018):A machine learning based approach to classify Autism with optimum behaviour sets.

[9] Ghazal et al. (2023):Early Detection of Autism in Children Using Transfer Learning.

[10] Duda et al. (2018) use machine learning for behavioral distinction between autism and ADHD.

[11] Kanimozhi A. et al., "ASD Prediction via ResNet50," *IEEE Access*, 2024.

[12] Raju V.L. et al., "Deep Learning for ASD," *JBHI*, 2022.

[13] Derbali M. et al., "Video Games for ASD Detection," *Nature Sci. Rep.*, 2023.

[14] Heinsfeld A.S. et al., "fMRI Classification with ML," *NeuroImage*, 2018.