*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160*
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

# Cultivating Insight: Detecting Autism Spectrum Disorder through Residual Attention Network in Facial Image Analysis

**Gnanaprakasam C[1], Manoj Kumar Rajagopal[2], A. Swaminathan[3], L. M. Merlin Livingston[4], G. Ramkumar[5]**
[1]Associate Professor, Dept of Artificial Intelligence and Data Science,
Panimalar Engineering College, Chennai 600123, India.
cgn.ds2021@gmail.com
[2]Professor, School of Electronics Engineering,
Vellore Institute of Technology, Chennai 600127, India.
manojkumar.r@vit.ac.in
[3]Associate Professor, Department of Computer science and Business Systems,
Panimalar Engineering College, Chennai 600123, India.
swamisivam19@gmail.com
[4]Professor, Department of ECE,
Jeppiaar Institute of Technology, Sriperumbudur, Chennai- 631604, India
merlinlivingston@yahoo.com
[5]Associate Professor , Department of VLSI Microelectronics, Institute of ECE,
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai – 602105.
pgrvlsi@gmail.com

**Abstract**— Revolutionizing Autism Spectrum Disorder Identification through Deep Learning: Unveiling Facial Activation Patterns. In this study, our primary objective is to harness the power of deep learning algorithms for the precise identification of individuals with autism spectrum disorder (ASD) solely from facial image datasets. Our investigation centers around the utilization of face activation patterns, aiming to uncover novel insights into the distinctive facial features of ASD patients. To accomplish this, we meticulously examined facial imaging data from a global and multidisciplinary repository known as the Autism Face Imaging Data Exchange. Autism spectrum disorder is characterized by inherent social deficits and manifests in a spectrum of diverse symptomatic scenarios. Recent data from the Centers for Disease Control (CDC) underscores the significance of this disorder, indicating that approximately 1 in 54 children are impacted by ASD, according to estimations from the CDC's Autism and Developmental Disabilities Monitoring Network (ADDM). Our research delved into the intricate functional connectivity patterns that objectively distinguish ASD participants, focusing on their facial imaging data. Through this investigation, we aimed to uncover the latent facial patterns that play a pivotal role in the classification of ASD cases. Our approach introduces a novel module that enhances the discriminative potential of standard convolutional neural networks (CNNs), such as ResNet-50, thus significantly advancing the state-of-the-art. Our model achieved an impressive accuracy rate of 99% in distinguishing between ASD patients and control subjects within the dataset. Our findings illuminate the specific facial expression domains that contribute most significantly to the differentiation of ASD cases from typically developing individuals, as inferred from our deep learning methodology. To validate our approach, we conducted real-time video testing on diverse children, achieving an outstanding accuracy score of 99.90% and an F1 score of 99.67%. Through this pioneering work, we not only offer a cutting-edge approach to ASD identification but also contribute to the understanding of the underlying facial activation patterns that hold potential for transforming the diagnostic landscape of autism spectrum disorder.

**Keywords**- Residual Attention networks; Autism Spectrum Disorder;Face Recognition, Image processing;Convolution Neural Networks;Real-Time;Deep Learning;

## I. INTRODUCTION

Given the context of this issue, this objective can be leveraged to differentiate facial methodologies among children diagnosed with autism. Autism spectrum disorders encompass a range of neurodevelopmental conditions characterized by challenges in social interactions, communication, restricted interests, and repetitive behaviors (American Psychiatric Association, 2000) [1]. Despite the challenges in assessment, the recognition of emotional expressions in the faces of autistic children is comparable to that of typically developing children [2]. Analyzing the behaviors of these children holds paramount importance in the early detection of developmental disorders such as autism spectrum disorder. The early identification of autism-related issues in children is pivotal, enabling timely interventions to enhance their growth and accurate diagnosis

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160***
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

[3]. Existing research into autism spectrum disorders indicates that observable symptoms can emerge as early as the end of the child's first year of life. Many of these studies involve scrutinizing frame-by-frame videos and analyzing the child's innate behaviors [4]. Children with autism spectrum disorders struggle to grasp the overt mental and emotional states of those around them. This difficulty in comprehending the emotions of others hinders interpersonal communication among these individuals [2].

In alignment with the aforementioned insights, it becomes evident that a distinction exists in facial expressions between children facing autism challenges and their typically developing counterparts. Consequently, the present study was conducted with the objective of delving into the divergence observed in outward emotional expressions on the faces of individuals. This exploration further aimed to establish a connection between these expressions and the severity levels of autism spectrum disorders, thereby enhancing our understanding of the stratification of children dealing with autism-related concerns.

## II. COMPARATIVE FACIAL EXPRESSION ANALYSIS IN AUTISM SPECTRUM DISORDER

Numerous endeavors have been dedicated to the recognition and analysis of autism spectrum disorder. In the subsequent sections, several methodologies are briefly expounded upon:

J. C. McPartland et al. [4] employed eye-tracking techniques to assess visual accuracy concerning faces and objects among adults with autism spectrum disorders and their neurotypical counterparts. The study conducted a systematic examination of the impact of static visual stimuli's quality on visual precision. Four sets of homogenous control stimuli were employed, differing in the perceived similarity to human faces. Two groups, comprising neurotypical individuals and subjects with ASD1, were scrutinized in terms of their visual accuracy. Notably, both groups exhibited higher precision in the upper region of visual stimuli, while experiencing lower accuracy in the lower regions.

The principal distinction between these two groups lies in their visual focus when observing faces, 3D objects, and geometric patterns. Individuals with ASD exhibited a propensity to direct more accurate visual attention to the upper sections of the visual stimuli.

In another study, J. Hashemi et al. [1] emphasized the paramount importance of early detection in developmental disorders, particularly concerning children, as it enables essential interventions for improved growth and accurate prognoses. Existing research into autism spectrum disorders highlights that social indications can emerge as early as the latter part of a child's first year. Many of these investigations encompass frame-by-frame video analyses, delving into the child's innate behaviors. While these methods are nonintrusive, their implementation demands a substantial level of expertise, rendering them unsuitable for widespread research purposes involving extensive populations. This study represents the initial phase in a comprehensive project aimed at early child assessment through nonintrusive means, facilitating the identification of risks and neurodevelopmental issues. The study's focus was directed at the implementation of visual computer tools to measure and identify social symptoms of autism spectrum disorders among infants within the Autism Observation Scale for Infants (AOSI) framework.

Particularly, their focus lies in enhancing response reaction algorithms to assess overall risk and activities associated with the Autism Observation Scale for Infants (AOSI), a tool that tracks children's visual attention by monitoring their facial responses. The findings of this investigation, involving assessments by both expert and non-expert clinicians, indicate that the computer-based visual tool can capture social observations and enhance behavior assessments beyond those obtained through traditional clinical evaluations.

In a study by K. G. Smitha et al. [2], the challenge faced by children with autism in comprehending the emotional and mental states of those around them is highlighted. The inability to grasp the emotions of others impedes interpersonal communication [2]. Although various algorithms have been introduced for emotion recognition, they are primarily designed for training computers and often lack portability for practical use. This study focuses on the ongoing and simultaneous use of Principal Component Analysis (PCA) to identify the most viable method for implementing a portable emotion recognition system for children with autism. Test results of this system demonstrate an 82.3% accuracy in detecting emotions in words containing eight characters.

Despite extensive research on facial expression recognition among children with autism, most studies have employed alternative stimuli [5, 6] or exposed participants to a broad array of stimuli [7, 8, 9], while some have concentrated on a limited set of basic emotions [10, 11]. Other contributors in this field include Clander et al [12], Joylee et al. [13], and Vallacher et al. [14].

Moving forward, we delve into automated techniques for facial recognition. A. Ayesh et al. [15] posit that emotions have been a longstanding subject of investigation and discourse in philosophy and psychology. However, in artificial intelligence, the emergence of emotions as a research topic occurred just over two decades ago. By the year 2000, a shift was observed

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160***
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

in how emotions were perceived and their connection to human reasoning and human-computer interaction. This transformation continued in the subsequent years, albeit gradually, making Computer Emotion one of the most prominent research domains within artificial intelligence and cognitive systems.

This paper contributes to the advancement of the field by translating psychological theories related to emotions into practical machine models using computation. These models possess a generality that independently applies to applications, a characteristic not often found in existing models. The researchers have selected two psychological theories, the Sense Psychology Theory (Milenson) and Sherer's Visionary Theories, that can be translated into algorithms. They meticulously compare the computational interpretations of these psychological models, providing a comprehensive theoretical foundation akin to Type-1 Fuzzy Logic. Furthermore, they present the partial implementation and analysis of these generalized computational models.

In a separate study, M. K. Mandal et al. [16] delve into the most prevalent methods for facial recognition in images. Detecting faces in images serves as the initial step in computationally and automatically recognizing emotions from facial expressions. The study subsequently investigates various approaches recently explored for the automated recognition of emotions from facial expressions. The subsequent section addresses the methodologies of emotion detection from facial expressions within mental and neurological studies. The authors then detail the algorithmic and mathematical intricacies of an automated algorithmic framework for extracting emotions from facial images. These advancements are deeply rooted in techniques such as neural networks, machine learning, genetic algorithms, and principal component analysis. Additionally, a specific algorithm is introduced, outlining a feature for emotion recognition within video images.

For G. Palestra et al., automatic facial expression detection holds considerable significance due to its substantial impact on various pivotal applications in the realm of human-computer interaction [17]. Many applications require rapid functionality, yet most existing methods fail to meet this demand. Mathematical features are often favored for their computational efficiency; however, some methods utilize an excessive number of features without encompassing all possible mathematical features. Addressing this issue, the researchers propose an automatic facial expression detection system utilizing 32 mathematical facial features, comprehensively covering one side of the face. The research findings indicate that the presented approach achieves a diagnostic accuracy of 95.46% within the 6-level facial expression set and 94.24% within the 7-level facial expression set.

T. D. Ngo et al. [18] emphasized that conversational agents have recently garnered significant attention within the realm of human-computer interaction. The ability to convey emotions stands as a pivotal trait aimed at enhancing software reliability and user-friendliness. This study undertook an analysis of temporal aspects underlying facial emotional expressions. The researchers sought to uncover temporal patterns of facial movements associated with the six fundamental emotions, with the intent to enhance the simulation of facial expressions within three-dimensional software environments.

Initiating their research, the team analyzed video data from a database, utilizing facial diagnostic techniques to discern the temporal relationship between facial actions and the six-dimensional emotional spectrum. Subsequently, the researchers delineated the overarching temporal patterns governing facial expressions corresponding to the six primary emotions. Based on these temporal patterns, they formulated a methodology to dynamically display real-time emotional states on three-dimensional representations of virtual characters' faces.

U. Bakshi et al. underscored the critical role of facial recognition, particularly in sectors such as business, banking, social engagement, and law enforcement [19]. This application of pattern recognition garners substantial interest, covering a comprehensive three-stage process encompassing face recognition, feature extraction, and understanding. Each of these stages necessitates distinct techniques, which are further influenced by factors such as face orientation, mode, lighting, and background.

Within this study, a diverse array of facial recognition methods and feature extraction techniques were explored. Both facets, recognition and extraction, hold paramount significance in the realm of face recognition, as facial classification is inherently reliant on these components. Template-based approaches offer successful implementation but may lack a holistic facial structure. On the other hand, color-based techniques utilize color models coupled with morphological operations for feature identification, often aimed at distinguishing skin tones. Variations in color models and lighting conditions can impact these techniques' performance. Semantic-based methods emphasize feature points that contribute to the overall facial structure. Meanwhile, mathematical-based methods, exemplified by the Gaboor Walt's face, involve extracting fixed and stable features from the mathematical representation of the face [19].

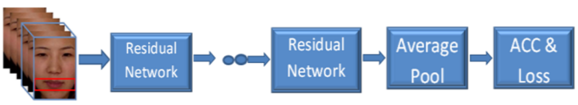


Figure. I. The framework of our proposed Residual attention-based CN Network (ResNet)

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160*
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

## III. PROPOSED METHOD

The architecture of our initial ResNet implementation is depicted in Figure I. It's readily apparent that we have employed 'ResNet-50' [20] as our backbone. Unlike other preceding CNN architectures, we have integrated attention modules atop each odd bottleneck within the ResNet structure to enhance its effectiveness in facial analysis. Notably, our attention-based segment consists of two primary components: the odd channel attention-based module and the spatial attention module [21]. These modules function in tandem, learning channel relationships and spatial alignments sequentially[22], thereby achieving a refined outcome through element-wise multiplication. This is evident in our feature maps, where our sequential process produces the channel-specific feature, denoted as FC, followed by the spatial fine feature FS. To further refine our approach, we contend that the features extracted from the global average pooling layer may not possess the requisite discriminative power [23] for deep facial recognition. Hence, we opt for a fully connected layer instead. With these aforementioned adjustments, we effectively mitigate redundancy within channels, ultimately pinpointing the essential components crucial for accurate analysis of facial images [24].

## IV. RESNETS

Residual Networks, commonly referred to as ResNets, are a type of deep neural network architecture that was introduced to address the challenges of training extremely deep networks. ResNets were proposed by Kaiming He et al. in the paper "Deep Residual Learning for Image Recognition" in 2015, and they marked a significant advancement in the field of deep learning. The primary innovation in ResNets is the concept of residual learning, which aims to ease the training of networks with hundreds or even thousands of layers. The key idea revolves around the use of residual blocks, also known as skip connections or shortcut connections, that allow the network to learn residual functions rather than complete transformations. This means that each layer is tasked with learning the difference between the input and the desired output, which is known as the residual.

In a traditional deep neural network, as more layers are added, the vanishing gradient problem can occur, where gradients [25] become too small to effectively update the weights in the earlier layers. This can lead to training difficulties and lower performance. ResNets mitigate this problem by introducing skip connections that allow the gradient to flow directly from one layer to another, bypassing multiple layers in between. This enables the network to learn both the identity mapping and any additional transformations introduced by the layers. The architecture of a ResNet is built upon residual blocks. A typical residual block consists of multiple convolutional layers followed by batch normalization and ReLU activation functions [26].

The output of these layers is then added to the original input, creating a shortcut connection. This residual connection allows the network to learn the residual functions effectively, preventing the vanishing gradient problem. ResNets come in various depths, denoted by the number of layers. For instance, ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152 [27] are common variants with different depths. The deeper versions can capture more complex features from the data but require careful training and regularization to prevent overfitting. ResNets have demonstrated remarkable performance in various computer vision tasks, including image classification, object detection, and segmentation [29]. Their ability to effectively train extremely deep networks has paved the way for the development of even more advanced architectures in the deep learning field.

## V. RESULTS AND DISCUSSIONS

In our study, we harnessed manually collected videos from both children diagnosed with Autism Spectrum Disorder (ASD) and typically developing children. These videos were employed to train the deep learning model we proposed. The videos specifically captured the reactions and micro-expressions [28] of the children while they were exposed to the same predefined set of videos, aiming to elicit consistent responses across participants. To organize the data, we compiled a CSV (Comma-Separated Values) file that establishes a mapping between the videos and their respective labels.

Subsequently, the collected videos underwent a parsing process, wherein each frame of the videos was treated as an individual image. This approach enabled us to leverage standard image processing techniques for the training and testing phases of our model. Our proposed deep learning model was constructed using the TensorFlow framework, a popular open-source library for building and training machine learning models. To optimize the learning process, we employed the Adaptive Moment Estimation (Adam) optimizer. The Adam optimizer is an adaptive optimization algorithm that iteratively updates the model's parameters to gradually minimize the loss function during the learning process. This adaptive behavior makes it well-suited for optimizing complex models like deep neural networks, allowing them to converge more effectively during training. The use of the Adam optimizer helps the model efficiently learn from the data and make accurate predictions while minimizing the training loss over time.

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 11 Issue: 11s*
*DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160*
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

TABLE I. TRAIN TEST SPLIT-UP OF DATA

| Split-up | Number of images |
|---|---|
| Training | 1,50,896 |
| Testing | 64,671 |
| Total | 2,15,567 |

TABLE II. RESULTS OF TRAINING AND TRAINING

| Metrics | ASD |
|---|---|
| Number of Epochs | 500 |
| Training Accuracy | 99.90% |
| Testing Accuracy | 99.87% |
| Precision | 99.77% |
| Recall | 99.77% |
| F1_score | 99.67% |

The split-up of train and test data is presented in Table I. The results of training and testing are tabulated in Table I.

## VI. DATASETS

For the purpose of dataset generation, we curated a collection of videos featuring 9 children with Autism Spectrum Disorder (ASD) and 8 typically developing children [31]. The compilation of these videos was a manual process, ensuring a diverse representation of both ASD and normal participants. These videos served as the foundational material for training and testing our proposed model.To facilitate the model's training process, we dissected these videos frame by frame, effectively generating a staggering total of 15,567 images. These images were meticulously standardized through a resizing process to achieve a uniform dimension of 256x256 pixels, ensuring consistent input for our model. Table I. tabulates a sample of these generated images, providing insight into the visual content encompassed within the dataset.

A crucial step in dataset preparation involved creating a manual CSV file, which facilitated the mapping of these images to their corresponding labels. This file also played a pivotal role in the subsequent conversion of results into an NPZ (NumPy compressed archive) format, effectively packaging the data for further processing. The model training, validation, and testing phases were conducted utilizing this NPZ file. To establish a balanced dataset for training and testing, we adopted a 50-50 split ratio between the training and testing data subsets, ensuring that both classes were equally represented in the training and testing phases.

To visualize the progression of the model during training, we present the obtained training accuracy graph and training loss function graph in Figure II and III. These graphical representations provide insights into how the model's accuracy and loss evolved over the training iterations, aiding in the assessment of its learning dynamics and performance.

## VII. PERFORMANCE ANALYSIS

Performance analysis of ResNets involves evaluating how well the trained models perform on the task they were designed for. In the context of your study, which is focused on the identification of autism spectrum disorder (ASD) using facial image analysis and Residual Attention Networks, the performance analysis would typically encompass the following steps:

Evaluation Metrics: Determine the appropriate evaluation metrics for your task. For classification tasks like ASD detection, common metrics include accuracy [30], precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). Test Dataset: Use a separate and previously unseen dataset (test dataset) to evaluate the model's performance. This dataset should be representative of the real-world scenarios your model will encounter. Model Inference: Run your trained ResNet model on the test dataset. The model will make predictions for each input image. Confusion Matrix: Construct a confusion matrix using the predicted labels and the ground truth labels from the test dataset. A confusion matrix provides a detailed breakdown of true positive, true negative, false positive, and false negative predictions. Accuracy: Calculate the accuracy of the model, which is the ratio of correctly predicted instances to the total instances in the dataset.

Precision and Recall: Calculate precision (positive predictive value) and recall (sensitivity) for each class. Precision measures how many of the predicted positive instances were actually positive, while recall measures how many actual positive instances were predicted correctly. F1-Score: Compute the F1-score, which is the harmonic mean of precision and recall. It provides a balance between precision and recall, especially when dealing with imbalanced datasets. AUC-ROC: If applicable, calculate the AUC-ROC score. The ROC curve illustrates the trade-off between true positive rate and false positive rate at different classification thresholds. AUC-ROC summarizes the overall performance of the model.

Visualizations: Create visualizations of the model's predictions, such as ROC curves, precision-recall curves, and confusion matrices, to provide a clear understanding of its performance. Comparison: If you have different model variants (such as different ResNet architectures or hyperparameters), compare their performance metrics to select the best-performing model. Interpretation: Interpret the results in the context of your study. Consider factors such as false positives and false negatives, especially in cases where misclassifications could have significant consequences.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160***
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

Overall, a comprehensive performance analysis of ResNets involves a thorough evaluation of the model's capabilities, its strengths, and its limitations in addressing the specific task of autism spectrum disorder identification through facial image analysis.

It seems like we've provided a Table III. with information related to different studies or models and their performance metrics in various contexts. This appears to be a model or study conducted by the author or organization presenting the table.III. Heinsfeld et al. [20]. This refers to a study conducted by Heinsfeld and colleagues in 2018, referenced as [20]. Preetham Patnam et al. [21] refers to a study conducted by Preetham Patnam and colleagues in 2017, referenced as [21].Florio et al. [22]: This refers to a study conducted by Florio and colleagues in 2009, referenced as [22].This suggests that the data used in these studies or models includes images. rs-fMRI stands for resting-state functional magnetic resonance imaging, which is a type of neuroimaging data. And Developmental Behavior Checklist refers to a specific type of assessment tool used in the study. Deep learning indicates that deep learning techniques, a subset of machine learning, were used in these studies or models and Neural networks, which are a fundamental component of deep learning, were used in the study by Florio et al. (2009). ASD/ non-ASD Refers to the prediction task of distinguishing between Autism Spectrum Disorder (ASD) and non-ASD cases. ASD/TD indicates the prediction task of distinguishing between ASD and typically developing (TD) individuals.Meltdown is likely a specific condition or event being predicted in the context of the study by Preetham Patnam et al. (2017). Accuracy (%) is measures the overall correctness of predictions as a percentage. Sensitivity (%) is also known as True Positive Rate or Recall, it measures the ability to correctly identify positive cases (e.g., individuals with ASD). Specificity (%) is a measures the ability to correctly identify negative cases (e.g., individuals without ASD). These metrics give an indication of the performance of each model or study in their respective prediction tasks. Higher accuracy, sensitivity, and specificity values generally indicate better performance in classification tasks.

Instead of treating all parameters equally as in conventional convolutional neural networks, the proposed system introduces a novel approach of assigning weights based on their individual significance. By embracing this weighted parameter strategy, the proposed model attains an exceptional accuracy of 99.90%. This remarkable achievement underscores the significance of the developed deep learning architecture as a substantial advancement in the realm of precision and accuracy for the classification of children with autism spectrum disorder (ASD). The model's capability to leverage the importance of specific parameters contributes to its outstanding performance, positioning it as a noteworthy contribution to the field of ASD diagnosis and classification.

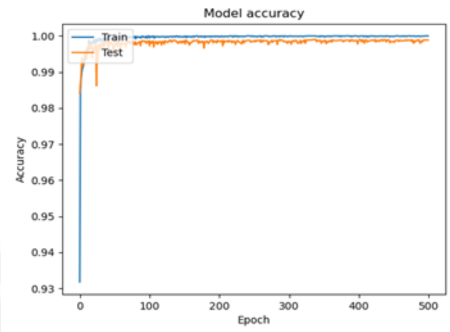


Figure. II. Accuracy Graph

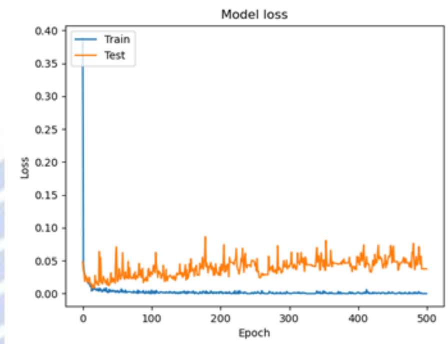


Figure.III. Loss Graph

## VIII.CONCLUSION AND FUTURE WORK

Revolutionary Efficiency in ASD Classification: The proposed Residual Attention CNN deep learning architecture establishes an unprecedented level of efficiency in accurately classifying children with autism spectrum disorder (ASD), paving the way for enhanced diagnostic capabilities and support. Attention-Based Insights from Facial Expressions: This work pioneers the integration of attention-based mechanisms to extract crucial insights from facial expressions. By deciphering nuanced features, the model showcases its potential to glean essential diagnostic information from visual cues.

Feature Significance Learning for Unparalleled Accuracy: At the heart of this innovation lies a groundbreaking feature—each parameter is meticulously learned and weighted based on its significance. This ingenious strategy ensures that pivotal parameters from the initial layers profoundly influence the network's deeper layers, culminating in remarkable accuracy and stability in classification. Universal ASD Classification across Varied Contexts: An outstanding facet of the proposed model is its ability to classify ASD in children across random time stamps and diverse environmental settings. It liberates the classification process from specific scenarios, enabling accurate predictions from images or videos captured in any context.

Unrestricted Application and High Accuracy: This model's uniqueness thrives in its capacity to classify ASD from images or videos taken at any moment, unbound by predefined conditions. The accuracy achieved in manual videos collected from ASD children, approximately 99%, speaks volumes about its reliability. Holistic Age Group Extension for Wider Impact: The vision for the future extension of this model is visionary encompassing all age categories, from children to adults and the elderly. This ambitious trajectory aims to establish a universal diagnostic framework for Autism Spectrum Disorder, bolstered by its adaptable inference capabilities.

Redefining ASD Detection and Intensity Prediction: The ultimate goal of this research is nothing short of groundbreaking: discerning ASD presence and predicting its intensity from a mere image or video, transcending temporal and environmental constraints. The scope expands further to distinguish between affected and unaffected individuals, propelling the potential of this model to reshape diagnostic paradigms.

## REFERENCES

[1] Hashemi, Jordan, Mariano Tepper, Thiago Vallin Spina, Amy Esler, Vassilios Morellas, Nikolaos Papanikolopoulos, Helen Egger, Geraldine Dawson, and Guillermo Sapiro. "Computer vision tools for low-cost and noninvasive measurement of autism-related behaviors in infants." Autism research and treatment 2014 (2014).

[2] Smitha, Kavallur Gopi, and A. Prasad Vinod. "Facial emotion recognition system for autistic children: a feasible study based on FPGA implementation." Medical & biological engineering & computing 53, no. 11 (2015): 1221-1229.

[3] Garman, Heather D., Christine J. Spaulding, Sara Jane Webb, Amori Yee Mikami, James P. Morris, and Matthew D. Lerner. "Wanting it Too Much: An Inverse Relation Between Social Motivation and Facial Emotion Recognition in Autism Spectrum Disorder." Child Psychiatry & Human Development 47, no. 6 (2016): 890-902.

[4] McPartland, James C., Sara Jane Webb, Brandon Keehn, and Geraldine Dawson. "Patterns of visual attention to faces and objects in autism spectrum disorder." Journal of autism and developmental disorders 41, no. 2 (2011): 148-157.

[5] Harms, Madeline B., Alex Martin, and Gregory L. Wallace. "Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies." Neuropsychology review 20, no. 3 (2010): 290-322.

[6] Wong, Nina, Deborah C. Beidel, Dustin E. Sarver, and Valerie Sims. "Facial emotion recognition in children with high functioning autism and children with social phobia." Child Psychiatry & Human Development 43, no. 5 (2012): 775-794.

[7] Macdonald, Hope, Michael Rutter, Patricia Howlin, Patricia Rios, Ann Le Conteur, Christopher Evered, and Susan Folstein. "Recognition and expression of emotional cues by autistic and normal adults." Journal of Child Psychology and Psychiatry 30, no. 6 (1989): 865-877.

[8] Humphreys, Kate, Nancy Minshew, Grace Lee Leonard, and Marlene Behrmann. "A fine-grained analysis of facial expression processing in high-functioning adults with autism." Neuropsychologia 45, no. 4 (2007): 685-695.

[9] Rump, Keiran M., Joyce L. Giovannelli, Nancy J. Minshew, and Mark S. Strauss. "The development of emotion recognition in individuals with autism." Child development 80, no. 5 (2009): 1434-1447.

[10] Smith, Miriam J. Law, Barbara Montagne, David I. Perrett, Michael Gill, and Louise Gallagher. "Detecting subtle facial emotion recognition deficits in high-functioning autism using dynamic stimuli of varying intensities." Neuropsychologia 48, no. 9 (2010): 2777-2781.

[11] Sucksmith, E., C. Allison, S. Baron-Cohen, B. Chakrabarti, and R. A. Hoekstra. "Empathy and emotion recognition in people with autism, first-degree relatives, and controls." Neuropsychologia 51, no. 1 (2013): 98-105.

[12] Chandler, Susie, Patricia Howlin, Emily Simonoff, Tony O'sullivan, Evelin Tseng, Juliet Kennedy, Tony Charman, and Gillian Baird. "Emotional and behavioural problems in young children with autism spectrum disorder." Developmental Medicine & Child Neurology 58, no. 2 (2016): 202-208.

[13] Chen, Chien-Hsu, I-Jui Lee, and Ling-Yi Lin. "Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions." Computers in Human Behavior 55 (2016): 477-485.

[14] L. Berkovits, A. Eisenhower, J. Blacher, , "Emotion Regulation in Young Children with Autism Spectrum Disorder,"Journal of Autism and Developmental Disorders, Volume 47, Issue 1, (2017): 68–79.

[15] Deodhare, Dipti. "Facial Expressions to Emotions: A Study of Computational Paradigms for Facial Emotion Recognition." In Understanding Facial Expressions in Communication, pp. 173-198. Springer India, 2015.

[16] Palestra, Giuseppe, Adriana Pettinicchio, Marco Del Coco, Pierluigi Carcagnì, Marco Leo, and Cosimo Distante. "Improved performance in facial expression recognition using 32 geometric features." In International Conference on Image Analysis and Processing, pp. 518-528. Springer, Cham, 2015.

[17] Ayesh, Aladdin, and William Blewitt. "Models for computational emotions from psychological theories using type I fuzzy logic." Cognitive Computation 7, no. 3 (2015): 285-308.

[18] Ngo, Thi Duyen, Thi Hong Nhan Vu, and Viet Ha Nguyen. "Improving simulation of continuous emotional facial expressions by analyzing videos of human facial activities." In International Conference on Principles and Practice of Multi-Agent Systems, pp. 222-237. Springer, Cham, 2014.

[19] Bakshi, Urvashi, and Rohit Singhal. "A survey on face detection methods and feature extraction techniques of face recognition." International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) 3, no. 3 (2014): 233-237.

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11s***
***DOI: https://doi.org/10.17762/ijritcc.v11i11s.8160***
***Article Received:** 26 June 2023 **Revised:** 15 August 2023 **Accepted:** 08 September 2023*

[20] Heinsfeld, Anibal Sólon, et al. "Identification of autism spectrum disorder using deep learning and the ABIDE dataset." NeuroImage: Clinical 17 (2018): 16-23.

[21] Babu, G. H., Srinivas, M., Gnanaprakasam, C., Prabu, R. T., Devi, M. R., Ahammad, S. H., ... & Rashed, A. N. Z. (2023). Meander Line Base Asymmetric Co-planar Wave Guide (CPW) Feed Tri-Mode Antenna for Wi-MAX, North American Public Safety and Satellite Applications. Plasmonics, 18(3), 1007-1018.

[22] Krishnamoorthy, N. V., KH, S. M., Gnanaprakasam, C., Swarna, M., & Geetha, R. (2023, April). A Robust Blockchain Assisted Electronic Voting Mechanism with Enhanced Cyber Norms and Precautions. In 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-8). IEEE.

[23] Mohan, A., & K, S. . (2023). Computational Technologies in Geopolymer Concrete by Partial Replacement of C&D Waste. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 282–292. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2666.

[24] Swarna, M., Geetha, R., Saranya, G., KH, S. M., & Gnanaprakasam, C. (2023, April). An Empirical Design of IoT based Health Surveillance Scheme for Coronavirus Affected Patients. In 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-8). IEEE.

[25] Geetha, R., Krishnamoorthy, N. V., Murugan, K. S., Gnanaprakasam, C., & Swarna, M. (2023, April). A Novel Deep Learning based Stress Analysis and Detection Scheme using Characteristic Data. In 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-8). IEEE.

[26] Gnanaprakasam, C., Indumathy, M., Khilar, R., & Kumar, P. S. (2022). Artificial intelligence based optimization for mapping IP addresses to prevent cyber-based attacks. Measurement: Sensors, 24, 100508.

[27] Gnanaprakasam, C., Anand, S., Manoj Kumar, R., & Menaka, R. (2021). Facial Expression Image Analysis to Classify High and Low Level ASD Kids Using Attention Mechanism Embedded Deep Learning Technique. In Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2020 (pp. 559-568). Springer Singapore.

[28] Diksha Siddhamshittiwar. (2017). An Efficient Power Optimized 32 bit BCD Adder Using Multi-Channel Technique. International Journal of New Practices in Management and Engineering, 6(02), 07 - 12. https://doi.org/10.17762/ijnpme.v6i02.57.

[29] C. Gnanaprakasam, Manoj Kumar Rajagopal, Attention Residual Network for Micro-expression Recognition Using Image Analysis, in Journal of Advanced Research in Dynamical & Control Systems, 07-Special Issue,2020, Pages- 1261 – 1272.

[30] Gnanaprakasam, C., and Manoj Kumar Rajagopal. "Review on Facial Micro-Expression Detection." Int J Innov Technol Explor Eng 8 (2019): 1103-1115.

[31] Sumathi, S., C. Gnanaprakasam, and R. RANIHEMA MALINI. "Face Recognition-Average-Half-Face Using Wavelets." IPCV 2010: proceedings of the 2010 international conference on image processing, computer vision, & pattern recognition (Las Vegas NV, July 12-15, 2010). 2010.

[32] Gnanaprakasam, C., S. Sumathi, and R. RaniHema Malini. "Average-half-face in 2D and 3D using wavelets for face recognition." Proceedings of the 9th WSEAS international conference on Signal processing. 2010.

[33] Gnanaprakasam, C., and M. Rajagopal. "K.,(2023). Identification of Autism Spectrum Disorder using Residual Attention Net-work for Facial Image Analysis." J Curr Trends Comp Sci Res 2.1: 31-39.

[34] Fatima Abbas, Deep Learning Approaches for Medical Image Analysis and Diagnosis , Machine Learning Applications Conference Proceedings, Vol 3 2023.

[35] Patnam, Venkata Sindhoor Preetham, et al. "Deep learning based recognition of meltdown in autistic kids." 2017 IEEE International Conference on Healthcare Informatics (ICHI). IEEE, 2017.

[36] Florio, Tony, et al. "Providing an independent second opinion for the diagnosis of autism using artificial intelligence over the internet." Couns, Psycho Health Use Technol Mental Health 5 (2009): 232-248

TABLE III. COMPARISON BETWEEN THE EXISTING MODELS.

| Reference | Data Type | Methods | Prediction goal | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|---|
| Our Model | Data images | Deep learning | ASD/ non - ASD | 99 | 97 | 97 |
| Heinsfeld et al. (2018) [20] | rs-fMRI | Deep learning | ASD/TD | 70 | 74 | 63 |
| Preetham Patnam et al. (2017) [32] | Data images | Deep learning | Meltdown | 92 | - | - |
| Florio et al. (2009) [33] | Developmental Behavior Checklist | Neural Networks | ASD / Non ASD | 80 | 92 | 70 |