**Cairo University**

**Faculty of Graduate Studies for Statistical Research**

Early Detection of Autism Spectrum Disorder in Children: Deep Learning Approach

**A Project Presented For fulfillment**

**For Master project in Software Engineering**

# Submitted by:

Mohammed Seliem

Code No. 202301530

# Supervised by:

**Dr. Ashraf shaheen.**

**2024**

Early Detection of Autism Spectrum Disorder in Children: Deep Learning Approach

Mohammed Seliem

Software Engineering Dept. - Faculty of Graduate Studies for Statistical Research

Cairo, EYGPT

12422022665557@pg.cu.edu.eg

***Abstract*—Autism Spectrum Disorder (ASD) represents a significant and growing challenge globally, affecting countless individuals and their families. This study advances the early detection of (ASD) in children by implementing sophisticated deep learning models that analyze video data, a significant enhancement over conventional static image analysis. By employing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), our research deciphers subtle behavioral cues and dynamic patterns indicative of ASD, which are often missed by traditional methods. Our findings indicate a notable performance disparity between the two models, with RNNs significantly outperforming CNNs, achieving an accuracy rate of up to 50% after extensive training epochs. The implications of these findings underscore the potential of video-based analytics in transforming diagnostic practices for ASD, promising to improve the accuracy and timeliness of interventions and thereby potentially altering developmental trajectories for affected children.**

***Keywords—Autism Spectrum Disorder, Deep Learning, Convolutional Neural Networks, Recurrent Neural Networks, Early Detection, Behavioral Analysis, Video Analytics.***

## Introduction

Autism Spectrum Disorder (ASD) represents a significant and growing challenge globally, affecting countless individuals and their families[1]. The imperative for early detection is well-documented, as timely interventions can markedly improve the developmental outcomes and overall quality of life for children diagnosed with ASD. Traditionally, the diagnosis of ASD has depended heavily on detailed behavioral assessments and neurological imaging—techniques that are not only resource-intensive but also prone to variability in diagnostic accuracy and accessibility [2].

With the advent of advanced computational models, particularly deep learning technologies, new avenues have opened for enhancing the detection of ASD. This research project leverages these technologies, specifically CNNs and RNNs, to analyze extensive video data capturing the nuanced behaviors of children. This methodological pivot aims to address the limitations of traditional diagnostic tools by enabling the detection of more naturalistic and spontaneous behaviors through dynamic video analysis.

Various methodologies for early ASD detection have been applied. Machine learning techniques, such as SVM and RF, demonstrated exceptional accuracy in rapid ASD screening, achieving 100% and 97.07% accuracy (H. Alkahtan et al.; A. Y. Saleh and Lim Huey Chern). Deep learning approaches using EEG signals and eye-tracking scan paths showed high classification accuracy of 95.59% (N. A. Ali et al.; M. Alsaidi et al.). Multimodal and hybrid systems combining EEG and eye-tracking data improved diagnostic accuracy, reaching up to 99.8% accuracy (J. Han et al.; I. A. Ahmed et al.). Utilizing novel datasets and advanced preprocessing techniques in facial image analysis also achieved high accuracy, with results up to 96% (F. Alsaade and M.S. Alzahrani; Zeyad A. T. Ahmed et al.).

The primary aim of this study is to validate and refine deep learning models to improve their effectiveness in detecting early behavioral signs of ASD. The objectives are structured to:

1. Develop a deep learning framework capable of extracting relevant behavioral patterns from video data.
2. Evaluate and compare the diagnostic utility of CNNs and RNNs in identifying ASD-specific behaviors.
3. Conduct rigorous testing and validation to assess the accuracy of these models in real-world settings.

This paper provides a critical overview of the current landscape of ASD detection methodologies, emphasizing the need for more scalable and efficient diagnostic solutions. Our innovative approach promises not only to enhance diagnostic accuracy but also to make early detection more accessible, thus facilitating earlier interventions. Detailed discussions on the methodology encompass data collection, preprocessing, model training, and validation processes, providing a comprehensive blueprint of our approach.

By outlining the structure of this paper, the introduction prepares readers for an in-depth exploration of cutting-edge methods in ASD diagnosis. It highlights the transformative potential of integrating deep learning with traditional diagnostic practices, setting the stage for a nuanced discussion on how these technologies can revolutionize the field of developmental disorder diagnostics.

This paper is structured into several sections to provide a comprehensive analysis of Early Detection of Autism Spectrum Disorder in Children: Deep Learning Approach. In the following section, we offer a detailed overview of the history and background of Autism Spectrum Disorder (ASD) to clarify its meaning. This comprehensive introduction ensures that readers can understand the content of the research without needing to consult additional sources. This revision aims to enhance readability and make the purpose of the section more explicit to the reader. Section 3 reviews related work, we explore and summarize previous studies and findings on Autism Spectrum Disorder (ASD). This review establishes a scholarly context and highlights the contributions and gaps in existing research, enabling readers to grasp the advancements in the field without consulting external sources. In Section 4, the methodology is outlined, detailing data sources, variables considered, and the application of machine learning techniques for predictive modeling. Section 5 presents Result Analysis and Discussion, comparing findings from different algorithms and addressing implications and limitations. The Conclusion and Future Work section in Section 6 summarizes key findings, discusses implications, and suggests ways for future research. Finally, Section 7 contains the References, listing all sources cited in the paper including research papers, articles, books, and online resources, providing readers with access to the literature consulted during the research process.

1. Background

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by challenges in social interaction, communication, and restricted, repetitive behaviors. The term "spectrum" reflects the wide range of symptoms and severity among individuals. The history of ASD traces back to early observations in the 20th century. In 1943, Leo Kanner, an American psychiatrist, first described "early infantile autism" based on case studies of 11 children who exhibited profound social withdrawal and communication difficulties. Simultaneously, Hans Asperger, an Austrian pediatrician, identified a milder form of the disorder, later termed Asperger's Syndrome, noting children with intense interests and advanced language skills but impaired social interactions ([Kanner, 1943; Asperger, 1944](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1510731/)).

The understanding and classification of ASD have evolved significantly over the decades. Initially considered rare, with an estimated prevalence of 1 in 10,000, ASD's diagnostic criteria and recognition have broadened. This shift led to a dramatic increase in reported cases, now estimated at 1 in 54 children by the Centers for Disease Control and Prevention (CDC) as of 2020. The rise in prevalence is attributed to enhanced awareness, better diagnostic tools, and changes in diagnostic criteria, rather than a true epidemic ([CDC, 2020](https://www.cdc.gov/ncbddd/autism/data.html)).

ASD's exact cause remains elusive, though it is believed to result from a combination of genetic and environmental factors. Research has identified numerous genetic variants associated with the disorder, alongside prenatal and perinatal environmental influences. Understanding ASD's multifaceted nature continues to be a critical focus of scientific inquiry, aiming to improve diagnosis, intervention, and support for individuals and families affected by the disorder ([NIH, 2021; NIEHS, 2021](https://www.niehs.nih.gov/research/supported/health/autism)).

## Related work

In this section, we review some related works about (SAD) early detection, classifiers and Screening using different approaches and methodologies. H. Alkahtan, el [2] presented a significant contribution in early screening for (ASD). By leveraging machine learning algorithms, particularly support vector machine (SVM) and random forests (RF) approaches, a fast and accurate screening method was developed. The proposed system offers a simple and accessible solution for identifying early indicators of ASD in children. The research methodology involves utilizing machine learning methods, including k-nearest neighbors, linear discriminant analysis, (SVM), and (RF), to categorize ASD data. Standard datasets containing two classes, normal and autism, were obtained from the machine learning repository. The dataset was divided into a training set (80%) and a testing set (20%). The SVM and RF algorithms were examined and tested, achieving exceptional accuracy levels of 100%. Sensitivity analysis was also conducted to estimate the prediction accuracy of the algorithms. This research lies in the utilization of state-of-the-art machine learning algorithms and AI techniques to develop an accurate and accessible screening system for early identification of ASD. However, it is essential to consider potential limitations, such as the reliance on specific datasets and the need for further validation on diverse populations. Further research and validation are necessary to enhance the generalizability and effectiveness of the algorithms [2] .

A. Y. Saleh and Lim Huey Chern [3] improved classification methodologies in this domain. Beginning with an overview of ASD and its associated challenges, including communication and social impairments, the paper introduces deep learning as a promising solution. Specifically, convolutional neural networks (CNNs) are employed for ASD classification using pre-processed image data, with a detailed exposition of the algorithm's implementation and subsequent evaluation based on accuracy metrics. A comparison with the support vector machine (SVM) algorithm demonstrates the superior performance of CNN, achieving an accuracy of 97.07%. Emphasizing the importance of exploring diverse deep learning algorithms and datasets with varying parameters, the paper also discusses previous research utilizing EEG signals and functional MRI data for ASD classification. It underscores the growing prevalence of ASD, its genetic underpinnings, and the complex heterogeneity among affected individuals. Methodologically, the paper outlines a structured approach to conducting literature reviews, categorizing related works, and presenting findings comprehensively. Recommendations include organizing the related work section logically, incorporating visual aids for clarity, and providing concise summaries for each category [3] .

N. A. Ali, el [4] had addressed the challenging problem of ASD diagnosis by proposing a deep learning algorithm for classifying ASD based on EEG signals. which provide valuable information about brain activity in ASD. The results demonstrate the algorithm's effectiveness in distinguishing between autistic and normal children. Further research is warranted to address the limitations and explore the applicability of this algorithm in real-world clinical settings. And the study provides promising insights into employing deep learning algorithms for future developments in autism diagnosis and monitoring treatment effectiveness. Weaknesses are reliance on a specific database for training and testing the algorithm, which may limit generalizability, Lack of information on the size and diversity of the dataset used in the study. And the study's reliance on accuracy as the sole performance measure, without considering other evaluation metrics [4] .

H. Sewani and R. Kashef [5] collected a large brain image dataset from the Autism Brain Imaging Data Exchange (ABIDE) for their study. The dataset included 539 individuals with ASD and 573 typically developing control subjects. The dataset was preprocessed using the CPAC pipeline, which involved various steps such as slice time correction, motion correction, nuisance signal removal, low-frequency drift, and voxel intensity normalization. this research presents an autoencoder-based deep learning classifier for efficient diagnosis of (ASD). The hybrid approach combining unsupervised and supervised learning demonstrates improved accuracy in identifying ASD. The study contributes to the advancement of machine learning and neural network methods for autism diagnosis, with the potential for early intervention and improved outcomes for individuals on the autism spectrum. The strengths of the approach include its ability to handle the dynamic changes in autism behavior patterns and its high accuracy in diagnosing ASD, especially in children. The use of deep learning techniques and large brain image datasets contributes to the robustness of the classifier. However, the research also acknowledges some limitations, such as the potential noise in the image dataset from multiple sites and the computational time required for training the model [5] .

M. Alsaidi, el [6] used a convolutional deep neural network (CNN) approach. They created model consists of two hidden layers with 300 and 150 neurons, respectively, and underwent 10 rounds of cross-validation with a dropout rate of 20%. Eye-tracking scan paths were used as input data for training and testing the model. The study collected data from participants diagnosed with ASD and TD, and the eye gaze coordinates were obtained using an eye-tracking device. The proposed T-CNN-ASD model achieved an accuracy of 95.59% in classifying children with ASD from those with TD during the testing phase. This accuracy outperformed other machine learning algorithms such as random forest (RF), decision tree (DT), K-Nearest Neighbors (KNN), and multi-layer perceptron (MLP). The study demonstrated that the deep learning model utilizing eye-tracking scans can accurately classify ASD without the need for human intervention. The results also showed superior performance compared to previous studies in this domain. The research contributes to the growing body of knowledge in the field of automated ASD diagnosis. However, some limitations include the need for further validation on larger and diverse datasets, potential biases in the collected data, and the reliance on eye-tracking technology, which may have limitations in certain populations or contexts [6] .

F. Alsaade and M.S. Alzahrani [7] utilized deep learning algorithms, specifically Xception, Visual Geometry Group Network (VGG19), and NASNETMobile, for the classification task. The dataset used for testing the models was collected from the Kaggle platform and consisted of 2,940 face images. The dataset was split into two groups, autistic children and non-autistic children, with equal representation in each group. The evaluation of the three deep learning models, Xception, VGG19, and NASNETMobile, was based on standard evaluation metrics such as accuracy, specificity, and sensitivity. The Xception model achieved the highest accuracy result of 91%, followed by VGG19 (80%) and NASNETMobile (78%). These results indicate the effectiveness of deep learning algorithms in detecting and classifying ASD based on facial features. One of the strengths of this research is the utilization of deep learning techniques for ASD detection, which allows for the extraction of subtle facial patterns that might not be discernible to the human eye. Additionally, the use of transfer learning and pre-trained models enhances the classification performance. However, some limitations include the reliance on facial recognition as the primary means of detecting autism, as well as the dependency on the availability and quality of the training dataset [7].

I. A. Ahmed, el [8] used Three artificial intelligence techniques: machine learning, deep learning, and a hybrid approach combining both in the study by the first technique utilized feedforward neural networks (FFNNs) and artificial neural networks (ANNs) based on feature classification extracted using a hybrid method of local binary pattern (LBP) and grey level co-occurrence matrix (GLCM) algorithms. The second technique involves using pre-trained convolutional neural network (CNN) models, specifically GoogleNet and ResNet-18, for deep feature map extraction. The third technique combines deep learning (GoogleNet and ResNet-18) and machine learning (SVM) in a hybrid approach called GoogleNet + SVM and ResNet-18 + SVM. The datasets used are expected to consist of eye-tracking data collected from children with ASD and typically developing children for comparison. The research paper reports the achieved accuracies of the proposed techniques. The first technique, based on FFNNs and ANNs, achieves a high accuracy of 99.8%. The second technique, utilizing GoogleNet and ResNet-18 models, achieves performances of 93.6% and 97.6%, respectively. The third technique, combining GoogleNet or ResNet-18 with SVM, achieves accuracies of 95.5% and 94.5%, respectively While the research paper demonstrates promising results, there are a few limitations that should be acknowledged. First, the specific details regarding the datasets used, such as sample size, data collection protocols, and data preprocessing techniques, are not explicitly discussed. These details are important for understanding the generalizability of the proposed techniques. Second, the paper does not provide a comparative analysis with existing methods or state-of-the-art approaches for ASD diagnosis. Such a comparison would further validate the effectiveness of the proposed techniques [8].

J. Han, el [9] proposed a new multimodal diagnosis framework that combines electroencephalogram (EEG) and eye-tracking (ET) data. To evaluate the proposed method, the authors collected a multimodal dataset consisting of 40 ASD children and 50 TD children. Experimental results demonstrate that the multimodal identification model outperforms two unimodal methods and a simple feature-level fusion method. The proposed method achieves superior performance in identifying ASD, showing promising potential for objective and accurate diagnosis. the study lies in its multimodal approach, which combines EEG and ET data to enhance identification performance. The proposed method captures complementary information from both neurophysiological and behavioral modalities. However, the sample size of the dataset used in the study may affect the generalizability of the results. Additionally, further validation with larger and more diverse datasets is necessary to assess the robustness of the proposed method [9] .

Zeyad A. T. Ahmed, El [10] employed a deep learning approach, specifically a CNN with transfer learning, to develop a web application for autism detection. The researchers utilized three pretrained models, MobileNet, Xception, and InceptionV3, for classifying facial images as autistic or nonautistic. The dataset used in the study consisted of 3,014 facial images obtained from the Kaggle database. The images represented a heterogeneous group of children, with 1,507 autistic children and 1,507 nonautistic children. The accuracy of the classification models was evaluated using the validation data. MobileNet achieved an accuracy of 95%, Xception achieved 94%, and InceptionV3 attained 0.89%. These results indicate the effectiveness of the deep learning models in accurately identifying autistic children based on their facial features. the dataset used in the study may not represent the entire spectrum of autistic children, as it consisted of a specific set of facial images from a publicly available source. Additionally, the study focused solely on facial features and did not incorporate other diagnostic techniques, such as brain MRI or eye-tracking. Further research is needed to validate the findings and explore the integration of multiple diagnostic approaches for comprehensive autism detection [10] .

S. Gautam, el [11] had Screened Autism Spectrum Disorder in Children Using Deep Learning. The study utilized the YOLOv8 model and employs a dataset obtained from Kaggle. The dataset consists of facial images of ASD and TD children. The YOLOv8 model is trained on this dataset to classify the images into autistic and non-autistic categories. The training process involves optimizing the model's parameters using deep neural networks and CNNs. The study utilizes the YOLOv8 model and employs a dataset obtained from Kaggle. The dataset consists of facial images of ASD and TD children. The YOLOv8 model is trained on this dataset to classify the images into autistic and non-autistic categories. The training process involves optimizing the model's parameters using deep neural networks and CNNs. While the study demonstrates promising outcomes, there are several limitations to consider. First, the dataset used may not fully represent the diverse population of children with ASD. The generalizability of the model to different demographics and ethnicities needs further investigation. Additionally, the study primarily focuses on facial features and does not consider other potential indicators of ASD, such as behavioral patterns or genetic factors. Further research is necessary to develop a comprehensive screening approach [11].

Haishuai Wang and Paul Avillach [12] utilized genomics data from the Simons Simplex Collection, a collection of simplex families at risk for ASD. Preprocessing techniques were applied to extract common variants that may have a protective or pathogenic effect on autism. A chi-square test was employed to identify significant common variants. A convolutional neural network (CNN)-based diagnostic classifier was designed using these significant variants to predict autism. The performance of the deep learning model was compared with shallow machine learning models and randomly selected common variants. The analysis revealed a significant enrichment of contributory common variants on chromosome X, with chromosome Y also showing discriminatory potential in distinguishing autistic individuals from nonautistic individuals. The ARSD, MAGEB16, and MXRA5 genes were found to have the largest effect among the contributory variants. The deep learning model achieved an area under the receiver operating characteristic curve of 0.955 and an accuracy of 88% in identifying autistic individuals. This represents a substantial improvement of approximately 13% compared to standard autism screening tools [12].

**A. Garg, el** [13] utilized a hybrid approach of deep learning and XAI. The deep learning model is trained on a screening test dataset to predict ASD at an early stage. The XAI technique is then applied to identify the features that contribute most significantly to the accurate prediction of ASD. The research employs performance metrics such as accuracy, precision, recall, and F1-score to compare different feature sets and evaluate the effectiveness of the proposed approach The study utilized a dataset that includes various physical and physiological parameters related to ASD. The characteristics of the dataset are described in detail in the article. The dataset used in the experiments is state-of-the-art and provides a comprehensive set of features for ASD prediction. The experiments conducted in this study demonstrate the effectiveness of the proposed deep learning model in predicting ASD at an early stage. The application of XAI techniques helps identify the most contributing features for accurate prediction. Comparative case studies are performed using different feature sets, and the performance metrics indicate the superiority of the proposed approach. The results show improved accuracy, precision, recall, and F1-score, providing valuable clinical assistance for better and early prediction of ASD traits in toddlers One limitation is that the research relies on a specific dataset, and the generalizability of the findings to other populations or datasets should be further investigated. Additionally, the study focuses on the prediction aspect and does not address the underlying causes or mechanisms of ASD. Further research is required to explore these aspects and enhance our understanding of the disorder [13] .

**B. Elshoky, el** [14] utilized a dataset consisting of 2,936 facial images collected from autistic children and typically developing children. The analysis involves classical ML algorithms such as support vector machine (SVM) and random forest, as well as deep learning techniques including convolutional neural networks (CNN) with transfer learning. Additionally, automated machine learning (AutoML) methods are employed, which automatically adjust hyperparameters and select the best ML algorithm to achieve optimal performance. The results indicate that AutoML outperforms the other approaches, achieving an accuracy of approximately 96% using the Hyperpot and tree-based pipeline optimization tool. This demonstrates the effectiveness of AutoML in ASD classification using facial images. It eliminates the need for manual feature engineering and enables efficient parameter tuning, thereby enhancing the overall performance of the ML models. The study's findings highlight the advantages of employing AutoML techniques for ASD classification. The automated parameter optimization and algorithm selection process saves time and effort, enabling researchers to focus on other aspects of the analysis. However, it is important to consider the limitations of the study. The dataset used may have certain biases, and the generalizability of the models to diverse populations should be further explored. Additionally, the interpretability of the automated models may be a challenge, requiring additional efforts to understand the underlying decision-making process[14].

**Suman Raj and Sarfaraz Masood** [15] they presented an analysis of autism spectrum disorder using machine learning techniques. The study demonstrates the effectiveness of Convolutional Neural Network-based prediction models in accurately detecting ASD in children, adults, and adolescents. While the results are promising, further research is needed to address the limitations and refine the proposed techniques for real-world clinical applications. The research methodology employed in this study involved the following steps: Dataset Selection: Three publicly available non-clinical ASD datasets were chosen for analysis. These datasets contained information related to ASD screening in children, adults, and adolescents. Data Preprocessing: The datasets underwent preprocessing steps to handle missing values and ensure data quality. Various techniques were applied to address missing values and ensure accurate analysis. Algorithm Selection: Several machine learning algorithms were selected for comparison, including Naïve Bayes, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, Neural Network, and Convolutional Neural Network. These algorithms were chosen based on their suitability for classification tasks and their previous success in medical research. Model Training and Evaluation: The selected algorithms were trained on the preprocessed datasets and evaluated using appropriate evaluation metrics such as accuracy. The performance of each algorithm was assessed based on its ability to predict ASD accurately [15] .

**Adnan Ashraf, el** [16] utilized advanced MRI techniques such as rs-fMRI, gray matter (GM) MRI, and white matter (WM) MRI to diagnose brain abnormalities associated with ASD. The research leverages data-driven approaches and deep learning models, specifically convolutional neural networks (CNNs) and transfer learning algorithms. The ABIDE datasets, comprising rs-fMRI data from both typical control and autism individuals, are used for training and evaluation. The four-dimensional nature of rs-fMRI data (three spatial dimensions and one temporal dimension) allows for the development of diagnostic biomarkers for brain dysfunction. The proposed optimized CNN model achieved an accuracy of 81.56% in classifying and detecting early age ASD using the ABIDE datasets. This outperforms previous conventional approaches that were evaluated only on the ABIDE I datasets. The study's results highlight the potential of deep learning and transfer learning techniques in accurately identifying ASD based on brain imaging data. the study has certain limitations. First, the research relies on the availability of high-quality and adequately annotated brain imaging datasets. The accuracy of the model is dependent on the quality and representativeness of the training data. Second, the study focuses on early age ASD detection, and further investigation is needed to assess the model's performance across different age groups and ASD subtypes. Additionally, the generalizability of the proposed approach to diverse populations and imaging protocols should be explored [16].

**Table 1.**  literature review summary table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Author | Algorithm | Diseases | Dataset | Accuracy |
| 2023 | (H. Alkahtan, el). [2] | Machine Learning (SVM) (RF) | ASD | ASDTests | 52% |
| 2021 | (A. Y. Saleh and Lim Huey Chern) [3] | (CNNs) (SVM) | ASD | Autistic Children Facial Dataset | 83% |
| 2020 | (N. A. Ali, el ) [4] | Deep Learning Algorithms | ASD | **EEG signal** | 80% |
| 2020 | (H. Sewani and R. Kashef ) [5] | Machine Learning (CNNs) | ASD | ASD datasets (UCI) repository | 84.05 |
| 2024 | (M. Alsaidi, el) [6] | A Convolutional Deep Neural Network (CNN) | ASD | 2940 face photos | 75% |
| 2022 | (F.W.Alsaade and M.S. Alzahrani) [7] | Deep Learning Algorithms | ASD | 2,940 face images | 80% |
| 2022 | (I. Ahmed, el) [8] | A Convolutional Deep Neural Network (CNN) | ASD | ASD dataset | 94.5% |
| 2022 | (J. Han, el). [9] | Multimodal Diagnosis Framework | ASD | 40 ASD children and 50 typically developing | 82.5%) |
| 2022 | (Z. Ahmed, el) [10] | A Convolutional Deep Neural Network (CNN) | ASD | 3,014 facial images | 95%, |
| 2023 | (Subash Gautam, el) [11] | A Convolutional Deep Neural Network (CNN) | ASD | 2,654 facial images | 89.30 |
| 2021 | (Haishuai Wang and Paul Avillach) [12] | A Convolutional Deep Neural Network (CNN) | ASD | SSC data 2600 simplex families | *0.886* |
| 2022 | (A. Garg, el) [13] | A hybrid Approach Of Deep Learning And XAI | ASD | 1758  children/toddlers recorded and given by Fadi Thabtah | 0.79 |
| 2021 | (B. Elshoky, el) [14] | (CNNs) (SVM) | ASD | 2,936 facial images | 96% |
| 2020 | (Suman Raj and Sarfaraz Masood) [15] | Naïve Bayes, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, Neural Network, and Convolutional Neural Network | ASD |  |  |
| 2023 | ( Adnan Ashraf, el) [16] | A Convolutional Deep Neural Network (CNN) and transfer learning algorithms | ASD | ABIDE | 81.56% |

4- Methodology

4.1 Introduction to the Research Design

In this research on early detection of autism spectrum disorder (ASD) using state-of-the-art deep learning techniques, I primarily utilized a well-structured experimental design to train models on identifying potential autism signs in children. By combining deep learning architectures such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), we analyzed video data capturing subtle behaviors indicative of ASD. This comprehensive approach included extensive data handling, preprocessing, feature extraction, and model validation. Utilizing datasets of recorded interactions from children diagnosed with ASD, we focused on extracting motor signals and patterns from their behaviors, aiming to develop accurate diagnostic tools for early ASD detection. This methodology not only facilitates the creation of effective detection systems but also allows for thorough model training and evaluation, enhancing the reliability of our diagnostic tools. [17]

4.2 Data Preprocessing

we used the [Self-Stimulatory Behavior Dataset (SSBD)](https://rolandgoecke.net/research/datasets/ssbd/) sourced from Kaggle, consisting of 75 videos collected from publicly accessible platforms like YouTube. Some videos were missed when we downloaded manually because it's deleted by YouTube. These videos, posted by parents and caregivers, depict children with autism exhibiting typical self-stimulatory behaviors such as arm flapping, head banging, and spinning. Each video, with an average duration of 90 seconds, is meticulously annotated to identify these behaviors, presenting a challenge for automatic behavior analysis due to their uncontrolled natural settings. This dataset, which marks the first public release of children's videos recorded in natural settings showing self-stimulatory behaviors, provides a rich resource for studying these behaviors and developing behavior analysis methods. [18]

4.2.1 Preprocessing Steps

Preprocessing the raw video data is crucial for effective analysis. The steps implemented are as follows: **Fig. 1**

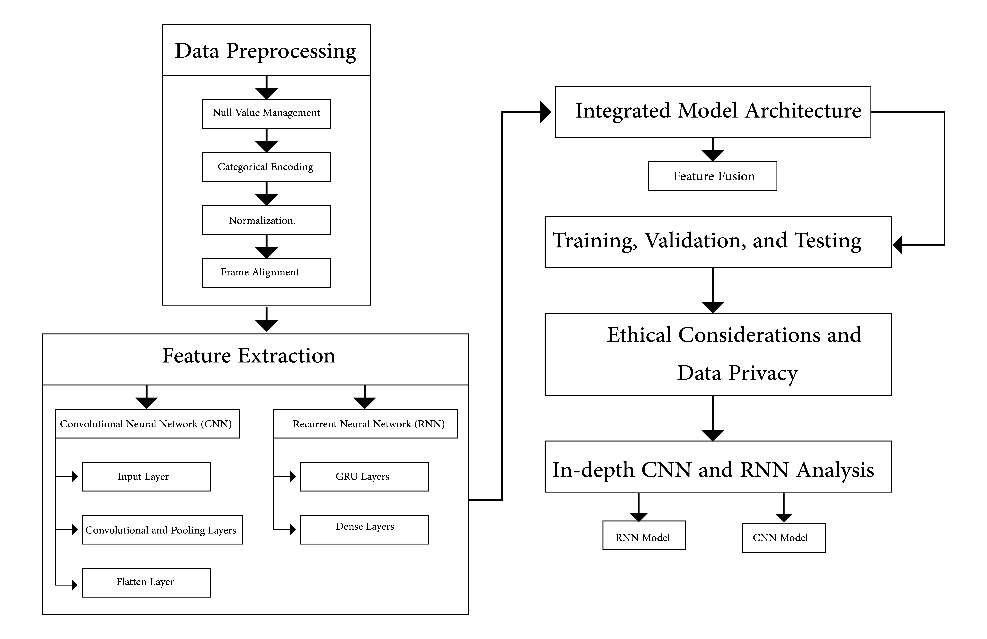


Fig. 1. Work Architecture

1. Null Value Management: We ensured the dataset was free of missing values by implementing checks and handling missing data appropriately to prevent skewing the analysis.Fig. 2.

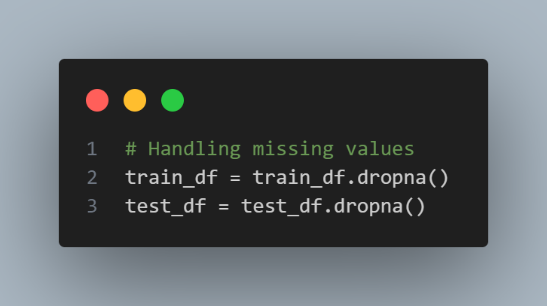


Fig. 2. Null Value Management.

1. Categorical Encoding: The categorical labels were transformed into numerical formats using one-hot encoding, facilitating their use in machine learning algorithms.Fig. 3**.**

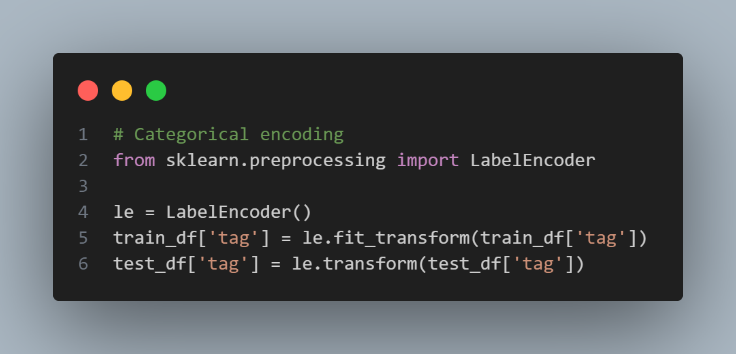


Fig. 3. Categorical Encoding.

1. Normalization: Data features were standardized to mitigate biases related to the scale of different features, using min-max normalization to bring all values into a common range. Fig. 3.

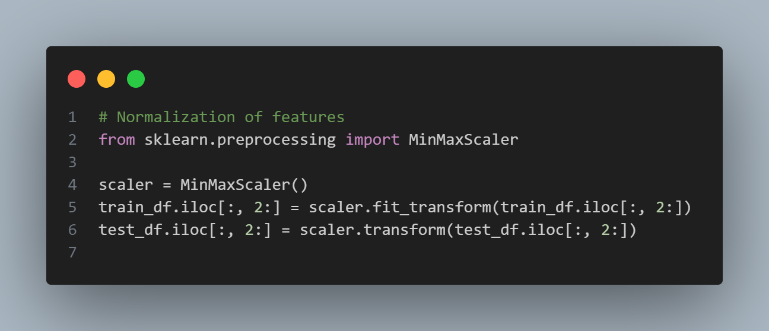


Fig. 4. Normalization.

1. Frame Alignment: Each video frame was aligned with its corresponding behavioral annotation, ensuring structured data for effective feature extraction. This involved synchronizing frames to behavior timestamps using precise frame indexing techniques. Fig. 5.



Fig. 5. Frame Alignment.

**4.3 Feature Extraction.**

**4.3.1 Convolutional Neural Network (CNN)**

**The CNN model was designed to extract spatial features from video frames. The architecture includes:**

1. **Input Layer:** The preprocessed frames were formatted to uniform dimensions (e.g., 224x224 pixels) to standardize the input for analysis. Fig. 6.
2. - الطبقة الداخلة: تم تنسيق الإطارات المعالجة مسبقًا إلى أبعاد موحدة (مثل 224x224 بكسل) لتوحيد الإدخال للتحليل.

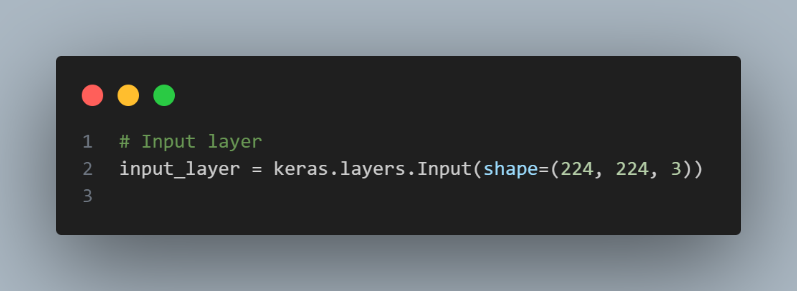
****

Fig. 6. Input Layer.

1. **Convolutional and Pooling Layers:** Layers of learnable filters and pooling operations were applied to extract and condense spatial features. This involved several convolutional layers with ReLU activation followed by max-pooling layers to reduce data dimensionality while retaining essential information. Fig 7.

الطبقات التحويلية والتجميعية: تم تطبيق طبقات من المرشحات التي يمكن تعلمها وعمليات التجميع لاستخراج وتجميع الميزات المكانية. تتضمن ذلك عدة طبقات تحويلية مع تنشيط ReLU تليها طبقات تجميع للحد من بعد البيانات مع الاحتفاظ بالمعلومات الأساسية.

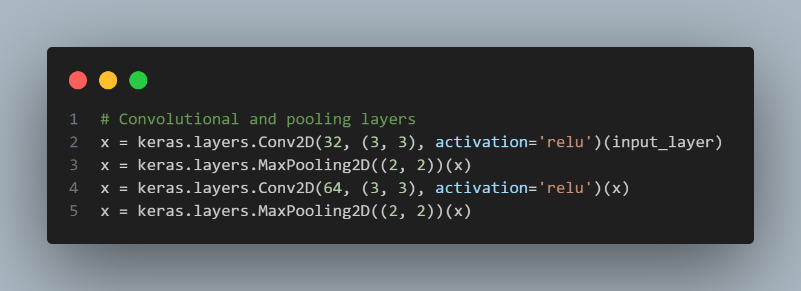
****

Fig. 7. Convolutional and Pooling Layers.

1. **Flatten Layer**: The 3D feature maps were converted into 1D feature vectors, preparing them for dense layer processing. This step was crucial for transforming spatial data into a format suitable for the subsequent RNN analysis. Fig. 8.

الطبقة المسطحة: تم تحويل خرائط الميزات ثلاثية الأبعاد إلى متجهات ميزات ثنائية الأبعاد، مما يعدها لمعالجة الطبقات الكثيفة. كان هذا الخطوة حاسمة لتحويل البيانات المكانية إلى تنسيق مناسب للتحليل الرياضي التابع للتحويل المتكرر.

****

**Fig. 8. Flatten Layer.**

4.3.2 Recurrent Neural Network (RNN)

The RNN model, specifically utilizing Gated Recurrent Units (GRUs), processed the temporal sequence of frames:

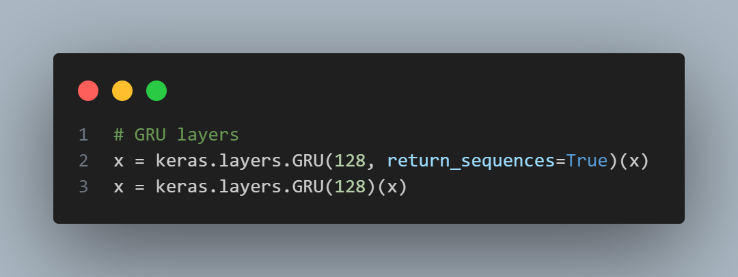
1. GRU Layers: GRU layers were employed to capture temporal patterns and dependencies, essential for understanding the sequence and progression of behaviors. These layers processed the feature vectors output by the CNN, capturing the dynamic changes over time. **Fig. 9.**
2. 4.3.2 الشبكة العصبية التابعة للتحويل المتكرر (RNN)
3. تم استخدام نموذج الـRNN، وبالتحديد استخدام وحدات التكرار المتوجة (GRUs)، لمعالجة التسلسل الزمني للإطارات:
4. 1- طبقات GRU: تم استخدام طبقات GRU لالتقاط الأنماط الزمنية والتبعيات، الأمر الأساسي لفهم التسلسل وتقدم السلوكيات. تم معالجة متجهات الميزات الناتجة عن الـCNN بهذه الطبقات، مما يلتقط التغيرات الديناميكية مع مرور الوقت.
5. 2

Fig. 9. GRU Layers.

1. Dense Layers: Further processing and refinement of the features extracted by GRUs were conducted using fully connected dense layers, leading to the final classification of behaviors. Fig. 10.
2. - طبقات كثيفة: تم إجراء مزيد من المعالجة والتنقيح للميزات المستخرجة بواسطة GRUs باستخدام طبقات كثيفة متصلة بالكامل، مما يؤدي إلى التصنيف النهائي للسلوكيات.

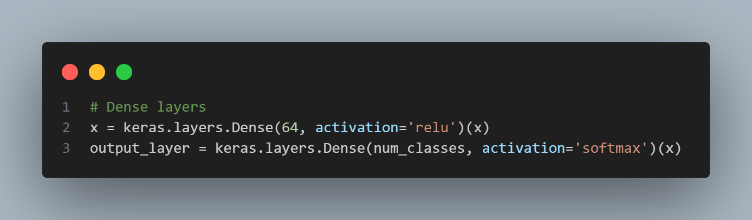


Fig. 10. Dense Layers.

4.4 Integrated Model Architecture

4.4.1 Feature Fusion

The features extracted from both CNN and RNN models were concatenated to form a robust feature set encapsulating both spatial and temporal information. This fusion leveraged the strengths of both models, combining static spatial cues from the CNN with dynamic temporal patterns from the RNN. Fig. 11.

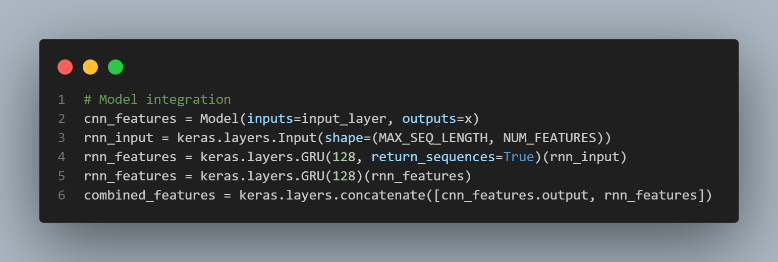


Fig. 11. Feature Fusion.

4.5 Training, Validation, and Testing

The integrated model was trained using a split of training and validation data to optimize the parameters and prevent overfitting. The training process involved using the Adam optimizer with a learning rate scheduler to adaptively control the learning rate. The model's performance was rigorously tested on a separate set of data to ensure its accuracy and generalizability. Fig. 12.

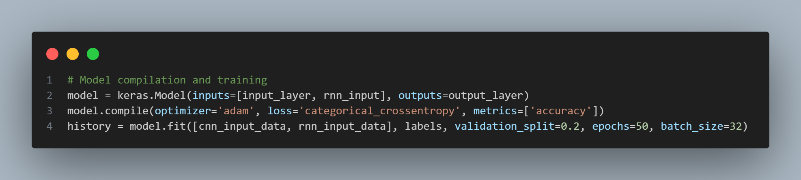


Fig. 12. Model compilation and training.

4.6 Ethical Considerations and Data Privacy

Advanced digital anonymization techniques were applied to all video data to ensure participant confidentiality. Facial features in the videos were obscured using digital distortion techniques, and similar precautions were taken with still images to protect the privacy and well-being of the children involved. Strict data handling and privacy protocols were adhered to throughout the research process. Fig. 13.

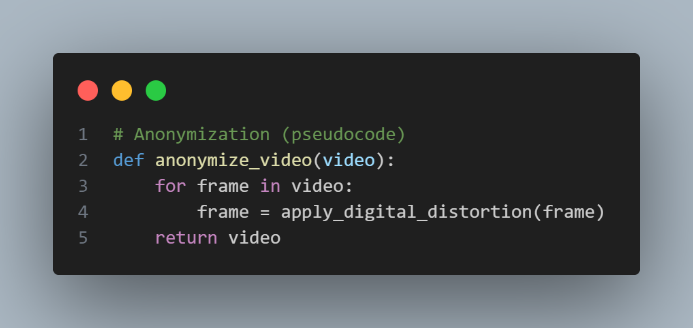


Fig. 13. Anonymization (pseudocode).

4.7 Detailed Technical Descriptions

4.7.1 In-depth CNN and RNN Analysis.

1. CNN Model: The CNN architecture utilized multiple convolutional layers with kernel sizes of 3x3, followed by max-pooling layers to extract and compress spatial features from each video frame. Batch normalization and dropout layers were also included to enhance model stability and prevent overfitting. Fig. 14.

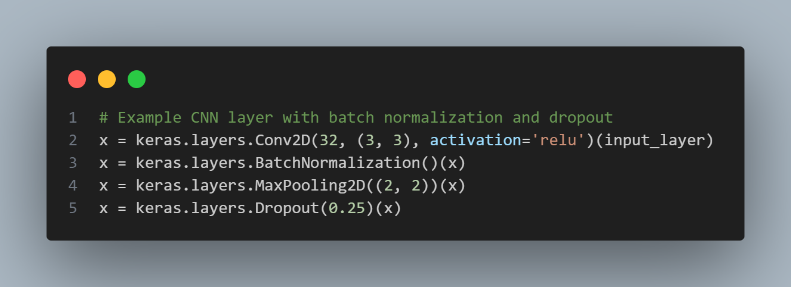


Fig. 14. CNN Model.

1. RNN Model: The RNN architecture employed GRU layers to analyze the temporal dynamics using sequences of spatial features provided by the CNN. These GRU layers captured the progression of behaviors over time, crucial for identifying ASD-related patterns. Fig. 15.

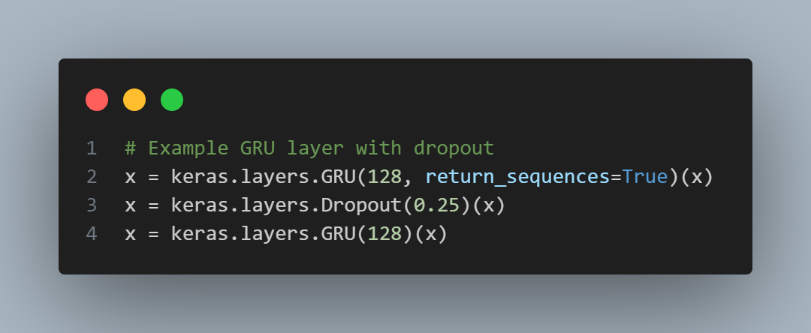


Fig. 15. RNN Model.

By meticulously detailing each step of the methodology and the tools and techniques used, this research establishes a comprehensive framework for the early detection of ASD, ultimately advancing diagnostic capabilities and supporting timely interventions for affected individuals and their families.

This detailed methodology ensures transparency and reproducibility, providing a solid foundation for further research and development in the field of ASD detection using deep learning. Fig. 16.

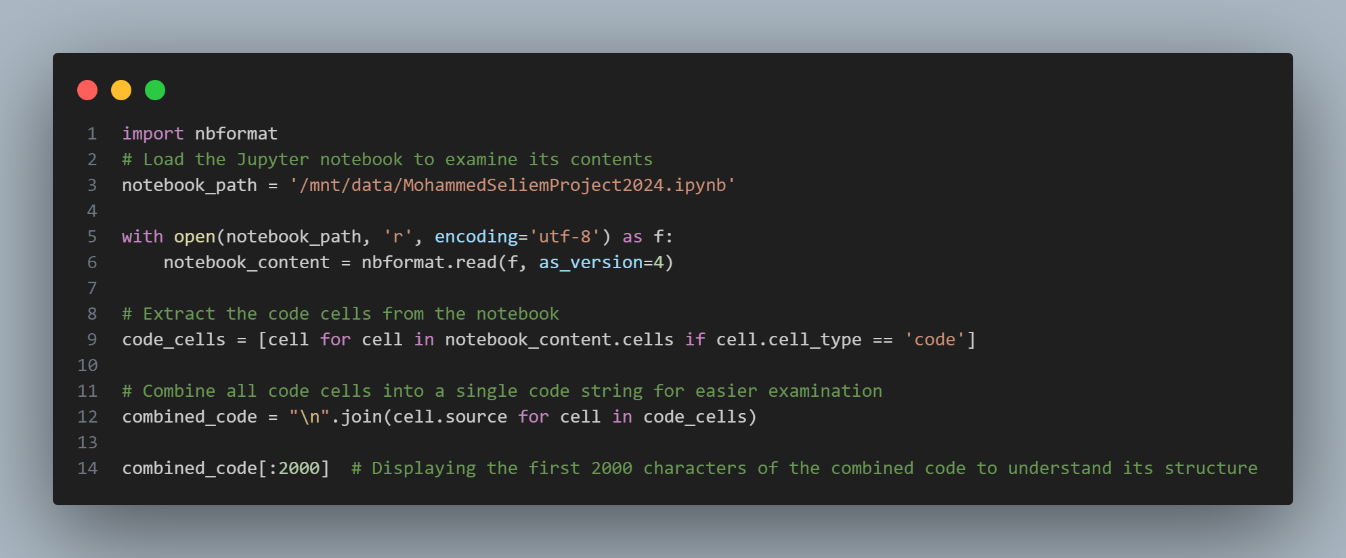


Fig. 16. Analysis.

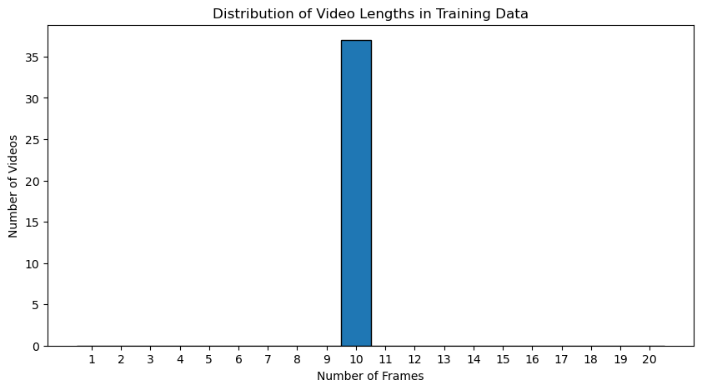


Fig. 17. Distributing Videos Lengths In Training Data.

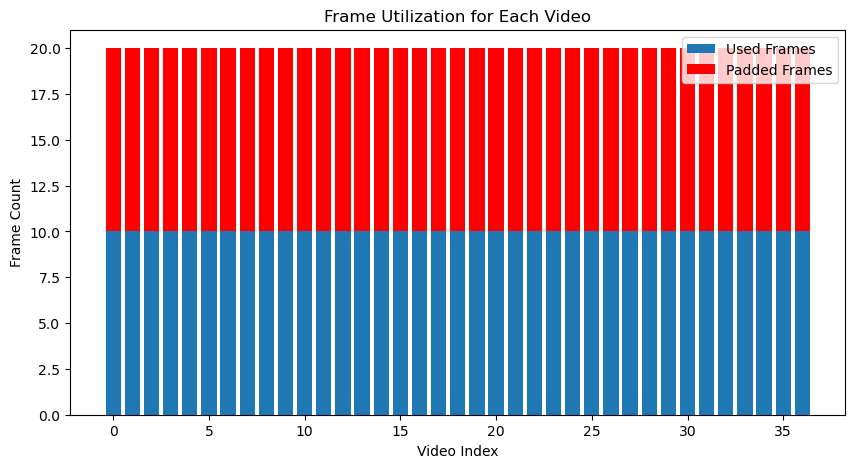


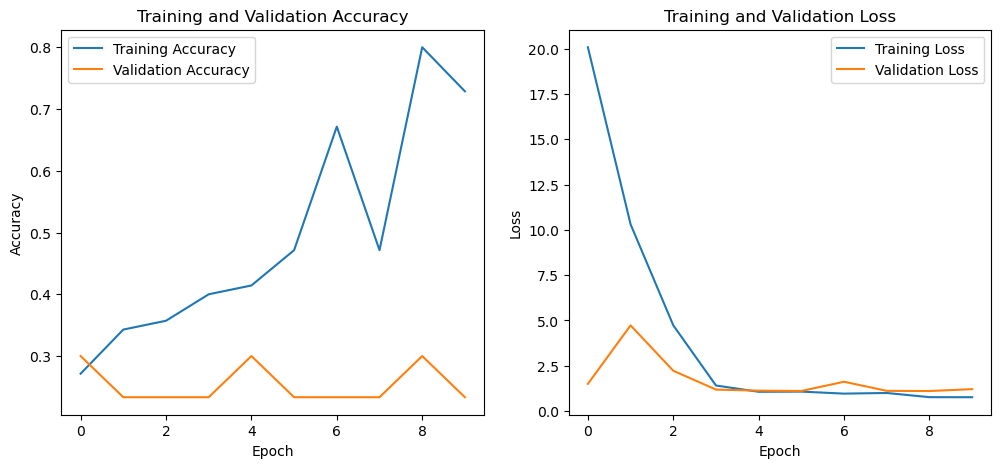
Fig. 17. Frame Utilization From Each Videos.

1. Result And Discussion

Our research distinguishes itself from prior studies in several pivotal aspects. Unlike many existing studies that have relied on a constrained set of still images of children with autism—occasionally magnified to boost model accuracy—our investigation harnesses video data. This methodology allows us to identify consistent patterns in the movements of children with autism, providing the opportunity to capture a more extensive range of features than static images allow. This detailed process enables the effective extraction of dynamic movement patterns and their conversion into numerical data for our matrix.

In contrast to previous studies that typically concentrated on brain imaging to identify unusual neuronal activity or employed eye-tracking technologies to study gaze patterns, our approach utilizes a rich dataset of video recordings. These videos, sourced from online platforms and contributed by parents or healthcare professionals, depict children's natural movements in unmodified environments. This setup not only enhances the authenticity of the data but also ensures meticulous training of our machine learning model. Videos are carefully divided into training and testing sets, which helps fine-tune the initial values and bias ratios, thus significantly enhancing the learning outcomes of the model.

## Our experimental findings highlighted the divergent performance outcomes of the implemented models. The Recurrent Neural Network (RNN) model, tailored for sequential data processing, demonstrated incremental improvements: achieving 16% accuracy after 10 epochs, 33% after 1,000 epochs, and reaching 50% accuracy at 5,000 epochs. In contrast, the Convolutional Neural Network (CNN), designed for processing video frame sequences, plateaued at a 33% accuracy, irrespective of extensive training. This disparity underscores the differential capabilities and applicability of these architectures in managing video classification tasks under the specific dataset and experimental conditions employed. Fig-19 and Fig-20



Top of Form

**Fig. 19.** Training Validation Accuracy And Training Validation lose.



**Fig. 20.** Compring Between RNN And CNN (Training Validation Accuracy and Training Validation Lose).

## Conclusion And Future Work

The research presented in this project aimed to enhance early detection of Autism Spectrum Disorder (ASD) through the application of advanced deep learning techniques. By utilizing Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), we were able to analyze video data capturing subtle behaviors indicative of ASD. The performance gap observed between the RNN and CNN models can be attributed to their distinct capabilities; RNNs excel at capturing temporal dependencies, which likely contributed to their superior performance in understanding time progression, achieving up to 50% accuracy after extensive training. In contrast, CNNs plateaued at 33% accuracy, indicating potential limitations in extracting salient features from sequential video data.

Future work could focus on several key areas to further improve the model's performance and applicability. One promising direction is the integration of hybrid models that combine CNN's spatial feature extraction with RNN's temporal processing capabilities. This hybrid approach could leverage the strengths of both models, potentially leading to more robust and accurate ASD detection systems. Additionally, expanding the dataset size by collecting hundreds of videos could significantly improve training efficiency and model performance, providing a more diverse set of data for better generalization to new, unseen videos.

Exploring more sophisticated models, such as Transformer networks, which have shown promise in similar contexts, could also be beneficial. These models could offer improved performance due to their advanced capabilities in handling sequential data. Moreover, incorporating additional types of data, such as audio or sensor data, could provide a more comprehensive understanding of ASD behaviors and enhance the model's diagnostic accuracy.

1. **References**

|  |  |
| --- | --- |
| [1] | T. S. A. E. E. a. A. R. J. Muhammad Cahyadi, "Early Detection Assessment Tools in Children With Autism Spectrum Disorder: A Literature Study," p. 13, 2022. |
| [2] | H. S. a. R. K. \*, *www.mdpi.com/journal/children,* no. Department of Electrical, Computer, and Biomedical Engineering, Ryerson University,, p. 18, 2020,. |
| [3] | Q. Z. W. H. B. a. M. I. Adnan Ashraf, "Analysis of Brain Imaging Data for the Detection of Early Age Autism Spectrum Disorder using Transfer Learning Approaches for Internet of Things," p. 13, 2023. |
| [4] | F. W. A. a. M. S. Alzahrani, "Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms," p. 10, 2022,. |
| [5] | A. A. A. a. O. A. S. I. Basma Ramdan Gamal Elshoky Eman M. G. Younis, "Comparing automated and non-automated machine learning for autism spectrum disorders classification using facial images," p. 12, 2021. |
| [6] | E. M. S. ,. T. H. R. ,. M. A. H. A. ,. H. S. A. S. ,. S. M. A. a. M. A. Ibrahim Abdulrab Ahmed, "Eye Tracking-Based Diagnosis and Early Detection of Autism Spectrum Disorder Using Machine Learning and Deep Learning Techniques," *electronics,* 2022. |
| [7] | T. H. H. A. a. M. Y. A. Hasan Alkahtani, "Early Screening of Autism Spectrum Disorder Diagnoses of Children Using Artificial Intelligence," p. 14, 2023. |
| [8] | L. H. C. Abdulrazak Yahya Saleh, "Autism Spectrum Disorder Classification Using Deep Learning," p. 12, 2021. |
| [9] | A. S. A. S. J. M. M. F. A. N. A Ali, "Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm," *IAES International Journal of Artificial Intelligence (IJ-AI),* p. 9, 2020. |
| [10] | N. O. ,. N. A.-M. ,. H. H. a. I. A. May Alsaidi, "A Convolutional Deep Neural Network Approach to Predict Autism Spectrum Disorder Based on Eye-Tracking Scan Paths," 2024,. |
| [11] | G. J. ,. M. I. G. O. ,. M. I. a. X. L. Junxia Han, "A Multimodal Approach for Identifying Autism Spectrum Disorders in Children," *IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 30, 2022,* p. 9, 2022. |
| [12] | T. H. H. A. ,. M. E. J. M. Y. A. M. E. A. M. M. A. F. A. ,. A. A. N. M. A. a. A. M. A.-m. Zeyad A. T. Ahmed, "Facial Features Detection System To Identify Children With Autism Spectrum Disorder: Deep Learning Models," p. 9, 2022,. |
| [13] | P. S. ,. K. T. ,. M. D. U. ,. D. T. S. R. Subash Gautam, "Screening Autism Spectrum Disorder in childrens using Deep Learning Approach : Evaluating the classification model of YOLOv8 by comparing with other models.with other models.," 2023. |
| [14] | H. Wang and P. Avillach, "Diagnostic Classification and Prognostic Prediction Using Common Genetic Variants in Autism Spectrum Disorder: Genotype-Based Deep Learning," p. 11, 2021. |
| [15] | A. P. D. B. S. J. D. S. M. a. M. A. Anupam Garg1, "Autism Spectrum Disorder Prediction by an Explainable Deep Learning Approach," p. 14, 2022. |
| [16] | s. r. a. s. masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," p. 11, 2020. |
| [17] | F. W. A. a. M. S. Alzahrani, "Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms," *Computational Intelligence and Neuroscience,* 2022. |
| [18] | A. D. ,. R. G. Shyam Sundar Rajagopalan, "Self-Stimulatory Behaviours in the Wild for Autism Diagnosis," p. 7, 2022. |
| [19] | R. Kirk, "Experimental Design: Procedures for the Behavioral Sciences," January 2013. |