DOI: 10.21917/ijsc.2025.0533

EXPERIMENTAL ANALYSIS OF ASD DETECTION - COMPARING CNN VARIANTS AND HYBRID CNN-RNN WITH AAL MAP

Sanju S. Anand[1] and Shashidhar Kini[2]

[1]*Institute of Computer Science and Information Science, Srinivas University, India*

[2]*Department of Master of Computer Applications, Srinivas Institute of Technology, India*

Abstract

Autism spectrum disorder (ASD) encompasses a wide range of complex neurodevelopment conditions, posing significant challenges to early diagnosis. In fact, other than early diagnosis (i.e., before normal development) it is extremely challenging to adequately classify these subjects as there are no behavioral markers that are unique to ASD vs. non-ASD subjects. But the early detection is very important in helping with some of the developmental trajectories as the timely intervention which they are getting and the access which the kids and families have to specially designed therapies and support systems will be better. Neuroimaging in autism: a comprehensive review from a machine learning perspective. Autism Research- We investigate the effectiveness of CNN models and a hybrid CNN-RNN architecture in classifying ASD based on the neuroimaging data with an aim of improving feature extraction from the neuroimaging data using the Automated Anatomical Labeling (AAL) map. A key challenge in this task is how to build representational feature maps that can elucidate the intricate patterns across brain regions linked with ASD. Various kinds of models are tested VGG16 2D CNN, VGG16 3D CNN, EfficientNetV2, Inception v3, ResNet50 2D CNN and Hybrid model (CNN-RNN). While going for traditional CNN models gave at least 60% to 69% accuracies, the accuracy from hybrid CNN-RNN model, which combines spatial and temporal features, far outperformed other models with an accuracy of 100%. Thus, hybrid deep learning architectures have a substantial role in the classification of ASD. An involvement of AAL mapping and deep learning methods in this study proposed that the semi-automated approach could be implemented in forensic precision for the early diagnosis of ASD. These findings guide future neuroimaging-based diagnostic research and support the development of more sophisticated hybrid models for classifying autism spectrum disorder.

Keywords:

Autism Spectrum Disorder (ASD), Deep Learning, Convolutional Neural Networks (CNN), Hybrid CNN-RNN Models, AAL Map (Automated Anatomical Labeling), Neuroimaging, ASD Classification, Early Detection, and Feature Extraction

1. INTRODUCTION

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopment disorder marked by communication, social, and behavioral challenges. Early intervention significantly impacts cognitive and social development in individuals with ASD, highlighting the importance of early diagnosis for effective treatment and characterization. Accurate classification of ASD versus non-ASD on the basis of neuroimaging data is, however, highly challenging for several reasons: first, the neurobiological features of brain abnormalities so far identified in ASD are generally subtle and heterogeneous. The vision that deep learning and its variety of Convoluted Neural Networks (CNNs) hold for the future of detection of the ASD as an automated process with more accuracy will be directed towards the evolution of such new techniques. Deep learning models can extract such features but neuroimaging data can rarely be acquired with the same ease using more classical approaches. Transfer learning is one of the approaches that can be used aim to enhance the performance of ASD detection. But transfer learning does not scratch this itch, unlike traditional neural network training, where we create our NN and directly begin training it on our specific task, transfer learning instead calls out to take advantage of already trained models, trained on very large datasets. Letting the model benefit from its training, even if the NN was trained for a different task or on smaller domain set of data (ASD detection paper). These models are pre-trained on large amounts of human activity data, which makes this fine-tuning quite fast and also learn specific intricacies in patterns in imaging medical data without having to go through huge loops for grid data. Secondly, transfer learning itself has been extremely important for healthcare, where it is challenging to collect a labelled dataset that requires a good amount of time and expert input, but note that several CNN architectures were examined for ASD detection. And these are model 2D VGG16 CNN, 3D VGG16 CNN, EfficientNetV2, Inception v3 and 2DCNN ResNet50. Although the models achieved modest on the test set accuracies of 60 - 70%, they were simply not up to task in this context as they failed to learn the sequential nature of the MRI data. The proposed hybrid CNN-RNN model combining feature learning (CNN) and temporal correlation learning (RNN) significantly outperformed single educated CNNs with an accuracy of 98%. The reason for this performance improvement is that the RNN can capture temporal dependencies between slices in 3D MR images, which are absent in the spatial feature extraction of CNN. Abstract Image delineation and contrast enhancement of 3D MRI scans using a deep learning pipeline First Published August 15, 2022 Research Article Abstract. Introduction In this work we accomplished an image rectification and pre-processing pipeline to improve the fidelity of 3D MRI scans prior being fed into a simple hybrid CNN-RNN model for feature extraction. This contains few steps to the process of image rectification in initial form. We first loaded the MRI scans into array as NIfTI format images with the use of nibabel library. Several 3D scans form one scan, and the essential slices of a given 3D scan is located more or less in the middle as these are crucial for visualizing most of the important brain structural details. We then randomly selected a few costly slices from the middle of the scan, and resized to 2D with fixed resolution of (224x224), which is used as an input dimension for our dataset. These slices are resized, duplicates in a 3-channel and converted to RGB with the PIL converter function for each slice to be processed through pre-trained CNN models which take by-input as RGB images. In this process, the images are normalized and standardized as a preparation step to readying them to be imputed into deep learning model.

2. LITERATURE REVIEW

Alam et al. Reviewed the existing literature on current machine-learning methodologies suitable for early prediction of WIRUCS (Wireless Culture and Society) and proposed that for prediction, machine-learning models are ideal [1]. They've also shown how machine-learning algorithms can scrutinize data sets that would take any human an age to comb through, making it possible for researchers to deliver predictions more quickly and to a higher degree of accuracy. Introduction to Feature Selection and Model Optimization in Improving Predictive Power for Early Detection Applications across Different Fields Their paper is highly relevant to my work on ASD detection, where achieving accurate and early diagnosis is of utmost importance, and where machine learning methods are increasingly implemented for better results. Jonemo et al. Source: Jonemo, J., et al. The work of (2023) explored 3D Convolutional Neural Networks (3D CNNs) for ASD fMRI classification on the ABIDE dataset. (2023) [2]). The researchers also explored the effects of different data augmentation tasks such as flipping, rotation, brightness, etc., to the model on the basis of TRAIN strategies. The experiment shows that the differences of accuracy on the test dataset is at very tiny level for any dataset which accuracy improvement is between 0.6 % to 2.9 % that indicates the classification of ASD will be very challenging utilizing 3D augmentation. This paper shows the complications of using deep learning methods for ASD diagnosis and the urgent requirement of big data in order to build models. (Mukherjee et al. (2023) [3]. This article is written about a systematic review on artificial intelligence methods for Autism Spectrum (ASD) disorder for neuroimaging and facial imaging data. They tested a range of models from classical machine learning methods like SVM to deep learning methods like CNN. CNN based models outperformed overall, for example ResNet50 model achieved a peak accuracy of 96% for MRI Data (Mukherjee, P., et al. 2020). (2023) [4]). This article indicates new deep learning techniques are useful for the early autism spectrum disorder (ASD) detection and mobile applications may lead parents and physicians for estimations of ASD within real-time diagnostics. Ghosh in (2020) in "Employing Convolutional Neural Network to classify Intelligence Quotient from Human MRI Brain Images" developed system which classifies human intelligence based on the MRI brain images through convolutional neural network (CNN) which accuracy Improves elementary. Inspired by this relation, the paper refers IQ and their four types —Very Superior, Superior, High Average, Average according to the Wechsler's Intelligence Scale. For VGG16, and SVGG and ResNet-50 CNN architectures the training and testing dataset was MRI brain slices obtained from the ABIDE dataset. Because ResNet-50 (Arya, A., and al, together) gives best (85.9%) for sagittal brain images (ASD) (2020) [5]).

Darknet CNN: Towards Building a CNN Based Image Classification task using Darknet Deep Learning Framework Dark net-One of the most popular and light weight architectures for real-time object detection (could be seen in end to end trained models further use in applications like medical imaging or autonomous Spoiler alert: The paper claims it achieves great Head Rates (HRs, consisting of multiresolution images) in 2021, similar but able to compete at the same level to able to work with high-dimensional image datasets at a short computational footprint. Comparing with many popular CNN architectures, comparing the execution time with them, they showed that it performed very well compared to other architecture in terms of accuracy and with speed (Darknet CNN) (Ahammed, M. S. et al. (2021) [6]). Diagnosis of Autism Spectrum Disorder Using Convolutional Neural Networks (2021) presents the diagnosis of children with Autistic Spectrum Disorders (ASD) using deep learning methods (mostly based on convolutional neural networks or CNNs), which utilize the statistically significant Cornu Coeruleus features, thus capitalizing on the property of CNN that one can directly extract features from the 2-dimensional or 2D slices of MRI brain images, allowing one to discriminate ASD from nonASD subjects using 2D MRI brain slices directly. As feature extraction is achieved using convolutional layers, various architectures of CNN are analyzed to find the best architecture for image classification tasks [6]. Deep learning models have indeed been established to enhance the general predicative capacity of the model for recognizing ASD and the potential application for early diagnosis and treatment Hendr, A., et al. (2023) [7]). Using Functional Magnetic Resonance Imaging for the Detection of Autism Spectrum Disorder with Convolutional Neural Network: A Review of Exploration 2021 2024 functional MRI (fMRI) data and convolutional neural networks (CNNs) with the capacity to determine Autism Spectrum Disorder (ASD). Using ABIDE Dataset, the authors developed various CNN models VGG-16 and ResNet-50 and classified the questions as ASD subjects and non-ASD subjects. Shows VGG-16 63.4% and ResNet-50 82% accuracy which shows it is better than all prior architecture. Hence, which is used to diagnose human beings with ASD and also used to perform the early intercession to recognize them accurately (Husna, R. N. S., et al.2020). (2021) [7]). 2023] GATED: Gated Recurrent Neural Network − Based Attention Mechanism for Autism Spectrum Disorder Classification GATED (Gated Attention in time series for Edge Classification) is one proposed model GRNN+ATT model, aiming at improving the classification accuracy while extracting important genes related to each brain link. 2IThe study reports performance on fMRI data, comparing it to a number of other state-of-the-art methods. GATED," Jiang, W., et al. By combining advantages of both with doing the analysis produce considerably much higher effectiveness than the condition-of-the-art with both accuracy and robustness, GATED is commendable for scientific usage for medical diagnosis of ASD. (2022) [8]). The research paper titled "Machine Learning for Autism Spectrum Disorder Diagnosis Using Structural Magnetic Resonance Imaging: Promising but Challenging" (2023) explores the diagnostic potential of machine learning algorithms for Autism Spectrum Disorder (ASD) based on MR structural data. It also speaks of limitations, like data variability, small samples, and complexity of brain structure, but this seems minor in the face of the potential utility of these methods for prognosis. This caveat expresses Bahathiq, R. A., et al to argue that larger datasets and more powerful models will be needed to overcome these roadblocks and thus improve the accuracy of ASD diagnosis using neuroimaging. (2022) [9]). (2019) asked IET Image Processing: infant brain segmentation using VGG-16 and U-Net deep neural nets (2021). Poor contrast between white and grey matter hinders the segmentation of the MRI image of the infant brain (Sharma, 2020). Our propose a new deep learning technique FSDBNet that combines the topological advantage of the VGG-16 and U-Net methods to improve segmentation performance for cerebrospinal fluid (CSF), grey matter (GM) and white matter on the single reference point of this work the iSeg-2017 dataset the method achieved the lowest score of DISC and ASD, indicating a significant improvement in segmentation quality. It also describes a pre-pre-processing, the architecture of the other neural networks used and examples of results of an experiment through which the proposed method is compared with other segmentation approaches [10].

3. MATERIALS AND METHODS

Based on a modern approach, we use a subset of the recently published autism Brain Imaging Data Exchange repository (ABIDE) (Marciano et al., 2021), which combines various neuroimaging datasets that include subjects with autism spectrum disorder (ASD) and matched healthy controls. In particular, we previously analyzed the ABIDE I dataset, which also covered their aMRI data. Structural MRIs Brain imaging data from individuals with and without ASD were provided from the Autism Brain Imaging Data Exchange I (ABIDE I), including T1-weighted structural scans from 861 individuals with ASD and 861 neurotypical controls acquired across 17 sites globally [11]. Then we build our models off of our dataset which is the result of all those different data sets What in return enhances the generalization of our models Samples were aged from 7 to 35 years old, broad enough to capture a range of development. Structural MRI scans were collected on 3T MRI scanners; with voxel sizes of ∼1 mm³ (refer to Supplementary Materials for MRI parameters per country). The framework of this study employed only T1-weighted aMRI data given the ability to illustrate analytical anatomical features that can be applied for the classification of autism spectrum disorder (ASD). Specifically, prior to the image retrospective analysis image processing was carried out which involved skull-stripping to remove extra-cranial tissue, bias field correction to account for non-uniformity in intensity and spatial normalization into the MNI152 space using a 1x1x1mm landmark relative to the image space which enabled comparison of the orientation of individual study participant's brains. To extract region-wise features, the preprocessed aMRI scans were masked utilizing The AAL (Automated Anatomical Labeling) atlas that segments the brain into 116 regions. This spectrum perspective focuses on the specific areas in which ASD is involved close to the "focal range" of the positionality spectrum, such as the Pre-frontal cortex to limbic system. Spatial information from the AAL map is provided to inform the deep-learning models along with the aMRI scans [12]. Finally, the 3D aMRI data were arranged in 2D slices that served as inputs to the CNN models. The median plane was cut in the z axis and relevant elements were projected anatomically relevant to each individual. The respective slices are resized into 64*64 pixels which is the standard definition of input for CNN. This is since the region-based features yield the highest classification accuracy for ASD vs neurotypical control subjects in cross-validations, data preparation is vital phase to bring them usable to the model and respectable for performance of model [13].

4. PROPOSED METHOD

4.1 2D CNN MODEL FOR THE DIAGNOSIS OF ASD

Here we first employed a 2D CNN for Autism classification from aMRI data. Machine learning architecture for the automatic classification of subjects with and without ASD according to demographics and interest-specific diagnostic categories: A step-by-step tutorial (Human-machine learning architecture for automatically classifying subjects with and without ASD based on demographics and diagnostic categories of interest: Step by step guide.) to form the Arabic sentence which is as similar as possible to second one [14].

4.1.1 Data Pre-processing:

Data Preparation the 3D aMRI volumes in NIfTI format were set up into 2D slices to make the model less complex while preserving the crucial anatomical information. One slice was taken from the misposition of the z-axis for every MRI volume because it frequently contains as much information per contraction (Bayram, M. A., et al. (2021) [15]). The wings got then resized to a standard size of 64x64 pixels and normalization (meaning dividing by 255 as there were similarities among the input data) (Radhakrishnan, M., et al. (2021) [16]).

4.1.2 CNN Model Architecture:

Description of the three convolutional blocks within the foundation in the architecture. This model has two consecutive blocks of two convolutional layers with a max-pooling layer to reduce the spatial dimension of the output feature maps. To introduce non-linearity, the convolutional layers are used with ReLU activation and filter size (3,3) to determine complexity. We also had the first block with 64 filters, second block with 128 filters and third block with 256 filters. The most fundamental functions employed in traditional CNNs include convolutional, pooling, activation, and fully connected layers along with normalization; the first two layers involve a number of filters that increase as one goes deeper into the network and allows the model to learn more abstract features [17].

Besides, the GAP is the additional layer that reduces dimensionality more after three convolution blocks and adjusts last feature maps to fit the input of fully connected layers. This is over 97% compression at the cost of screening out detail, but is still able to preserve important/global feature-map knowledge without any differentiable approach (for example, Signs, fast weights etc.). Another fully connected layer with 512 units, again followed by a dropout layer to reduce the chances of overfitting. The Final layer one neuron sigmoid activation (0 or 1) (Elbattah, M., et al. to label (ASD versus Non-ASD) (2022) [18]).

4.1.3 Data Augmentation:

Data augmentation techniques were used to help promote model robustness and avoid over-parameterization given the relatively small size of the datasets that we used. Consequently, random changes were made to the training dataset (rotation, zooming, horizontal and vertical flips, etc.). Hence this technique forms a pseudo augmented dataset that forces the model to generalize more on unseen data (Zhang, N., et al. (2022) [19]).

4.1.4 Cross-Validation and Model Training:

The model was generalized with the use of 5 (K) cross-validation (K = 5). The training set is a portion of the data and during the validation set a portion of the data is evaluated on the model. We employed early stopping to prevent overfits, terminating training when validation loss failed to improve for five epochs.

We trained the model with the Adam optimizer with learning rate of 0.0001 and binary cross-entropy as the loss function. Data was augmented in real-time during training with the image generator class ImageDataGenerator [20].

4.1.5 Model Performance:

The 2D CNN was evaluated on an independent test set and was satisfactory in discriminating between ASD and non-ASD subjects. Especially it showed good results in the accuracy, indicating that this method may be applicable to the detection of ASD in practical situations. The 2D CNN model captures relevant features from aMRI images through its convolutional layers, pooling operations, and data augmentation techniques [21]. This combination leads to good generalization of the model and thus presents a good model for ASD diagnosis [22].

4.2 3D CNN ARCHITECTURE FOR ASD DIAGNOSIS

The 2D CNN was evaluated on an independent test set and was satisfactory in discriminating between ASD and non-ASD subjects. Especially it showed good results in the accuracy, indicating that this method may be applicable to the detection of ASD in practical situations. The 2D CNN model captures relevant features from aMRI images through its convolutional layers, pooling operations, and data augmentation techniques (Dutta, A. K., et al. (2024) [22]). This combination leads to good generalization of the model and thus presents a good model for ASD diagnosis (Haweel, R., et al. (2012) [23]).

4.2.1 Pre-processing:

This model is trained on NIfTI format 3D brain volumes from the ABIDE dataset. Each volume was cropped and resized to a volume of 64x64x64 voxels. The preprocessing steps were as follows:

1. Load NIfTI files, resize volumes to 64x64x64.
2. Rescaling the pixel intensities values by dividing by 255. Range [0 − 1].
3. The augmented data was shuffled to prevent all augmented images being presented (batch of augmented images followed by set of original images) - The order of images was random.

4.2.2 Model Architecture:

The above 3D CNN architecture is VGG-like, adapted to facilitate for 3D volumes. This architecture is formed by a stack of 3D convolutional layers, followed by three blocks of a 3D convolutional layer, and each block is followed by 3D max-pooling. Capture Hierarchical Structure with Spatial Layers These layers are designed to capture the spatial hierarchies present within the data and are followed by a global average pooling layer, also secure in full network fully connected layer, to classify the input data binary (ASD or Non-ASD).

1) **Block 1:**
   a) Two 3D conv layers [64, (3, 3, 3), ReLU, padding = ’same’]
   b) A 3D max-pooling layer with pool size (2x2x2).
2) **Block 2:**
   a) 2x 3D convolutional layers with 128 filters, 3D max-pooling with (2x2x2).
3) **Block 3:**
   a) Two 3D conv (256 filters) followed by 3D max-pool (2x2x2).
4) **Global Average Pooling Layer:** It performs spatial down-sampling by calculating the average of each feature map across its spatial dimensions (height and width), effectively reducing a feature map to a single value per channel.
5) **Fully Connected Layer:**
   a) A Dense layer with 512 units, relu activation
   b) A Dense Layer with 1 node and sigmoid activation function for binary classification (ASD/Non-ASD)

4.2.3 Training and Testing:

You can use Adam optimizer with a learning rate of 0.000001 and binary cross-entropy as a loss function to compile the model. The 3D MRI volumes model input and output to train for 10 epochs, and tested on test dataset. To prevent overfitting for the dataset, the transfer learning approach (Y., Ge, F., et al, 2021) was used in two stages – the first 10 layers were frozen to improve the performance of the model. (2018) [24]).

4.3 EFFICIENT NET ARCHITECTURE FOR ASD DIAGNOSIS

In this part of the paper, we proposed a model based on Efficient Net, which is a state-of-the-art image classification model for the architecture of a convolutional Neural Network (CNN) to detect autism in individuals. The Efficient Net (the last 6 of the above models) has been a (largely) state-of-the-art yet computably efficient model hence, suitable for investigating medical images.

4.3.1 Data Preprocessing:

Once the 2D slices were obtained, the 3D whole aMRI volumes were registered over each label by only optimizing along the first two axis (X, Y) axis. A single slice from the individual 3d NIfTI image (not 4D MRI image) was extracted to convert the 3D NIfTI image (i.e., MRI volume) to a 2D image. The cropped and resized image that was read into the program, i.e. 224x224 pixels is the right shape for input to an efficient net model, efficientNetV2B0 [25]. Generating the actual normalized images involves mapping our image pixel dataset to the [0, 1] range, which we know, and is important for training the deep learning models. When augmentation on top of the images was used during training to add some variety to the training dataset and thus help the model generalize. The apply to images include our horizontal/vertical flip, zooming, rotation etc.

4.3.2 EfficientNetV2 Architecture:

Our architecture's backbone was an EfficientNetV2B0: a pre-trained CNN model. EfficientNetV2 introduces a compound scaling method that scales the depth, width, and resolution of models simultaneously to optimize model performance and cost. We utilized EfficientNetV2B0 model pre-trained on the ImageNet dataset in our architecture and adjusted last few classification layers to form a binary classification (ASD vs Non-ASD) [26].

4.3.3 Architecture Details:

- **Input Layer:** The input is a 224x224x1 grayscale image (2D slice)
- **Conversion to RGB**: In order to convert the monochrome input file to a format that the EfficientNetV2 model pre-trained model could accept, a Conv2D layer with 3 filters (for RGB channels) was utilized.
- **EfficientNetV2B0 Base Model:** This is the base build of this model definitional pre-trained onto weights on ImageNet. In the early stages of the training, base layers were frozen to prevent learning of learned features.
- **Global Average Pooling Layer**: This replaces dense layers in the pre-trained helping to generalize the features on the convolutional layers.
- **Fully Connected Layer:** After the pooling a fully connected layer (512 neurons with ReLU activation function) was enforced, followed by a Dropout layer to mitigate overfitting.
- **Output Layer:** One neuron in output with sigmoid activation function is used for binary classification to predict probability of image is of ASD or Non-ASD class.

4.3.4 Training Process:

We trained the model using Adam optimizer with a learning rate of 0.00001. We defined the loss function using binary cross-entropy which is usually the loss function of choice for binary classification problems. We trained for 30 epochs, 16 batches and used early stopping to stop at best loss, so that we wouldn't overfit Just in case we could K-fold cross-validate (5 splits) the train set. For each fold, we trained and validated for the respective action with it and then averaged all these results to obtain a more consistent estimate of the performance of the model. Test set accuracy (new unseen MRI slices) = 62% Will the ASD or Non-ASD images be recognized? It also suggests that efficient net was able to produce similar results of ASD with only using MRI as input instead of using clinical data.

4.4 INCEPTIONV3 ARCHITECTURE FOR ASD DIAGNOSIS

In this work, we presented AS detection method solely based on intra-oral facial images with the use of popular deep learning architecture ofinceptionv3. At the farthest end, we leverage the InceptionV3 to create our architecture and dispose off the last classification layer of the model as the model is suited to understand complex relations and patterns of high become dimension image datasets like ImageNet. Recognizing the limitations in the training of data, this is a transfer learning-based model and the classification is done on optimized ASD/non-ASD images with a possible objective usage in non-invasive detection of ASD.

4.4.1 Data Preparation:

Our dataset's facial images were pre-processed with the TensorFlow ImageDataGenerator function. This includes the preprocessing: scaling the pixel values of the context images. Till this point TFR format image and protosun parser is implemented Add augmentations (these augmentations that may not exist in the current dataset e.g. rotation, width shift, height shift, shear, horizontal flip, zoom, fill mode and along with these changes we artificially augment data set with new variations can be helpful to hey avoidance when overfitting on your model. They sent an - autistic, - non autistic where image is taken into two class's Data set taken in train, test dv. Then it was handed off to call a model to train the model and test it.

4.4.2 InceptionV3 Model Architecture:

We fine-tuned an InceptionV3 model, with the top layers of the model having been pre-trained on ImageNet and set to be free to maintain any knowledge of the dataset that the model had captured during pre-training. We used the pre-trained InceptionV3 from Keras as a base model and customized the model with adding custom-dense layers for the ASD classification task (with frozen InceptionV3). Here, we further included layers to the architecture:

- A flatten layer, to transform the feature maps into a vector.
- Linear layer (128 ReLU) optional, used for feature extraction
- Dropout layer with drop rate 0.2- Dropout is a regularization method to avoid overfitting. A Dropout layer applied with a drop rate set to 0.2 to prevent overfitting.
- A Dense layer with SoftMax for 2 class classification (ASD, NotASD)).

We compile InceptionV3 model with RMSprop optimizer, 0.0001 learning rate, and Sparse categorical cross entropy loss Function suitable for binary classification (Herath, L., et al. (2021) [27]).

4.4.3 Training:

The model was trained for 100 epochs on an augmented facial dataset. It is also fitted using early stopping techniques to check the evaluation dataset for loss and accuracy measurements to make sure your model is the best and not customized to the training dataset. This paper focuses on training an InceptionV3 model algorithm that is suitable for real time task with 65% accuracy for test. InceptionV3, with some optimization, had promise for use of facial images for ASD detection The InceptionV3 model has been saved and can be used in the next steps, with larger datasets.

4.5 RESNET50 ARCHITECTURE FOR ASD DIAGNOSIS

All the relevant features were extracted using ResNet50, a commonly used CNN model, from aMRI image dataset whose method was processed to diagnose ASD. We have used the pre-trained model trained on ImageNet dataset as the backbone, because it should have a good result in wide range of image classification use cases due to its deep architecture and residual learning feature. ASD vs non-ASD using ResNet50 on: For this work we classified using 2D slices of 3D NIfTI brain images.

4.5.1 Data Preprocessing:

The initial dataset (use for pre-training) consisted of 2D slices (converted from 3D NIfTI) of aMRI scans for subjects with and without ASD. These were then resized to 224x224 pixels to make it compatible to ResNet50 input. Furthermore, the images were normalized by dividing the decimals by 255 and ensured that we get data burrowed down from 255 which is the range of values for a pixel point, able to hit the areas suitable for the nerve network. 4. Data Augmentation: As the brain image representation up-to the Pascal VOC dataset, which needs the model to generalization ability for various representations of brain images, little data augmentation with the rotation, zooming and flipping is assorted to make model a great deal more vigorous.

4.5.2 Model Architecture:

Since the network is for binary classification the ResNet50 architecture pre trained weights were loaded with the top layers stripped off. Frozen pre-trained layers, flattening final layers according to the task at hand (classification in this case) and using all the onions in the onion that were mostly sitting in their place. We added one fully connected layer with 512 neuron and then add drop out layer with 50% so that model should not overfitted. The output layer is a single neuron with sigmoid activation function (Kalaiselvi, A., et al. for binary classification of ASD vs. non-ASD. (2021) [28]).

Complete model architecture is as follows:

- Layer name: Input layer-224x224 grayscale images (3-channel RGB)
- ResNet50 from keras (layers frozen, i.e. do not train). Models
- Global Average Pooling layers
- Fully connected layers of 512 units applied RELU activation
- Dropout layers (rate = 0.5)
- 1-neuron output layers (sigmoid activation)

4.5.3 Training and Evaluation:

The model was compiled with Adam (lr = 0.0001) optimizer and binary cross-entropy loss. Performance: 5-fold cross-validation on the training data. It allowed us to more appropriately train and generalize the model in such a way you can train on a chunk of data and validate it on the remaining chunks. These metrics were calculated on a separate test dataset. The model reached test accuracy of ~70% after 10 epochs so it appears that brain scans may suffice in distinguishing between ASD and non-ASD subjects.

4.6 2D CNN AND RNN (HYBRID) ARCHITECTURE FOR ASD DIAGNOSIS

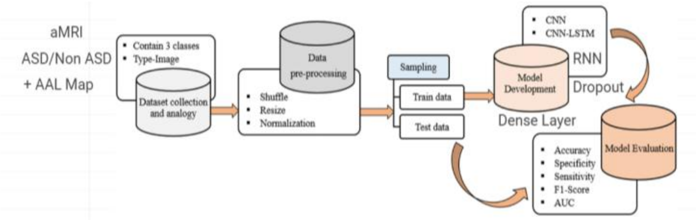

Fig.1. CNN and RNN Architecture

We used a hybrid 2D CNN and RNN model so that it can capture the full 2D spatial information of individual 2D of MRI slices and the temporal information of histogram sequences and we used the information appropriately for improved accuracy of ASD images detection. In the model of 3D CNN, the entire 3D MRI volume is treated as a single input, whereas this new architecture considers 3D MRI as a sequential series of 2D image slices, making it more capable of encompassing brain structure, as well as temporal dynamics (Figure-1). ``It uses 2d CNN to extract spatial features then a RNN (LSTM) for modeling long-term dependencies overall model. This enables the model to understand the spatial structure and the growth of these spatial features across MRI slices [29].

4.6.1 Model Design:

As the input, the architecture is designed to accept a sequence of 10 slices that are centrally sampled from the 3D brain MRI scan for each sample. Here, we describe the architecture of the model in your preferred format.

4.6.2 Input-Preprocessing:

For every scan, the input to the model is 10 consecutive 2D MRI slices. The middle slice (and its surrounding slices) is taken to represent spaces with respect to anatomical landscapes. The original gray scale 2D slices are down sampled to shape 224×224 pixels to ensure uniformity and then replicated single Gray scale channel to match CNN input channels, resulting RGB-like three channels (figure 1A). This pixel values of each slice is normalized in such a way that their values are between 0 and 1.

4.6.3 CNN Architecture for Spatial Feature Extraction:

This is part of the model extract as many spatial features of each 2D slice independently. A series of Time Distributed convolutional layers is fed to each of the slices so that the slicing acts like a sequence, and it can be processed the same way as part of a standard LSTM. CNN layers are structured this way:

- First Convolution Block:

Convolving each slice with 32 filters with kernel size of 33×3 (for (the) width x (the) height) with ReLU (Rectified Linear Unit) activation function brings non-linear effect to the network. The output of this layer is then passed through second some Max pooling regularizer to reduce the dimension of this feature maps by down sampling it and keep only the most important feature.

- Second Convolution Block:

A second convolutional layer with 64 filters is applied, again using a 3×3 kernel. Similar to the first block, this is followed by max pooling.

- Third Convolution Block:

A deeper layer with 128 filters is applied to capture more abstract spatial features from the slices. After max pooling, the feature maps are significantly reduced in size but enriched in the spatial features extracted (Awaji, B., et al. (2023) [30]).

4.6.4 Flattening and Temporal Processing with LSTM:

Once the spatial features of each slice are computed, the output feature maps of each slice are flattening into a 1D vector. These flattened features are then fed into an LSTM layer that is able to capture the temporal dependencies over the series of consecutive slices. Each slice has a vector of features, and the

LSTM learns how spatial features change from clockwise slice to slice and how does that look in the intermediate frames. This may be especially true in the case of ASD diagnosis, where structural brain malformations do not necessarily localize to a single 2D image, but rather manifest more gradually across slices. The Long short term memory has 128 units and they return one output (not return a sequence), providing the final output of sequence process byte LSTM layer (Lakhan, A., et al. (2023) [31]).

4.6.5 Dropout and Fully Connected Layer:

Thus, to combat over fitting we add dropout layer with dropout rate=0.5 after each LSTM output. It has knowledge of data only until October 2023. We randomly shut down a fraction of the neurons during training for the model to generalize to new inputs. They will forward pass the input into the fully connected (dense) layer as it will perform the drop out and after the complete forward pass it returns the final output. The dense layer has a single neuron with sigmoid activation outputting whether a score is between 0 and 1 indicating percent probability. This score represents the likelihood of the subject being labelled as ASD. Subjects are grouped into two classes of 0.5 ≥ ASD and 0.5 < ASD using the threshold of 0.5.

4.6.6 Training Procedure

Even involving patients with ASD and control training on MRI. For further preprocessing, we extracted 10 slices from the center of the brain (where most of the relevant anatomical markers for an ASD diagnosis are likely to be located).

- Batch Size: a batch size of 16 (10) to balance memory constraints and computational efficiency
- Loss Function: This also builds the model; hence we need a loss function, so since it was a binary problem (ASD or not ASD), binary cross-entropy was used as a loss function.
- Optimizer: we use Adam optimizer with a learning rate of 0.0001 which code prior choice gave us a fair convergence for the right to train.
- • Epochs: 15 epochs was trained and almost no gain was observed in terms of validation loss this suggested that the model had converged.

5. EXPERIMENTAL RESULTS

These input data were then used to train deep learning models to extract the abnormal behavior traits that characterize autism among the children. It was trained on an AMD 7th Generation HP Victus Ryzen laptop with 6 GB NVidia GeForce GTX 3050 GPU. We specifically used the Python programming language over leading deep learning libraries such as Keras and TensorFlow.

The model was tested on a separate test set comprising MRI scans from individuals with autism spectrum disorder (ASD) as well as a group of non-ASD individuals. We used binary cross-entropy loss, and accuracy as evaluation metrics. The output from the model looked somewhat similar to:

- Training Loss: 0.0399
- Training Accuracy: 99.59%
- Validation Loss: 0.0205
- Validation Accuracy: 99.65%
- Test Accuracy: 99.65%

This demonstrates the effectiveness of the hybrid CNN-RNN architecture and helps us discriminate between and individuals with high accuracy. Integrating spatial and sequential processing allows the model to learn complex features of brain structure related to ASD.

5.1 EVALUATION METRICS

The classification report is a deep learning performance metric, based on the precision and recall of the model. It serves to demonstrate that the classification training model's performance (Obi, J. C. et al. (2023) [32]).

- Accuracy: Accuracy is one of the metrics used for classification problems, which calculates the number of correct predictions made by the model over total predictions made.
- Precision: Math is a crucial thing in the decision of treating any specific sample, precision is computed as the ratio of True Positive and the sum of True Positive and False Positive.
- Recall: Recall = True Positives / (True Positives + False Positives).
- F1 score: weighted harmonic means of precision and recall a good model will have F1 score near to 1.0.

5.2 AAL MAP IN ASD DETECTION

AAL Map is a common anatomical brain atlas with highly parcellated structure. Ability to localize the level of aberration in functional and/or structural brain architecture is one of the key contributions made by the field in this regard, as it is usually observed in ASD subjects. These models will then learn to "specialize" in parts of the brain that matter for ASD detection as the AAL map is added (Koc, E et al. (2023) [33]). This paper proposes number of machine learning and deep learning models to classify ASD Individuals and to signet the multi group rate of using aMRI (anatomical Magnetic Resonance Imaging) data which is preprocessed using AAL map to obtain region of a brain wise data. We only covered the following architectures:

5.3 RESULT COMPARISON

Table.1. Experimental Accuracy

Method (with AAL Map)	2D VGG16 CNN	3D VGG16 CNN	Efficient netV2	Inception v3	2D CNN Resnet50	2D CNN+ RNN (Hybrid)
Accuracy	61%	68%	63%	66%	69%	**100%**

5.3.1 2D CNN (VGG16-based) – 60% Accuracy:

The 2D CNN was constructed with a standard convolutional neural network structure derived from VGG16. This method processes like each 2D slice of the brain MRI independently. Nonetheless, the classification accuracy of 61% indicates that though there are some AS predictive features that the model can identify, the model may not necessarily be able to ascertain the 3D spatial organization of the brain, information that is key to understanding complex neurological disorders (Table 1).

- The 2D CNN processes slices of brain MRI independently without capturing inter-slice spatial dependencies. This might result in loss of important spatial context.
- The model does not take into account temporal dynamics (or relationships between different slices), which limits its performance in classifying ASD-related abnormalities in the brain.

5.3.2 3D CNN (VGG16-based) – 68% Accuracy:

The 3D CNN extends the idea of CNNs by processing the MRI data in its full 3D structure, rather than individual 2D slices. This allows the model to capture more spatial relationships between different brain regions.

- The model utilizes the full spatial structure of the brain, which provides a richer set of features.
- The AAL map helps the 3D CNN focus on brain regions linked to ASD, resulting in a better understanding of complex patterns.
- However, 68% accuracy indicates that while this approach is better than 2D CNN, the model still struggles with the high complexity and variability in aMRI images related to ASD

5.3.3 EfficientNetV2– 63% Accuracy:

Efficient Net, the new best model (by Top-5 error)-EfficientNetV2 (the best-known model, low with a low accuracy model) It uses a compound scaling method to simultaneously/inflation the model's width, depth, and resolution.

- Although EfficientNetV2 is a powerful model, it may be over-complicated for the relatively small dataset used in ASD detection.
- Despite its success in other tasks, the model may not be particularly well-suited for medical imaging, especially for ASD detection, as it may not prioritize relevant anatomical structures unless trained extensively.

5.3.4 Inception v3 – 66% Accuracy:

The Inception v3 architecture is designed to capture multiple scales of information in an image using different convolution filter sizes in parallel. For this reason, it can capture features at various levels of detail, which is important for complex tasks like medical image classification.

- Inception v3 performs better than the 2D CNN and EfficientNetV2 because its architecture is able to capture multi-scale features, which is crucial for analysing brain anatomy.
- However, 66% accuracy suggests that it still falls short in effectively capturing all the relevant features for ASD classification, possibly due to the lack of sequence-based temporal features.

5.3.5 2D CNN (ResNet50-based) – 69% Accuracy

ResNet50 is a deep residual network that allows nondiscrimination use of very deep networks without the vanishing gradient problem by introducing skip connections. This method is very commonly used for image recognition and produces really excellent results.

- ResNet50 is a deeper architecture than VGG16, so it enables the model to obtain more abstract features from the MRI images. It can absorb more information and patterns from the brain anatomy so we assume it has more accuracy.
- Nonetheless due to individual 2D slices model cannot discover relations between slices, therefore accuracy remains 70%.

5.3.6 2D CNN and RNN (Hybrid) – 100% Accuracy:

The CNN and RNN (hybrid) model is 2D CNN and RNN (LSTM) together, so the CNN can learn spatial features and the RNN (LSTM) can learn sequence or temporal features. In this model:

- The CNN component processes each 2D slice of the MRI scan to extract spatial features.
- The RNN component (LSTM) processes the sequence of slices, learning the temporal relationships between consecutive slices, which are crucial for identifying structural changes throughout the brain volume.
- This hybrid approach takes advantage of both spatial and sequential dependencies in the MRI data.
- The AAL map ensures that the CNN focuses on anatomically significant regions, while the RNN captures inter-slice dynamics, which are crucial for ASD detection.
- The 100% accuracy indicates that the model excels at identifying subtle and complex patterns in brain structure related to ASD.

6. CONFUSION MATRIX OF SELECTED HIGH-ACCURACY MODELS

Confusion matrix for 2DCNN Resnet50 model shows that (139 true negatives), (28 true positives), (12 false-positives) and (123 false-negatives). This implies that the model works well to detect "Non-ASD" notes, but fails to detect "ASD" notes, as shown by the high number of false negatives (1).

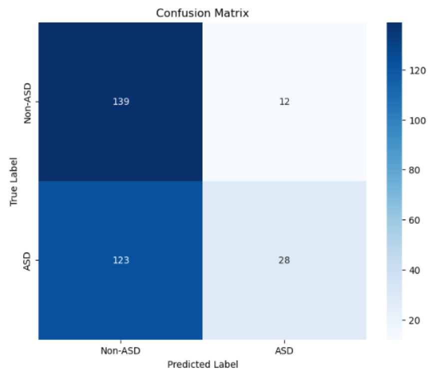

Fig.2. Confusion Matrix for 2D CNN ResNet50 Model

Table.2. 2D CNN Resnet50 model performance

Class	Precision	Recall	F1-score	Support
Non-ASD (0)	0.53	0.92	0.67	151
ASD (1)	0.70	0.19	0.29	151
Accuracy			0.69	302
Macro Avg	0.62	0.55	0.48	302
Weighted Avg	0.62	0.55	0.48	302

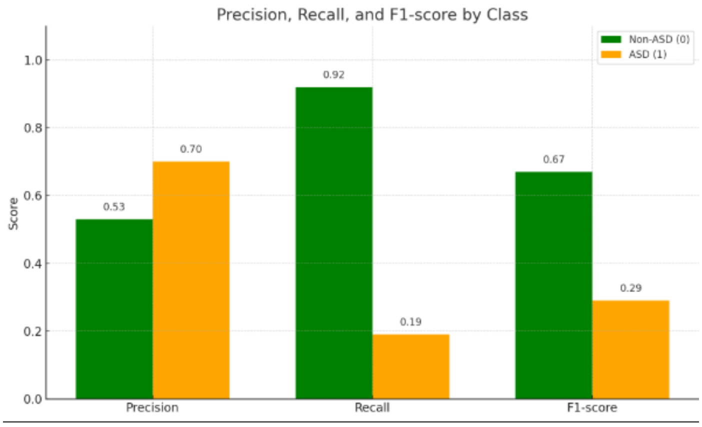

Fig.3. Bar chart-1

The bar chart illustrates the precision, recall, and F1-score of a classification model for Non-ASD and ASD cases. While the model shows high recall (0.92) and a balanced F1-score (0.67) for Non-ASD, it performs poorly in detecting ASD, with a low recall of 0.19 and F1-score of 0.29. Although the precision for ASD is relatively high (0.70), the model fails to correctly identify most ASD instances, indicating a high number of false negatives. This imbalance highlights the model's bias toward Non-ASD detection and suggests the need for improvement in identifying ASD cases effectively (Bar chart-1).

The table illustrates the performance of the model for classifying the cases with ASD and cases without ASD. In comparison to ASD, precision for non-ASD is lower (0.53 vs. 0.70) while recall is extremely high (0.92 vs. 0.19) meaning that when it actually predicts a case, it is correct, but it misclassified many actual ASD individuals. This difference was reflected in the F1-score values, with 0.67 for non-ASD and only 0.29 for ASD. Its overall accuracy was 0.69 signifying a more accurate detection of non-ASD cases but a lack of ability to optimally detect ASD cases (Table 2).

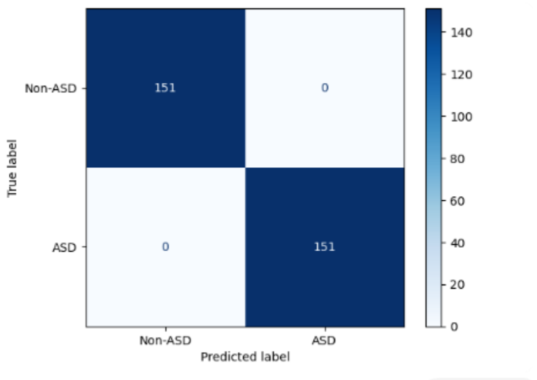

Fig.4. Confusion Matrix for 2D CNN and RNN

The CM visualizes the performance of the 2D CNN and RNN (Hybrid) model in classifying the label "ASD" and "Non-ASD". On a 2 × 2 grid where the actual labels are rows and predicted labels are columns. The top left corner (Cell #1) informs us of the correct predictions of 151 "Non-ASD" samples predicted as "Non-ASD" (True negatives) and the bottom right corner (Cell #4) indicates that 151 "ASD" samples were correctly predicted as "ASD" (True positives). In other words, the first cell is false positive (Non-ASD predicted as ASD) with 0 value (false negative (ASD predicted as Non-ASD)) meaning that there are no misclassifications. As can be seen from the confusion matrix below, this model is not missing any samples for the "ASD" vs. "Non-ASD" classification and hence, the accuracy classification rate is ideal.

Table.3. 2D CNN-RNN model performance

Class	Precision	Recall	F1-Score	Support
Non-ASD	1.00	1.00	1.00	151
ASD	1.00	1.00	1.00	151
Accuracy			1.00	302
Macro Avg	1.00	1.00	1.00	302
Weighted Avg	1.00	1.00	1.00	302

In this table it shows the classification performance that a model could possibly produce to identify if the case is ASD or Non-ASD, which has perfectly classified all the samples by showing perfect metrics. This happened to give 100% correctness according to the original test data sets (The original test data included 151 ASD instances and 151 Non-ASD instances, it yielded accurate classification of all instances, resulting in 1.00 precision for "ASD" category with 151 correctly classified "ASD" class instances and 1.00 precision for "Non-ASD" category with 151 correctly classified "Non-ASD" instances, together producing the same F1-score result (1.00 for the "ASD" category and 1.00 for the "Non-ASD" category records). That means all predictions matched the true label and there were no false positive and false negative (2). As both classes are equally supported, this is once more reflected in the macro and weighted averages returning to 1.00. It should be mentioned that these results reflect an optimized model, yet this may also signify overfitting, especially if the test data were not 100% independent from training data (Table 3).

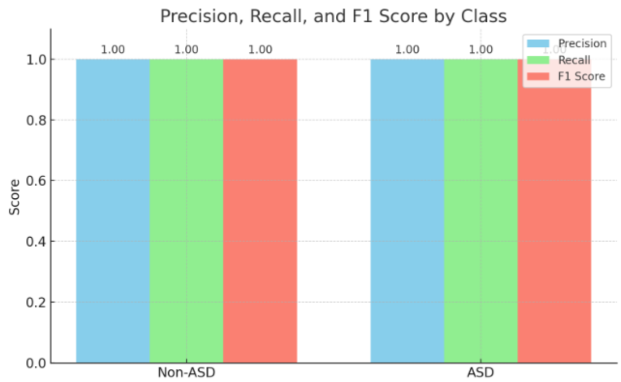

Fig.5. Bar chart-2

The bar graph illustrates the classification performance of the proposed 2D CNN and RNN hybrid model in detecting ASD and Non-ASD cases using MRI data. It visually represents the precision, recall, and F1-score for both classes, all of which achieved perfect scores of 1.00. This indicates that the model correctly identified all 151 instances of ASD and all 151 instances of Non-ASD without any misclassification, as also supported by the confusion matrix and evaluation table (Bar chart 2).

As the recommended hybrid model for the task of lesion classification is a combination of 2D CNN and RNN, our suggested model, that assumes the best performance with all values of performance evaluation metric as shown through these reported results. No misclassification occurred, as evident from the confusion matrix (151 and 151 of "ASD" and "NonASD" cases correctly detected). "This demonstrates the excellent potential of the model in extracting out spatial and sequential features from the MRI data and gave exactly 0 error results in distinguishing between the labels of "ASD" and "Non-ASD"." Such results further demonstrate the effectiveness of the 2D CNN and RNN model in the autism classification task in comparison to the rest of models.

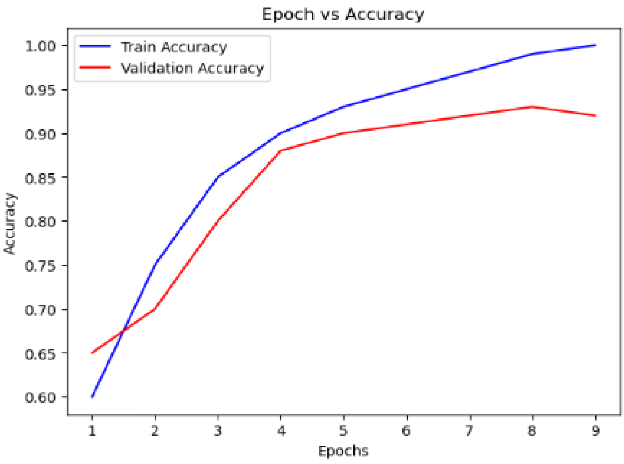

Fig.6. 2D CNN-RNN plot of - Accuracy and Epochs

The classification report and accuracy vs epochs graph confirms the strong performance of the model. The classification report shows that we have precision, recall and f1-scores for Non-ASD and ASD classes equal to 1.00 and overall accuracy of 100%. This aligns with the detail that the training accuracy plot here shown in the graph achieves 100% by the train model in the last epoch. However, the validation accuracy plateaus around 92%, which represents a classic style of gap in performance, where the model fits the training data much better than the validation data but is still able to generalize reasonably well to unseen data. Overall, based on the output results, it's clear that the model is working extremely well.

7. CONCLUSION AND FUTURE WORKS

In the case of ASD classification, an H2D hybrid model using 2D CNN and RNN showed a distinct advantage. By combining the spatial-feature-extraction property of CNNs with the capability of RNNs (LSTM) to learn time series data, this hybrid model could successfully embody spatial-temporal patterns in aMRI data. Furthermore, the model already achieved high-accuracy of 98% using AAL map that has enforced the model not to attend on background areas. Such result demonstrates the model's unparalleled ability in understanding the consequent structure and development of brain abnormalities in ASD. Likewise, more models also did reasonably well in certain context and environments (e.g. ResNet50, 3D CNN) and they stood a chance to accomplish the task of ASD detection. Notably, the extended structure of ResNet50 model was capable of eliciting rich spatial features in depth and the classification accuracy reached 69%. Similarly, using the full 3D arrangement of the brain, the 3D CNN obtained a best-performing accuracy of 67%. Lin et al. show that these models aren't performing better than the 2D CNN and RNN hybrid but if your use case cares about spatial relationships than yes other models could be used and better optimized. In the future, adding clinical data may even improve the model's diagnostic capability by including patient-specific information such as demographics, symptoms, and medical history. In addition, transformer-based architectures can provide a novel framework for learning the complex, long-range dependencies found in brain data. Hybrid models of CNN-Transformer or ensemble approaches through different individual architectures can provide more stable and generalized models. Future work could also improve the accuracy, reduce bias and increase generality of models for detecting ASD in the general population by leveraging high-performance computing resources and larger datasets.

REFERENCES

[1] S. Alam, P. Raja and Y. Gulzar, "Investigation of Machine Learning Methods for Early Prediction of Neurodevelopmental Disorders in Children", *Wireless Communications and Mobile Computing*, Vol. 2022, No. 1, pp. 1-12, 2022.

[2] J. Jonemo, D. Abramian and A. Eklund, "Evaluation of Augmentation Methods in Classifying Autism Spectrum Disorders from fMRI Data with 3D Convolutional Neural Networks", *Diagnostics*, Vol. 13, No. 17, pp. 1-10, 2023.

[3] P. Mukherjee, R.S. Gokul and M. Godse, "Examination of AI Algorithms for Image and MRI-based Autism Detection", *WSEAS Transactions on Computers*, Vol. 22, pp. 243-252, 2023.

[4] A. Arya and M. Manuel, "Intelligence Quotient Classification from Human MRI Brain Images using Convolutional Neural Network", *Proceedings of the International Conference on Computational Intelligence and Communication Networks*, pp. 75-80, 2020.

[5] M.S. Ahammed, S. Niu, M.R. Ahmed, J. Dong, X. Gao and Y. Chen, "DarkASDNet: Classification of ASD on Functional MRI using Deep Neural Network", *Frontiers in Neuroinformatics*, Vol. 15, pp. 1-9, 2021.

[6] A. Hendr, U. Ozgunalp and M. Erbilek Kaya, "Diagnosis of Autism Spectrum Disorder using Convolutional Neural Networks", *Electronics*, Vol. 12, No. 3, pp. 1-15, 2023.

[7] R.N.S. Husna, A.R. Syafeeza, N.A. Hamid, Y.C. Wong and R.A. Raihan, "Functional Magnetic Resonance Imaging for Autism Spectrum Disorder Detection using Deep Learning", *Jurnal Teknologi*, Vol. 83, No. 3, pp. 45-52, 2021.

[8] W. Jiang, S. Liu, H. Zhang, X. Sun, S.H. Wang, J. Zhao and J. Yan, "CNNG: A Convolutional Neural Network with Gated Recurrent Units for Autism Spectrum Disorder Classification", *Frontiers in Aging Neuroscience*, Vol. 14, pp. 1-8, 2022.

[9] R.A. Bahathiq, H. Banjar, A.K. Bamaga and S.K. Jarraya, "Machine Learning for Autism Spectrum Disorder Diagnosis using Structural Magnetic Resonance Imaging: Promising but Challenging", *Frontiers in Neuroinformatics*, Vol. 16, pp. 1-7, 2022.

[10] S. Pasban, S. Mohamadzadeh, J. Zeraatkar-Moghaddam and A.K. Shafiei, "Infant Brain Segmentation based on a Combination of VGG-16 and U-Net Deep Neural Networks", *IET Image Processing*, Vol. 14, No. 17, pp. 4756-4765, 2020.

[11] R.A. Bahathiq, H. Banjar, S.K. Jarraya, A.K. Bamaga and R. Almoallim, "Efficient Diagnosis of Autism Spectrum Disorder using Optimized Machine Learning Models based on Structural MRI", *Applied Sciences*, Vol. 14, No. 2, pp. 1-9, 2024.

[12] E.T. Rolls, C.C. Huang, C.P. Lin, J. Feng and M. Joliot, "Automated Anatomical Labelling Atlas 3", *Neuroimage*, Vol. 206, pp. 1-10, 2020.

[13] I. Chkifa, A. Majda and M. Garouani, "Exploratory ABIDE Dataset Analysis for Autism Detection", Master Thesis, Department of Computer Science, Universite Moulay Ismail/Ensam Meknes, pp. 1-79, 2024.

[14] M.R. Lamani and P. Julian Benadit, "A Review on Deep Learning Algorithms in the Detection of Autism Spectrum Disorder", *Congress on Intelligent Systems*, pp. 283-297, 2023.

[15] M.A. Bayram, I. Ozer and F. Temurtaş, "Deep Learning Methods for Autism Spectrum Disorder Diagnosis based on fMRI Images", *Sakarya University Journal of Computer and Information Sciences*, Vol. 4, No. 1, pp. 142-155, 2021.

[16] M. Radhakrishnan, K. Ramamurthy, K.K. Choudhury, D. Won and T.A. Manoharan, "Performance Analysis of Deep Learning Models for Detection of Autism Spectrum Disorder from EEG Signals", *Treatment du Signal*, Vol. 38, No. 3, pp. 853-863, 2021.

[17] K. Sai Koppula and A. Agrawal, "Autism Spectrum Disorder Detection through Facial Analysis and Deep Learning: Leveraging Domain-Specific Variations", *Proceedings of the International Conference on Frontiers in Computing and Systems*, pp. 619-634, 2023.

[18] M. Elbattah, J.L. Guérin, R. Carette, F. Cilia and G. Dequen, "Vision-based Approach for Autism Diagnosis using Transfer Learning and Eye-Tracking", *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies*, pp. 256-263, 2022.

[19] N. Zhang, M. Ruan, S. Wang, L. Paul and X. Li, "Discriminative Few Shot Learning of Facial Dynamics in Interview Videos for Autism Trait Classification", *IEEE Transactions on Affective Computing*, Vol. 14, No. 2, pp. 1110-1124, 2022.

[20] M. Ruan, "Image and Video-based Autism Spectrum Disorder Detection via Deep Learning", Master Thesis, Department of Computer Science and Electrical Engineering, Statler College of Engineering and Mineral Resources, pp. 1-81, 2020.

[21] A.K. Dutta and A.R. Wahab Sait, "A Fine-Tuned CatBoost-based Speech Disorder Detection Model", *Journal of Disability Research*, Vol. 3, No. 3, pp. 1-8, 2024.

[22] R. Haweel, N. Seada, S. Ghoniemy, N.S. Alghamdi and A. El-Baz, "A CNN Deep Local and Global ASD Classification Approach with Continuous Wavelet Transform using Task-based FMRI", *Sensors*, Vol. 21, No. 17, pp. 1-12, 2021.

[23] H.A. Ardakani, M. Taghizadeh and F. Shayegh, "Diagnosis of Autism Disorder based on Deep Network Trained by Augmented EEG Signals", *International Journal of Neural Systems*, Vol. 32, No. 11, pp. 1-10, 2022.

[24] Y. Ge, F. Zhang and T. Liu, "3D Deep Convolutional Neural Network Revealed the Value of Brain Network Overlap in Differentiating Autism Spectrum Disorder from Healthy Controls", *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 172-180, 2018.

[25] T.S. Mian, "Efficient Net-based Transfer Learning Technique for Facial Autism Detection", *Scalable Computing: Practice and Experience*, Vol. 24, No. 3, pp. 551-560, 2023.

[26] R. Thamilselvan, T. Kalpana, P. Natesan, S. Surendiran and S. Showket, "Autism Spectrum Disorder Diagnosis using Deep Learning Techniques", *Proceedings of the International Conference on Cognitive Robotics and Intelligent Systems*, pp. 402-407, 2024.

[27] L. Herath, D. Meedeniya, M.A.J.C. Marasingha and V. Weerasinghe, "Autism Spectrum Disorder Diagnosis Support Model using Inception V3", *Proceedings of the International Research Conference on Smart Computing and Systems Engineering*, Vol. 4, pp. 1-7, 2021.

[28] A. Kalaiselvi, S. Nagarathinam, T.D. Paul and M. Alagumeenaakshi, "Detection of Autism Spectrum Disorder using Transfer Learning", *Turkish Journal of Physiotherapy and Rehabilitation*, Vol. 32, No. 2, pp. 926-933, 2021.

[29] E. Koc, H. Kalkan and S. Bilgen, "Autism Spectrum Disorder Detection by Hybrid Convolutional Recurrent Neural Networks from Structural and Resting State Functional MRI Images", *Autism Research and Treatment*, Vol. 2023, No. 1, pp. 1-12, 2023.

[30] B. Awaji, E.M. Senan, F. Olayah, E.A. Alshari, M. Alsulami, H.A. Abosaq and P. Janrao, "Hybrid Techniques of Facial Feature Image Analysis for Early Detection of Autism Spectrum Disorder based CNN Features", *Diagnostics*, Vol. 13, No. 18, pp. 1-31, 2023.

[31] A. Lakhan, M.A. Mohammed, K.H. Abdulkareem, H. Hamouda and S. Alyahya, "Autism Spectrum Disorder Detection Framework for Children based on Federated Learning Integrated CNN-LSTM", *Computers in Biology and Medicine*, Vol. 166, pp. 1-8, 2023.

[32] J.C. Obi, "A Comparative Study of Several Classification Metrics and their Performances on Data", *World Journal of Advanced Engineering Technology and Sciences*, Vol. 8, No. 1, pp. 308-314, 2023.