

Facial image analysis for autism spectrum disorder detection in toddlers using deep learning and transfer learning

Anupam Das, Prasant Kumar Pattnaik, Anjan Bandyopadhyay

School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Bhubaneswar, India

Article Info

Article history:

Received Dec 26, 2024

Revised Jun 1, 2025

Accepted Jul 3, 2025

Keywords:

Autism spectrum disorder

Deep learning

Facial image analysis

Machine learning

Transfer learning

VGG-16

ABSTRACT

Autism spectrum disorder (ASD) is a neurological illness that manifests itself through restricted and repeated activity patterns, frivolous or recidivist interests or hobbies and consistent handicaps to social interactions and exchanges. Better results and early intervention are dependent upon the early identification of people with ASD. Doctors employ a variety of techniques to anticipate autism, including genetic testing, neuropsychological testing, hearing and vision screenings, and diagnostic interviews. In addition to requiring more time and money, the traditional diagnosis approach makes the parents of children with extensive developmental abnormalities feel too inadequate to disclose their condition. So, we need a tool that can detect autism early in less time and money. Machine learning methods can be used to fulfill this criterion. In this study, deep learning with transfer learning (VGG-16) is used to detect autism through facial images of children and achieved almost 97% accuracy. The suggested model significantly improves accuracy and saves time and money by using face features in photos of children to identify early autism tendencies in children.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Anupam Das

School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University

Bhubaneswar, Pin-751024, Odisha, India

Email: anukiit23@gmail.com

1. INTRODUCTION

Autism spectrum disorder (ASD) [1] act as a neurological developmental disorder that affects socialization as well as communication. Early diagnosis is crucial as ASD can impact social, academic, and professional aspects of life. Many children show signs within the first year, such as reduced eye contact, lack of interest in caregivers, or delayed response to names [2], [3]. Some may regress between 18–24 months, losing acquired skills. Symptoms vary in severity and impact on functioning, making assessment complex. Artificial intelligence (AI) and machine learning (ML) [4], [5] are revolutionizing ASD diagnosis and treatment, offering faster, more accurate, and scalable approaches by analyzing large datasets and identifying subtle patterns [6], [7]. Globally, ASD affects about 1 in 100 children, influenced by genetic and environmental factors [8]. Diagnosis relies on observing behavior and developmental milestones, with specialists able to provide reliable assessments by age two [9]. Early intervention significantly improves developmental outcomes [10]–[12] emphasizing the need for prompt treatment to maximize potential.

ASD which was first identified in 2013, is a developmental illness characterized by limited and repetitive behavioral patterns, interests, or hobbies in addition to persistent challenges with social engagement and communication [13]. Key symptoms of Autism can be seen in Figure 1. It has superseded the earlier nomenclature for disorders like Asperger's syndrome and autism disorder that were considered to

be on “the great continuum” of autism [14], [15]. Even though autism has probably been around for a while, Dr. Leo Kanner provided the first clinical description of the condition in 1943 [16]. Eleven children, eight boys and three girls, were diagnosed by Dr. Kanner, the creator of the nation's first pediatric psychiatric program, with what he called “autistic disturbances of affective contact” [17]. Over the Atlantic, at about the same time, a pediatrician from Austria named Hans Asperger was treating a similar set of kids. Later, a milder version of autism was referred to as “Asperger syndrome” in his honor.

Researchers have not determined the specific factors causing autism since they believe that multiple genetic elements alongside environmental factors play a combined role. The odds of developing autism increase in cases with either genetic abnormalities or a family medical background. Validating a diagnosis of autism becomes possible when children reach early childhood through systematic behavioral assessments supported by developmental history review. Behavioral diagnosis for Autism relies on multiple expert professionals including pediatricians along with psychologists and speech-language pathologists. Early treatment and therapies such as social skills training together with speech therapy and occupational therapy and behavioral interventions help enhance the life experience of the affected persons with autism spectrum disorder despite the condition being incurable. The Centers of Disease Control and Prevention (CDC) published new data on the frequency of autism in the population of children: 1 in 36 in the United States, 1 in 36 in Arizona [18], [19].

The autism and developmental disabilities monitoring network sent updated data to the CDC and prevention on March 23. According to the latest data, 1 in 36 American children aged 8 received a diagnosis of ASD in 2020 [20]. Compared to the previously stated prevalence of 1 in 44 in 2018, this number indicates an increase as shown in Figure 2.

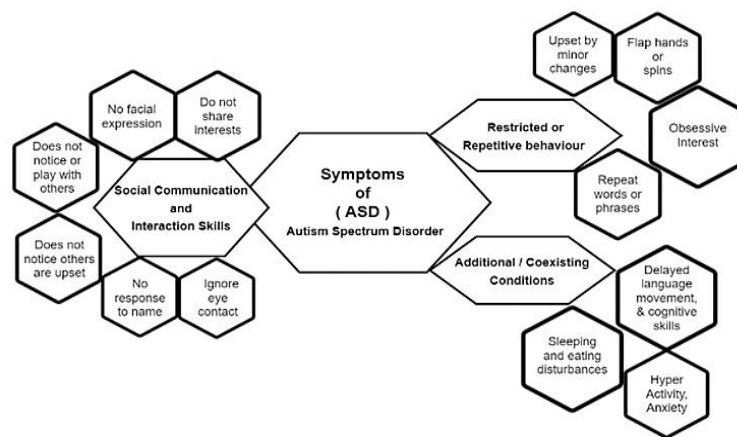


Figure 1. Key symptoms of Autism

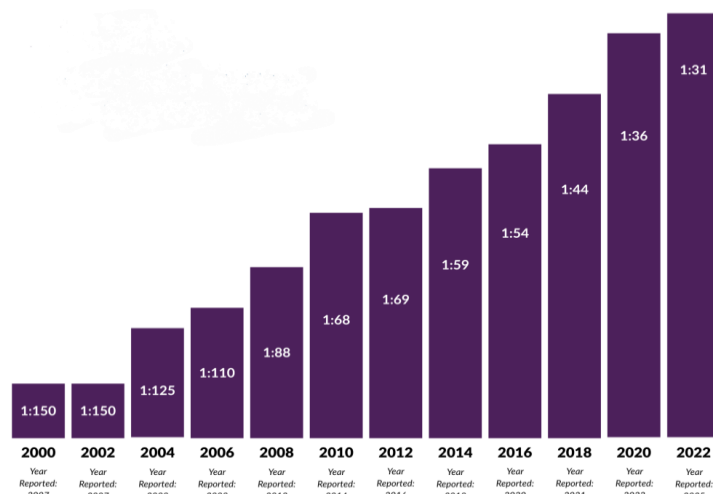


Figure 2. Prevalence rate of Autism

The current techniques for autism diagnosis require extensive costs yet produce subjective results thus causing early intervention to be delayed. Researchers require an automated and cost-effective detection system that identifies ASD during its initial stages. This work fills a crucial knowledge gap through VGG-16 deep learning analysis of facial images to develop an efficient, non-contact, and precise model for detecting ASD, which removes conventional diagnosis techniques' requirements.

This research presents an AI algorithm based on deep learning (CNNs) along with transfer learning (VGG-16) for analyzing pictures of toddler faces as an ASD detection tool. Using a pre-trained VGG-16 model allows the approach to effectively extract ASD-related facial features which leads to 99.50% classification accuracy. Using this method healthcare providers gain a faster and more affordable diagnostic technique that needs no surgery. An online ASD screening tool derived from this solution will help detect ASD early for intervention purposes without needing extensive clinical evaluations.

2. METHOD

Analyzing ASD is crucial, but diagnosing it can be challenging since it lacks a clinical benchmark, e.g. blood test, to determine the issue. To come up with a conclusion, professionals take into consideration of the developmental history of the teenager. Analyzing ASD is crucial because, in the absence of a diagnosis, it can cause a great deal of distress and confusion for the unidentified person about a variety of daily difficulties. This may result in disruptive behaviors and social disengagement.

In response to the need, the accessible data associated with autism and its analysis have been deconstructed. A model has been developed to manage questionnaires in a quick and easy way for diagnosing. In order to analyze autism through a child's facial picture, a dataset of facial images has been used that will help the model to prepare, test, and approve. Next, using CNNs and added transfer learning techniques, a model has been developed.

2.1. Data set

In this research, the data was taken off the publicly available Kaggle [21]. Two different types of datasets have been used in this analysis. Consolidated is the name of the workout set. Autistic and non-autistic are its two sub-indices. Details of the training, validation, and test data are considered as per Tables 1, 2, and 3 accordingly. Figure 3 illustrates the data pre-processing pipeline used in this study. DL along with the method of transfer learning is also applied to recognizing autism in the given study by using child facial photos [22]–[25].

Table 1. Details of the training data used for classification

Classes	No. of images	Type of images
Autistic	1470	jpeg
Non-Autistic	1470	jpeg

Table 2. Details of the validation data used for classification

Classes	No. of images	Type of images
Autistic	100	jpeg
Non-Autistic	100	jpeg

Table 3. Details of the test data used for classification

Classes	No. of images	Type of images
Autistic	100	jpeg
Non-Autistic	100	jpeg



Figure 3. Data pre-processing images

2.2. Convolutional neural network (CNN)

The foundation of DL, a crucial branch of ML, is neural networks. An input layer, an output layer, and one or more hidden layers are all present in these networks. Nodes are connected by thresholds and weights in each layer. Data is passed to the next layer when a node's output surpasses its threshold; otherwise, it stays dormant. Three primary layers are pooling, convolutional and fully connected. The central constituent of a CNN is where the main processing is done. In order to identify particular features within an input image, this layer applies a filter, occasionally called a kernel, which is a tiny matrix of weights that moves over the receptive field. The pooling layer comes after the convolutional layer in a CNN, is an essential component. Similar to the convolutional layer, the pooling layer also carries out operations that involve sweeping across the input image, but for a different reason.

Fully connected layer which classifies images using the properties collected from the prior layers, is crucial to the final stages of a CNN. Figure 4 depicts the architecture of the CNN used in this study. A neuron in a layer above indicates that all other neurons in the layer below it, are connected when it is considered to be fully interconnected.

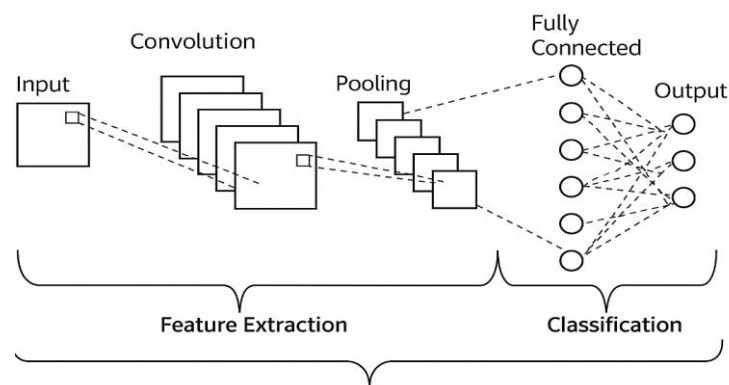


Figure 4. CNN architecture

2.3. Transfer learning

By using information from one task or dataset to enhance a model's performance on a separate but related task, transfer learning is a ML technique. To put it another way, transfer learning makes better use of knowledge acquired in one context to enhance generalization in another. As shown in Figure 5, applications for transfer learning are numerous and range from deep learning model training to data science regression problem-solving. That is especially attractive for the latter, considering the volume of data required to build deep neural networks.

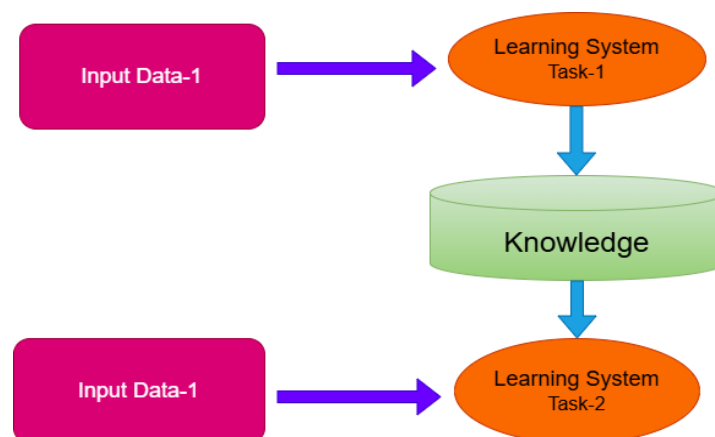


Figure 5. Transfer learning

2.3.1. Advantages

Shortened training less training time transfer learning can save a vast amount of training time by relying on the knowledge acquired in the past. Instead of starting from scratch, models fine-tune existing weights, saving computational resources and speeding up development. Improved performance with less data. One of the biggest advantages is improved performance, especially when dealing with small datasets. Pre-trained models can generalize better and avoid overfitting since they have already been trained on large datasets. Lower computational costs since the model does not need to learn from the ground up, transfer learning reduces the need for powerful hardware and long training periods, making it more cost-effective and accessible. Effective with limited labelled data transfer learning is particularly helpful in scenarios where labeled data is hard to come by. The knowledge from a related domain helps the model learn useful representations even when the target domain lacks sufficient labeled examples.

2.4. VGG16

VGG16 functions as a widely adopted deep learning model which successfully identifies 927 images within its 1,000 category dataset. Stakeholders choose VGG16 for many deep learning tasks because its defined design together with its simple programming and ability to perform transfer learning. The network unfolds into 16 composition blocks that include 13 convolutional operations along with 5 max-pooling operations and 3 fully connected operations. The system analyzes 224×224 RGB photos while using 3×3 convolutional kernel structures with stride set at 1 to obtain precise features from images. Two types of layers with parallel operations are applied to reduce dimensions through 2×2 max-pooling with a stride of 2. The filter count increases systematically from Conv-1 (64 filters) to Conv-2 (128 filters), Conv-3 (256 filters) and ends at Conv-4 and Conv-5 (512 filters each). The model applies 4096 neurons to each of its first two fully connected layers followed by the final 1000-class SoftMax activated layer. The alternation pattern between pooling and convolution layers leads to enhanced feature extraction abilities while increasing the classification precision. Transfer learning allows VGG16 to obtain domain-specific accuracy through brief additional training processes.

The combination of VGG16 model's excellent accuracy performance together with its adaptable pre-trained weights makes it ideal for use in medical imaging besides acting as a tool for object detection and facial recognition and autonomous systems applications. The benchmark status of the deep learning framework keeps its place as a standard model that provides a solid combination between depth and operational efficiency with high accuracy across different computer vision applications.

2.5. Evaluation of model

The performance assessment of classification models uses accuracy together with precision and recall expressed through (1), (2) and (3). Model accuracy measures the balance between correctly forecasted results against all predictions made on the testing data. Accuracy determines the performance metrics by counting the total number of forecasted results while accounting for correctly predicted cases.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Samples} \quad (1)$$

Precision stands for the quotient between actual positive matches for all of the positive forecasts. Model shows its capability of recognizing valid examples from a particular field.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (2)$$

Recall demonstrates the relationship between actual positive instances and true positive predictions to total class instances. The measure indicates whether the model properly identifies all relevant examples from an assigned class.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (3)$$

3. RESULTS AND DISCUSSION

Convolutional neural networks that also benefit from TL method: the model is trained on 1,470 facial photos of children who are autistic and 1,470 who are not; selected the features for determining explicitness, affectability, and accuracy of the predicted model. CNN works with pre-trained VGG16 version of ImageNet, the activator of sigmoid, Adam optimizer, and a 8-epsilon binary loss function, which are shown in Figure 6.

Epoch 1/8	92/92	793s	9s/step	accuracy: 0.6194	loss: 1.4575	val_accuracy: 0.8050	val_loss: 0.4319
Epoch 2/8	92/92	866s	9s/step	accuracy: 0.8309	loss: 0.3898	val_accuracy: 0.8600	val_loss: 0.3389
Epoch 3/8	92/92	819s	8s/step	accuracy: 0.8718	loss: 0.3081	val_accuracy: 0.9000	val_loss: 0.2338
Epoch 4/8	92/92	808s	9s/step	accuracy: 0.9019	loss: 0.2471	val_accuracy: 0.9450	val_loss: 0.1813
Epoch 5/8	92/92	855s	9s/step	accuracy: 0.9342	loss: 0.1892	val_accuracy: 0.9350	val_loss: 0.1853
Epoch 6/8	92/92	798s	9s/step	accuracy: 0.9589	loss: 0.1327	val_accuracy: 0.9850	val_loss: 0.0778
Epoch 7/8	92/92	763s	8s/step	accuracy: 0.9624	loss: 0.1178	val_accuracy: 0.9850	val_loss: 0.0722
Epoch 8/8	92/92	769s	8s/step	accuracy: 0.9667	loss: 0.1011	val_accuracy: 0.9950	val_loss: 0.0391

Figure 6. Epochs

3.1. Rectified linear unit function (ReLU)

This is expressed as (4):

$$ReLU(x) = \frac{x + |x|}{2} \quad (4)$$

where: $ReLU(x)$ is the output of the ReLU function; x is the input variable; $|x|$ represents the absolute value of x , which ensures that the output is either x (if x is positive) or 0 (if x is negative). Because of its ease of use and capacity to successfully add non-linearity to the model, ReLU is frequently used as an activation function in neural networks.

3.2. Sigmoid function

Because it compresses input values into this range, this function is helpful for binary classification tasks, producing values between 0 and 1 in (5):

$$A = \frac{1}{1+e^{-x}} \quad (5)$$

The value range of A is from 0 to 1. where: A is the output of the function; e is the base of the natural logarithm; x is the input variable.

The value range of A is from 0 to 1, referring to that A goes to 0 as x goes to negative infinity and A goes to 1 as x goes to positive infinity. The following lists of the prediction model's outcomes, along with its performance metrics and expectations for learning and adapting are displayed. Accuracy, losses, and confusion matrix are shown in Figures 7, 8 and 9 respectively: accuracy: 0.9667, loss: 0.1011, val-accuracy: 0.9950, val-loss: 0.0391. Table 4 displays the classification performance metrics.

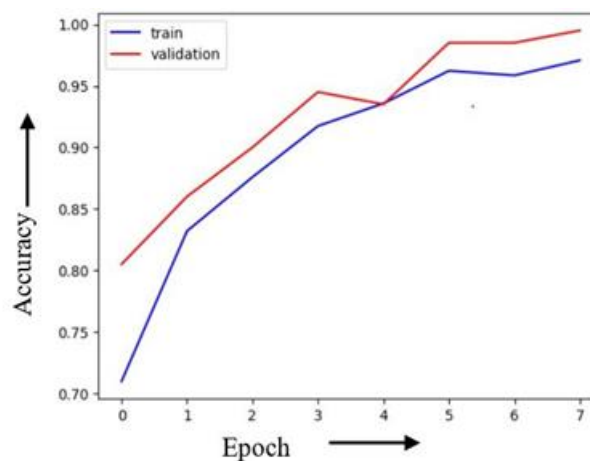


Figure 7. Accuracy

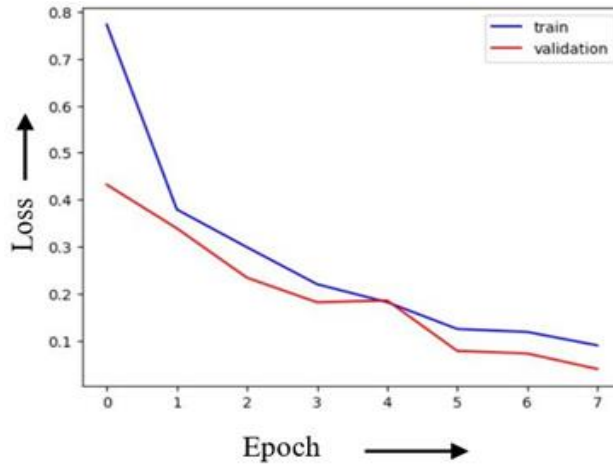


Figure 8. Losses

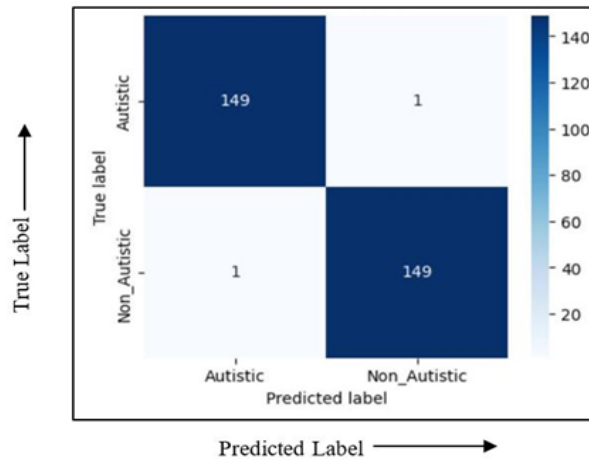


Figure 9. Confusion matrix

Table 4. Classification performance metrics

Classes	Precision	Recall	F1-Score	Support
0	0.98	0.99	0.99	150
1	0.99	0.98	0.98	150
Accuracy			0.985	300
Macro Avg	0.985	0.985	0.985	300
Weighted Avg	0.985	0.985	0.985	300

The proposed model for detecting ASD using facial images with deep learning and transfer learning techniques (VGG-16) demonstrated high efficiency and accuracy. The model accuracy in the training and validation was 96.67% and 99.50, respectively with corresponding loss values of 0.1011 and 0.0391, respectively. These results highlight the model’s strong predictive capability and generalization across datasets. The low loss values suggest minimal error in prediction, and the high validation accuracy implies resilience to overfitting. The confusion matrix as shown in Figure 9, shows minimal misclassifications, indicating the model’s ability to discern subtle facial feature variations indicative of autism.

Leveraging transfer learning through the pre-trained VGG-16 model reduced computational cost and improved accuracy. The model effectively handled binary classification tasks using the binary cross-entropy loss function and sigmoid activation, yielding high sensitivity and specificity. Compared to prior research as shown in Table 5, the model outperformed other approaches regarding accuracy, sensitivity, in addition to efficiency of computation, demonstrating its applicability for real-world scenarios.

This approach reduces the time and cost of traditional diagnostic methods, offering a non-invasive, deployable tool for early detection of ASD. However, challenges related to dataset diversity remain, and expanding the dataset and incorporating multimodal data could enhance the model's generalizability. Further research into model interpretability could increase trust among clinicians and families, solidifying its role as a valuable tool for early diagnosis.

Table 5. Comparison with existing work: model training and validation accuracy

Model	Training accuracy	Validation accuracy
Jahanara and Padmanabhan [23]	0.9610	0.8467
Proposed model with VGG-16	0.9667	0.9950

4. CONCLUSION

Early detection of autism is crucial for several reasons, as it can lead to significant positive outcomes for individuals on the autism spectrum. If a machine learning model can do that in less time in an efficient manner, then this will be a great achievement for the medical field. In this study, a ML classification model with a facial image dataset through CNN-style deep learning and transfer learning (VGG-16) to detect autism has been proposed. In the future, using this model, an online application can be developed where parents can upload images of their children and find out the possibility of autism at an early stage.

ACKNOWLEDGMENTS

Our sincere gratitude goes out to the MeitY, Government of India, for their steadfast assistance through the scholarship offered by the Visvesvaraya Ph.D. Scheme.

REFERENCES




- [1] D. Singh, M. Rakhra, and S. Aggarwal, "Autism spectrum disorder detection using the Deep learning approaches," in *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, Oct. 2022, pp. 761–766, doi: 10.1109/ICTACS56270.2022.9988442.
- [2] F. Zhu, S. Wang, W. Liu, H. Zhu, M. Li, and X. Zou, "A multimodal machine learning system in early screening for toddlers with autism spectrum disorders based on the response to name," *Frontiers in Psychiatry*, vol. 26, no. 14, p. 1039293, Jan. 2023, doi: 10.3389/fpsy.2023.1039293.
- [3] S. M. Hasan, M. P. Uddin, M. Al Mamun, M. I. Sharif, A. Ulhaq, and G. Krishnamoorthy, "A machine learning framework for early-stage detection of autism spectrum disorders," *Industrial Engineering Journal*, vol. 54, no. 02, pp. 142–156, 2025, doi: 10.36893/iej.2025.v52i2.015.
- [4] F. Thabtah and D. Peebles, "A new machine learning model based on induction of rules for autism detection," *Health Informatics Journal*, vol. 26, no. 1, pp. 264–286, Mar. 2020, doi: 10.1177/1460458218824711.
- [5] O. P. Jena, B. Bhushan, and U. Kose, "Machine learning and deep learning in medical data analytics and healthcare applications," *Machine Learning and Deep Learning in Medical Data Analytics and Healthcare Applications*, pp. 1–279, 2022, doi: 10.1201/9781003226147.
- [6] K. K. Hyde *et al.*, "Applications of supervised machine learning in autism spectrum disorder research: A review," *Review Journal of Autism and Developmental Disorders*, vol. 6, no. 2, pp. 128–146, Jun. 2019, doi: 10.1007/s40489-019-00158-x.
- [7] B. Twala and E. Molloy, "On effectively predicting autism spectrum disorder therapy using an ensemble of classifiers," *Scientific Reports*, vol. 13, no. 1, p. 19957, 2023, doi: 10.1038/s41598-023-46379-3.
- [8] J. Zeidan *et al.*, "Global prevalence of autism: A systematic review update," *Autism Research*, vol. 15, no. 5, pp. 778–790, May 2022, doi: 10.1002/aur.2696.
- [9] C. Lord, S. Risi, P. S. DiLavore, C. Shulman, A. Thurm, and A. Pickles, "Autism from 2 to 9 years of age," *Archives of General Psychiatry*, vol. 63, no. 6, p. 694, Jun. 2006, doi: 10.1001/archpsyc.63.6.694.
- [10] J. S. Handleman and S. Harris, *Preschool education programs for children with autism*, TX: Pro-Ed. 2001.
- [11] National Research Council; Division of Behavioral and Social Sciences and Education; Committee on Educational Interventions for Children with Autism, *Educating children with autism*. Washington, D.C.: National Academies Press, 2001.
- [12] A. Das, P. Pattanaik, S. Mukherjee, S. Turjya, and A. Bandopadhyay, "Early autism spectrum disorder screening in toddlers: A comprehensive stacked machine learning approach," *International Journal of Computing and Digital Systems*, vol. 17, no. 1, pp. 1–12, Jan. 2025, doi: 10.12785/ijcds/1571010950.
- [13] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders*, 5 ed. Arlington, VA.: American Psychiatric Publishing, 2013.
- [14] T. Grandin, *Thinking in pictures : My life with autism*. 2006.
- [15] B. Khan, S. M. Bhatti, and A. Akram, "Autism spectrum disorder detection in children via deep learning models based on facial images," *Bulletin of Business and Economics (BBE)*, vol. 13, no. 1, pp. 557–572, 2024, doi: 10.61506/01.00241.
- [16] F. R. Volkmar and A. Klin, *Issues in the classification of autism and related conditions*. 2005.
- [17] Kanner L, "Autistic disturbances of affective contact," *Acta Paedopsychiatrica*, vol. 35, no. 4, pp. 100–136, 1968.
- [18] S. Zauderer, "69+ autism statistics: how many people have autism?" Accessed: May 02, 2022. [Online]. Available: <https://www.crossrivertherapy.com/autism-statistics>.
- [19] Y. Loftus, "Autism statistics you need to know in 2022," *Autism Parenting Magazine*. p. 1, 2024, Accessed: Dec. 29, 2023. [Online]. Available: <https://www.autismparentingmagazine.com/autism-statistics/>.
- [20] SARRC, "CDC releases new autism prevalence: 1 in 31 children in the U.S. Identified with autism," 2025.

<https://autismcenter.org/autismprevalence/> (accessed May 20, 1BC).




- [21] C. Senol, "Autism_image_data." Kaggle <https://www.kaggle.com/datasets/cihan063/autism-image-data>.
- [22] F. W. Alsaade and M. S. Alzahrani, "Classification and detection of autism spectrum disorder based on deep learning algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–10, Feb. 2022, doi: 10.1155/2022/8709145.
- [23] S. Jahanara and S. Padmanabhan, "Detecting autism from facial image," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 7, no. 2, pp. 219–225, 2021, doi: 10.13140/RG.2.2.35268.35202.
- [24] A. Sukumaran, "A study on autism detection from facial images," *International Journal for Research in Applied Science and Engineering Technology*, vol. 11, no. 11, pp. 2344–2348, Nov. 2023, doi: 10.22214/ijraset.2023.56908.
- [25] A. Das, P. K. Pattnaik, and A. Bandyopadhyay, "Early intervention in autism spectrum disorder: Hybridizing convolutional neural networks with ensemble methods for enhanced detection using facial imagery," in *2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC)*, Feb. 2025, pp. 781–786, doi: 10.1109/ISACC65211.2025.10969308.

BIOGRAPHIES OF AUTHORS






Anupam Das    is engaged as a full-time research scholar as part of the Visvesvaraya Ph.D. scheme, supported by MeitY, Government of India. Holding an M.Tech. degree from West Bengal University of Technology, he boasts an extensive publication record in both national and international journals and conferences. With over a decade of experience in teaching, his expertise spans various domains, including AI, machine learning, social network analysis, graph databases, and optimization techniques. He can be reached at anukit23@gmail.com.



Prasant Kumar Pattnaik    is currently serving as a professor at KIIT University, Bhubaneswar, India. He holds a Ph.D. in computer science and is a Senior Member of the IEEE. He has also contributed to the fields of cloud computing, mobile computing, brain-computer interface, and privacy protection over the past fifteen years. In addition to becoming a fellow of the Institution of Electronics and Telecommunication Engineers (IETE), Dr. Pattnaik has a number of research publications in prestigious international journals and conferences. He can be reached at patnaikprasant@gmail.com.



Anjan Bandyopadhyay    assistant professor at Kalinga Institute of Industrial Technology, completed his Ph.D. at NIT, Durgapur, under the MHRD's Visvesvaraya Ph.D. Fellowship. His research focuses on algorithmic game theory, particularly mechanism design, with interests spanning cloud, fog computing, IoT, healthcare and metaverse. He holds an M.Tech. in Information Security from NIT, Durgapur. With 42 research papers published in various domains. He has garnered several best paper awards, including at 3PGCIC. Dr. Bandyopadhyay currently serves as the nodal officer of the Visvesvaraya Ph.D. Scheme under the MeitY, Government of India. He can be reached at anjan.bandyopadhyayfcs@kiit.ac.in.