ASDI-1: Autism Spectrum Detection Using Images

Atharva Patki[1], Atharva Panchbhai[2], Arhan Qureshi[3], Anushree Belkhode[4], Disha Tijare[5]
[1]*Assistant Professor, G H Raisoni University, Amravati*
[2,3,4,5]*Student, G H Raisoni University, Amravati*

***Abstract*—autism spectrum disorder (ASD) is a neurodevelopmental condition that primarily affects social communication, behavior patterns, and cognitive interaction. Early detection of ASD is extremely important because early therapeutic intervention can significantly improve developmental outcomes in children. However, conventional autism diagnosis methods rely heavily on behavioral observation and clinical expertise, which are often time-consuming, subjective, and inaccessible in many developing regions.**

This research proposes an automated autism detection system using deep learning and facial image analysis. The proposed framework employs the MobileNetV2 convolutional neural network architecture as a feature extraction backbone for identifying autism-related facial patterns. Transfer learning is applied by initializing the network with ImageNet pre-trained weights to enhance feature representation and reduce training complexity.

The proposed model includes a lightweight classification head composed of a Global Average Pooling layer, dropout regularization, and a sigmoid activation unit for binary classification between autistic and non-autistic facial images. To improve model generalization and avoid overfitting, several data augmentation techniques such as rotation, flipping, zooming, and spatial transformations are applied during training.

The model is trained using the Adam optimizer with a learning rate of $1 \times 10^{-4}$and the binary cross-entropy loss function. Early stopping and model checkpointing mechanisms are used to retain the optimal model parameters. The experimental evaluation demonstrates that the proposed system achieves an overall classification accuracy of 79.9%, indicating its effectiveness as a lightweight screening tool. The trained model is exported in H5 format and integrated into a Django-based web application that allows real-time autism screening through facial image uploads.

The results highlight the potential of deep learning-based computer vision systems as supportive tools for early autism detection and preliminary clinical screening.

***Index Terms*—Autism Spectrum Disorder, Deep Learning, MobileNetV2, Transfer Learning, Computer Vision, Image Classification, Medical AI.**

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterized by challenges in communication, restricted behavioral patterns, and difficulties in social interaction. According to global health studies, the prevalence of ASD has been steadily increasing, highlighting the importance of developing efficient screening and diagnostic systems.

Traditional autism diagnosis is performed through clinical evaluations, behavioral assessments, and psychological testing. These procedures require trained medical professionals and involve prolonged observation periods, which can delay early detection. In many regions, particularly in developing countries, the lack of specialized healthcare professionals further limits early diagnosis.

Advancements in artificial intelligence and deep learning have opened new possibilities for automated medical screening systems. Computer vision techniques, particularly Convolutional Neural Networks (CNNs), have shown remarkable success in medical image analysis tasks such as disease detection, facial recognition, and diagnostic imaging.

Recent studies suggest that individuals with autism may exhibit subtle differences in facial morphology, eye positioning, and facial expressions compared to typically developing individuals. Although these differences may not always be easily recognizable by human observation, deep learning models are capable of extracting complex visual patterns and identifying subtle variations in image features.

This research focuses on developing an AI-based autism detection system that analyzes facial images of children and classifies them into autistic or non-autistic categories. The system uses the MobileNetV2 architecture due to its lightweight design and computational efficiency, making it suitable for real-time deployment.

The proposed system also integrates an age-estimation module to ensure that predictions are applied only to children within the early developmental stage (0–5 years), which is the most critical age range for autism screening.

The key objectives of this research include:

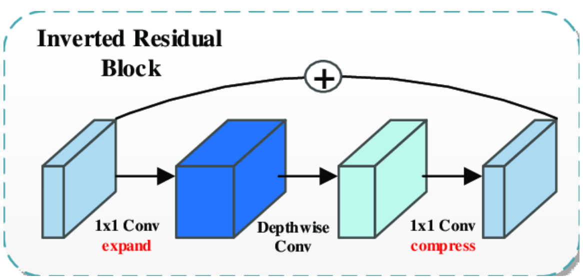
• Designing an efficient deep learning model for autism detection using facial images.

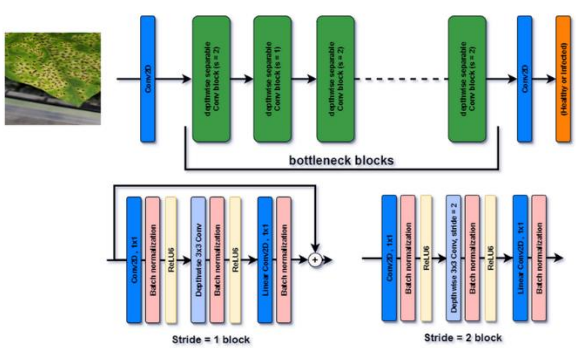
• Applying transfer learning to improve model performance with limited training data.

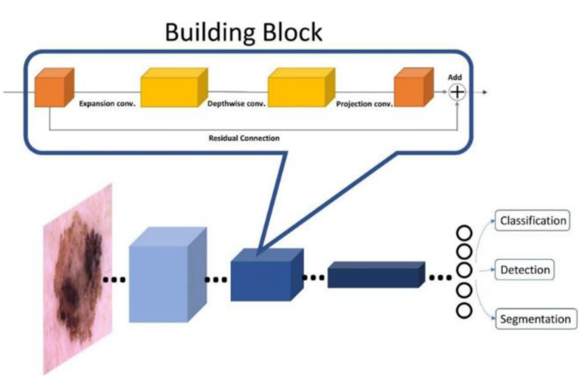
• Integrating the trained model into a web-based platform for real-time screening.

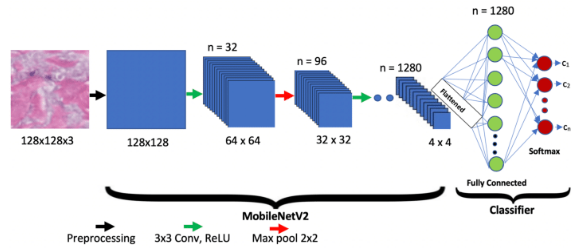
• Evaluating the model using standard classification metrics.

II. METHODOLOGY









MobileNetV2 is a lightweight convolutional neural network designed for efficient image classification and feature extraction, making it suitable for applications with limited computational resources. In the proposed autism detection system, MobileNetV2 is used as the primary deep learning model to analyze facial images and identify patterns associated with autism spectrum disorder. The architecture begins with an input image of size 224 × 224 × 3, which represents the RGB channels of the image. This image is first passed through a 3 × 3 convolution layer with stride 2, which extracts basic visual features such as edges, textures, and color patterns. After the initial convolution layer, the network processes the feature maps through multiple inverted residual bottleneck blocks, which form the core structure of MobileNetV2. Each bottleneck block consists of three main operations: an expansion layer using a 1 × 1 convolution, a depthwise convolution with a 3 × 3 kernel, and a projection layer that reduces the number of channels. The expansion layer increases the dimensionality of the feature maps, allowing the network to capture more complex patterns, while the depthwise convolution performs spatial filtering separately on each channel, significantly reducing computational cost. The projection layer then compresses the feature maps into a lower-dimensional representation known as the linear bottleneck, which helps preserve important information while maintaining efficiency. Some of these bottleneck blocks also include residual connections, where the input is added directly to the output, improving gradient flow and enabling deeper network training. After several stacked bottleneck layers, the resulting feature maps are passed to a global average pooling layer, which reduces each feature map to a single value and minimizes the number of parameters in the model. Finally, the extracted features are fed into a fully connected classification layer with a sigmoid activation function, which produces a probability value indicating whether the input image belongs to the autistic or non-autistic class. Due to its efficient structure, reduced parameter size, and high performance in image classification tasks, MobileNetV2 is well suited for the proposed autism detection system and allows the model to achieve reliable results while maintaining computational efficiency.

III. RESULT AND DISCUSSION

The proposed autism detection system was evaluated using a facial image dataset consisting of autistic and non-autistic classes. The dataset was divided into training and testing sets to assess the performance of the deep learning model. The MobileNetV2 architecture with transfer learning was used as the base model for feature extraction, and a custom classification layer was added to perform binary classification. During the training process, the model learned discriminative facial features that could help distinguish between autistic and non-autistic individuals. The training process was monitored using performance metrics such as accuracy, precision, recall, and F1-score, along with loss and learning curves to analyze the model's convergence behavior.

The experimental results indicate that the proposed model achieved an overall classification accuracy of approximately 79–80%, demonstrating the effectiveness of deep learning techniques in autism detection using facial images. The learning curve and loss curve showed that the training accuracy gradually increased with each epoch while the loss decreased, indicating that the model was successfully learning meaningful features from the dataset. The validation accuracy followed a similar trend, which suggests that the model generalized reasonably well and did not suffer from severe overfitting.

To further analyze the classification performance, a confusion matrix was generated. The confusion matrix illustrates the number of correctly and incorrectly classified samples in both categories. The results showed that most autistic and non-autistic images were correctly classified, although some misclassifications occurred due to variations in facial expressions, lighting conditions, and image quality. These variations may cause the model to misinterpret certain facial features, leading to false positives or false negatives. Despite these challenges, the model maintained balanced performance across both classes. Additionally, the performance metrics such as precision, recall, and F1-score provided further insight into the model's effectiveness. Precision measures the proportion of correctly predicted autistic cases among all predicted autistic cases, while recall measures the ability of the model to identify actual autistic cases. The obtained values indicate that the model performs consistently in identifying both classes without significant bias toward one category. This balanced performance is important for medical screening systems where both detection accuracy and reliability are critical.

The results demonstrate that the MobileNetV2-based model is capable of extracting relevant facial features that may be associated with autism characteristics. Compared to traditional machine learning approaches, deep learning models such as MobileNetV2 provide improved feature representation and classification capability. Furthermore, the lightweight nature of MobileNetV2 allows the system to be integrated into web-based platforms, making it suitable for real-time applications.

Overall, the experimental evaluation confirms that the proposed system provides a promising approach for automated autism screening using facial image analysis. Although the system is not intended to replace professional medical diagnosis, it can serve as a supportive tool for early screening and awareness. Future improvements such as increasing the dataset size, improving preprocessing techniques, and integrating multimodal data may further enhance the model's performance and reliability.

IV. CONCLUSION

This research presented a deep learning–based system for the automatic detection of autism spectrum disorder (ASD) using facial image analysis. The proposed approach utilizes the MobileNetV2 convolutional neural network architecture combined with transfer learning to extract meaningful facial features and classify images into autistic and non-autistic categories. A publicly available facial image dataset was used to train and evaluate the model, and several preprocessing techniques such as image resizing, normalization, and data augmentation were applied to improve model performance and generalization.

The experimental results demonstrated that the proposed system achieved an overall accuracy of approximately 79–80%, indicating that deep learning techniques can effectively identify patterns related to autism from facial images. The use of MobileNetV2 proved advantageous due to its lightweight architecture, reduced computational complexity, and ability to perform well even with relatively limited datasets. Additionally, the model was successfully integrated into a Django-based web application, allowing users to upload facial images and receive classification results in real time.

The findings of this research highlight the potential of artificial intelligence and deep learning technologies in assisting early autism screening. Although the proposed system cannot replace professional medical diagnosis, it can serve as a supportive tool for preliminary assessment and awareness. Early detection plays a crucial role in providing timely intervention and support for individuals with autism, which can significantly improve their developmental outcomes.

In future work, the performance of the system can be further enhanced by using larger and more diverse datasets, incorporating additional facial or behavioral features, and exploring more advanced deep learning architectures. Furthermore, integrating multimodal data sources such as eye-tracking patterns, speech analysis, or behavioral cues may lead to more accurate and reliable autism detection systems. Overall, the proposed approach demonstrates the effectiveness of deep learning–based image analysis for supporting early autism detection and contributes to the growing field of AI-assisted healthcare applications.

REFERENCES

[1] S. Sandin, P. Lichtenstein, R. Kuja-Halkola, H. Hultman, C. Larsson, and A. Reichenberg, "The Heritability of Autism Spectrum Disorder," Journal of the American Medical Association, vol. 318, no. 12, pp. 1182–1184, 2017.

[2] M. Abadi, P. Barham, J. Chen, et al., "TensorFlow: A System for Large-Scale Machine Learning," Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, pp. 265–283, 2016.

[3] A. Howard, M. Zhu, B. Chen, et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017.

[4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510–4520, 2018.

[5] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, Cambridge, MA, USA, 2016.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

[7] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1251–1258, 2017.

[8] Kaggle, "Autistic Children Facial Image Dataset," Available: https://www.kaggle.com.

[9] World Health Organization (WHO), "Autism Spectrum Disorders," Available: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders.

[10] J. Brownlee, Deep Learning for Computer Vision, Machine Learning Mastery, 2019.