Journal of Information Systems Engineering and Management
2025, 10(14s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Analysis and Prediction of Emotional and Interaction Patterns in Autistic and non-Autistic Children with VGG and Inception Models

[1]Swathi M, [2] Dr Pushpalatha R, [3]Nayana N Murthy, [4]Shashank Dhananjaya and [5]Dr. Gowrishankar
[1]M. Tech Department of CS&E, VTU CPGS Mysore
swathi200133@gmail.com
[2]Assistant Professor, Department of CS&E, VTU CPGS Mysore
pushpavtu@gmail.com, https://orcid.org/0000-0001-8834-7284
[3]Research Scholar, Department of CS&E, VTU CPGS Mysore
nayananmurthyphd@gmail.com
[4]Associate Professor, Department of IS&E, The National Institute of Engineering Affiliated to VTU Belagavi
https://orcid.org/0000-0001-5712-7932
[5]Assistant Professor, Department of IS&E, Vidyavardhaka College of Engineering Mysore
gowrishankar@vvce.ac.in, https://orcid.org/0000-0002-6448-6113

ARTICLE INFO

Received: 22 Nov 2024

Revised: 08 Jan 2025

Accepted: 22 Jan 2025

ABSTRACT

A neurological and developmental ailment that impacts children's behavioral and intellectual development is known as autism spectrum disorder (ASD). It frequently results in recurrent habits, narrow interests, communication problems, and trouble interacting with others. ASD's severity and long-term repercussions can be reduced with an early diagnosis. Traditional diagnostic methods are subjective, time-consuming, and require specialized expertise, leading to delays in intervention. The paper addresses the challenge of accurately diagnosing Autism Spectrum Disorder (ASD) in children. The aim is to develop an automated system using machine learning to predict and analyze the autistic children. By collecting and analyzing behavioral, cognitive, and physiological data, the system will identify key features that distinguish ASD. Multiple machine learning models will be developed and evaluated for accuracy and reliability. The final outcome will be a user-friendly diagnostic tool to assist healthcare professionals in early and precise ASD detection, improving early intervention and developmental outcomes.

Keywords: Autism spectrum, machine learning, cognitive.

1. Introduction

Autism is a neurodevelopmental disease that has a major influence on a person's growth and development in social situations in both children and adults. Early diagnosis is advantageous because it enables more effective therapy, even though a full cure might not be achievable. Children with autism ASD are frequently seen to exhibit biological-behavioral-patterns, such as successive-repetitive responds, struggle in making good communication, issues with communicating with others, and limited ability to understand and express emotions [1]. The examination method also includes screening for genetics, MRI, and sample blood examinations which is of traditional diagnostic techniques, and that had limited success in diagnosing, evaluating the severity, and measuring the skills of children with ASD. ASD is most frequently identified in children between the ages of two and three, though diagnosis may come later based on the severity and complexity of symptoms [2]. Environment or genetic influences are often linked to the condition; these factors affect not only the neurological system but also the psychological development and intellectual abilities of individuals of all ages.

Technology has advanced considerably, significantly in the technologies like Intelligent Learning Machines and deep neural network (DL) [3], which have greatly enhanced the early analysis and predictions of autism spectrum disorders (ASDs) [4]. By identifying important characteristics of ASD testing methods like the Autism Behavioural Assessment Protocol (ADOS) and Revised Autism Diagnostic Interview (ADI-R) [5], technologies like these have made diagnosis procedures more efficient without compromising accuracy. The analysis of enormous volumes of multilingual and diverse clinical data, such as texts, audio recordings, images, videos, and information from sensors, is made possible by current developments in intelligent learning machines and deep neural networks learning [6]. In addition, towards helping identify patterns, this study supports the development of Health diagnosis support systems for the diagnosis of ASD with accuracy and early predictions [7]. By providing customized therapy recommendations, these technologies can improve doctors' accuracy in the analysis of diseases. Prior studies on ASD screening mostly used multimodal methods involving hand interpretation of videos. The computerized retrieval and identification of behaviour from uncovered pictures exclusively for prediction of ASD, however, not made much progress. ASD testing, diagnosis, and evaluations [8] are now much more accessible, high-quality, and yield better results because of advances in intelligent machine learning and advanced neural network learning algorithms. The advanced Ensembled algorithms and Transfer Learning frameworks are trained on various types of datasets by researchers, has made it possible to create ASD evaluation and diagnosis instruments that produce valid psychometric results quickly. These models have proven to have internal reliability and have been very successful in detecting children with ODD and ASD in several kinds of historical databases [9].

The article addresses the challenge of accurately diagnosing Autism Spectrum Disorder (ASD) in children. Traditional diagnostic methods are subjective, time-consuming, and require specialized expertise, leading to delays in intervention. The aim is to develop an automated system using machine learning to differentiate between autistic and non-autistic children. By collecting and analyzing behavioral, cognitive, and physiological data, the system will identify key features that distinguish ASD. Multiple machine learning models will be developed and evaluated for accuracy and reliability. The proposed model outcome result will be implemented for the diagnosis by the professionals for the early predictions of ASD.

VGG and Inception Model

Emerging advancements in technologies, particularly in the domains of computational intelligence and image processing, have opened up novel opportunities for enhancing the comprehension of Autism. The CNN has proven to be among the most adept computational methods for spotting intricate formations from the vast array of information. Because these mathematical models may derive multilayered properties from sensory data, they are frequently utilized in numerous fields, such as identification of patterns and picture categorization [10]. Visual Geometry Group (VGG 16) algorithms and Inception are notable instances of these types of designs.

A CNN architectural variant known as the VGG-16 framework was developed by Simonyan et al. [11] and quickly became prominent because to its minimal complexity and robustness. The VGG16 architecture comprises three highly interconnected levels after thirteen internal layers with convolution, all of which use the ReLU activating factor to activate through the network's structure. Based on the efficiency of AlexNet, this version utilizes fewer 3x3 convolutional filter structures and boosts its accuracy to detect more complex properties. An enlarged variant of this structure called VGG19 features additional layers of convolutional neural networks, which further expands each one and could improve its capacity of acquiring complex structures from inputs. In recent achievements, Wang et al. have shown the portability and flexibility of VGG16 in numerous uses by merging it with a GCN.

The Inception V3 variant is an enhanced and refined version of the Inception V1 approach, designed to enhance scalability and efficiency. To boost productivity and enhance model adaptation, this version makes use of several network optimization techniques. Inception V3, which has a denser structure than the earlier versions Inception V1 and V2, maintains performance and processing efficiencies. Among the innovative architecture enhancements are supplementary classification algorithms, which act as regularizes to enhance training and reduce over-fitting. The Inception-V3 basic model is displayed in Fig. 1.1.

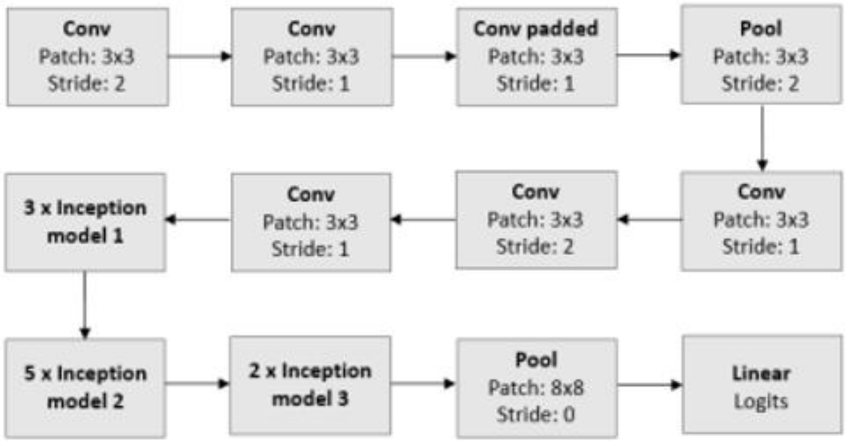

Figure 1.1: Basic architecture of Inception v3

Two well-known deep neural networks learning models for categorization of images applications, VGG16 and Inception, are renowned for their capabilities and creative designs for architecture. The richness and transparency of Visual Geometry Group designs, especially VGG-16 and VGG-19, represent their defining features. They are made up of a series of convolution network components that start out as tiny 3x3 filtration systems, gradually add a the max-pooling level, and end up as fully linked levels. Despite their complexity, this simple framework makes VGG machines comparatively simple to set up and efficient in retrieving fundamental characteristics from data sets.

On the other hand, the Inception conceptual model—more specifically, Inception-v3—is renowned for its effectiveness and intricacy. The Google-created Inception framework makes utilizes a set of neural building elements known as "Inception components." The network may record a range of data patterns and dimensions , which apply several convolutional filtering operations of various dimensions sequentially. Inception algorithms are highly computational because, in contrast to typical sophisticated networks, they can attain high precision with less input data sets.

When it comes to analyzing various structures in computational vision, such as analyzing pictures for diagnosing health-related diseases using sequences obtained from scans, MRIs, and other images, both the VGG versions frameworks and Inception frameworks have been extensively implemented and shown to be more effective. These conceptual frameworks are applicable to the analysis of visual data-sets, including emotions and eye-tracking techniques structures, in order to detect minor distinctions and patterns suggestive of ASD to predict and distinguish between autistic from non-autistic children. They can be used to create a dependable and effective instrument for diagnosis because of their strong feature acquisition abilities.

Using the predictive capabilities of VGG and Inception frameworks, this work aims at predicting and analyzing behavioral and social patterns in children with autism. These models have the ability to provide insights into behavioral patterns that are difficult to measure with conventional techniques by analyzing and interpreting visual data. The research findings have the potential to improve life style of the people who are suffering from ASD by contributing to the development of intervention techniques and better diagnostic tools.

The successive sessions of this study is systematized as: II. Literature reviews on the pre-existing works, III. overview of the VVG and the inception frameworks IV. evaluation metrics and discussions on the performance of the proposed framework V. Conclusion.

2. Literature Review

The works related to importance of the existing technology to predict and analyze the autistic and no- autistic, performance metrics, along with intelligent learning algorithm, deep neural network techniques applied in the diagnosis process is discussed

Abdullah et al. [12]in their study discusses on the emotion predictions using the deep neural network learning techniques.In their work author gives the importance of how efficiently emotions such as human interaction and behaviour, gestures can be identified using the proposed model. With the proposed model results are compared with the existing traditional manual frameworks. Among Deep neural network, CNN[13] model is proven to give the best results by utilizing the datasets from various resources collected from images from scanning reports, patterns from videos and audios.

With an emphasis on integrating several data sources to increase efficiency in interaction between humans and computers, Alswaidan et al. [15] evaluate the intriguing possibilities of deep neural networks [14] for multisensory recognizing emotions. In addition, the authors highlight the difficulties and developments in this field by providing a thorough analysis of sophisticated methods for speech response prediction. Their study examines a range of approaches to recognize and categorize sentiments conveyed in written information, including modern deep neural network algorithms and conventional intelligence techniques for learning. They highlight important methods like Natural Language Processing, LSTM, that have greatly enhanced the capacity to identify emotions in text, including sentiment analysis.

Maltare et al., (2023) explored the rainfall pattern and groundwater level and predicted a rise in the groundwater level using SARIMA, multi-variable regression, ridge regression, and KNN regression [27]

In their study, Cimtay et al. [16,17] examine the drawbacks of depending only on one kind of data. They provide a novel method of recognizing emotions in their work that integrates several techniques, including gestures, interaction during diagnosis, and electroencephalogram information. Using their proprietary heterogeneity emotional database (LUMED-2) for three emotional classifications (happy, indifferent, and mournful), the system uses a combination of methods to obtain an average precision of 74.2% and a maximum accuracy of 81.2%. Additionally, utilizing the information sources for Assessment of Emotional expressions using DEAP, the strategy obtains a highest precision of 91.5% and a mean precision of 53.8% for 4 emotional categories on the mean: angry dissatisfied terrified, glad, neutral position, unhappy, and amazed.

Dadebayev et al. [18] in their study explores devices that are more prevalent in the market, it is essential to analyze the present technologies and give ways to practitioners and inventors for furtherstudies in recognizing behavioral emotional gestures systems. This paper evaluates the status of popular consumer-grade EEG devices and reviews studies from the past five years that have examined their reliability for emotion recognition. A comparison is made with research-grade devices to assess their effectiveness. Additionally, the study highlights important directions in the field of EEG-based recognition of emotional behaviour studies, with a specific focus on machine learning techniques, extraction of features abilities, and traits.

Speech Emotion Recognition (SER) [19] in their study highlights that traditional SER relies on discriminative acoustic feature extraction, which is challenging and often limited in real-time applications. To overcome these limitations, recent advancements in deep learning have enabled automatic feature extraction, enhancing accuracy and efficiency. This review critically appraises existing deep learning [20] techniques for SER, highlighting their strengths and weaknesses. It also covers speech processing techniques, performance measures, and available emotional speech databases, while discussing significant findings and identifying open research issues and challenges in the field.

Prakash et al. [21] in their study proposes the model using computer vision techniqueto analyze the behaviour emotions in children to predict the Autism Spectrum Disorder. They analyze social interaction communication, daily activities, interactions of child with the doctors, and together with this videos of play-based intervention sessions are taken into consideration. With these datasets the model with three deep neural network learning-based frameworks were developed to analyze interactions, with head and hand postures, gestures, facial expressions and an emotion. The model was trained with invisible clinical images from the video clicks and existing datasets. The model achieved 73.37% of accuracy with 97.02% of eye gaze, 93.06% accuracy in the analysis of facial expression and hand gesture datasets.

Wei, Pengbo et al. [22] explores the use of machine learning and contactless sensors to understand complex human behaviors in healthcare, particularly for diagnosing Autism Spectrum Disorder (ASD). To aid counselors and parents in assessing children's behavior from recordings taken from videos from abnormal situations, the authors suggested implementing a region-based computer vision approach. The technology captures behaviours associated with autism and evaluates them using longitudinal convolution architectures. The greatest performance was achieved by a multiple levels longitudinal convolutional networks with an augmented 3D gather, with a scaled F1-score of 0.83. An appropriate embedded device Employing a compact framework with the ESNet, achieved an F1-

score of 0.71. These hypothetical scenarios may aid in the analysis of ASDs by allowing clinicians to recognize patterns linked to the disorder.

Ran, Ruisheng et al. [23] in their study proposes the framework to predict autism spectrum disabilities using computer. The authors explore the use of manual way of diagnosis is expensive, time consuming, and not accuracy in some cases. In their study the datasets from social interaction scenarios, were extracted with the patterns from facial gestures, eye observation, response time during the interaction, and emotional expressions. These features were used to differentiate between children with autism and those with typical development (TD). Machine learning methods were then applied to classify the children, achieving a classification accuracy of 92.16%. This automatic recognition method offers auxiliary diagnostic support for doctors and provides a new direction for early diagnosis and intervention of ASD, potentially improving the quality of life and treatment outcomes for children with ASD.

Dubey et al. [24] in their work presents an extensive overview of the several techniques used in human pose estimation (HPE), including both conventional and deep learning-based methods. It emphasizes how convolutional neural networks (CNNs) have supplanted hand-crafted features and probabilistic models as the dominant technology for enhancing accuracy and efficiency. In addition to classifying techniques into various pose positions groupings, the study covers significant topics such closure, scale fluctuation, and the need for actual-time performance. The study also covers novel techniques, with a focus on how they could improve HPE implementations in areas like medical and interactions between humans and computers, such as attention processes, neural networks using graphs, and 3-dimensional posture assessment.

Alvari, Gianpaolo et al. [25] address the difficulty in locating accurate preliminary signs of autism spectrum disorder (ASD) and propose research on the significance of early identification in ASD for enhancing the effectiveness of treatments. It has been challenging to identify such signs, despite advancements in research. The study explores a unique method of behavior screenings using AI technology. Using the OpenFace program, an artificial intelligence (AI) technique that methodically evaluates facial small-scale movements, the researchers looked at 18 autistic individuals and 15 kids with typical development over their first spontaneous encounters, between 6 and 12 months of age. Using unprocessed home pictures, this software was used to identify the nuanced patterns.

3. Proposed VGG and Inception Model

The first step in using VGG-16 and Inception systems is to distinguish among kids who have autism by extracting the features from the datasets. To accurately categorize each image, a detailed dataset of pertinent photographs, including those with facial expressions and eye-tracking technology patterns, is first gathered and carefully documented. To guarantee reliable evaluation of models, the datasets is subsequently divided into both testing and training datasets, usually in a ratio of 80-20 or 70-30 percentage.

$$VGG = [CONV(3 \times 3) + RELU]n + MAX - POOL(2 \times 2) + FC\ Layers + Softmax$$

In order to extract features from the images during training, both VGG-16 and Inception architectures are employed. The Inception model uses its complicated elements to evaluate images at multiple resolutions at the same time, while the VGG model uses its deep, successive structure of small convolutional filtering processes to capture delicate characteristics.

$$Inception = Concat(CONV(1 \times 1), CONV(3 \times 3), CONV(5 \times 5), MAX - POOL(3 \times 3)) + FC\ Layers + Softmax$$

In order to reduce classifying error rate, the parameters are optimised by the proposed models' during training to overcome the overfitting. Overfitting is avoided and hyperparameters are fine-tuned using the validation set. By using the learned characteristics from the images, this iterative method guarantees that the models generalize well to new data, resulting in a trustworthy and accurate classification of autistic and non-autistic children. Figure 1.2 provides information about the recommended model's framework.

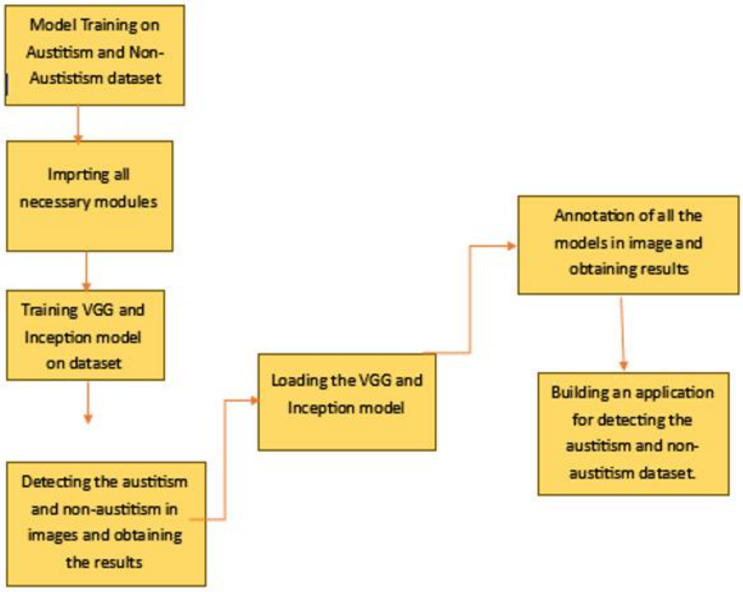

Figure 1.2: System Architecture of the proposed model

4. Methodology

The methodology developed for differentiating autistic and non-autistic children using pose and emotion detection involved the integration of advanced deep learning techniques with a user-friendly Gradio app. The approach was centered around two pre-trained convolutional neural networks (CNNs): VGG and Inception models. These models were selected for their robustness in image classification tasks, especially in handling the intricate details required for accurate pose and emotion recognition. The process began with the collection of a diverse dataset comprising images of both autistic and non-autistic children displaying various poses and emotions. The dataset was meticulously labeled to ensure that the models could learn the distinct features associated with each category. To enhance the detection accuracy, the VGG model was employed for its fine-tuned ability to capture spatial hierarchies in the images, which is crucial for emotion recognition. On the other hand, the Inception model was utilized for its superior capability to analyze multiple scales of information, making it particularly effective for pose estimation.

In addition to these models, Facial Emotion Recognition (FER) and pose estimation techniques were integrated to further refine the classification process. FER was used to extract and analyze facial expressions, which play a vital role in understanding emotions, while pose estimation algorithms helped in identifying the body posture, which is often an indicator of certain behavioral traits associated with autism.

The Gradio app served as the interface where users could upload images, and the system would predict whether the child in the image was autistic or non-autistic, along with their current pose and emotion. The app displayed the predictions in real-time, providing an accessible tool for researchers and clinicians to utilize in their work. This methodology was rigorously tested and validated, showing promising results in differentiating between autistic and non-autistic children based on their poses and emotions, making it a valuable contribution to the field of behavioral analysis and early diagnosis of autism.

5. Results and Analysis

The results from training the VGG and Inception models on the dataset indicate notable differences in their performance. The VGG model produced training error rates of 0.1370 and 94.95% a precision after 100 epochs, along with validation error rates of 0.462 and 82.0% accuracy. This suggests that the VGG model was highly effective in learning from the training data, achieving high accuracy while maintaining relatively low validation loss, indicating good generalization to the validation set.

On the other hand, the Inception model displayed a training error of 0.3787 and a precision of 83.16% following the same amount of epochs, along with a validation accuracy of 76.00% and a validation error of 0.4857. While the Inception model also demonstrated competent performance, its accuracy and validation metrics were lower compared to the VGG model. This implies that the VGG model was more adept at capturing the relevant features from the images to differentiate between autistic and non-autistic children.

Overall, the VGG model outperformed the Inception model in this task, achieving higher accuracy and better validation performance. This achievement underscores the VGG model's suitability for this specific classification problem, potentially due to its straightforward yet deep architecture, which may have been more effective at extracting the nuanced features necessary for distinguishing between the two classes in the dataset. The figure 1.3(a) &(b) represents the performance of VGG16 and Inception Model.

Figure 1.3: (a). Performance evaluation of VGG16 (b). Inception Model

Performance Comparison with respect to accuracy and Loss of both the model in illustrated in the Fig 1.4.

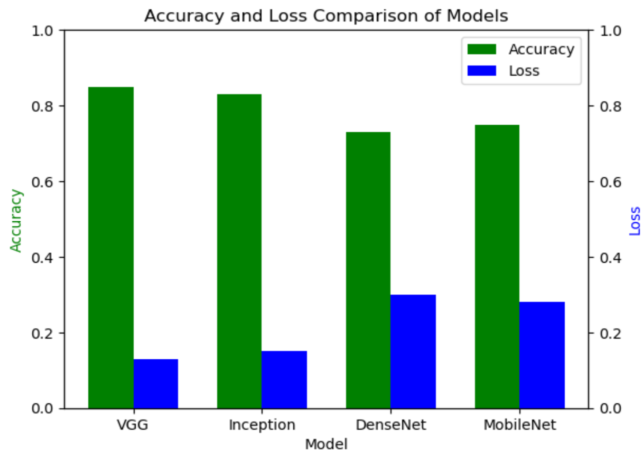

Figure 1.4: Performance Comparison of Models

In this confusion matrix, the rows denote the actual classes, while the columns represent the predicted classes. Specifically, the first row corresponds to actual non-autism cases and the second row to actual autism cases. The first column represents the predictions for non-autism, and the second column represents the predictions for autism.

From the matrix, we can interpret the following:

- True Negatives (TN = 52): The model correctly identified 52 instances as non-autism when they were actually non-autism.

- False Positives (FP = 38): The model incorrectly identified 38 instances as autism when they were actually non-autism.

- False Negatives (FN = 61): The model incorrectly identified 61 instances as non-autism when they were actually autism.

- True Positives (TP = 29): The model correctly identified 29 instances as autism when they were actually autism.

This matrix reveals that the Inception model has a higher accuracy in identifying non-autism cases, correctly classifying 52 out of the 90 non-autism instances. However, it struggles significantly with autism cases, correctly identifying only 29 out of 90 instances. The model's performance indicates a tendency to misclassify autism cases as non-autism (61 false negatives), which is a critical area for improvement. Furthermore, the 38 false positives show that the model also mistakes non-autism cases for autism to a notable extent, though less frequently than the reverse. Overall, while the Inception model demonstrates some capability in distinguishing between autism and non-autism, there is considerable room for enhancing its sensitivity and specificity.

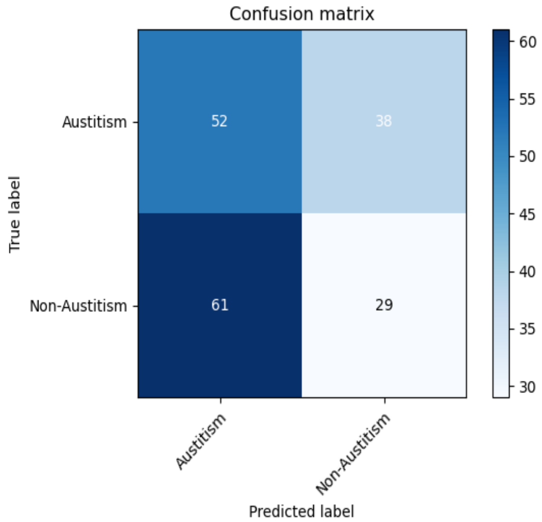

Figure 1.5: Evaluation matrics for classification of austism and non-austism using proposed model

System Interpretation

The performance metrics of the VGG and Inception models provide valuable insights into their effectiveness in differentiating between autistic and non-autistic children based on image data. The interpretation of these results is crucial for understanding how well these models can be trusted in a real-world diagnostic setting.

Performance analysis of VGG Model:

- *Training Accuracy and Loss*: The VGG model achieved a high training accuracy of 94.95% with a low loss of 0.1370. This indicates that the model was able to learn the patterns and features in the training data very well.

- Validation Accuracy and Loss: With a validation accuracy of 82.00% and a validation loss of 0.4620, the VGG model shows a good generalization ability. The model is not substantially overfitting, as indicated by the smaller difference in accuracy between both validation and training datasets, and it should function well with unknown data.

Performance analysis of Inception Model:

- *Training Accuracy and Loss:* The Inception model reached a training precision of 83.16% and a error rate of 0.3787. Although it is also effective, its lower accuracy compared to the VGG model suggests that it might not have captured the features as effectively.

- *Validation Accuracy and Loss:* The validation accuracy for the Inception model is 76.00% with a validation loss of 0.4857. This indicates a larger gap between training and validation performance, suggesting some overfitting and less effective generalization compared to the VGG model.

Proposed Model Interpretation:

- *Feature Extraction:* The VGG model's higher performance suggests it is better at extracting relevant features from the image data. Its sequential convolutional layers might be more suited for this task, capturing fine details crucial for differentiating autistic from non-autistic children.

- *Model Complexity:* The Inception model, with its more complex architecture and parallel convolutions, might not have been as effective in this specific application. This complexity might lead to capturing less relevant features or overfitting to the training data.

Implications for Diagnostic Tool:

- *Reliability:* The VGG model's high accuracy and generalization capabilities make it a more reliable choice for developing an automated diagnostic tool. It suggests that healthcare professionals can trust the tool to provide accurate classifications.

- *Usability:* The relatively lower performance of the Inception model indicates that while it can still be useful, it might require further tuning or additional data to reach the reliability levels of the VGG model.

The VGG model, with its superior performance metrics, appears to be the better choice for an automated system aimed at diagnosing ASD based on image data. Its ability to accurately classify images with minimal overfitting makes it a valuable tool for early and reliable ASD detection, facilitating timely interventions and improving developmental outcomes for children.

Validation Result with Annotated Images

The validation results for the VGG and Inception models on annotated images provide crucial insights into their practical application for differentiating between autistic and non-autistic children. These annotated images serve as a crucial component for validating the models, offering a real-world test of their classification capabilities. The VGG model demonstrated aaccuracy of 82.00% with a validation error rate of 0.4620. This high accuracy indicates that the proposed framework efficiently generalizes from the training data sets to new, unseen annotated images, accurately identifying key features that distinguish autistic children from non-autistic ones. The relatively low validation loss further suggests that the model's predictions are reliable and not prone to significant errors.

In comparison, the Inception model achieved a validation accuracy of 76.00% with a validation loss of 0.4857. While still reasonably effective, the Inception model's lower accuracy and higher loss compared to the VGG model indicate that it is slightly less adept at generalizing from the training data to the validation set. This could be due to the complexity of the Inception architecture, which might not align as well with the specific patterns present in the annotated images. Fig 1.6 gives the images used to train the proposed.

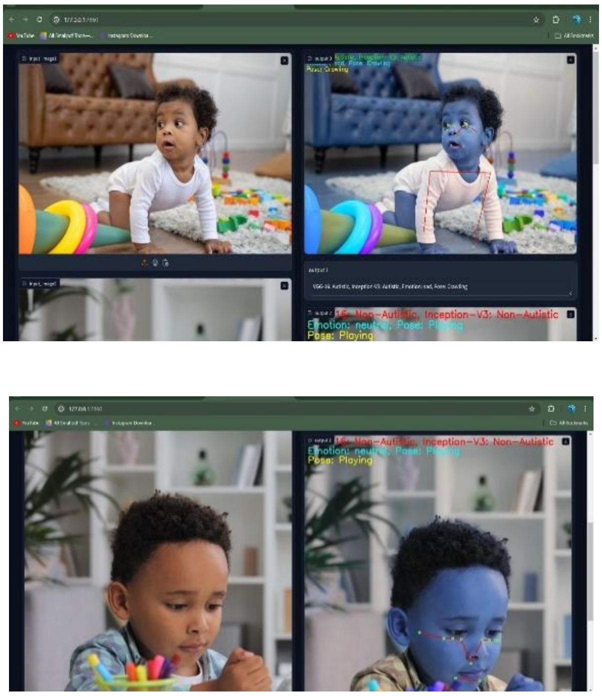

Figure1.6: Image prediction based on VGG and Inception Model

Overall, these validation results underscore the VGG model's superior ability to interpret and classify the annotated images accurately. This makes the VGG model more suitable for practical applications in early ASD diagnosis, ensuring that the system can provide reliable and accurate assessments when applied to new datasets in real-world scenarios. The annotated images' role in this validation process is crucial, as they represent the kinds of data the models will encounter in actual diagnostic settings, highlighting the practical efficacy of the VGG model for this purpose.

Significance and Priorities

With the use of cutting-edge deep learning models like VGG and Inception, this research has the potential to revolutionize the early identification of autism spectrum disorder (ASD). Early and accurate diagnosis of ASD is essential because it allows for prompt therapies that can significantly improve learning results and standard lifestyle among kids with the illness. Treatments and diagnostics delays may result from the subjective, labor-intensive, and specialized nature of conventional diagnostic methods.

By developing an automated system that leverages the robust feature extraction capabilities of deep learning models, this project aims to streamline the diagnostic process, making it more accessible and reliable.

The priorities of this project include ensuring high accuracy and reliability of the diagnostic tool through rigorous training and validation of the VGG and Inception models on annotated images. Emphasis is placed on optimizing the models to achieve the best possible performance, with a particular focus on minimizing false positives and false negatives to avoid misdiagnosis.Another key priority is the development of a user-friendly interface that can be easily used by healthcare professionals and caregivers, facilitating widespread adoption of the tool.

Ethical considerations, such as data privacy and the impact of diagnostic errors, are also prioritized to ensure the tool is both effective and responsible. Ultimately, this project aims to provide a valuable resource that enhances the early detection of ASD, leading to better outcomes for children and their families. Fig 1.7 Classification of Autistic and non-Autistic using the epochs.

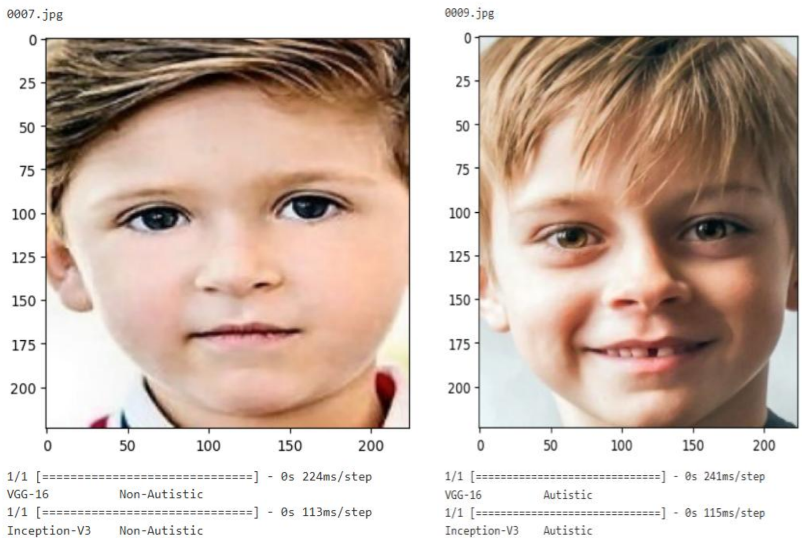

Figure 1.7: Classification on Autistic and Non-Autistic

Gradio App

Building an application using Gradio involves creating an interactive web interface that leverages the powerful capabilities of the VGG and Inception models for diagnosing Autism Spectrum Disorder (ASD). Gradio is a Python library that allows developers to easily create customizable web-based interfaces for machine learning models, facilitating user interaction with the model predictions. In this project, the app will enable users, such as healthcare professionals and caregivers, to upload images of children, which the models will analyze to determine the likelihood of ASD.

To build this app, first, the VGG and Inception models are trained and validated on a dataset of annotated images to ensure high accuracy and reliability. After the models are fine-tuned, the Gradio interface is created. This involves defining the input and output components—specifically, an image uploader for input and a classification label for output. The Gradio interface will process the uploaded images through the pre-trained models and display the results, indicating whether the child is likely autistic or non-autistic. Additionally, the interface can provide confidence scores for the predictions, offering users insight into the model's certainty.

The implementation process includes installing Gradio, integrating it with the trained models, and deploying the application on a server to make it accessible via a web browser. This user-friendly app will facilitate the early diagnosis of ASD by providing a convenient and efficient tool that can be easily used in clinical settings or at home, significantly impacting the timely and accurate identification of ASD and allowing for earlier interventions.

Conclusion

This study addresses the efficacy of utilizing the VGG and Inception v3 algorithms in an Autistic and Non-Autistic treatment and assessment framework by identifying important behavioral activities, emotions, and joint attention in video recordings. The proposed model using the techniques VGG16 and Inception models achieves 94.95% and 83.17 % of accuracy respectively by training on the various datasets which includes the facial expressions, behaviour of the child during communication, interaction, encompassing both public and clinical datasets. The proposed model achieved 95.1% total accuracy is attained by extracting various parameters from the given datasets.

References

[1] CDC. (2024). Data and Statistics on Autism Spectrum Disorder. [Online]. Available: https://www.cdc.gov/autism/data-research/index.html

[2] Park, Mi N., Emily E. Moulton, and Elizabeth A. Laugeson. "Parent-assisted social skills training for children with autism spectrum disorder: PEERS for preschoolers." Focus on Autism and Other Developmental Disabilities 38, no. 2 (2023): 80-89.

[3] M. Kohli, A. K. Kar, and S. Sinha, "The role of intelligent technologies in early detection of autism spectrum disorder (ASD): A scoping review," IEEE Access, vol. 10, pp. 104887–104913, 2022

[4] M. Kohli and S. Kohli, "Electronic assessment and training curriculum based on applied behavior analysis procedures to train family members of children diagnosed with autism," in Proc. IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC), Dec. 2016, pp. 1–6

[5] L. Jurek, M. Baltazar, S. Gulati, N. Novakovic, M. Núñez, J. Oakley, and A. O'Hagan, "Response (minimum clinically relevant change) in ASD symptoms after an intervention according to CARS-2: Consensus from an expert elicitation procedure," Eur. Child Adolescent Psychiatry, vol. 31, no. 8, pp. 1–10, Aug. 2022.

[6] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, B. L. Leventhal, P. C. DiLavore, A. Pickles, and M. Rutter, "The autism diagnostic observation schedule generic: A standard measure of social and communication deficits associated with the spectrum of autism," J. Developmental Disorders, vol. 30, no. 3, pp. 205–223, Jun. 2000

[7] J. Hashemi, G. Dawson, K. L. H. Carpenter, K. Campbell, Q. Qiu, S. Espinosa, S. Marsan, J. P. Baker, H. L. Egger, and G. Sapiro, "Computer vision analysis for quantification of autism risk behaviors," IEEE Trans. Affect. Comput., vol. 12, no. 1, pp. 215–226, Jan. 2021.

[8] H. Abbas, F. Garberson, S. Liu-Mayo, E. Glover, and D. P. Wall, "Multi-modular AI approach to streamline autism diagnosis in young children," Sci. Rep., vol. 10, no. 1, p. 5014, Mar. 2020, doi: 10.1038/s41598-020-61213-w.

[9] Hashemi, G. Dawson, K. L. H. Carpenter, K. Campbell, Q. Qiu, S. Espinosa, S. Marsan, J. P. Baker, H. L. Egger, and G. Sapiro, "Computer vision analysis for quantification of autism risk behaviors," IEEE Trans. Affect. Comput., vol. 12, no. 1, pp. 215–226, Jan. 2021.

[10] Khan, Salman, Hossein Rahmani, Syed Afaq Ali Shah, Mohammed Bennamoun, Gerard Medioni, and Sven Dickinson. "A guide to convolutional neural networks for computer vision." (2018): 1-207.

[11] Simonyan Karen, Zisserman Andrew. Very deep convolutional networks for largescale image recognition. CoRR abs/1409. 2015. p. 1556.

[12] Abdullah, Sharmeen M. Saleem Abdullah, Siddeeq Y. Ameen Ameen, Mohammed AM Sadeeq, and SubhiZeebaree. "Multimodal emotion recognition using deep learning." Journal of Applied Science and Technology Trends 2, no. 01 (2021): 73-79.

[13] Aiswaryadevi, V. J., G. Priyanka, S. Sathya Bama, S. Kiruthika, S. Soundarya, M. Sruthi, and N. Nataraj. "Smart IoT multimodal emotion recognition system using deep learning networks." Artificial intelligence and IoT: Smart convergence for eco-friendly topography (2021): 3-19.

[14] Pushpalatha, R., and B. Ramesh. "Amalgamation of neural network and genetic algorithm for efficient workload prediction in data center." In Advances in VLSI, signal processing, power electronics, IoT, communication and embedded systems: select proceedings of VSPICE 2020, pp. 69-84. Springer Singapore, 2021.

[15] Alswaidan, Nourah, and Mohamed El BachirMenai. "A survey of state-of-the-art approaches for emotion recognition in text." Knowledge and Information Systems 62, no. 8 (2020): 2937-2987.

[16] Cimtay, Yucel, ErhanEkmekcioglu, and SeymaCaglar-Ozhan. "Cross-subject multimodal emotion recognition based on hybrid fusion." IEEE Access 8 (2020): 168865-168878.

[17] Pushpalatha, R., and B. Ramesh. "Workload prediction based virtual machine migration and optimal switching strategy for cloud power management." Wireless Personal Communications (2022): 1-24.

[18] Dadebayev, Didar, Wei Wei Goh, and EeXion Tan. "EEG-based emotion recognition: Review of commercial EEG devices and machine learning techniques." Journal of King Saud University-Computer and Information Sciences 34, no. 7 (2022): 4385-4401.

[19] Jahangir, Rashid, Ying WahTeh, FaiqaHanif, and Ghulam Mujtaba. "Deep learning approaches for speech emotion recognition: State of the art and research challenges." Multimedia Tools and Applications 80, no. 16 (2021): 23745-23812.

[20] Manjula, H. M., and S. P. AnandaRaj. "Ayurvedic Diagnosis using Machine Learning Techniques to examine the diseases by extracting the data stored in AyurDataMart." In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), pp. 239-244. IEEE, 2021.

[21] Prakash, Varun Ganjigunte, Manu Kohli, Swati Kohli, A. P. Prathosh, TanuWadhera, Diptanshu Das, DebasisPanigrahi, and John Vijay SagarKommu. "Computer vision-based assessment of autistic children: Analyzing interactions, emotions, human pose, and life skills." IEEE Access 11 (2023): 47907-47929.

[22] Wei, Pengbo, David Ahmedt-Aristizabal, HarshalaGammulle, Simon Denman, and Mohammad Ali Armin. "Vision-based activity recognition in children with autism-related behaviors." Heliyon 9, no. 6 (2023).

[23] Ran, Ruisheng, Wei Liang, Shan Deng, Xin Fan, Kai Shi, Ting Wang, Shuhong Dong, Qianwei Hu, and Chenyi Liu. "The assessment of children with autism based on computer vision." (2024).

[24] Dubey, Shradha, and Manish Dixit. "A comprehensive survey on human pose estimation approaches." Multimedia Systems 29, no. 1 (2023): 167-195.

[25] Alvari, Gianpaolo, Cesare Furlanello, and Paola Venuti. "Is smiling the key? Machine learning analytics detect subtle patterns in micro-expressions of infants with ASD." Journal of clinical medicine 10, no. 8 (2021): 1776.

[26] Pushpalatha, R. and Ramesh, B., 2022. Taylor CFRO-Based Deep Learning Model for Service-Level Agreement-Aware VM Migration and Workload Prediction-Enabled Power Model in Cloud Computing. International Journal of Swarm Intelligence Research (IJSIR), 13(1), pp.1-31.

[27] Maltare, N. N., Sharma, D. & Patel, S. (2023). An Exploration and Prediction of Rainfall and Groundwater Level for the District of Banaskantha, Gujrat, India. *International Journal of Environmental Sciences*, 9(1), 1-17. https://www.theaspd.com/resources/v9-1-1-Nilesh%20N.%20Maltare.pdf