electronics

MDPI

Article

RETRACTED: Autism Spectrum Disorder Diagnosis Based on Attentional Feature Fusion Using NasNetMobile and DeiT Networks

Zainab A. Altomi [1,†], Yasmin M. Alsakar [2,†], Mostafa M. El-Gayar [2,3,*], Mohammed Elmogy [2,*,‡] and Yasser M. Fouda [1,‡]

[1] Computer Science Division, Mathematics Department, Faculty of Science, Mansoura University, Mansoura 35516, Egypt; zainabaltomi@std.mans.edu.eg (Z.A.A.); ymafouda@mans.edu.eg (Y.M.F.)
[2] Information Technology Department, Faculty of Computers and Information, Mansoura University, Mansoura 35516, Egypt; yasminmahmoud@mans.edu.eg
[3] Department of Computer Science, Arab East Colleges, Riyadh 11583, Saudi Arabia
* Correspondence: mostafa_elgayar@mans.edu.eg (M.M.E.-G.); melmogy@mans.edu.eg (M.E.)
† These authors contributed equally to this work.
‡ These authors also contributed equally to this work.

Abstract: Autism spectrum disorder (ASD) is a neurodevelopmental condition that affects social interactions, communication, and behavior. Prompt and precise diagnosis is essential for prompt support and intervention. In this study, a deep learning-based framework for diagnosing ASD using facial images has been proposed. The methodology begins with logarithmic transformation for image pre-processing, enhancing contrast and making subtle facial features more distinguishable. Next, feature extraction is performed using NasNetMobile and DeiT networks, where NasNetMobile captures high-level abstract patterns, and the DeiT network focuses on fine-grained facial characteristics relevant to ASD identification. The extracted features are then fused using attentional feature fusion, which adaptively assigns importance to the most discriminative features, ensuring an optimal representation. Finally, classification is conducted using bagging with a support vector machine (SVM) classifier employing a polynomial kernel, enhancing generalization and robustness. Experimental results validate the effectiveness of the proposed approach, achieving 95.77% recall, 95.67% precision, 95.66% F1-score, and 95.67% accuracy, demonstrating its strong potential for assisting in ASD diagnosis through facial image analysis.

Keywords: autism spectrum disorder; NasNetMobile; DeiT; attention feature fusion

check for updates

Academic Editors: Pawel Strumillo and Hyunjin Park

Received: 24 March 2025
Revised: 18 April 2025
Accepted: 25 April 2025
Published: 29 April 2025
Retracted: 14 October 2026

Citation: Altomi, Z.A.; Alsakar, Y.M.; El-Gayar, M.M.; Elmogy, M.; Fouda, Y.M. RETRACTED: Autism Spectrum Disorder Diagnosis Based on Attentional Feature Fusion Using NasNetMobile and DeiT Networks. *Electronics* **2025**, *14*, 1822. https://doi.org/10.3390/electronics14091822

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by challenges in social interactions, communication, and the presence of repetitive or restrictive behaviors. The prevalence of ASD has been on the rise, with recent estimates suggesting that approximately 1 in 36 children in the United States (Centers for Disease Control and Prevention, 2023) and around 1 in 100 children globally (World Health Organization) are diagnosed with the condition [1].

Research indicates that individuals with ASD exhibit disrupted neurodevelopmental pathways, deviating from typical brain development patterns. The complexity of ASD arises from the interplay of genetic factors, environmental influences, epigenetic mechanisms, cognitive processes, and behavioral components, contributing to a wide range of symptoms and comorbid conditions [2]. With a strong genetic foundation, ASD has

This paper has been retracted. A retraction notice was published on 14 October 2026 in *Electronics* **2026**, *15*, 4715. https://doi.org/10.3390/electronics15204715

RETRACTED

been increasingly studied using advanced genomic technologies such as microarrays and next-generation sequencing (NGS). These help identify genetic variations and enhance understanding of its genetic framework [3].

ASD was first identified by Kanner in 1943 when he documented 11 cases, primarily in male children, exhibiting severe social and language impairments. The global prevalence of ASD is approximately 1%, with a male-to-female ratio of 4:1. Around 50% of individuals with ASD also have intellectual disabilities (IDs) and often experience comorbid neurodevelopmental and psychiatric conditions, such as depression, anxiety, sleep disturbances, and gastrointestinal problems. Moreover, over 35% of individuals with ASD are affected by epilepsy, with EEG abnormalities frequently observed even in those without seizures [4].

An increasing number of observational studies have emphasized the link between specific endogenous environmental factors, such as parental age, and a heightened ASD risk [5–7]. Additionally, research indicates that even slight increases in maternal prenatal stress are associated with a greater risk of developing ASD and ADHD [8,9]. One proposed biological factor contributing to the onset and progression of ASD involves immune system abnormalities that can result in atypical neuroimmune responses [10].

Children with autism may show added comorbid conditions alongside the core symptoms, such as social difficulties, language impairments, and repetitive behaviors. Finding these medical conditions is crucial, as they can often trigger or worsen the abnormal behaviors seen in children with autism. Treating these conditions can lead to a resolution of the associated behaviors. Additionally, when children with autism are unwell, their performance may decline, and they may struggle to keep or acquire skills due to the impact of these medical issues [11,12].

Many behaviors and symptoms typically associated with autism may be indicative of other underlying medical conditions. For example, headbanging could result from headaches or pain due to frustration, especially when a child is unable to express these feelings. Frequent fidgeting may be a sign of discomfort related to constipation. Aggressive or self-injurious behavior could stem from undiagnosed pain that the child cannot communicate. Pica, the tendency to eat non-food items, might show nutrient deficiencies, especially iron, which is common in children with autism. Similarly, food refusal may not only be linked to the selective eating habits often seen in autism. However, it could also be due to food allergies, intolerances, or even dental issues [13].

Research indicates that anxiety and sensory over-responsivity (SOR) share common neurobiological mechanisms. Studies examining the neurobiological basis of SOR in ASD have shown that SOR is associated with heightened neural responses to unpleasant sensory stimuli, particularly in brain regions involved in sensory processing. Autistic individuals with high SOR levels exhibit reduced amygdala habituation and weaker top-down regulation from the prefrontal cortex over the amygdala during sensory processing compared to those with lower SOR levels [14].

Anxiety is a common co-occurring condition in individuals with ASD, affecting around 40% of children compared to 10% in the general population. Its diagnosis is often complicated by atypical symptom presentations, such as fear of change, unusual phobias, or sensory-related anxieties, which do not always align with standard diagnostic criteria like those in the DSM. For children with ASD, clinical anxiety can exacerbate challenges in learning, relationships, and emotional well-being, increasing risks of self-injury, depression, and disruptive behavior, ultimately leading to significant life disruptions [15,16]. Anger outbursts (AOs) are often linked to more severe symptoms, greater impairment, and poorer treatment outcomes in children with anxiety. However, there is limited research examining AO in youth with both ASD and anxiety disorders [17].

RETRACTED

Autism symptoms vary, but some are notably common. Over 90% of children with ASD have at least one co-occurring condition, such as GI disorders (up to 70%), movement disorders (79%), sleep issues (50–80%), and intellectual disabilities (45%). GI problems are particularly prevalent, affecting 9–91% of individuals, and are linked to cognitive and behavioral challenges. Though their exact causes remain unclear, factors like the gut–brain axis, genetics, microbiota, and immune responses may play a role. A meta-analysis found that children with ASD are 2–4 times more likely to experience GI issues, including constipation, diarrhea, and stomach discomfort [18].

Achieving complete recovery from ASD is challenging, as Piven et al. [19] determined that it is a lifelong condition with evolving characteristics. Their study of 38 individuals with ASD found that all but five continued to meet DSM-IV criteria in adulthood, while the remaining five still exhibited persistent autistic traits. However, most participants showed progress from childhood to adolescence and adulthood, with 82% improving in communication, 82% in social interaction, and 55% in reducing ritualistic and repetitive behaviors. Overall, while not universal, improvement is a common trend among individuals with autism.

A study conducted in Japan by Seltzer [20] found that many autism symptoms tend to improve over time; however, adults with autism continue to experience challenges in various aspects of daily life. Similarly, a British study by Beadle-Brown et al. [21] observed significant progress in self-care skills, communication, and educational achievements over 11 years.

The current process for detecting ASD has several limitations. Clinicians need extensive training and considerable time to apply diagnostic tools effectively. However, artificial intelligence (AI) advancements have accelerated ASD diagnosis, enhanced clinicians' capabilities, and improved access to early intervention programs. These AI-driven technologies have notably increased during the COVID-19 pandemic [22]. The emergence of machine learning (ML) and deep learning (DL) techniques has revolutionized medical diagnostics by enabling automated and high-precision decision-making. These technologies have shown promising results in analyzing complex data such as medical images, leading to faster and more accurate diagnoses. Data augmentation techniques have been widely adopted to synthetically expand training datasets while preserving key features to enhance model robustness and generalization [23]. In addition, contrastive learning frameworks, such as those utilizing stochastic pseudo-neighborhoods, have shown great potential in unsupervised representation learning, allowing models to better distinguish subtle patterns in medical images without relying on large labeled datasets [24]. Recent advancements, including those presented by Wang et al. in their multi-scale three-path network (MSTP-Net) and Zhao et al. in their review of cancer data fusion methods, have further underscored the importance of multi-scale and data fusion techniques in improving model performance in complex domains like medical imaging [25,26]. To ensure the robustness and generalizability of the proposed model, replication techniques were applied across multiple experimental runs using different data splits. This approach minimizes the risk of biased evaluation and confirms the model's stability across varying subsets of the dataset.

DL, in particular, has emerged as a pivotal technique in diagnosing ASD, owing to its ability to automatically learn hierarchical feature representations from raw data [27,28]. In the context of ASD, deep learning models, especially convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated remarkable success in identifying patterns in facial images, behavioral data, and other diagnostic indicators. Recent studies have highlighted how deep learning approaches, such as transfer learning and ensemble models, have significantly improved the accuracy and efficiency of ASD

RETRACTED

diagnosis. Furthermore, the ability of deep learning models to process large volumes of complex, high-dimensional data without manual feature extraction has made them invaluable in clinical settings.

Table in Abbreviations presents the abbreviations utilized in this manuscript. This research proposes an innovative deep learning-based framework for ASD diagnosis through facial images, incorporating various advanced techniques to improve diagnostic accuracy. The main contributions of this study include the following:

- High-Level Representation via NasNetMobile: NasNetMobile is employed to extract high-level abstract patterns from the input images, utilizing its deep architecture to generate robust and discriminative features.
- Fine-Grained Feature Extraction using DeiT Network: The DeiT network focuses on capturing fine-grained facial characteristics, enriching the overall representation with detailed local information essential for precise analysis.
- Attentional Feature Fusion: The extracted features are fused using an adaptive attention-based mechanism that assigns importance to the most discriminative features, leading to improved feature representation and robustness.
- Robust Classification Strategy: We utilize a bagging-based SVM classifier with a polynomial kernel, enhancing generalization capabilities and mitigating overfitting issues.

The remainder of this paper is organized as follows: Section 2 reviews related work on autism spectrum disorder diagnosis. Section 3 describes the proposed classification method. Section 4 presents the experimental results and comparisons, followed by a detailed discussion of the findings in Section 5. Lastly, Section 6 concludes the paper and highlights potential future research directions.

2. Related Work

Facial expressions are crucial in diagnosing ASD, offering valuable insights into emotional and social processing patterns. Individuals with ASD often struggle with recognizing or expressing emotions, which affects their ability to engage in social interactions. AI, particularly through machine learning (ML) and deep learning (DL) technologies, is increasingly used to analyze facial expressions using advanced recognition algorithms. These AI systems can detect subtle emotional cues that may be overlooked by human evaluators, thereby enhancing diagnostic accuracy. By using AI, clinicians can ease earlier detection and develop more personalized, effective intervention strategies for individuals with ASD. Table 1 compares previous research according to paper, methodology, strengths, limitations, and accuracy (%).

Several studies have investigated this approach. For example, Akter et al. [29] proposed a transfer learning-based framework for autism face recognition. This framework utilizes an enhanced MobileNet-V1 model, which surpasses other advanced ML and DL models in distinguishing between control and autistic children from various sources. The model achieved 83% accuracy on the validation set and 91% on the test set. Additionally, the k-means clustering algorithm was employed to classify autism faces into different subtypes based on varying k values. The enhanced MobileNet-V1 model demonstrated an impressive accuracy of 92.10% in predicting binary subtypes (k = 2). This system holds great potential for aiding early autism detection, offering a valuable tool for physicians and healthcare professionals.

Li et al. [30] proposed a method that improves the area under the curve (AUC) while utilizing smaller image sizes compared to previous studies. Their approach integrates two-phase transfer learning with multi-classifier fusion to enhance system performance. They apply two-phase transfer learning to MobileNetV2 and MobileNetV3-Large models optimized for mobile devices to improve their effectiveness. Subsequently, a multi-classifier

RETRACTED

fusion strategy combines these models, further enhancing accuracy. Additionally, they introduce a technique for integrating classifier outputs to achieve more precise final predictions. The integrated classifier attained 90.5% accuracy and a 96.32% AUC, marking a 3.51% improvement over the 92.81% AUC reported in previous studies.

Melinda et al. [31] introduced an innovative approach to evaluating the performance of ResNet-50 combined with DeepLabV3+ segmentation for classifying facial images of children with and without ASD. The study aims to enhance classification accuracy by minimizing noise and eliminating irrelevant features. Initially, ResNet-50 alone achieved an accuracy of 83.7%. However, integrating DeepLabV3+ segmentation significantly improved accuracy to 85.9%.

Ahmed et al. [32] suggested that CNNs hold great potential for diagnosing ASD. In their study, various pre-trained CNN models, such as AlexNet, ResNet34, VGG16, ResNet50, MobileNetV2, and VGG19, were utilized to diagnose ASD, and their performances were compared. Transfer learning was applied to each model to enhance the results. Among all the models, the ResNet50 model demonstrated the highest accuracy, achieving 92%, surpassing the performance of the other deep learning models.

Fahaad et al. [33] proposed a deep learning-based approach leveraging vision transformer (ViT) models to classify facial images for early ASD detection in children. Their method autonomously segments facial images into patches and processes them through transformer blocks, effectively capturing fine-grained and holistic features. Experimental results demonstrate that the ViT model achieves a validation accuracy of 77%, surpassing conventional models like VGG-16. This non-invasive technique shows significant potential as a reliable tool for early ASD diagnosis, facilitating timely interventions and enhancing clinical outcomes.

Reddy et al. [34] proposed an approach for diagnosing ASD in children using facial image analysis. They utilized three pre-trained CNN models—VGG16, VGG19, and EfficientNetB0—on a Kaggle dataset of 3014 images. Among these, EfficientNetB0 achieved the highest accuracy of 88.33%, surpassing VGG16 (84.67%) and VGG19 (87.66%). This research improves early ASD detection and better support for affected children. Mahmoud et al. [35] introduced a sequencer-based patch-wise Local Feature Extractor combined with a Global Feature Extractor to enhance ASD classification. The extracted features from both modules are integrated to form a comprehensive representation for final classification. Experimental evaluations on a publicly Autism Facial Image Dataset proved and achieved an accuracy of 94.7%, precision of 94.0%, recall of 95.3%, and an F1-score of 94.6%.

Mujeeb et al. [36] investigated static facial features from photographs of autistic children as potential biomarkers to distinguish them from typically developing children. Their study applied five deep CNN models such as EfficientNetB0, MobileNet, Xception, EfficientNetB1, and EfficientNetB2—as feature extractors, with a DNN classifier employed for binary autism classification. The models were trained on a public dataset containing, among the evaluated models, Xception which achieved the highest performance, with 90% for AUC, 88.46% for sensitivity, and 88% for negative predictive value (NPV).

Alam et al. [37] conducted a pioneering study on two facial image datasets—Kaggle and YTUIA—leveraging federated learning to address domain variations effectively. Their approach ensures the confidentiality of sensitive medical data while facilitating robust feature learning, leading to enhanced evaluation performance across various datasets. Using Xception as the backbone of their federated learning framework, the study achieved an impressive accuracy of nearly 90% across all test sets. Notably, this represents a significant improvement of over 30% in classification performance for test sets from different domains. Hossain et al. [38] introduced a non-invasive and cost-effective method for ASD identification using facial images. Their study systematically evaluated twelve deep learning

RETRACTED

models, such as MobileNetV2, ResNet-50, MobileNetV3, ResNet-101, AlexNet, ResNet-152, DenseNet201, InceptionV1, EfficientNetB0, SqueezeNet, DenseNet121, and VGG16. Among these, DenseNet121 achieved the best performance, with 90.33%, 92%, 92%, and 90%, for accuracy, precision, recall, and F1-score, respectively.

Despite previous studies developing methods for detecting ASD based on children's faces, several limitations hinder their reliability and effectiveness. One major issue is the use of limited datasets, which may lack diversity in terms of ethnicity, age, and severity of ASD traits, leading to biased models and reduced generalizability. Additionally, low-quality images—affected by lighting, resolution, and pose variations—can compromise feature extraction and analysis. Moreover, existing methods often fail to capture subtle ASD-related features due to the reliance on shallow or single-path feature extraction techniques that are not robust to variations across individuals. They also tend to overlook the multi-scale nature of facial characteristics, which are crucial for distinguishing developmental disorders, such as ASD. Age-related changes in facial structures further complicate the identification of consistent diagnostic markers. Additionally, the influence of non-ASD factors (e.g., genetic disorders or environmental factors) can lead to high false positive or negative rates.

To address these limitations, our proposed system introduces a generalized and robust diagnostic framework. It includes a carefully designed image pre-processing stage to enhance image quality, followed by deep feature extraction using a hybrid architecture that captures local and global facial cues. We further enhance feature discrimination through an attention-guided fusion module before classification. These components improve the system's resilience to data variability and generalization ability across diverse populations.

Table 1. A comparison of previous studies for detecting and classifying various ASD images.

Paper	Methodology	Dataset	Strengths	Limitations	Evaluation Metric%
Akter et al. [29] (2021)	MobileNet-V1	Autism Image Data [39]	Improved MobileNet-V1 outperforms other methods with higher accuracy	Limited images and low quality	Accuracy: 92.10
Li et al. [30] (2023)	MobileNetV2 and MobileNetV3	Autism Image Data [39]	Suitable for mobile devices	Low accuracy	Accuracy: 90.5 Recall: 92.33 F1-score: 90.67
Melinda et al. [31] (2024)	DeepLabV3	Autism Image Data [39]	The integration of DeepLabV3 improves accuracy	Limited dataset	Accuracy: 85.9 Recall: 90 Precision: 85.9 F1 score: 87
Ahmed et al. [32] (2024)	ResNet34, ResNet50, AlexNet, MobileNetV2, VGG16, and VGG19	Autism Image Data [39]	Efficient use of transfer learning	Low accuracy	Accuracy: 92
Fahaad et al. [33] (2024)	ViT model	Autism Image Data [39]	ViT models capture both local and global features	Limited dataset	Accuracy: 77
Reddy et al. [34] (2024)	EfficientNetB0	Autism Image Data [39]	Lightweight deep learning	Low accuracy	Accuracy: 87.9
Mahmoud et al. [35] (2023)	A sequencer-based patch wise Local Feature Extractor along with a Global Feature Extractor.	Autism Image Data [39]	Combines local and global features for improved classification.	Limited dataset	Accuracy: 94.7 Recall: 95.3 Precision: 94 F1-score: 94.6
Mujeeb et al. [36] (2022)	Used five CNN models as MobileNet, Xception, EfficientNetB0, EfficientNetB1, EfficientNetB2 for FE and a DNN for classification.	Autism Image Data [39]	Strong features	Limited dataset	Accuracy: 90 Recall: 88.46 Precision: 92
Alam et al. [37] (2025)	Xception	Autism Image Data [39]	Effectively handles domain differences	Limited dataset	Accuracy: 91 Recall: 91 Precision: 91 F1-score: 91
Hossain et al. [38] (2025)	DenseNet121	Autism Image Data [39]	Used explainable AI techniques for interpretability	Low accuracy	Accuracy: 90.33 Recall: 92 Precision: 92 F1-score: 90

RETRACTED

3. Proposed Methodology

The proposed methodology for ASD diagnosis begins with the input of facial images of individuals. These images undergo pre-processing using logarithmic transformation, which enhances contrast and highlights subtle facial features that may be crucial for ASD detection. Next, feature extraction is performed using NasNetMobile and DeiT Network, two DL models that capture intricate patterns and facial characteristics relevant to ASD diagnosis. The extracted features from both networks are then fused using Attentional Feature Fusion, which adaptively assigns importance to the most discriminative features, ensuring an optimal representation. Finally, the fused features are used for classification, where a bagging ensemble with an SVM classifier employing a polynomial kernel is applied to enhance robustness and improve diagnostic accuracy. This approach effectively leverages deep feature learning, attention-based fusion, and ensemble classification to support ASD diagnosis using facial image analysis. Figure 1 indicates the framework for Autism Spectrum Disorder diagnosis based on children's faces.

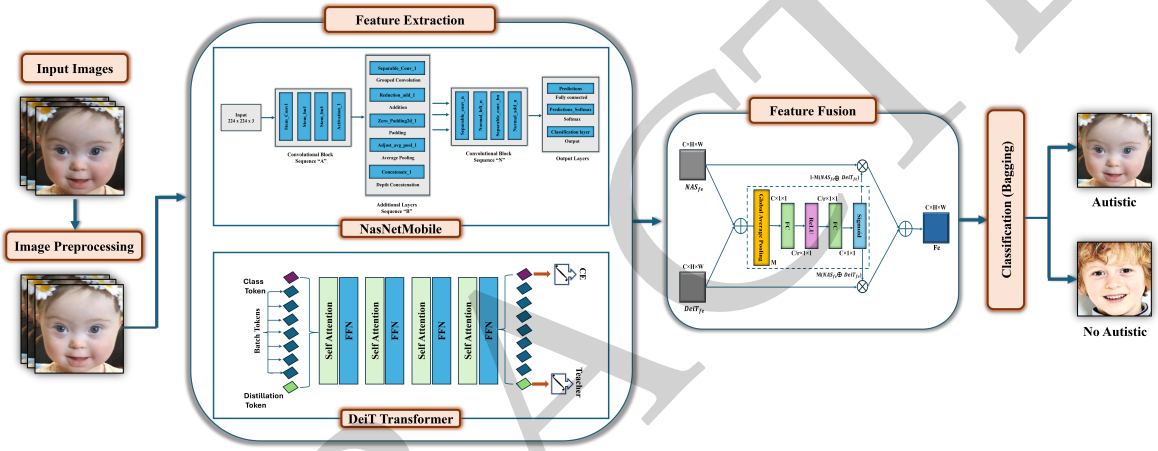


Figure 1. Framework of the proposed ASD diagnosis system based on facial image analysis.

3.1. Input Images

The input images in this study consist of facial images collected from individuals for ASD diagnosis. These images serve as the primary data source for the proposed methodology, where deep learning models analyze facial patterns and features that may indicate ASD-related characteristics. The dataset includes images captured under varying conditions to ensure facial expressions, lighting, and angle diversity, making the model more robust to real-world variations. Each image undergoes pre-processing to enhance its quality and highlight essential feature extraction and classification features. A detailed description of the dataset, including its sources, characteristics, and pre-processing steps, will be provided in Section 4.1 of this paper.

3.2. Images Pre-Processing

Image pre-processing is a crucial step in computer vision and image analysis, aimed at enhancing image quality and improving the performance of subsequent processing tasks. It involves various techniques to refine raw images by reducing noise, adjusting contrast, normalizing intensity values, and enhancing essential features. Common pre-processing methods include grayscale conversion, histogram equalization, normalization, and filtering techniques such as Gaussian or median filtering. Additionally, transformations like logarithmic and gamma correction help adjust brightness and contrast levels. In deep learning and machine learning applications, image pre-processing ensures consistency in input data, improving feature extraction and classification accuracy.

RETRACTED

Logarithmic (Log) Transformation A fundamental image enhancement technique for improving image contrast is presented in [40]. This method expands narrow-range, low frames into a wider range of output levels. It brightens darker intensities, enhancing image characteristics and increasing their visibility to the human eye by brightening darker intensities. Figure 2 illustrates this enhancement process. Initially, dataset images are normalized to achieve narrow-range pixel values by dividing each pixel value by the maximum value of 255, as presented by Equation (1). Subsequently, the Log transformation approach is used as presented by Equation (2).

$$D_{\text{tNorm}} = D_t / 255 \tag{1}$$

where D_{tNorm} is the dataset used after applying normalization

$$D_{\text{tLog}} = c * \log(1 + D_{\text{tNorm}}) \tag{2}$$

where D_{tLog} represents the dataset after applying log transformation for image enhancement, and c is a scaling constant whose value varies depending on the specific application [41]. In this study, the value of c is set to 2, as referenced in [42,43].

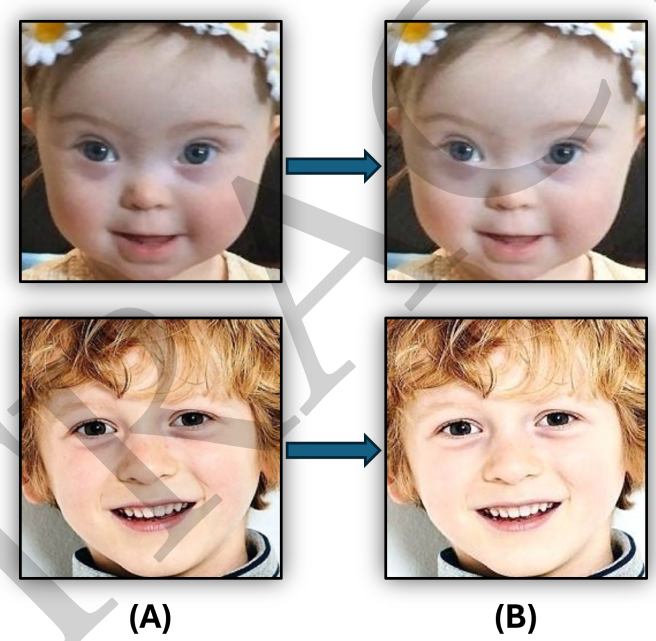


Figure 2. Log transformation: (**A**) the original image and (**B**) the enhanced image.

3.3. Features Extraction

Feature extraction is a critical step in image analysis and classification, where meaningful representations are derived from raw images to improve the ML model's performance [44–46]. We employ the DeiT (Data-efficient Image Transformer) and NASNetMobile networks as feature extractors in this process. DeiT, a vision transformer, efficiently captures long-range dependencies and global contextual information, making it highly effective for extracting rich and robust features. On the other hand, NASNetMobile, a lightweight convolutional neural network optimized through neural architecture search, provides high-quality hierarchical features with reduced computational cost. By leveraging DeiT and NASNetMobile, we integrate transformer-based and CNN-based feature representations, ensuring a comprehensive and discriminative feature set for improved classification accuracy.

RETRACTED

3.3.1. NASNetMobile DL Model

Neural architecture search (NAS) is an advanced DL technique in artificial neural networks (ANNs). Introduced by the Google Brain team in 2016, NAS consists of three key components: search strategy, search space, and performance estimation [47]. The search space defines various architectural elements, including convolutional layers, fully connected layers, and max-pooling. It also determines their connections to form feasible network architectures, as shown in Figure 3. The search strategy employs random search and reinforcement learning methods to explore potential network architectures by evaluating their performance based on metrics like accuracy and computational efficiency. Performance estimation focuses on minimizing computational costs and optimizing time management, ensuring that network performance is assessed within the search strategy framework when evaluating candidate architectures [48,49].

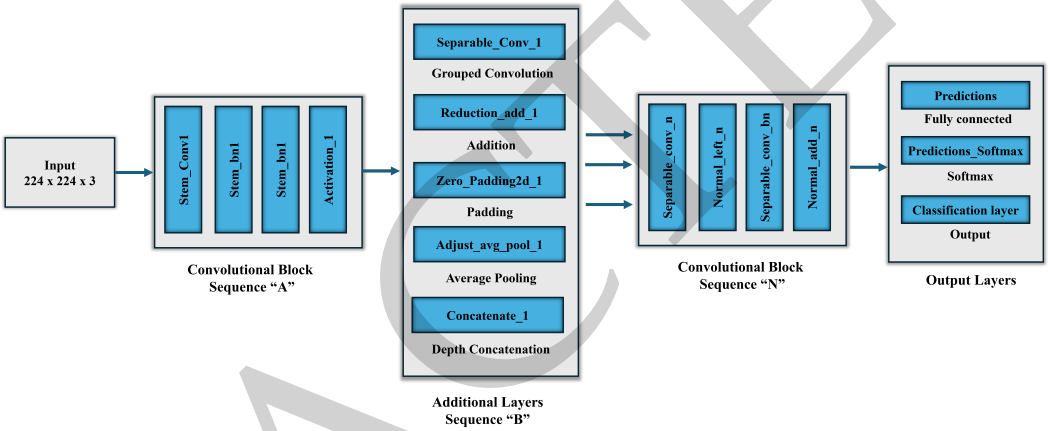


Figure 3. The NASNetMobile architecture.

3.3.2. Data-Efficient Image Transformer (DeiT)

DeiT is an optimized version of the ViT that enhances data efficiency through a unique teacher–student distillation approach. It processes images by dividing them into a series of patches, which are then represented as tokens. These tokens undergo embedding and are analyzed using self-attention mechanisms to extract spatial relationships and contextual details [50]. Figure 4 indicates the DeiT architecture.

Let the input image be denoted as $\mathbf{x} \in \mathbb{R}^{H \times W \times C}$, where H, W, and C represent the height, width, and number of channels, respectively. The process of tokenization in DeiT involves reshaping and embedding patches of size $p \times p$ as follows:

- Patch Embedding: The image is split into a sequence of N patches, where

$$N = \frac{H \times W}{p^2}.$$

- Linear Embedding: Each patch is linearly embedded into a vector of dimension d, resulting in

$$\mathbf{z}_0 = [\mathbf{x}_{cls}; E\mathbf{x}_1; E\mathbf{x}_2; \ldots; E\mathbf{x}_N] + E_{pos}$$

  where $\mathbf{x}_{cls}$ is a class token, E is the embedding matrix, and E_{pos} represents the position embeddings.
- Self-Attention Mechanism: The self-attention layer is defined as

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right)\mathbf{V}$$

RETRACTED

where Q, K, and V are query, key, and value matrices, and d_k is the dimension of the keys. After multiple layers of self-attention and feed-forward transformations, DeiT generates the final feature representation, z out z_{out}, which captures a comprehensive understanding of the input image [51]. Figure 4 indicates the DeiT network architecture.

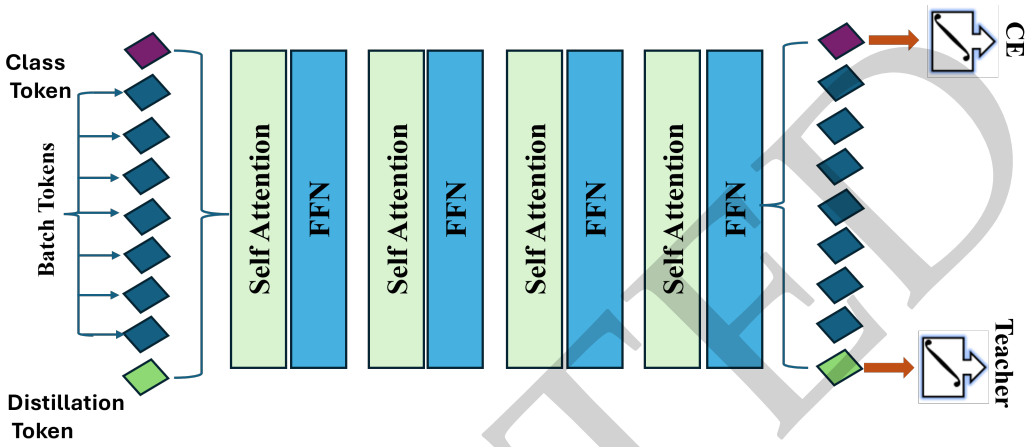


Figure 4. The DeiT architecture.

3.4. Feature Fusion

Feature fusion is a crucial step in many deep learning models, especially in tasks involving multi-source or multi-level feature representations [52–54]. Generally, feature fusion techniques can be categorized into conventional methods such as concatenation, addition, and multiplication, and more advanced strategies based on attention mechanisms [55–57]. Among the attention-based approaches, two popular methods are stochastic attention-based fusion [58] and attention feature fusion [59].

Stochastic attention introduces an element of randomness into the attention weights during the fusion process. This allows the model to explore different feature combinations and potentially improve generalization. However, the inherent randomness may cause instability during inference, making it less predictable and sometimes harder to interpret. While stochastic attention can lead to more robust models in certain cases, it may come at the cost of consistent and stable performance.

On the other hand, attention feature fusion uses a deterministic approach to compute attention weights based on the relevance of different features. This method focuses on the most informative parts of the features, assigning higher attention to the important areas and maintaining consistency and stability throughout training and inference. Attention feature fusion is generally preferred when the goal is to achieve stable performance and interpretability, especially when dealing with large datasets or when the model's reliability and transparency are crucial. In our approach, we opt for attention feature fusion because it provides stable and interpretable results, ensuring that the model focuses on the most critical features while maintaining consistent performance across different tasks.

After the feature extraction stage, two feature maps were obtained: one for DeiT features and the other from NASNetMobile. Feature fusion, as a key element of modern architecture, integrates features from various layers. While summation and concatenation are common, attention-based fusion improves the process [59]. This approach, aided by skip connections, captures information from shallow and deep layers. The Feature Fusion M module combines features of different resolutions, representing complex structures and

RETRACTED

asymmetrical cloud shadow patterns. Equation (3) presents the fusion of DeiT features and those from NASNetMobile.

$$\begin{aligned} Fe = M(NAS_{FE} \oplus DeiT_{FE}) \otimes DeiT_{FE} \\ + (1 - M(NAS_{FE} \oplus DeiT_{FE})) \otimes NAS_{FE} \end{aligned} \tag{3}$$

where NAS_{FE} and $DeiT_{FE}$ be the two input feature sets, and $Fe \in \mathbb{R}^{C \times H \times W}$ the fused feature. The weight function $M(NAS_{FE} \oplus DeiT_{FE})$, derived from the channel attention module M, takes values between 0 and 1. Similarly, $1 - M(NAS_{FE} \oplus DeiT_{FE})$, shown by the dotted arrow in Figure 5, has values in the same range. Here, $\oplus$ denotes elementwise addition and $\otimes$ elementwise multiplication.

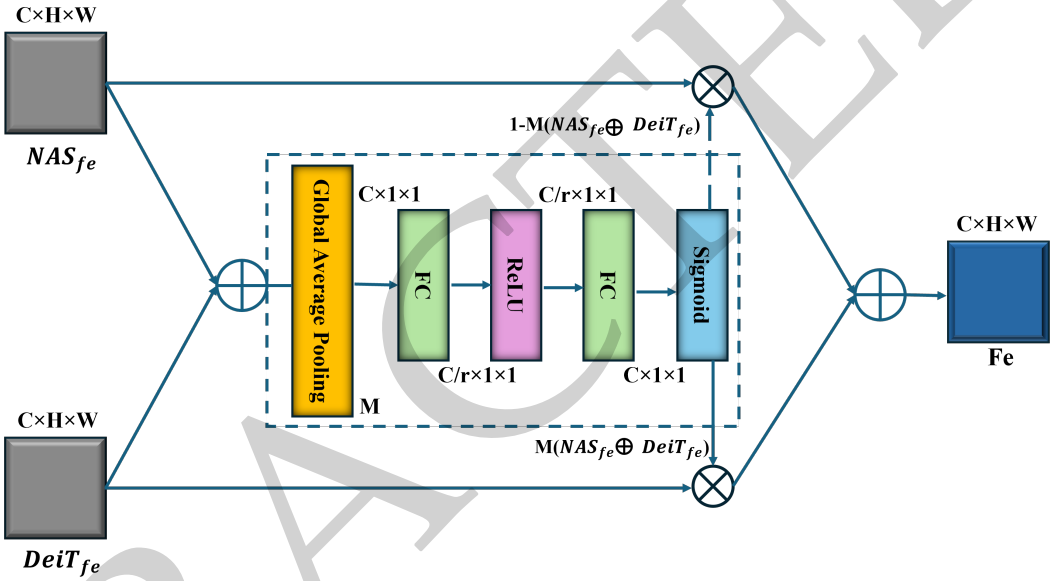


Figure 5. The attention feature fusion architecture.

3.5. Classification

The classification stage assigns input data to pre-defined categories based on the extracted features. Ensemble learning techniques like bagging (i.e., Bootstrap Aggregating) improve performance by reducing variance and increasing stability. Bagging trains various base classifiers on different subsets of data and aggregates their predictions [60]. When paired with a polynomial kernel SVM, bagging enhances robustness by capturing complex decision boundaries, allowing for better class separation in nonlinear datasets. This combination boosts model generalization and accuracy through ensemble diversity and the polynomial kernel's ability to capture intricate patterns.

The ensemble prediction method derives the final estimator through classification voting, where the class receiving the most votes has been chosen as the final prediction. Each base learner contributes their vote for each class, and the total number of votes for every class is accumulated, as shown by Equation (4).

$$F(x) = \arg\max_{y} \sum_{i=1}^{B} (f_i(x) == y) \tag{4}$$

where $F(x)$ denotes the predicted class label for input x, and $\arg\max_y$ identify the class y with the highest vote. The summation $\sum_{i=1}^{B}$ is overall base classifiers, where B is the total number of classifiers. $f_i(x)$ represents the prediction of the i-th classifier for input x, and $f_i(x) == y$ evaluates to 1 if the i-th classifier predicts class y, otherwise 0. The variable

RETRACTED

y iterates through all class labels. The key advantage of integrating SVM into the framework is its ability to handle high-dimensional feature spaces effectively.

4. Experimental Results

The proposed Autism Spectrum Disorder diagnosis using the facial images scheme was evaluated using Python, a widely used programming language known for its simplicity, versatility, and extensive library support. Python is commonly applied in various fields, including bioinformatics, machine learning, data science, and AI. In our implementation, we utilized several libraries, such as NumPy 2.2.4, Matplotlib 3.10.1, TensorFlow 2.18.0, scikit-learn 1.6.1, Keras 3.0.0, and OpenCV 4.11.0, to develop and assess the models. The computer specifications for the experiments are as follows: CPU – Intel(R) Core(TM) i7-9750H @ 2.60 GHz (Lenovo, Beijing, China); Memory—16 GB RAM; Operating System—Microsoft Windows 10 (Microsoft, Redmond, WA, USA); Programming Language—Python 3.10.5. The following section discusses the performance evaluation metrics used in this study.

4.1. Dataset Description

One of the key challenges in our research was the lack of a large, publicly available image dataset, which is crucial for developing ML-based image classification models. To construct our proposed models, we leveraged the autistic children dataset from the Kaggle repository [39], which, to our knowledge, is the first and only dataset of its kind. This dataset comprises 2936 colored 2D facial images of children aged 2 to 14, mostly between 2 and 8 years old. The gender ratio in the autistic class (male to female) was approximately 3:1, while in the typically developing (TD) class, it was around 1:1.

The dataset lacks essential details, including clinical history, ASD severity score, ethnicity, and socio-economic background. It is structured into three main folders: training, validation, and test, each containing two subfolders—autistic and non-autistic. The training set contains 2536 images, while the validation and test sets include 100 and 300 images evenly distributed across the subfolders.

For optimal accuracy and consistency, an ML model should ideally be trained on a diverse and extensive dataset representing the full spectrum of ASD. Machine learning-based image classifiers typically require tens of thousands of images for effective training. Compared to other image datasets, the current dataset is relatively small. Figure 6 indicates some samples of the evaluated dataset. Also, Table 2 indicates the numerical attributes of used dataset.

To ensure fairness and eliminate class-level bias, we examined the dataset distribution across training, validation, and test splits. The dataset was balanced in terms of class representation and image quality, making it suitable for unbiased evaluation of the proposed model.

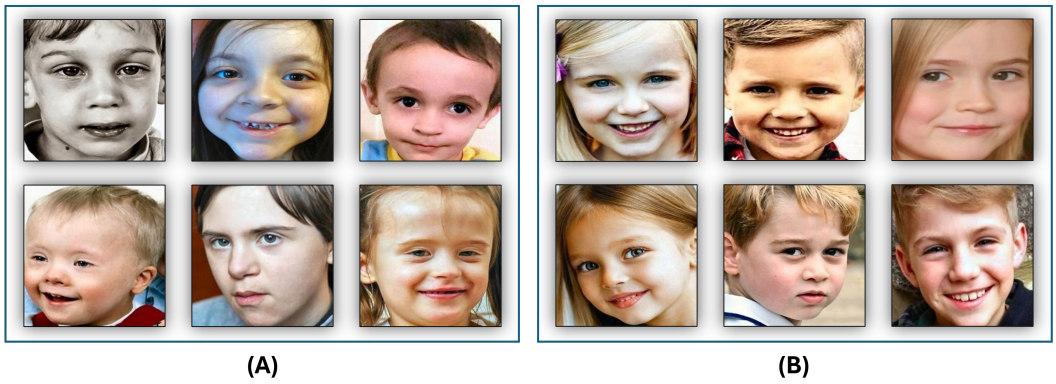


Figure 6. Samples of ASD images: (**A**) autistic and (**B**) not autistic.

RETRACTED

Table 2. The autistic children dataset details.

Attribute	Value
Total Number of Images	2936
Number of Training Images	2536
Number of Validation Images	100
Number of Test Images	300
Age Range	2 to 14 years old (mostly 2 to 8 years old)

4.2. Evaluation Metrics

To assess how well the suggested technique for diagnosing ASD using face images works, various performance metrics are utilized. This section presents the mathematical formulations for calculating these metrics, including accuracy, precision, recall, specificity, and F1-score. These metrics are derived from four key values: true positive (TP), false positive (FP), true negative (TN), and false negative (FN), which are defined as follows: TP: childrens with ASD correctly identified. TN: Neurotypical individuals correctly identified. FP: Neurotypical individuals are incorrectly classified as having ASD. FN: Individuals with ASD are incorrectly classified as neurotypical.

- Accuracy (ACC): This metric has been used to measure the general percentage of correct predictions of the system [61]. It is calculated by Equation (5).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

- Precision (PREC): Quantifies the percentage of accurately classified positive samples of individuals with ASD (TP) to the total predicted positive samples, including both correctly and incorrectly classified cases $(TP + FP)$ [62]. It is computed by Equation (6).

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

- Sensitivity (Recall (REC)): Measures the percentage of correctly identified individuals with ASD (TP) to the total number of individuals with ASD $(TP + FN)$ [63]. It is computed by Equation (7).

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

- Dice Similarity Coeffient (F1-Score): Estimates the system quality [64]. It is the balance between precision and recall. It is computed by Equation (8).

$$F1 - Score = \frac{2 * TP}{2 * TP + FN + FP} \tag{8}$$

- p-values: To quantify the probability that the observed improvement is due to chance. A p-value < 0.05 indicates statistical significance.
- Mann-Whitney U test: It is a non-parametric test to compare the distributions of our model's accuracy against the baseline, especially useful when data are not normally distributed.

RETRACTED

5. Results and Discussion

This section aims to systematically validate the superiority of the proposed framework at each stage, from feature extraction and fusion to feature selection and classification. It discusses the experimental results and comparisons tested on the public dataset, categorized into four main tracks: (1) Comparison of various ML classifiers applied to different DL feature extraction methods. (2) Comparison of ML classifiers applied to fused features from the DeiT transformer and various DL features. (3) Evaluation against state-of-the-art research. A detailed analysis and discussion of these comparisons are provided below.

5.1. A Comparison of ML Classifiers to Pre-Trained DL Models

Transfer learning is a robust deep learning technique that allows pre-trained models to function as feature extractors, eliminating the request for manual feature engineering [65]. Models such as NASNetMobile, DeiT, InceptionResNetV2, VGG16, EfficientNetB0, and MobileNetV2 are widely used for extracting hierarchical features from raw images. NASNetMobile is a lightweight neural architecture search-based model optimized for mobile applications [49]. DeiT (Data-efficient Image Transformer) is a vision transformer that efficiently processes image data without convolutional layers. InceptionResNetV2 integrates inception modules with residual connections to enhance feature learning [66]. VGG16, a deep CNN with 16 layers, is known for its simple yet effective architecture [67]. EfficientNetB0 optimizes accuracy and efficiency by scaling network depth, width, and resolution [68]. MobileNetV2 employs depthwise separable convolutions, making it suitable for mobile and edge computing [69]. These models, pre-trained on large datasets such as ImageNet, can extract robust features that improve classification accuracy.

Once features are extracted, ML classifiers, such as SVM, decision trees (DTs), random forest (RF), and ensemble-based methods, are applied for classification [70]. SVM is effective in high-dimensional spaces and finds an optimal hyperplane for separation [71]. DTs provide interpretable decision rules but are prone to overfitting [72]. RF, an ensemble of decision trees, mitigates overfitting by aggregating multiple decision trees for robust classification [73]. Ensemble-based methods such as bagging and boosting enhance performance by combining multiple weak learners [74,75]. While ML classifiers are computationally efficient and perform well in small datasets, deep learning models offer superior feature extraction capabilities, particularly in complex image classification tasks. However, DL models require significant computational resources. This study compares these approaches to assess their feature extraction and classification accuracy effectiveness.

Table 3 shows the experimental results of different ML classifiers, which have been applied to the features extracted using various pre-trained models (i.e., NASNetMobile, DeiT, InceptionResNetV2, VGG16, EfficientNetB0, and MobileNetV2). ML classifiers that have been applied include SVM (linear), SVM (RBF), SVM (Poly), DT, RF, and bagging classifiers based on SVM(RBF). This table shows the experimental results of precision, recall, F1-Score, and accuracy for the seven classes in the ASD benchmark dataset. The results presented in Table 3 indicate the superiority of the DeiT with the bagging based on the SVM(poly) with an accuracy of 92.67% and also the NASNetMobile DL-based model with the bagging based on the SVM(RBF), achieving the highest results with 91.52% for the overall accuracy against other combinations.

RETRACTED

Table 3. The experimental results of various pre-trained DL models with different ML classifiers.

Model	Classifier	Metric	Classes Name				Overall Accuracy (%)
			Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)	
NASNetMobile	SVM (Linear)	Non_Autistic Autistic	85 85	85 85	85 85	85 85	84.67
	SVM (Poly)	Non_Autistic Autistic	86 88	88 86	87 87	86 88	87
	SVM (RBF)	Non_Autistic Autistic	82 85	84 82	83 84	82 85	83.33
	KNN	Non_Autistic Autistic	85 71	75 82	79 76	85 71	78
	DT	Non_Autistic Autistic	83 80	81 82	82 81	83 80	81.33
	RF	Non_Autistic Autistic	89 92	92 89	90 90	89 92	90.33
	Bagging	Non_Autistic Autistic	91 92	92 91	92 92	91 92	91.67
DeiT	SVM (Linear)	Non_Autistic Autistic	85 85	85 85	85 85	85 85	85
	SVM (Poly)	Non_Autistic Autistic	91 88	88 91	90 89	91 88	89.67
	SVM (RBF)	Non_Autistic Autistic	85 87	87 86	86 86	85 87	86.33
	KNN	Non_Autistic Autistic	93 72	77 92	84 81	93 72	82.67
	DT	Non_Autistic Autistic	86 85	85 86	86 86	86 85	85.67
	RF	Non_Autistic Autistic	90 95	94 90	92 93	90 95	92.33
	Bagging	Non_Autistic Autistic	92 93	93 92	93 93	92 93	92.67
InceptionResNetV2	SVM (Linear)	Non_Autistic Autistic	89 87	88 89	88 88	89 87	88
	SVM (Poly)	Non_Autistic Autistic	87 87	87 87	87 87	87 87	87
	SVM (RBF)	Non_Autistic Autistic	79 81	81 80	80 81	79 81	80.33
	KNN	Non_Autistic Autistic	69 87	84 74	76 80	69 87	78
	DT	Non_Autistic Autistic	87 84	84 86	86 85	87 84	85.33
	RF	Non_Autistic Autistic	93 88	89 92	91 90	93 88	90
	Bagging	Non_Autistic Autistic	91 90	90 91	91 91	91 90	90.66
VGG16	SVM (Linear)	Non_Autistic Autistic	89 79	81 88	85 83	89 79	84
	SVM (Poly)	Non_Autistic Autistic	91 87	88 91	90 89	91 87	89.33
	SVM (RBF)	Non_Autistic Autistic	84 84	84 84	84 84	84 84	84
	KNN	Non_Autistic Autistic	92 65	73 89	81 75	92 65	78.67
	DT	Non_Autistic Autistic	83 85	84 83	84 84	83 85	83.67
	RF	Non_Autistic Autistic	85 87	86 86	86 86	85 87	86
	Bagging	Non_Autistic Autistic	91 87	88 91	90 89	91 87	89.33

RETRACTED

Table 3. *Cont.*

Model	Classifier	Metric	Classes Name				Overall Accuracy (%)
			Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)	
EfficientNetB0	SVM (Linear)	Non_Autistic Autistic	87 92	92 87	89 90	87 92	89.33
	SVM (Poly)	Non_Autistic Autistic	85 88	88 85	86 87	85 88	86.33
	SVM (RBF)	Non_Autistic Autistic	83 83	83 83	83 83	83 83	83.33
	KNN	Non_Autistic Autistic	66 89	85 72	74 80	66 89	77.33
	DT	Non_Autistic Autistic	83 86	86 84	84 85	83 86	84.67
	RF	Non_Autistic Autistic	89 86	86 88	88 87	89 86	87.33
	Bagging	Non_Autistic Autistic	85 88	88 85	86 87	85 88	86.33
MobileNetV2	SVM (Linear)	Non_Autistic Autistic	88 87	87 88	88 88	87 88	87.67
	SVM (Poly)	Non_Autistic Autistic	89 89	89 89	89 89	89 89	89.33
	SVM (RBF)	Non_Autistic Autistic	85 88	88 86	86 87	85 88	86.67
	KNN	Non_Autistic Autistic	91 75	79 90	85 82	91 75	83.33
	DT	Non_Autistic Autistic	87 86	86 87	86 86	87 86	86.33
	RF	Non_Autistic Autistic	85 91	91 86	88 88	85 91	88
	Bagging	Non_Autistic Autistic	89 89	89 89	89 89	89 89	89.33

5.2. A Comparison of ML Classifiers on Fusion Between DL Models with DeiT Transformer

The fusion of DeiT with other DL models, such as NASNetMobile, InceptionResNetV2, VGG16, EfficientNetB0, and MobileNetV2, leverages the strengths of both convolutional and transformer-based architectures for enhanced feature representation. DeiT, as a transformer-based model, excels in capturing long-range dependencies and global contextual information in images, making it highly effective for detailed feature extraction. On the other hand, CNN-based models like NASNetMobile and EfficientNetB0 provide efficient hierarchical feature learning, focusing on local spatial patterns. InceptionResNetV2 combines the power of Inception modules and residual connections, enhancing feature diversity and reducing vanishing gradient issues. VGG16, known for its deep yet simple architecture, extracts low- to high-level features through its sequential convolutional layers. MobileNetV2, optimized for efficiency, is particularly useful in lightweight applications while maintaining strong feature extraction capabilities. Applying feature fusion techniques, such as Attentional Feature Fusion, the most discriminative features from DeiT and CNN-based models are combined, ensuring a rich, multi-scale representation that enhances classification performance. This hybrid approach improves model robustness, leading to higher accuracy, better generalization, and superior adaptability in complex computer vision tasks.

Table 4 shows the experimental results of different ML classifiers, which have been applied to the features extracted using various pre-trained models (i.e., NASNetMobile, DeiT, InceptionResNetV2, VGG16, EfficientNetB0, and MobileNetV2) fused with the DeiT model.

RETRACTED

ML classifiers that have been applied include SVM (linear), SVM (RBF), SVM (Poly), DT, RF, and bagging classifiers based on SVM (RBF). This table shows the experimental results of precision, recall, F1-Score, and accuracy for the seven classes in the ASD benchmark dataset. The results presented in Table 3 indicate the superiority of the NASNetMobile fused with DeiT using the bagging based on the SVM (Poly) with an accuracy of 92.67% and also NASNetMobile DL-based model with the bagging based on the SVM (RBF), achieving the highest results with 95.67% for the overall accuracy against other combinations. The qualitative results shown in Figure 7 illustrate the visual impact of logarithmic enhancement on facial images. For each pair, the left image represents the original input, while the right image displays the result after applying the logarithmic transformation. As observed, the enhanced images exhibit improved brightness and contrast, particularly in low-intensity regions. This transformation helps to reveal subtle facial features that may be suppressed in the original images, making them more prominent and visually distinguishable. Such enhancement benefits downstream tasks such as facial recognition and feature extraction, especially under varying lighting conditions.

The performance comparison Table 5 highlights the strong effectiveness of the proposed model. It achieves high class-specific accuracies, with 0.9800 for the "Autistic" class and 0.9300 for the "Non-Autistic" class, demonstrating its capability to distinguish between the two categories. The model also exhibits impressive average metrics, with precision, recall, and F1-score values near 95.7, indicating robust performance. Furthermore, the overall accuracy of 95.67 showcases its consistency across different data points. Statistical significance is confirmed by a Mann–Whitney U p-value of less than 0.0001, suggesting meaningful improvements over the baseline. The 95% confidence interval for accuracy, ranging from 0.9300 to 0.9800, further assures the model's reliability and stability in various scenarios.

Table 4. The experimental results of various pre-trained DL models with different ML classifiers.

Model	Classifier	Metric	Classes Name				Overall Accuracy (%)
			Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)	
InceptionResNetV2 + DeiT	SVM (Linear)	Non_Autistic Autistic	93 90	90 92	91 91	93 90	91.33
	SVM (Poly)	Non_Autistic Autistic	93 92	92 93	93 93	93 92	92.67
	SVM (RBF)	Non_Autistic Autistic	91 90	90 91	91 91	91 90	90.67
	KNN	Non_Autistic Autistic	67 91	89 74	77 82	67 91	79.33
	DT	Non_Autistic Autistic	87 81	82 86	84 84	87 81	84
	RF	Non_Autistic Autistic	94 89	90 94	92 91	94 89	91.67
	Bagging	Non_Autistic Autistic	95 93	93 95	94 94	95 93	94
VGG16 + DeiT	SVM (Linear)	Non_Autistic Autistic	88 90	90 88	89 89	88 90	89
	SVM (Poly)	Non_Autistic Autistic	87 93	92 88	90 90	87 93	90
	SVM (RBF)	Non_Autistic Autistic	88 91	91 88	89 90	88 91	89.67
	KNN	Non_Autistic Autistic	88 81	82 87	85 84	88 81	84.33
	DT	Non_Autistic Autistic	87 87	87 87	87 87	87 87	87

RETRACTED

Table 4. *Cont.*

Model	Classifier	Metric	Classes Name				Overall Accuracy (%)
			Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)	
VGG16 + DeiT	RF	Non_Autistic Autistic	96 91	91 96	94 93	96 91	93.33
	Bagging	Non_Autistic Autistic	90 95	95 91	92 93	90 95	92.67
EfficientNetV2B0 + DeiT	SVM (Linear)	Non_Autistic Autistic	88 91	91 88	89 90	88 91	89.67
	SVM (Poly)	Non_Autistic Autistic	91 90	90 91	90 90	91 90	90.33
	SVM (RBF)	Non_Autistic Autistic	93 91	91 93	92 92	93 91	92
	KNN	Non_Autistic Autistic	67 97	96 75	79 85	67 97	82.33
	DT	Non_Autistic Autistic	89 87	87 88	88 88	89 87	87.67
	RF	Non_Autistic Autistic	96 89	90 96	93 92	96 89	92.67
	Bagging	Non_Autistic Autistic	93 93	93 93	93 93	93 93	93
MobileNetV2 + DeiT	SVM (Linear)	Non_Autistic Autistic	90 91	91 90	90 90	90 91	90.33
	SVM (Poly)	Non_Autistic Autistic	92 91	91 92	91 91	92 91	91.33
	SVM (RBF)	Non_Autistic Autistic	92 92	92 92	92 92	92 92	92
	KNN	Non_Autistic Autistic	72 93	91 77	80 84	72 93	82.33
	DT	Non_Autistic Autistic	87 86	86 87	86 86	87 86	86.33
	RF	Non_Autistic Autistic	97 90	91 96	94 93	97 90	93.33
	Bagging	Non_Autistic Autistic	95 93	93 95	94 94	95 93	94
NASNetMobile + DeiT	SVM (Linear)	Non_Autistic Autistic	93 90	90 92	91 91	93 90	91.33
	SVM (Poly)	Non_Autistic Autistic	93 91	91 93	92 92	93 91	92
	SVM (RBF)	Non_Autistic Autistic	91 89	90 91	90 90	91 89	90.33
	KNN	Non_Autistic Autistic	82 92	91 84	86 88	82 92	87
	DT	Non_Autistic Autistic	87 87	87 87	87 87	87 87	86.67
	RF	Non_Autistic Autistic	95 90	90 94	93 92	95 90	92.33
	Bagging	Non_Autistic Autistic	98 93	94 98	96 96	98 93	95.67

RETRACTED

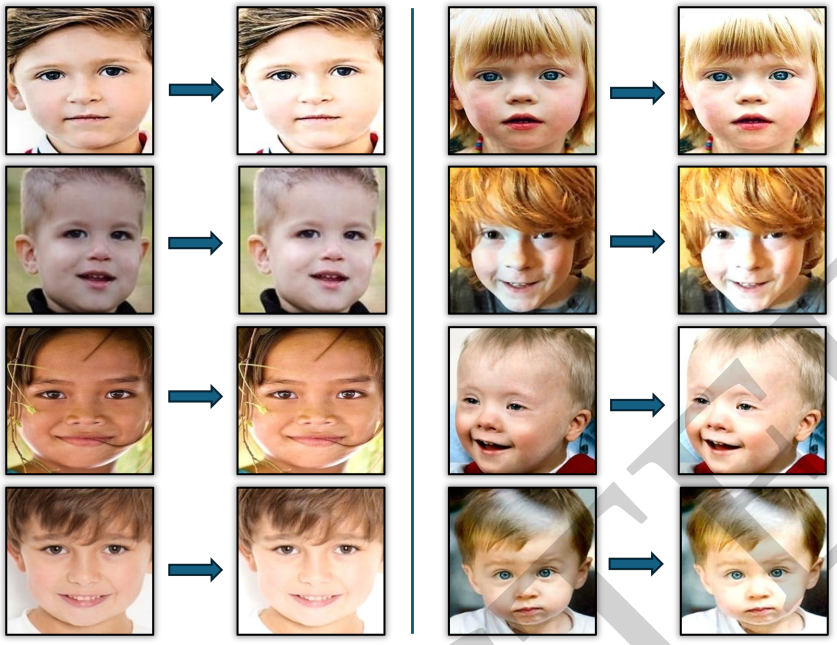


Figure 7. Some qualitative examples of applying logarithmic enhancement.

Table 5. The performance comparison between the proposed model and the baseline.

Metric	Value
Class-Specific Accuracy	
Autistic	0.9800
Non-Autistic	0.9300
Average Metrics	
Precision	0.9577
Recall	0.9567
F1-Score	0.9566
Overall Accuracy	0.9567
Statistical Significance	
Mann–Whitney U *p*-value	<0.0001
95% CI for Accuracy	[0.9300, 0.9800]

To further evaluate the performance of the proposed methodology, Figure 8 illustrates several qualitative examples where the model produced misclassifications. These samples highlight the model's challenges in distinguishing between autistic and non-autistic facial features. For instance, in the left panel of Figure 8, all images are of autistic individuals, yet the model incorrectly predicted them as non-autistic. These cases may be attributed to subtle facial expressions or lighting conditions that resemble those typically found in non-autistic samples. Conversely, the right panel presents non-autistic individuals who were misclassified as autistic. This confusion could stem from overlapping features such as gaze direction, facial symmetry, or expression nuances that are not easily separable by the model. These examples emphasize the importance of incorporating more robust feature extraction techniques and possibly leveraging attention mechanisms or multimodal data to improve classification in borderline or ambiguous cases.

RETRACTED

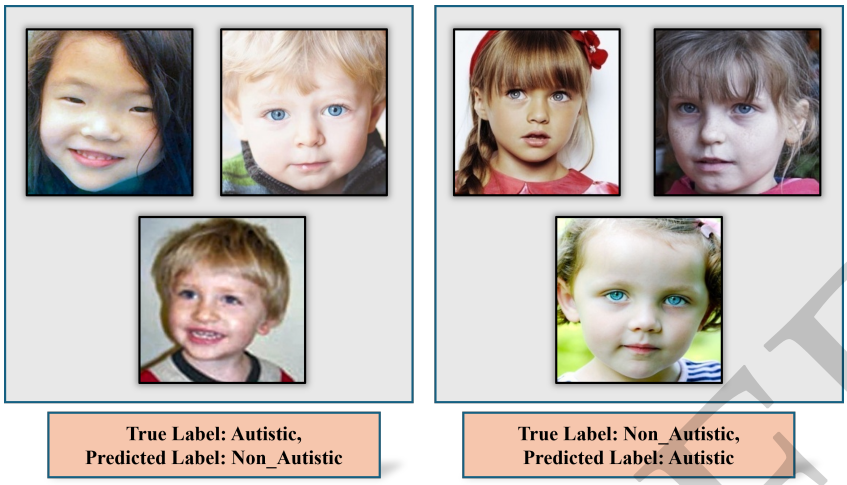


Figure 8. Some examples of misclassified images of the proposed methodology.

5.3. Comparison with the State-of-the-Art Techniques

The comparison of the proposed methodology with existing state-of-the-art methods demonstrates its superior performance in ASD diagnosis using facial images. As shown in Table 6, several previous studies, including those of Akter et al. [29], Li et al. [30], Melinda et al. [31], Ahmed et al. [32], Fahaad et al. [33], and Reddy et al. [34] achieved accuracy values ranging between 85.9% and 92.1%. At the same time, some more recent approaches, such as those by Mahmoud et al. [35], Mujeeb et al. [36], Alam et al. [37], and Hossain et al. [38] reached higher accuracy levels of 94.7%, 90% and 90.33%, respectively. However, the proposed methodology outperforms all prior methods, achieving the highest recall (95.77%), precision (95.67%), F1-score (95.66%), and accuracy (95.67%). These results highlight the effectiveness of feature fusion between DeiT and deep learning models (NASNetMobile, InceptionResNetV2, VGG16, EfficientNetB0, and MobileNetV2) combined with bagging-based SVM classification, which enhances robustness and generalization. The significant performance improvement underscores the potential of the proposed model for more accurate and reliable ASD diagnosis, contributing to the advancement of automated and non-invasive screening methods. Figure 9 indicates the confusion matrix of the proposed methodology for autistic and not autistic classes.

Table 6. The comparison of the proposed methodology with the state-of-the-art techniques.

Paper	Date	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Akter et al. [29]	2021	—	—	—	92.10
Li et al. [30]	2023	92.33	—	90.67	90.5
Melinda et al. [31]	2024	90	85.9	87	85.9
Ahmed et al. [32]	2024	—	—	—	92
Fahaad et al. [33]	2024	—	—	—	77
reddy et al. [34]	2024	—	—	—	87.9
Mahmoud et al. [35]	2023	95.3	94	94.6	94.7
Mujeeb et al. [36]	2022	88.46	92	—	90
Alam et al. [37]	2025	91	91	91	91
Hossain et al. [38]	2025	92	92	90	90.33
Prposed Methodology	2025	95.77	95.67	95.66	95.67

RETRACTED

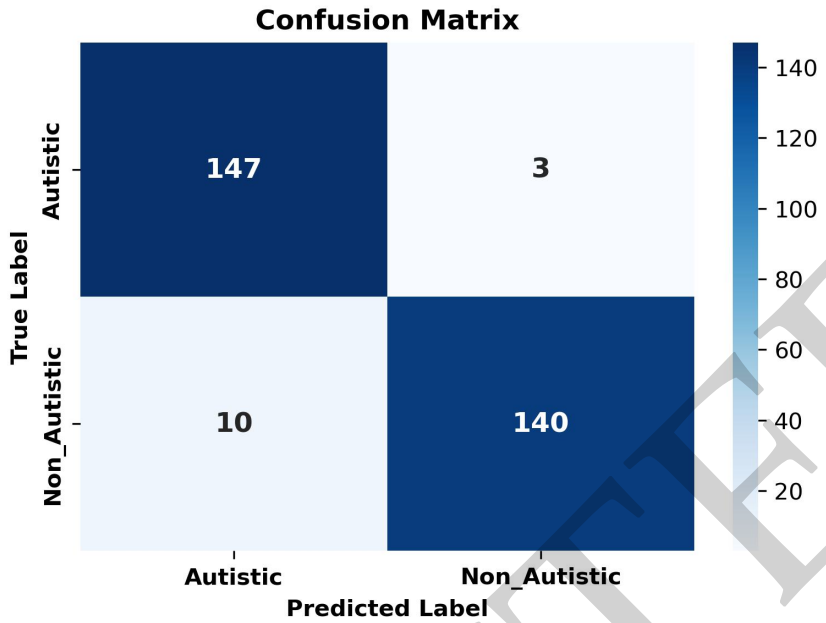


Figure 9. The autism confusion matrix for proposed methodology.

6. Conclusions

ASD is a complex neurodevelopmental condition that affects social interaction, communication, and behavior. Early and accurate diagnosis plays an essential role in providing timely intervention and support for individuals with ASD. Traditional diagnostic approaches typically depend on behavioral assessments, which can be time-consuming and subjective. To address these challenges, this study proposed a deep learning-based approach for ASD diagnosis using facial images, leveraging advanced feature extraction and machine learning techniques. The methodology involved logarithmic transformation for image pre-processing, feature extraction using NasNetMobile and DeiT networks, and feature fusion with attentional feature fusion, followed by classification using bagging with an SVM classifier (polynomial kernel). The experimental results demonstrate the effectiveness of our approach, achieving 95.77% recall, 95.67% precision, 95.66% F1-score, and 95.67% accuracy, highlighting the model's robustness and potential in ASD diagnosis. These findings indicate that facial image-based deep learning models can be a promising tool for early and automated ASD detection, offering a non-invasive, objective, and scalable diagnostic solution. Future advancements in deep learning and multi-modal data fusion may further enhance the accuracy and applicability of such models, contributing to the development of more efficient ASD screening systems. While the proposed model achieved high performance, several directions can be explored to further enhance its effectiveness. One key improvement is expanding the dataset by incorporating a larger and more diverse set of facial images to improve generalization across different populations. Enhancing accuracy by fine-tuning the model and applying advanced feature selection techniques can optimize performance. Another potential advancement is integrating additional pre-trained models, such as Vision Transformers (ViTs), EfficientNet, or Swin Transformer, to extract richer and more diverse features. Beyond DL, exploring multi-modal data fusion by combining facial image analysis with other diagnostic modalities, such as eye-tracking data, speech analysis, or behavioral assessments, could provide a more comprehensive ASD diagnosis. Moreover, improving explainability through attention maps or interpretability techniques can help highlight the most critical facial regions influencing predictions, increasing trust in the model.

RETRACTED

Author Contributions: Conceptualization, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; methodology, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; software, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; validation, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; formal analysis, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; investigation, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; resources, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; data curation, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; writing—original draft preparation, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; writing—review and editing, M.M.E.-G., M.E. and Y.M.F.; visualization, Z.A.A., Y.M.A., M.M.E.-G., M.E. and Y.M.F.; supervision, M.M.E.-G., M.E. and Y.M.F.; project administration, M.M.E.-G., M.E. and Y.M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The proposed methodology has been evaluated on public dataset [39].

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ASD	Autism Spectrum Disorder	NGS	Next-generation sequencing
ID	Intellectual disabilities	ADHD	Hyperactivity disorder
EEG	Electroencephalogram	SOR	Sensory Over-Responsivity
AOs	Anger outbursts	DSM	Diagnostic and Statistical Manual of Mental Disorders
GI	Gastrointestinal	AI	Artificial intelligence
DL	Deep learning	EGs	Experimental groups
CGs	Control groups	TD	Typically developing
LSTM	Long short-term memory	DWT	Discrete wavelet transform
KNN	k-nearest neighbors	XGB	Extreme gradient boosting
AUC	Area under curve	LSTM	Long Short-Term Memory
NLP	Natural language processing	Bi-LSTM	Bidirectional LSTM
ML	Machine learning	DeiT	Data-efficient Image Transformer
NAS	Neural architecture search	ViT	Vision transformer
ANNs	Artificial neural networks	SVMs	Support vector machines
DTs	Decision trees	RF	Random forest
TP	True positive	FP	False positive
TN	True negative	FN	False negative

References

1. Wankhede, N.L.; Kale, M.B.; Shukla, M.; Nathiya, D.; Roopashree, R.; Kaur, P.; Goyanka, B.; Rahangdale, S.R.; Taksande, B.G.; Upaganlawar, A.B.; et al. Leveraging AI for the diagnosis and treatment of autism spectrum disorder: Current trends and future prospects. *Asian J. Psychiatry* **2024**, *101*, 104241. [CrossRef] [PubMed]
2. Liloia, D.; Zamfira, D.A.; Tanaka, M.; Manuello, J.; Crocetta, A.; Keller, R.; Cozzolino, M.; Duca, S.; Cauda, F.; Costa, T. Disentangling the role of gray matter volume and concentration in autism spectrum disorder: A meta-analytic investigation of 25 years of voxel-based morphometry research. *Neurosci. Biobehav. Rev.* **2024**, *164*, 105791. [CrossRef]
3. Choi, L.; An, J.Y. Genetic architecture of autism spectrum disorder: Lessons from large-scale genomic studies. *Neurosci. Biobehav. Rev.* **2021**, *128*, 244–257. [CrossRef]
4. Khogeer, A.A.; AboMansour, I.S.; Mohammed, D.A. The role of genetics, epigenetics, and the environment in ASD: A mini review. *Epigenomes* **2022**, *6*, 15. [CrossRef]
5. Sandin, S.; Lichtenstein, P.; Kuja-Halkola, R.; Larsson, H.; Hultman, C.M.; Reichenberg, A. The familial risk of autism. *JAMA* **2014**, *311*, 1770–1777. [CrossRef]

RETRACTED

6. Sandin, S.; Schendel, D.; Magnusson, P.; Hultman, C.; Surén, P.; Susser, E.; Grønborg, T.; Gissler, M.; Gunnes, N.; Gross, R.; et al. Autism risk associated with parental age and with increasing difference in age between the parents. *Mol. Psychiatry* **2016**, *21*, 693–700. [CrossRef]
7. Frans, E.; Lichtenstein, P.; Hultman, C.; Kuja-Halkola, R. Age at fatherhood: Heritability and associations with psychiatric disorders. *Psychol. Med.* **2016**, *46*, 2981–2988. [CrossRef]
8. Ronald, A.; Pennell, C.E.; Whitehouse, A.J. Prenatal maternal stress associated with ADHD and autistic traits in early childhood. *Front. Psychol.* **2011**, *1*, 223. [CrossRef]
9. Rijlaarsdam, J.; Pappa, I.; Walton, E.; Bakermans-Kranenburg, M.J.; Mileva-Seitz, V.R.; Rippe, R.C.; Roza, S.J.; Jaddoe, V.W.; Verhulst, F.C.; Felix, J.F.; et al. An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics* **2016**, *11*, 140–149. [CrossRef] [PubMed]
10. Hughes, H.K.; Onore, C.E.; Careaga, M.; Rogers, S.J.; Ashwood, P. Increased monocyte production of IL-6 after toll-like receptor activation in children with autism spectrum disorder (ASD) is associated with repetitive and restricted behaviors. *Brain Sci.* **2022**, *12*, 220. [CrossRef]
11. Al-Beltagi, M. Autism medical comorbidities. *World J. Clin. Pediatr.* **2021**, *10*, 15. [CrossRef] [PubMed]
12. Hustyi, K.M.; Ryan, A.H.; Hall, S.S. A scoping review of behavioral interventions for promoting social gaze in individuals with autism spectrum disorder and other developmental disabilities. *Res. Autism Spectr. Disord.* **2023**, *100*, 102074. [CrossRef] [PubMed]
13. Summers, J.; Shahrami, A.; Cali, S.; D'Mello, C.; Kako, M.; Palikucin-Reljin, A.; Savage, M.; Shaw, O.; Lunsky, Y. Self-injury in autism spectrum disorder and intellectual disability: Exploring the role of reactivity to pain and sensory input. *Brain Sci.* **2017**, *7*, 140. [CrossRef] [PubMed]
14. Cummings, K.K.; Jung, J.; Zbozinek, T.D.; Wilhelm, F.H.; Dapretto, M.; Craske, M.G.; Bookheimer, S.Y.; Green, S.A. Shared and distinct biological mechanisms for anxiety and sensory over-responsivity in youth with autism versus anxiety disorders. *J. Neurosci. Res.* **2024**, *102*, e25250. [CrossRef]
15. Thiele-Swift, H.N.; Dorstyn, D.S. Anxiety prevalence in Youth with Autism: A systematic review and Meta-analysis of Methodological and Sample moderators. *Rev. J. Autism Dev. Disord.* **2024**, *11*, 1–14. [CrossRef]
16. Ambrose, K.; Adams, D.; Simpson, K.; Keen, D. Exploring profiles of anxiety symptoms in male and female children on the autism spectrum. *Res. Autism Spectr. Disord.* **2020**, *76*, 101601. [CrossRef]
17. Townsend, A.N.; Guzick, A.G.; Hertz, A.G.; Kerns, C.M.; Goodman, W.K.; Berry, L.N.; Kendall, P.C.; Wood, J.J.; Storch, E.A. Anger outbursts in youth with ASD and anxiety: Phenomenology and relationship with family accommodation. *Child Psychiatry Hum. Dev.* **2024**, *55*, 1259–1268. [CrossRef]
18. Wang, J.; Ma, B.; Wang, J.; Zhang, Z.; Chen, O. Global prevalence of autism spectrum disorder and its gastrointestinal symptoms: A systematic review and meta-analysis. *Front. Psychiatry* **2022**, *13*, 963102. [CrossRef]
19. Piven, J.; Harper, J.; Palmer, P.; Arndt, S. Course of behavioral change in autism: A retrospective study of high-IQ adolescents and adults. *J. Am. Acad. Child Adolesc. Psychiatry* **1996**, *35*, 523–529. [CrossRef]
20. Seltzer, M.M.; Krauss, M.W.; Shattuck, P.T.; Orsmond, G.; Swe, A.; Lord, C. The symptoms of autism spectrum disorders in adolescence and adulthood. *J. Autism Dev. Disord.* **2003**, *33*, 565–581. [CrossRef]
21. Beadle-Brown, J.; Murphy, G.; Wing, L.; Gould, J.; Shah, A.; Holmes, N. Changes in skills for people with intellectual disability: A follow-up of the Camberwell Cohort. *J. Intellect. Disabil. Res.* **2000**, *44*, 12–24. [CrossRef] [PubMed]
22. Kohli, M.; Kar, A.K.; Sinha, S. The role of intelligent technologies in early detection of autism spectrum disorder (asd): A scoping review. *IEEE Access* **2022**, *10*, 104887–104913. [CrossRef]
23. Zhang, X.; Quan, L.; Yang, Y. SAPDA: Significant areas preserved data augmentation. *Int. J. Mach. Learn. Cybern.* **2024**, *15*, 5107–5118. [CrossRef]
24. Biswas, M.; Buckchash, H.; Prasad, D.K. pNNCLR: Stochastic pseudo neighborhoods for contrastive learning based unsupervised representation learning problems. *Neurocomputing* **2024**, *593*, 127810. [CrossRef]
25. Wang, J.; Li, X.; Ma, Z. Multi-Scale Three-Path Network (MSTP-Net): A new architecture for retinal vessel segmentation. *Measurement* **2025**, *250*, 117100. [CrossRef]
26. Zhao, Y.; Li, X.; Zhou, C.; Peng, H.; Zheng, Z.; Chen, J.; Ding, W. A review of cancer data fusion methods based on deep learning. *Inf. Fusion* **2024**, *108*. [CrossRef]
27. Yin, W.; Mostafa, S.; Wu, F.X. Diagnosis of autism spectrum disorder based on functional brain networks with deep learning. *J. Comput. Biol.* **2021**, *28*, 146–165. [CrossRef]
28. Ding, Y.; Zhang, H.; Qiu, T. Deep learning approach to predict autism spectrum disorder: A systematic review and meta-analysis. *BMC Psychiatry* **2024**, *24*, 739. [CrossRef]
29. Akter, T.; Ali, M.H.; Khan, M.I.; Satu, M.S.; Uddin, M.J.; Alyami, S.A.; Ali, S.; Azad, A.; Moni, M.A. Improved transfer-learning-based facial recognition framework to detect autistic children at an early stage. *Brain Sci.* **2021**, *11*, 734. [CrossRef]

RETRACTED

30. Li, Y.; Huang, W.C.; Song, P.H. A face image classification method of autistic children based on the two-phase transfer learning. *Front. Psychol.* **2023**, *14*, 1226470. [CrossRef]
31. Melinda, M.; Aqif, H.; Junidar, J.; Oktiana, M.; Basir, N.B.; Afdhal, A.; Zainal, Z. Image segmentation performance using Deeplabv3+ with Resnet-50 on autism facial classification. *JURNAL INFOTEL* **2024**, *16*, 441–456. [CrossRef]
32. Ahmad, I.; Rashid, J.; Faheem, M.; Akram, A.; Khan, N.A.; Amin, R.u. Autism spectrum disorder detection using facial images: A performance comparison of pretrained convolutional neural networks. *Healthc. Technol. Lett.* **2024**, *11*, 227–239. [CrossRef] [PubMed]
33. Fahaad Almufareh, M.; Tehsin, S.; Humayun, M.; Kausar, S. Facial Classification for Autism Spectrum Disorder. *J. Disabil. Res.* **2024**, *3*, 20240025. [CrossRef]
34. Reddy, P. Diagnosis of Autism in Children Using Deep Learning Techniques by Analyzing Facial Features. *Eng. Proc.* **2024**, *59*, 198. [CrossRef]
35. Mahamood, M.N.; Uddin, M.Z.; Shahriar, M.A.; Alnajjar, F.; Ahad, M.A.R. Autism Spectrum Disorder Classification via Local and Global Feature Representation of Facial Image. In Proceedings of the 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Honolulu, HI, USA, 1–4 October 2023; pp. 1892–1897.
36. Mujeeb Rahman, K.; Subashini, M.M. Identification of autism in children using static facial features and deep neural networks. *Brain Sci.* **2022**, *12*, 94. [CrossRef]
37. Alam, S.; Rashid, M.M. Enhanced Early Autism Screening: Assessing Domain Adaptation with Distributed Facial Image Datasets and Deep Federated Learning. *IIUM Eng. J.* **2025**, *26*, 113–128. [CrossRef]
38. Hossain, S.S.; Al-Islam, F.; Islam, M.R.; Rahman, S.; Parvej, M.S. Autism Spectrum Disorder Identification from Facial Images Using Fine Tuned Pre-trained Deep Learning Models and Explainable AI Techniques. *Semarak Int. J. Appl. Psychol.* **2025**, *5*, 29–53. [CrossRef]
39. Gerry. Autistic Children Data Set. 2020. Available online: https://www.kaggle.com/cihan063/autism-image-data (accessed on 2 July 2021).
40. Chaudhury, S.; Raw, S.; Biswas, A.; Gautam, A. An integrated approach of logarithmic transformation and histogram equalization for image enhancement. In *Proceedings of the Fourth International Conference on Soft Computing for Problem Solving: SocProS 2014, Silchar, Assam, India, 27–29 December 2014*; Springer: New Delhi, India, 2015; Volume 1, pp. 59–70.
41. Manikpuri, U.; Yadav, Y. Image enhancement through logarithmic transformation. *Int. J. Innov. Res. Adv. Eng. (IJIRAE)* **2014**, *1*, 357–362.
42. Bhosale, A. Log Transform. MATLAB Central File Exchange. 2023. Available online: https://www.mathworks.com/matlabcentral/fileexchange/50286-log-transform (accessed on 22 October 2023).
43. Alsakar, Y.M.; Sakr, N.A.; Elmogy, M. An enhanced classification system of various rice plant diseases based on multi-level handcrafted feature extraction technique. *Sci. Rep.* **2024**, *14*, 30601. [CrossRef]
44. Guyon, I.; Elisseeff, A. An introduction to feature extraction. In *Feature Extraction: Foundations and Applications*; Springer: Berlin/Heidelberg,Germany, 2006; pp. 1–25.
45. Nader, N.; El-Gamal, F.E.Z.A.; Elmogy, M. Enhanced kinship verification analysis based on color and texture handcrafted techniques. *Vis. Comput.* **2024**, *40*, 2325–2346. [CrossRef]
46. Mutlag, W.K.; Ali, S.K.; Aydam, Z.M.; Taher, B.H. Feature extraction methods: A review. *J. Phys. Conf. Ser.* **2020**, *1591*, 012028. [CrossRef]
47. Addagarla, S.K.; Chakravarthi, G.K.; Anitha, P. Real time multi-scale facial mask detection and classification using deep transfer learning techniques. *Int. J.* **2020**, *9*, 4402–4408. [CrossRef]
48. He, X.; Zhao, K.; Chu, X. AutoML: A survey of the state-of-the-art. *Knowl.-Based Syst.* **2021**, *212*, 106622. [CrossRef]
49. Zoph, B.; Vasudevan, V.; Shlens, J.; Le, Q.V. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8697–8710.
50. Touvron, H.; Cord, M.; Douze, M.; Massa, F.; Sablayrolles, A.; Jégou, H. Training data-efficient image transformers & distillation through attention. In *Proceedings of the 38th International Conference on Machine Learning (ICML), Virtual Event, 18–24 July 2021*; PMLR: Birmingham, UK, 2021; Volume 139, pp. 10347–10357.
51. Wang, W.; Zhang, J.; Cao, Y.; Shen, Y.; Tao, D. Towards data-efficient detection transformers. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Cham, Switzerland, 2022; pp. 88–105.
52. Dai, Y.; Gieseke, F.; Oehmcke, S.; Wu, Y.; Barnard, K. Attentional feature fusion. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Virtual Event, 5–9 January 2021*; pp. 3560–3569.
53. Shahzad, I.; Khan, S.U.R.; Waseem, A.; Abideen, Z.U.; Liu, J. Enhancing ASD classification through hybrid attention-based learning of facial features. *Signal Image Video Process.* **2024**, *18*, 475–488. [CrossRef]
54. Dong, X.; Qin, Y.; Gao, Y.; Fu, R.; Liu, S.; Ye, Y. Attention-based multi-level feature fusion for object detection in remote sensing images. *Remote Sens.* **2022**, *14*, 3735. [CrossRef]

RETRACTED

55. An, L.; Wang, L.; Li, Y. HEA-Net: Attention and MLP hybrid encoder architecture for medical image segmentation. *Sensors* **2022**, *22*, 7024. [CrossRef]
56. Fan, X.; Li, X.; Yan, C.; Fan, J.; Chen, L.; Wang, N. Converging Channel Attention Mechanisms with Multilayer Perceptron Parallel Networks for Land Cover Classification. *Remote Sens.* **2023**, *15*, 3924. [CrossRef]
57. Dong, S.; Liu, J.; Han, B.; Wang, S.; Zeng, H.; Zhang, M. UMAP-Based All-MLP Marine Diesel Engine Fault Detection Method. *Electronics* **2025**, *14*, 1293. [CrossRef]
58. Li, W.; Deng, Y.; Ding, M.; Wang, D.; Sun, W.; Li, Q. Industrial data classification using stochastic configuration networks with self-attention learning features. *Neural Comput. Appl.* **2022**, *34*, 22047–22069. [CrossRef]
59. Du, W.; Fan, Z.; Yan, Y.; Yu, R.; Liu, J. AFMUNet: Attention Feature Fusion Network Based on a U-Shaped Structure for Cloud and Cloud Shadow Detection. *Remote Sens.* **2024**, *16*, 1574. [CrossRef]
60. Alsakar, Y.M.; Elazab, N.; Nader, N.; Mohamed, W.; Ezzat, M.; Elmogy, M. Multi-label dental disorder diagnosis based on MobileNetV2 and swin transformer using bagging ensemble classifier. *Sci. Rep.* **2024**, *14*, 25193. [CrossRef] [PubMed]
61. Arjunagi, S.; Patil, N. Texture based leaf disease classification using machine learning techniques. *Int. J. Eng. Adv. Technol. (IJEAT)* **2019**, *9*, 2249–8958. [CrossRef]
62. Bonidia, R.P.; Sampaio, L.D.H.; Lopes, F.M.; Sanches, D.S. Feature extraction of long non-coding rnas: A fourier and numerical mapping approach. In Proceedings of the Iberoamerican Congress on Pattern Recognition, Havana, Cuba, 28–31 October 2019; Springer: Cham, Switzerland, 2019; pp. 469–479.
63. Wang, B.; Zhang, C.; Du, X.x.; Zhang, J.f. lncRNA-disease association prediction based on latent factor model and projection. *Sci. Rep.* **2021**, *11*, 19965. [CrossRef]
64. Chowdhury, M.E.; Rahman, T.; Khandakar, A.; Ayari, M.A.; Khan, A.U.; Khan, M.S.; Al-Emadi, N.; Reaz, M.B.I.; Islam, M.T.; Ali, S.H.M. Automatic and reliable leaf disease detection using deep learning techniques. *AgriEngineering* **2021**, *3*, 294–312. [CrossRef]
65. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
66. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
67. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
68. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning (PMLR), Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.
69. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
70. Pereira, F.; Mitchell, T.; Botvinick, M. Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage* **2009**, *45*, S199–S209. [CrossRef]
71. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.
72. Safavian, S.R.; Landgrebe, D. A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 660–674. [CrossRef]
73. Sandika, B.; Avil, S.; Sanat, S.; Srinivasu, P. Random forest based classification of diseases in grapes from images captured in uncontrolled environments. In Proceedings of the 2016 IEEE 13th International Conference on Signal Processing (ICSP), Chengdu, China, 6–10 November 2016; pp. 1775–1780.
74. Chen, J.; Zeb, A.; Nanehkaran, Y.A.; Zhang, D. Stacking ensemble model of deep learning for plant disease recognition. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 12359–12372. [CrossRef]
75. Vo, H.T.; Quach, L.D.; Hoang, T.N. Ensemble of deep learning models for multi-plant disease classification in smart farming. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 1045–1054. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

RETRACTED