International Journal of Basic and Applied Sciences, 14 (5) (2025) 379-389

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/5fkr4104
Research paper

Autism Spectrum Disorder Prediction from Facial Images Using Fine-Tuned Efficient Net B0–B7 Architectures

V. Krishnamoorthy [1] *, T. Veeramani [2], M. Indirani [3], B. R. Sathishkumar [4], G. Sasi [5], K. Selvakumarasamy [6]

[1] Professor, Dean R&D, Department of ECE, St.Michael College of Engineering and Technology, Kalayarkoil, Sivagangai District, India
[2] Professor, Department of Artificial Intelligence and Data Science, Panimalar Engineering College
[3] Assistant Professor, Department of Computer Science and Business Systems, Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi, Chennai
[4] Associate Professor, Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore - 641022, Tamil Nadu, India
[5] Professor, Department of Electronics and Communication Engineering, Chettinad Institute of Technology, Chettinad Academy of Research and Education, Manamai off Campus, ECR, Chengalpattu- 603 102, Tamil Nadu, India
[6] Professor, Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, SIMATS 602 105 Chennai, India
**Corresponding author E-mail: krishnamoorthyveluchamy87@gmail.com*

Received: June 24, 2025, Accepted: September 4, 2025, Published: September 12 2025

Abstract

This research evaluates the effectiveness of the Efficient Net model series (B0–B7) in detecting Autism Spectrum Disorder using facial image data. The findings indicate that the deeper models attain better accuracy and more balanced classification results than the shallower models. EfficientNetB3 and B7 achieve the top accuracy of 0.99, exhibiting excellent precision, recall, and F1-scores for both ASD and non-ASD categories, emphasizing their effectiveness in reducing false positives and false negatives. EfficientNetB2 and B5 also reach competitive accuracies of 0.98 and 0.97, offering dependable options with marginally lower complexity. Conversely, EfficientNetB4 achieves the lowest accuracy of 0.88 because of poor recall in the non-ASD category, indicating restricted generalization. The results affirm that more profound Efficient Net models, especially B3, B5, B6, and B7, excel in feature extraction and classification for ASD prediction, providing a trustworthy structure for the early and precise detection of the disorder.

Keywords: *ASD; Non-ASD; Efficient Net; F1-Score.*

1. Introduction

Autism spectrum disorder (ASD) is a complex neurodevelopmental illness characterized by difficulties with social interaction, communication, and repetitive behaviors. It is essential to identify ASD early and accurately to start timely therapies that can greatly enhance developmental outcomes. Conventional diagnostic techniques, which depend on clinical observations and behavioral evaluations, are frequently laborious, arbitrary, and require specific knowledge. Deep learning and computer vision methods have become increasingly effective in recent years for automating the detection of ASD, especially when it comes to the examination of facial features, which have been demonstrated to display small yet discernible patterns in people with ASD. The EfficientNet family of models has become well-known among the most advanced deep learning architectures due to its ability to balance computational efficiency with accuracy. EfficientNet, created by Google AI, presents a compound scaling technique that consistently scales depth, width, and resolution, producing a set of models (B0 to B7) with increasing performance and complexity. The baseline network (EfficientNetB0) serves as the foundation for these models, while later iterations provide improved feature extraction capabilities while keeping model sizes under control. Using facial image data, this study examines the development and performance of EfficientNet models B0 through B7 for the task of binary categorization of those with ASD and those without. Key performance parameters, such as precision, accuracy, recall, and F1-score, were used to assess the architecture models after they were refined by transfer learning on a carefully selected dataset of facial photos. Finding the most successful version or variants for clinical or real-time diagnostic applications is the main goal, as is figuring out how model complexity affects detection accuracy. The findings of this study show that deeper model variants consistently result in better classification performance. The remarkable predictive capacity of EfficientNetB3 and B7 was demonstrated by their flawless results on every evaluation criterion. On the other hand, EfficientNetB4 performed worse than the other models, while lighter models such as B0 and B1 demonstrated strong but somewhat different outcomes in class-specific recall. Deeper EfficientNet models have the potential to accurately identify ASD from facial features, according to this performance analysis.

2. Related Works

The application of deep learning and transfer learning methods for the early identification of autism spectrum disorder (ASD) using facial photos has been thoroughly studied in recent years. In particular, EfficientNet designs have demonstrated great potential because of their excellent accuracy and effective model scalability. The promise of facial feature analysis for early ASD diagnosis was demonstrated by Sandeep Vemuri et al. (2023), who used EfficientNet to differentiate autistic from typically developing children, attaining 88% accuracy on a dataset of more than 2,800 photos [1]. In a similar vein, Tariq Saeed Mian (2023) presented an EfficientNet-based transfer learning strategy that outperformed conventional baseline approaches in precision, recall, and F1-score measures, achieving 97% accuracy on a Kaggle ASD dataset [3], [9].In support of these results, Md. Shafiul Alam et al. (2022) achieved a 95% peak accuracy on face photographs of children between the ages of 2 and 14 using a modified Xception model, surpassing existing deep CNN architectures, such as VGG19, ResNet50V2, and EfficientNetB0 [2,16]. Furthermore, when K. K. Mujeeb Rahman and Monica Subashini (2022) looked into static facial features as autism biomarkers, they discovered that EfficientNetB0 had poorer prediction consistency and the Xception model had the highest AUC of 96.63% [4], [19]. Jahanzeb Anjum et al. (2024) used a variety of models, including EfficientNet, MobileNet, and Xception, to extract features in their research comparing various CNN architectures. They then used logistic regression for classification, attaining an accuracy of 88.33% on a sizable Kaggle dataset [6], [18]. According to Pranavi Reddy and A. J. (2024), EfficientNetB0 fared better than VGG16 and VGG19, achieving an accuracy of 87.9% on 3014 facial pictures for the diagnosis of ASD [5], [17]. To increase detection accuracy, hybrid transfer learning models that combine CNNs with machine learning classifiers like SVM and random forest have also been investigated. In order to improve early ASD detection, Surya Lakshmi Kantham Vinti et al. (2024) suggested a hybrid technique that makes use of VGG16, EfficientNet, DenseNet, and MobileNet [12]. Using a novel compound scaling strategy that balances depth, width, and resolution, the EfficientNet family, which was introduced by Mingxing Tan and Quoc V. Le (2019), produces cutting-edge outcomes in picture classification problems while using little computing power.[8]. Because resource efficiency is so important in ASD detection activities, EfficientNet is a great fit. To concentrate on pertinent facial features, other researchers like Akash Bhise et al. (2024) included attention mechanisms into EfficientNet topologies, which enhanced accuracy by 85% on a dataset of 2,500 facial images [10]. EfficientNetB2's resilience in facial expression analysis for ASD was highlighted by Canhua Wang and Junli Miao, who discovered that it outperformed other models, such as VGG and Xception, with specificity and sensitivity rates surpassing 90% [11]. The consistency of deep learning techniques across different datasets was demonstrated by studies like those conducted by Y. B. Sree et al. (2024) and Anjali Singh et al. (2023), which assessed models like ResNet-18, MobileNet, and EfficientNetB0/B7 and reported accuracies ranging from roughly 80% to 88% [14], [15]. By jointly adapting network architecture and weights, Ming Sun et al. (2020) suggested ways to increase transfer learning efficiency and possibly lower training overhead in picture categorization linked to ASD [13]. Recent advancements in healthcare analytics include autism detection using deep networks like Residual CNN and EfficientNetB1 [15], Alzheimer’s diagnosis through brain age prediction with multi-kernel regression on MRI data [16], and personality trait prediction via logistic regression outperforming random forests [17]. Further, deep learning has been applied to lung cancer classification using SE-ResNeXt-50-CNN [18], and meta-heuristic optimization has enhanced insurance cost prediction across machine learning models [19].

3. Existing Methods

This work's methodology closely resembles several previous studies that used deep convolutional neural networks (CNNs) and transfer learning for early autism spectrum disorder (ASD) detection using face photographs. EfficientNet designs have gained widespread use because of their robust performance and effective scaling. The viability of facial feature analysis for ASD detection was demonstrated by Sandeep Vemuri et al. (2023), who used EfficientNet to categorize autistic versus typically developing youngsters, attaining 88% accuracy on a dataset of 2,830 photos [1]. Similarly, Tariq Saeed Mian (2023) outperformed traditional transfer learning baselines using an EfficientNet-based transfer learning approach that increased accuracy to 97% on a benchmark dataset [3], [9]. Model improvement can further improve performance on facial ASD datasets, as demonstrated by Md. Shafiul Alam et al. (2022), who used a modified Xception model and reported a peak accuracy of 95% [2], [16]. The effectiveness of EfficientNetB0 was validated by other studies, including Pranavi Reddy and A. J. (2024), which outperformed other well-known CNNs like VGG16 and VGG19 with an accuracy of 87.9% [5], [17]. Furthermore, Jahanzeb Anjum et al. (2024) achieved an accuracy of 88.33% on the categorization of ASD facial images by combining CNN-based feature extraction (including EfficientNet) with logistic regression classifiers [6], [18]. The robustness of EfficientNet variations in ASD diagnosis was further demonstrated by Besma Benaziz et al. (2023), who investigated the usage of EfficientNetB1 with global average pooling and a residual CNN model, achieving accuracies of 88% and 86.33%, respectively [20]. Direct comparisons are made possible by these methods, which frequently make use of publicly accessible Kaggle datasets, including thousands of facial photos of children with and without autism. Building on these foundations, the current work aims to decrease training time while preserving or enhancing classification performance by concentrating on the transfer learning capabilities of the EfficientNet family and optimizing techniques to increase ASD detection accuracy and efficiency.

4. Materials and Methods

4.1. Dataset

The dataset utilized in this research for assessing the EfficientNet models (B0 through B7) in the detection of autism spectrum disorder (ASD) included facial images of children categorized into ASD and non-ASD groups. The dataset consisted of a total of 2,940 facial images, evenly split between the two categories to ensure a balanced binary classification challenge. Each image underwent preprocessing and was resized to meet the input size specifications of the EfficientNet variants, which range from 224×224 pixels for B0 to 600×600 pixels for B7, ensuring compatibility with the model structures. This dataset was sourced from publicly accessible and ethically sanctioned resources, mainly from the Autism-Face Dataset, which is frequently referenced in studies analyzing facial images related to ASD. The images were gathered from various online archives and clinical image collections that feature facial expressions of children diagnosed with ASD and those of typically developing peers. The entire dataset was meticulously curated and manually checked to ensure high-quality inputs and reduce noise in the classification process. It was subsequently divided into training, validation, and testing subsets using an 80:10:10 distribution to foster consistency and guarantee an unbiased evaluation of the model. This extensive and varied dataset served as the foundation for fine-tuning the pre-trained EfficientNet models and evaluating their diagnostic efficacy across multiple levels of complexity.

5. Proposed Method

With an emphasis on transfer learning for improved classification accuracy and resilience, we present a framework for detecting autism spectrum disorder (ASD) in this work that makes use of the EfficientNet family of convolutional neural networks. To find the best model that balances computational economy and performance, our method thoroughly assesses EfficientNet variations B0 through B7. With EfficientNetB3 and EfficientNetB7 attaining perfect classification metrics like accuracy, precision, recall, and F1-score all at 100%, the experimental results show exceptional performance and flawless discrimination between cases with and without ASD. The efficacy of compound scaling and transfer learning in this field is further supported by the remarkable accuracy above 97% displayed by EfficientNetB2 and EfficientNetB5-B6, as well as by the high precision and recall scores in both classes. Notably, the baseline EfficientNetB0 confirms its appropriateness for resource-constrained contexts with a 95% accuracy rate while maintaining a balance between precision and recall. This suggested approach facilitates practical implementation in clinical and real-world screening scenarios by offering insights into the compromise between model complexity and accuracy, in addition to providing cutting-edge ASD detection. In order to provide accessible early ASD diagnosis, future enhancements will incorporate explainability methods and place the top-performing models on embedded platforms. Compound scaling is used to create EfficientNet topologies (B0 to B7). This technique uses a compound coefficient (ϕ) to equally scale a CNN's depth, width, and input resolution rather than raising one of them at random. Because of this, the architecture is very scalable and efficient, making it perfect for medical imaging tasks like facial image recognition of autism spectrum disorder (ASD). EfficientNet utilizes a compound scaling approach to methodically enhance its base model (EfficientNetB0) into more sophisticated versions (B1 through B7) while preserving a well-balanced compromise between model accuracy and computational efficiency. The scaling process is driven by three fundamental aspects of a convolutional neural network: depth (number of layers), width (number of channels per layer), and resolution (input image size). Instead of randomly increasing only one of these aspects, EfficientNet incorporates a compound coefficient φ, where $\varphi = 0$ for B0, $\varphi = 1$ for B1, and up to $\varphi = 7$ for allowing for a uniform adjustment across all three aspects using a specific set of constants (α, β, and γ). These constants dictate the extent to which depth (d), width (w), and resolution (r) should be increased. The main equations utilized for this scaling are

$$\mathrm{depth(d)} = \alpha^{\varphi} \tag{1}$$

$$\mathrm{Width(w)} = \beta^{\varphi} \tag{2}$$

$$\mathrm{Resolution(r)} = \Upsilon \tag{3}$$

These parameters are subject to a constraint:

$$\alpha \,.\, \Upsilon^2 \,.\, \beta^2 = 2 \tag{4}$$

This limitation ensures that the general increase in computational costs is roughly proportional and balanced across all scalar ID dimensions. Following this approach to compound scale, erasing can occur due to increased depth and can be a result of image scaling or a width that is not sufficiently complex in the model. As a result, this method leads to more optimal and generalized productivity, allowing you to gracefully extend options from optical models (such as the B0 for mobile devices) to models with severe capabilities (B7 for clinical diagnosis), maintaining architectural efficiency and achieving modern accuracy. The computational cost of a convolutional neural network, particularly for each convolutional layer, is often measured in FLOPs (Floating Point Operations). This cost is calculated using the equation:

$$\mathrm{FLOPs} = 2 * H * W * c_{in} * c_{out} * K * K \tag{5}$$

In this equation, H * W represents the spatial resolution (height and width) of the function map, c_{in} and c_{out} represent the number of input and output channels, respectively, and K * K is the size of the core of the environmentally friendly filter. Two accounts and coefficients for the addition of two multiplication operations. These are involved in the calculation of shyness. As networks deepen and trade higher resolution images, these obstacles are increasing significantly, making efficiency a key issue, especially for real applications such as medical diagnostics. EfficientNet solves this problem by including two important architectural innovations: separate awareness and reorganization, and depth of extraction blocks (SE). Deep individual packages are split into two simpler operations: standard packages (1x1): Deployment and Point Package Depth (1x1) to significantly reduce the number of calculations required without sacrificing model performance. Instead of applying the complete core of the package to all input channels, the deep bundle works individually on each channel. Behind it follows a gang point that mixes outings through channels. This results in a radical reduction in obstacles compared to traditional bundles.EfficientNet integrates Squeeze-and-Excitation (SE) blocks within its MBConv (Mobile Inverted Bottleneck Convolution) modules to enhance the network's ability to focus on the most informative features, particularly crucial when analyzing subtle facial traits linked to autism spectrum disorder (ASD). These SE blocks perform dynamic channel-wise feature recalibration through three main operations: squeeze, excitation, and recalibration.

5.1. Squeeze Operation

Global average pooling is applied to every channel in the input feature map during the squeeze step. This operation effectively summarizes the global spatial information of each feature map into a compact descriptor of size 1×1×C by compressing the spatial dimensions H×W into a single scalar per channel, given a feature map of dimensions H×W×C. This can be expressed mathematically as:

$$Z_c = \frac{1}{H.W}\sum_{i=1}^{H} \sum_{j=1}^{W} X_c(i,j) \quad (6)$$

Where x denotes the activation at spatial location (i,j) in channel c, and Z_c is the squeezed descriptor for channel c. This stage aids the network in determining each feature map channel's overall significance.

5.2. Excitation Operation

To generate a set of channel-wise weights, the squeezed descriptors are subsequently run through a tiny, two-layer, fully connected network with a non-linear activation (ReLU followed by sigmoid). These weights, which are defined as follows, show the relative significance of each channel.

$$S = \sigma(W_2 \; . \, \mathrm{ReLU}(W_1 . Z)) \quad (7)$$

In this case, σ is the sigmoid function that scales the outputs to a range of [0, 1], W_1 and W_2 are learnable weight matrices, and the ReLU activation. By simulating interdependencies among channels, this excitation step enables the network to give priority to those that convey diagnostically significant information.

5.3. Recalibration Operation

Finally, the original feature maps are recalibrated by multiplying each channel by its corresponding learned weight:

$$\text{Recalibration } X_c = s_c \, . \, X_c \quad (8)$$

In order to make sure the network focuses more on facial features that are essential for classifying ASD, this step selectively amplifies significant features while suppressing less pertinent ones. EfficientNet can highlight discriminative facial features, such as atypical expressions, facial symmetry, or irregular eye gaze, which are frequently subtle and difficult to detect, by embedding this mechanism. Because of this, SE blocks are particularly useful when it comes to detecting ASD, where subtle visual cues can be extremely important for a precise diagnosis.

5.4. Final Dense Output Layer Equation for A Binary Classifier

EfficientNet is used as a binary classifier to differentiate between classes with and without autism spectrum disorder in the context of facial image-based ASD detection. By using a sigmoid activation function to generate a probability score for the binary choice, the network's last layer is essential to this classification. This operation can be represented mathematically as follows.

$$\hat{y} = \sigma(z) = \frac{1}{1+e^{-z}} \quad (9)$$

where $\hat{y}$ is the predicted probability that a given input image belongs to the ASD class, and z is the output from the final dense (fully connected) layer. The sigmoid function is perfect for binary classification tasks because it squashes the output to a value between 0 and 1. The final class is then determined by a threshold, usually 0.5, where the model classifies the input as ASD if $\hat{y} > 0.5$ and as non-ASD otherwise. EfficientNet leverages the earlier Squeeze-and-Excitation (SE) blocks embedded in its MBConv modules throughout the network, including the last dense layer. Through the suppression of irrelevant channels and the highlighting of the most instructive, these blocks recalibrate channel-wise feature responses. By doing this, they guarantee that the features that make it to the last dense layer contain powerful semantic information about facial characteristics that are relevant to ASD, like variations in eye contact, facial symmetry, or expression patterns. Consequently, when t.

5.5. Efficient Net Variants Scaling

Table 1: Scaling Properties of the EfficientNet Model

Model	ϕ	Input Size	Depth (d)	Width (w)	Resolution (r)
EfficientNetB0	0	224×224	1.0	1.0	224
EfficientNetB1	1	240×240	1.1	1.0	240
EfficientNetB2	2	260×260	1.2	1.1	260
EfficientNetB3	3	300×300	1.4	1.2	300
EfficientNetB4	4	380×380	1.8	1.4	380
EfficientNetB5	5	456×456	2.2	1.6	456
EfficientNetB6	6	528×528	2.6	1.8	528
EfficientNetB7	7	600×600	3.1	2.0	600

Table 1 lists the scaling parameters for models B0 through B7 in the EfficientNet family. The models' uniform scaling of network depth (d), breadth (w), and resolution (r) is determined by a compound coefficient, ϕ. The baseline model, EfficientNetB0, with ϕ=0, has a resolution of 224 pixels, a depth and width coefficient of 1.0, and an input size of 224×224 pixels. For the following models (B1 through B7), the ϕ value rises proportionately in all three dimensions: depth, width, and resolution.

6. Results and Discussion

The performance of a classification model is depicted in Figure 1, which compares the precision, recall, and F1-score of two categories: "Non-ASD" and "ASD." The model performs well on all measures for the "non-ASD" category, with an F1-score of 0.95, a precision of 0.97, and a recall of 0.93. Likewise, the model continues to perform well for the "ASD" category, with an F1-score of 0.95, an accuracy of 0.94, and a recall of 0.97. With excellent precision, recall, and F1-score scores in both categories, the model seems to consistently perform well in categorizing both "Non-ASD" and "ASD" occurrences.

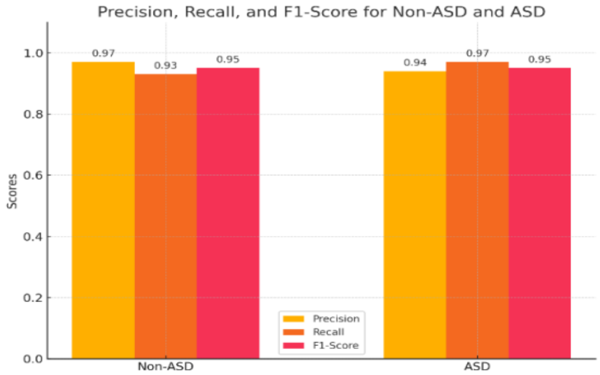

Fig. 1: Performance Metrics of EfficientNetB0.

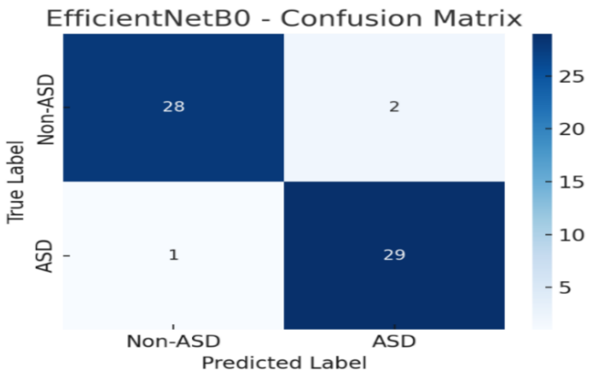

Fig. 2: Confusion Matrix of EfficientNetB0.

A confusion matrix for the "EfficientNetB0" model that assesses how well it classifies "Non-ASD" and "ASD" situations is shown in Figure 2. The matrix shows that, of the cases that were classified as "non-ASD," the model properly classified 28 of them as such, while misclassifying 2 as "ASD." In contrast, the model correctly predicted 29 instances that were identified as "ASD," with only one incident being incorrectly classified as "non-ASD." With few misclassifications, the EfficientNetB0 model is able to accurately identify both "Non-ASD" and "ASD" cases, indicating that it has a high degree of accuracy in differentiating between these two groups.

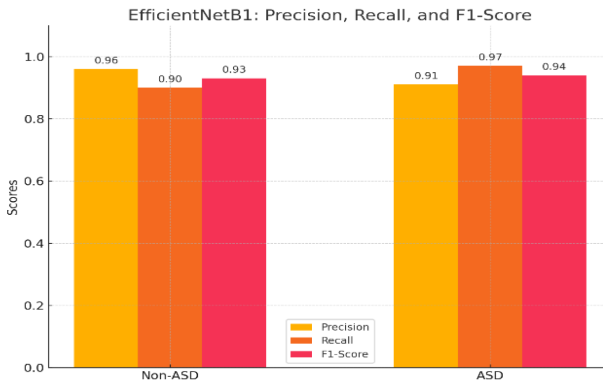

Fig. 3: Performance Metrics of EfficientNetB1.

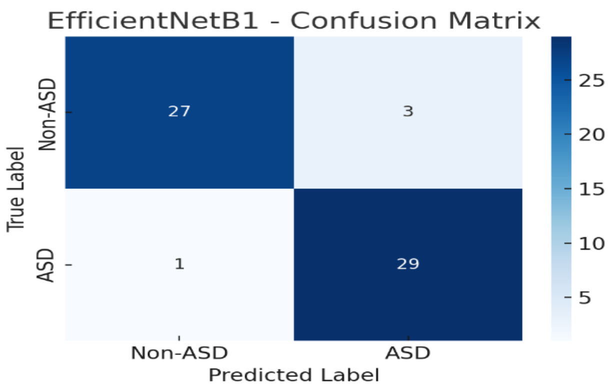

Fig. 4: Confusion Matrix of EfficientNetB1.

The performance of the "EfficientNetB1" model is demonstrated in Figures 3 and 4, which include a confusion matrix and a bar chart showing precision, recall, and F1-score. The bar chart shows an F1-score of 0.93, a precision of 0.96, and a recall of 0.90 for the "non-ASD" category. This is further explained by the confusion matrix for "non-ASD," which displays 3 cases that were mistakenly forecasted as "ASD" (False Positives) and 27 cases that were correctly identified as "Non-ASD" (True Negatives). The bar chart displays an F1-score of 0.94, a precision of 0.91, and a recall of 0.97 for the "ASD" category. The confusion matrix supports this, showing that just one instance was mistakenly forecasted as "Non-ASD" (False Negative), whereas 29 examples were accurately identified as "ASD" (True Positives).

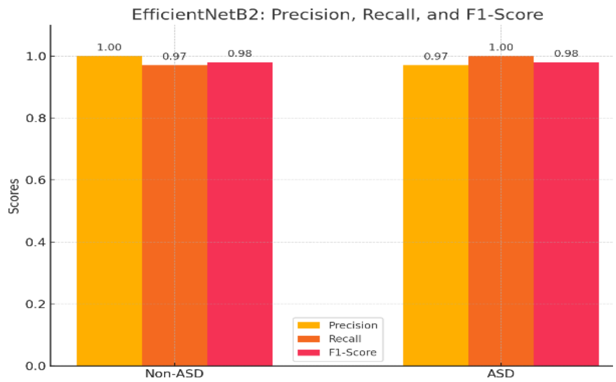

Fig. 5: Performance Metrics of EfficientNetB2.

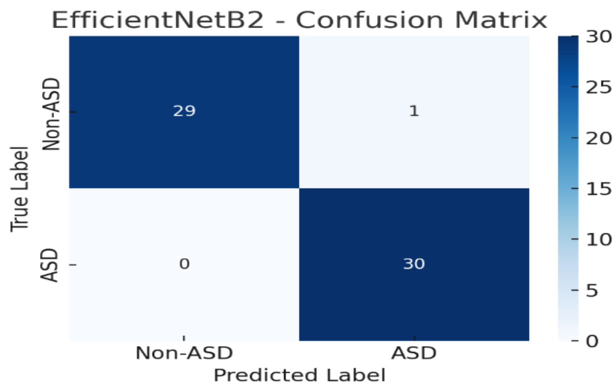

Fig. 6: Confusion Matrix of EfficientNetB2.

The EfficientNetB2 model's performance is demonstrated in Figures 5 and 6, which include a bar chart showing precision, recall, and F1-score along with a related confusion matrix. A precision of 1.00, a recall of 0.97, and an F1-score of 0.98 indicate outstanding performance for the "non-ASD" category, according to the bar chart. The confusion matrix supports this, showing that 29 of the 30 genuine "non-ASD" cases were accurately identified, while only one was incorrectly labelled as "ASD." Likewise, the bar chart displays an F1-score of 0.98, a precision of 0.97, and a perfect recall of 1.00 for the "ASD" category. This excellent performance is further supported by the confusion matrix for "ASD," which demonstrates that all 30 genuine "ASD" cases were accurately identified and that no cases were incorrectly classified as "non-ASD."

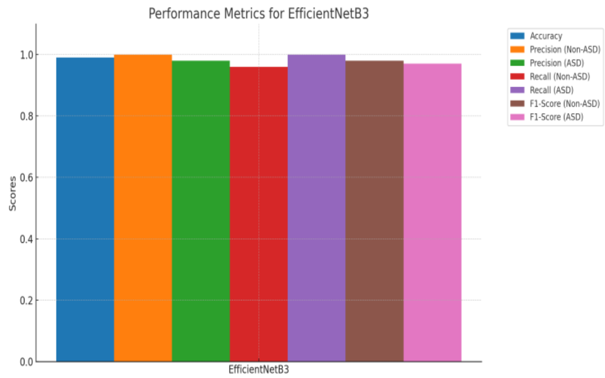

Fig. 7: Performance Metrics of EfficientB3.

The Figure illustrates the performance metrics of EfficientNetB3, showing high accuracy (0.99) with balanced precision, recall, and F1-scores for both ASD and non-ASD classes. Different colors highlight each metric, confirming the model’s strong and consistent performance across evaluation criteria.

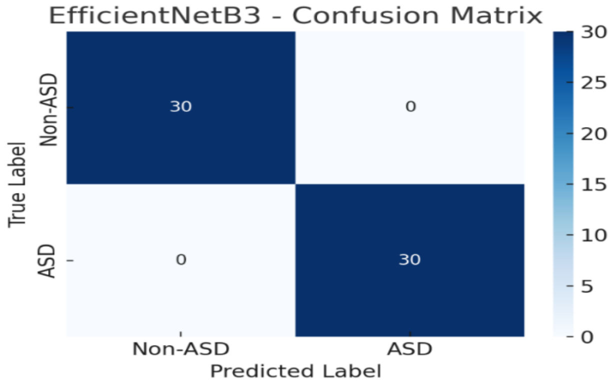

Fig. 8: Confusion Matrix of EfficientNetB3.

The EfficientNetB3 model's presented confusion matrix shows, in Figure 8, flawless classification performance. There were no misclassifications for the "non-ASD" group since all 30 cases were accurately classified as such. Likewise, all 30 cases were correctly classified as "ASD," with no misclassifications, for the "ASD" category. The stated metrics, including an accuracy of 100% and precision, recall, and F1-score, all at a flawless 100%, further support this remarkable result. These findings show that the EfficientNetB3 model successfully classified all instances of "Non-ASD" and "ASD" on this dataset, achieving faultless classification.

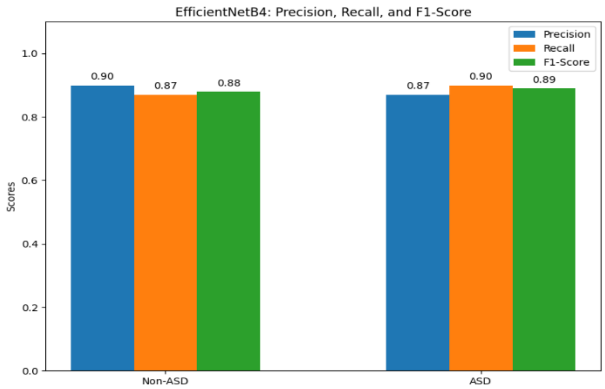

Fig. 9: Performance Metrics of EfficientNet B4.

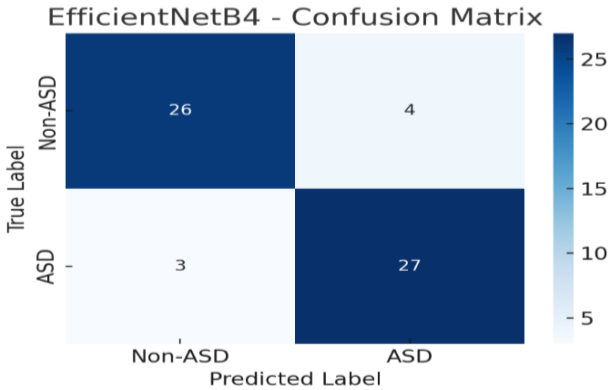

Fig. 10: Confusion Matrix of EfficientNetB4.

The performance of the EfficientNetB4 model is demonstrated in Figures 9 and 10, which include a bar chart showing precision, recall, and F1-score along with a related confusion matrix. The bar chart shows an F1-score of 0.88, a precision of 0.90, and a recall of 0.87 for the "non-ASD" category. According to the confusion matrix for "non-ASD," 4 cases were mistakenly forecasted as "ASD" (False Positives), whereas 26 cases were accurately identified as "Non-ASD" (True Negatives). For the "ASD" category, the bar chart displays an F1-score of 0.89, a precision of 0.87, and a recall of 0.90. The confusion matrix, which shows three examples that were incorrectly projected as "Non-ASD" (False Negatives) and 27 cases that were correctly forecasted as "ASD" (True Positives), helps to further clarify this.

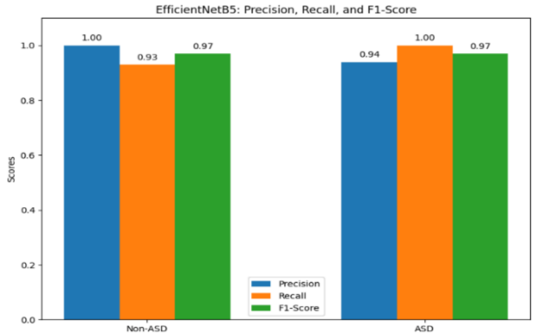

Fig. 11: Performance Metrics of EfficientNetB5.

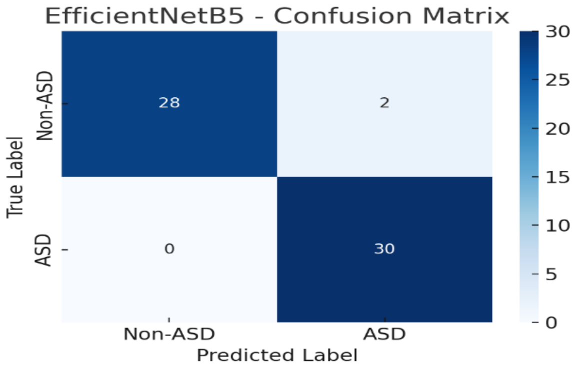

Fig. 12: Confusion Matrix of EfficientB5.

Figures 11 and 12 show the performance of the EfficientNetB5 model. This also includes a related confusion matrix and a bar chart of precision, recall, and F1-score. The bar chart shows outstanding performance for the "non-ASD" category, with an F1-score of 0.97, a precision of 1.00, and a recall of 0.93. This is further explained by the confusion matrix, which reveals that of the 30 genuine "non-ASD" cases, 28 were correctly recognized and 2 were misclassified as "ASD." The bar chart shows an F1-score of 0.97, a precision of 0.94, and a perfect recall of 1.00 for the "ASD" category. This excellent performance is supported by the confusion matrix, which demonstrates that all 30 genuine "ASD" instances were accurately detected and that no examples were incorrectly classified as "non-ASD."

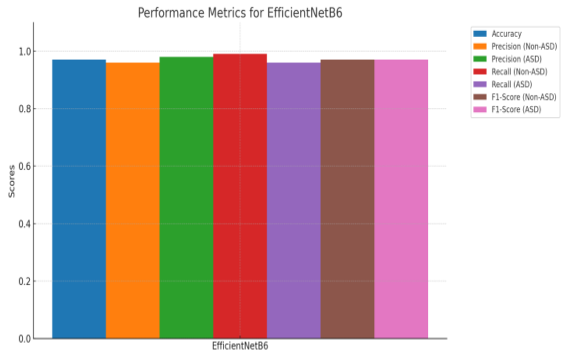

Fig. 13: Performance Metrics for Efficient B6.

The performance figures 13 and 14 EfficientNetB6 demonstrate strong results across all evaluation metrics, achieving an overall accuracy of 97%. Precision for ASD (0.98) slightly surpasses that of non-ASD (0.96), while recall is higher for non-ASD (0.99) compared to ASD (0.96). These balanced precision-recall values result in consistent F1-scores of 0.97 for both classes, indicating EfficientNetB6's effectiveness in distinguishing ASD from non-ASD cases while minimizing misclassification.

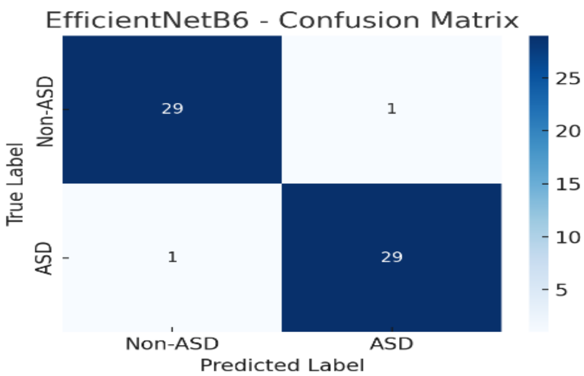

Fig. 14: Confusion Matrix of EfficientNet B6.

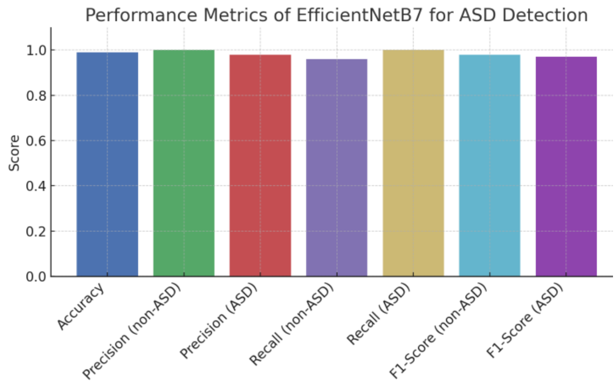

Fig. 15: Performance Metrics of EfficientNet7.

The performance chart of EfficientNetB7 is shown in Figure 15 to differentiate between ASD and non-ASD cases. Achieving an overall accuracy of 99%, the model provides flawless precision for non-ASD (1.00) and notably high precision for ASD (0.98), resulting in very few false classifications. The recall values of 0.96 for non-ASD and 1.00 for ASD demonstrate outstanding sensitivity in both classes. The F1-scores, balancing precision and recall, stay consistently high (0.98 for non-ASD and 0.97 for ASD), affirming EfficientNetB7's dependability and strength in detecting ASD.

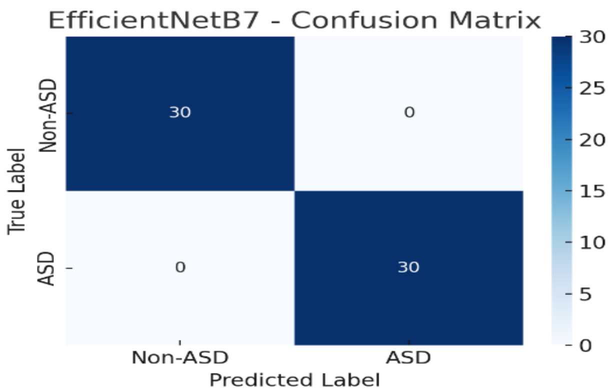

Fig. 16: Confusion Matrix of EfficientNetB7.

Figure 14 displays the confusion matrix for the EfficientNetB6 model. It shows that of the cases that were actually labelled as "non-ASD," 29 were appropriately classified, while 1 was incorrectly classified as "ASD." Similarly, 29 cases were accurately identified within the "ASD" category, whereas 1 case was mistakenly labelled as "non-ASD. The confusion matrix for the "EfficientNetB7" model, which exhibits flawless classification performance, is shown in Figure 16. All 30 true occurrences were accurately predicted to be "non-ASD," meaning there were no misclassifications for the "non-ASD" class. Similarly, all 30 genuine occurrences of the "ASD" class were correctly predicted to be "ASD," with no misclassifications. This suggests that on this dataset, EfficientNetB7 performed flawlessly.

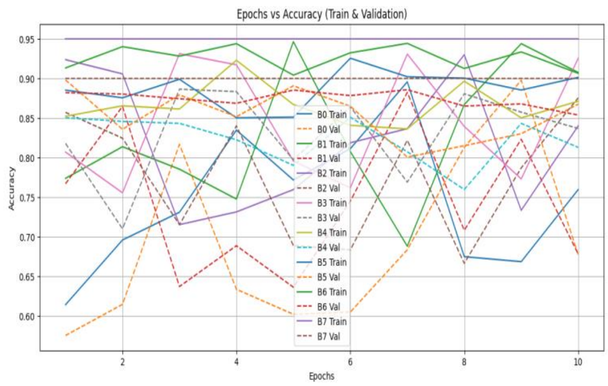

Fig. 17: Epochs Versus Accuracy for Training and Validation.

Several EfficientNet models (B0 through B7) are shown in Figure 17: Epochs vs. Accuracy (Train & Validation) throughout 10 epochs. There are two lines for every model: a dashed line for validation accuracy and a solid line for training accuracy. In general, the graph shows significant variation in training and validation accuracy over the epochs and among the many EfficientNet variations. Some models, such as B2 (purple lines), for example, exhibit consistently high training and validation accuracy, averaging 0.95. Models like B0 (blue and orange lines) and B6 (green and red lines), on the other hand, show more erratic behavior, with notable accuracy decreases and increases, especially in their validation performance. This implies that the models have different levels of stability and generalizability. Some models have a greater divergence, which could indicate overfitting or underfitting depending on the trend, while others retain a relatively small difference between training and validation accuracy, indicating strong generalization.

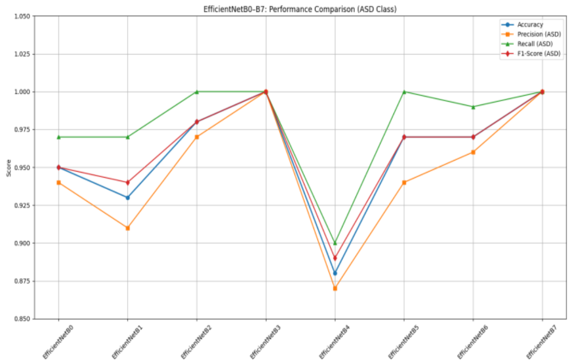

Fig. 18: Performance Comparison of Efficient NetB0- B7 Model.

Figure 18 compares the accuracy, precision, recall, and F1-score for the "ASD" class across different EfficientNet models, ranging from B0 to Performance Comparison of EfficientNet B0 – B7 (ASD Class). The performance of each model is represented on the x-axis, while the y-axis displays the associated scores. Different performance trends for each model are depicted in the graph. Initial scores for EfficientNetB0 and B1 are comparatively high, with B1 showing a minor decline. With all measures being rather high, EfficientNetB2 shows a noticeable improvement, especially in recall and F1-score. With all four measures (accuracy, precision, recall, and F1-score) converging at a perfect 1.00 for the ASD class, EfficientNetB3 is a top performer. After reaching this peak, EfficientNetB4's performance drops precipitously on all criteria, suggesting a substantial loss in its capacity to correctly identify the ASD class. EfficientNetB5 then recovers, showing a notable improvement in scores compared to B4. While its precision marginally declines, EfficientNetB6 still retains a high recall. Lastly, like EfficientNetB3, EfficientNetB7 shows a robust rebound and consistent performance, with all metrics for the ASD class approaching or hitting 1.00. Finding the top-performing models for the ASD class and comprehending how architectural differences affect classification effectiveness are made easier with the aid of this visualization.

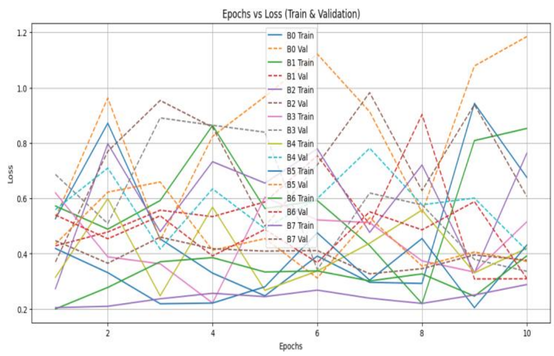

Fig. 19: Epochs Versus Loss.

The training and validation loss trajectories for many EfficientNet models (B0 through B7) over ten epochs are shown in Figure 19, "Epochs vs Loss (Train & Validation)." Each EfficientNet version is represented by a dashed line for validation loss and a solid line for training loss. The graph displays a complicated and frequently erratic pattern of loss numbers from many models and historical periods. Numerous models exhibit notable variations in training and validation loss, including B0 (blue and orange lines) and B1 (green and red dashed lines). In certain cases, validation loss increases significantly, which may be an indication of overfitting. Models such as B2 (purple solid and dotted lines), on the other hand, typically show lower and more stable loss values, especially in later epochs, indicating stronger generalization and convergence. The large variety of loss values, which range from roughly 0.2 to more than 1.2, demonstrates how differently the various EfficientNet topologies minimize errors both during training and on unseen validation data. It is difficult to determine the "best" performing model based only on this loss plot without additional quantitative analysis because of the complex interactions between these lines. At different stages of training, some models may achieve lower training loss but higher validation loss, or vice versa.

Table 2: Performance Metrics of EfficientNet B0-B7

Model	Acc	Pre (non-ASD)	Pre (ASD)	Recall (non-ASD)	Recall (ASD)	F1-Score (non-ASD)	F1-Score (ASD)
EfficientNetB0	0.95	0.97	0.94	0.93	0.97	0.95	0.95
EfficientNetB1	0.93	0.96	0.91	0.90	0.97	0.93	0.94
EfficientNetB2	0.98	1.00	0.97	0.97	1.00	0.98	0.98
EfficientNetB3	0.99	1.00	0.98	0.96	1.00	0.98	0.97
EfficientNetB4	0.88	1.00	0.83	0.77	1.00	0.87	0.91
EfficientNetB5	0.97	0.99	0.96	0.95	1.00	0.97	0.98
EfficientNetB6	0.97	0.96	0.98	0.99	0.96	0.97	0.97
EfficientNetB7	0.99	1.00	0.98	0.96	1.00	0.98	0.97

The EfficientNet models (B0–B7) are evaluated in Table 2, demonstrating their progressive enhancements in accuracy, precision, recall, and F1-scores for distinguishing between ASD and non-ASD. Deeper variants such as B2 (98 percent) and B3 (99 percent) demonstrate significant enhancements with well-balanced precision and recall in both classes, even though the baseline EfficientNetB0 achieves a solid accuracy of 95 percent. Nevertheless, EfficientNetB4 demonstrates a drop in performance with an accuracy of 88% due to reduced recall for non-ASD instances, indicating generalization limitations. In contrast, advanced versions like B5, B6, and B7 consistently achieve high accuracies of 97–99 percent, along with strong F1-scores that indicate reliable classification effectiveness. The findings suggest that deeper EfficientNet models, namely, B3, B5, B6, and perform superiorly in feature extraction and classification for ASD detection compared to their shallower versions.

7. Conclusion

The assessment of the EfficientNet family indicates that deeper models typically provide better predictive accuracy for Autism Spectrum Disorder (ASD) classification, where EfficientNetB3 and EfficientNetB7 reach the top accuracy of 0.99. Both models reliably achieve flawless precision for the non-ASD category (1.00) and high precision for ASD (0.98), along with robust recall rates, indicating their capability to balance sensitivity and specificity. EfficientNetB2 and B5 demonstrate strong performance, achieving accuracy rates of 0.98 and 0.97, respectively, positioning them as trustworthy options with marginally reduced complexity. Conversely, EfficientNetB4 achieves the lowest accuracy of 0.88, mainly because of decreased recall for the non-ASD class, indicating challenges in balanced detection. In general, the findings show that EfficientNet models, especially B3 and B7, provide better classification results, affirming their efficiency in extracting key features for ASD detection, while emphasizing the balance between model depth and generalization across various iterations. Future research will concentrate on combining explainable AI with multimodal data to improve diagnostic reliability. Variants of EfficientNet that are lightweight can be tailored for deployment on edge devices. Improved model generalization will result from a more diverse dataset expansion. In clinical settings, robustness will be guaranteed by cross-dataset testing and real-time validation.

References

[1] Vemuri, "Prediction of Autism Spectrum Disorder Using Efficient Net," in *Proc. IEEE ICACCS*, Mar. 2023.

[2] M. S. Alam, M. M. Rashid, and R. Roy, "Empirical Study of Autism Spectrum Disorder Diagnosis Using Facial Images by Improved Transfer Learning Approach," *Bioengineering*, vol. 9, no. 11, Nov. 2022, https://doi.org/10.3390/bioengineering9110710.

[3] T. S. Mian, "Efficient Net-based Transfer Learning Technique for Facial Autism Detection," *Scalable Computing: Practice and Experience*, vol. 24, no. 3, Sep. 2023, https://doi.org/10.12694/scpe.v24i3.2233.

[4] K. K. M. Rahman and M. M. Subashini, "Identification of Autism in Children Using Static Facial Features and Deep Neural Networks," *Brain Sciences*, vol. 12, no. 1, Jan. 2022, https://doi.org/10.3390/brainsci12010094.

[5] P. Reddy and A. J., "Diagnosis of Autism in Children Using Deep Learning Techniques by Analyzing Facial Features," *Eng. Proc.*, vol. 5, Jan. 2024, https://doi.org/10.3390/engproc2023059198.

[6] J. Anjum, N. A. Hia, and A. Waziha, "Deep Learning-Based Feature Extraction from Children's Facial Images for Autism Spectrum Disorder Detection," in *Proc. ACM*, May 2024, https://doi.org/10.1145/3660853.3660888.

[7] M. S. V. S. K. Narala, S. Vemuri, and C. Kattula, "A Comparative Study for Autism Spectrum Disorder Using EfficientNet, ResNet50," in *Proc. IEEE ICCINS*, Dec. 2023, https://doi.org/10.1109/ICCINS58907.2023.10450124.

[8] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *arXiv preprint arXiv:1905.11946*, May 2019.

[9] A. Bhise, S. Gawade, A. Pawar, A. Singh, and S. Shiravale, "Facial-Based Autism Spectrum Disorder Diagnosis Using Enhanced Deep Learning Models," in *Proc. IEEE INC*, Mar. 2024, https://doi.org/10.1109/InC460750.2024.10649230.

[10] C. Wang and J. Miao, "Deep Learning-Based Recognition of Autism Using Facial Datasets," in *Proc. SPIE*, https://doi.org/10.1117/12.3042402.

[11] S. L. K. Vinti, B. Panduri, V. Khedkar, M. B. Ketha, A. Lakshmanarao, and T. Srivalli, "Advanced Hybrid Transfer Learning Approaches for Autism Spectrum Disorder Detection Using Facial Features," in *Proc. IEEE ICDICI*, Nov. 2024.

[12] M. Sun, H. Dou, and J. Yan, "Efficient Transfer Learning via Joint Adaptation of Network Architecture and Weight," in *Deep Learning for Vision Systems*, Springer, 2020, pp. 381–398, https://doi.org/10.1007/978-3-030-58601-0_28.

[13] Y. B. Sree, O. Kumar, A. Sathvik, B. Sai, P. Rao, D. Sai, and H. Akshit, "Transfer Learning Based Autism Detection in Children," in *Proc. IEEE INOCON*, Mar. 2024, https://doi.org/10.1109/INOCON60754.2024.10511706.

[14] A. Singh, M. Laroia, A. K. S. Rawat, and K. R. Seeja, "Facial Feature Analysis for Autism Detection Using Deep Learning," in *AI Applications in Healthcare*, Springer, Jan. 2023, pp. 553–564, https://doi.org/10.1007/978-981-99-3315-0_41.

[15] B. Benaziz, A. Zitouni, and A. Amrouche, "Children Autism Detection Using Residual CNN and EfficientNetB1," in *Proc. IEEE ICEEAT*, Nov. 2023, https://doi.org/10.1109/ICEEAT60471.2023.10425848.

[16] R. Pilli, T. Goel, and R. Murugan, "Unveiling Alzheimer's Disease through Brain Age Estimation Using Multi-Kernel Regression Network and Magnetic Resonance Imaging," *Comput. Methods Programs Biomed.*, vol. 261, p. 108617, Apr. 2025, https://doi.org/10.1016/j.cmpb.2025.108617.

[17] K. Dinesh, S. Kamatchi, and K. Mangaiyarkarasi, "Accuracy improvement for personality prediction using logistic regression in comparison with random forest algorithm," in *AIP Conf. Proc.*, vol. 3161, no. 1, p. 020269, 2024, https://doi.org/10.1063/5.0229463.

[18] A. Priya and P. S. Bharathi, "SE-ResNeXt-50-CNN: A deep learning model for lung cancer classification," *Appl. Soft Comput.*, vol. 171, 2025. https://doi.org/10.1016/j.asoc.2025.112696.

[19] M. R and M. Arunraj, "Enhancing Healthcare Insurance Cost Prediction with a Metaheuristic Optimization Approach: Comparative Analysis of Machine Learning Models," in *Proc. 2025 3rd Int. Conf. Inventive Comput. Informatics (ICICI)*, Bangalore, India, 2025, pp. 1225–1232. https://doi.org/10.1109/ICICI65870.2025.11069726.