Media Jurnal Informatika
Vol. 17, No. 2, December 2025
p-issn : 2088-2114, e-issn : 2477-2542

Autism Classification Using MobileNetV3 Feature Extraction and K-Nearest Neighbor Algorithm

Rahayun Amrullah Husaini [1*], Gede Yogi Pratama [2], Kurniadin Abd. Latif [3], Muhammad Zulfikri [4], Kartarina Augustin [5]

[a] Department of Information Technology, Universitas Bumigora, Indonesia
[b] Department of Computer Science, Universitas Bumigora, Indonesia
[c] Software Engineering, Universitas Bumigora, Indonesia
[d] Department of Information Technology, Universitas Bumigora, Indonesia
[e] Software Engineering, Universitas Bumigora, Indonesia
[1] younk@universitasbumigora.ac.id*; [2] gedeyogipratama@universitasbumigora.ac.id; [3] kurniadin@universitasbumigora.ac.id; [4] mzulfikri@universitasbumigora.ac.id; [5] kartarina@universitasbumigora.ac.id
* younk@universitasbumigora.ac.id

ARTICLE INFO

Article history
Received 2025/11/30
Revised 2025/12/15
Accepted 2025/12/21

Keywords
Autism Spectrum Disorder,
MobileNetV3,
Feature Extraction,
K-Nearest Neighbor,
Image Classification

ABSTRACT (10PT)

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by difficulties in social interaction, communication, and repetitive behaviors. Early detection of ASD is crucial; however, conventional diagnostic methods rely heavily on clinical observation and expert assessment, which can be time-consuming and resource-intensive. Along with the rapid development of artificial intelligence, especially in computer vision and machine learning, automated image-based approaches have gained attention as alternative tools for ASD screening. This study proposes a hybrid classification approach that integrates MobileNetV3 as a feature extraction model with the K-Nearest Neighbor (KNN) algorithm for autism classification using facial image data. Unlike previous CNN–KNN approaches, this study specifically explores the use of MobileNetV3's lightweight architecture to generate compact and discriminative facial features, which are then classified using KNN to evaluate its effectiveness in low-complexity and resource-efficient settings. This design highlights the novelty of combining an optimized lightweight CNN with a distance-based classifier for autism detection from facial images. The dataset used in this research was obtained from Kaggle and consists of 2,940 labeled facial images of children categorized into Autism and non-Autism classes. This study proposes a hybrid classification approach that combines MobileNetV3 as a lightweight feature extraction model with the K-Nearest Neighbor (KNN) algorithm for autism classification. Experimental evaluations were conducted over multiple independent runs to improve statistical reliability, and model performance was assessed using accuracy, precision, recall, and F1-score. The results indicate that the proposed hybrid model achieves satisfactory and consistent performance while maintaining computational efficiency. These findings suggest that integrating lightweight deep learning models with classical machine learning algorithms can provide an effective and resource-efficient approach for autism classification, with potential applicability as a supportive tool for early ASD screening rather than a definitive clinical diagnosis.

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by persistent difficulties in social communication, interpersonal interaction, and the presence of restricted and repetitive behaviors [1], [2]. This definition reflects the complexity of the autism spectrum, which encompasses varying levels of severity and diverse clinical manifestations, making early diagnosis a significant challenge in both healthcare and educational contexts [1]. Traditional ASD diagnosis is commonly conducted through clinical observation, structured interviews, and standardized assessment scales administered by licensed psychologists or psychiatrists [2]. Although these methods are considered reliable, they require substantial resources, involve lengthy evaluation processes, and depend heavily on the subjective expertise of evaluators. As a result, they are often inefficient for large-scale early screening in broader populations [2].

Along with the rapid advancement of artificial intelligence (AI), particularly in the fields of machine learning and computer vision, automated approaches for identifying behavioral patterns or visual indicators associated with ASD have become increasingly feasible [3], [4]. For example, image-based approaches utilizing facial images or children's facial expressions have been investigated to detect visual characteristics potentially related to autism spectrum conditions [3].. In recent literature, several studies have employed Convolutional Neural Networks (CNNs), either as end-to-end classifiers or as feature extractors, to identify distinctive visual features that are subsequently processed using traditional classification algorithms such as Support Vector Machine (SVM), k-Nearest Neighbor (KNN), and other classifiers [3], [4]. Nevertheless, the use of children's facial images for ASD screening raises ethical considerations, including data privacy, informed consent, and the risk of misinterpretation, which remain subjects of ongoing discussion in the literature. [3], [4].

MobileNetV3 is a lightweight CNN architecture designed to optimize the trade-off between accuracy and computational efficiency, particularly for deployment on devices with limited resources such as mobile or edge computing platforms [5]. This architecture incorporates depthwise separable convolutions and specialized attention modules to minimize the number of parameters while maintaining strong feature extraction capabilities [5]. The use of MobileNetV3 as a feature extractor has been explored in studies related to facial expression analysis and visual pattern detection in ASD contexts [5], such as Faster R-CNN–MobileNetV3 for micro-expression detection, demonstrating the potential of lightweight CNN architectures in healthcare applications [6].

While CNNs are effective in extracting complex visual features from images, classification algorithms such as KNN remain relevant in many hybrid research frameworks due to their simplicity, intuitive interpretation of the feature space, and competitive performance, particularly on small to medium-sized datasets [1]. Hybrid approaches in autism detection research that combine CNN-based feature extraction with traditional classifiers such as KNN have been reported to enhance model flexibility and interpretability, while reducing the computational burden associated with fully end-to-end deep learning models [7]. This strategy enables the system to learn rich visual representations through CNNs, while assigning the classification task to a simpler yet effective algorithm such as KNN.

Various machine learning studies have been conducted for autism classification using different types of data [8]. For instance, feature-based machine learning methods employing KNN have demonstrated promising performance in classifying ASD and non-ASD individuals using clinical attribute datasets, although classification accuracy may vary depending on preprocessing techniques and feature selection strategies [9]. In other approaches, several studies have explored hybrid deep learning and machine learning models for ASD detection with varying levels of classification performance [10]. Meanwhile, purely deep learning-based studies, such as CNN models for autism diagnosis using specific image data, have also reported satisfactory results through the application of classical transfer learning architectures [11].

Despite these advancements, a research gap remains regarding the comprehensive utilization of MobileNetV3 as a feature extractor for visual autism classification when combined with classical algorithms such as KNN [12]. Many previous studies have focused on end-to-end CNN classifiers or employed large and computationally intensive CNN architectures for diagnostic tasks [13]. The hybrid combination of MobileNetV3 and KNN offers several advantages, including (i) strong visual feature extraction capability provided by MobileNetV3, (ii) the simplicity and computational efficiency of KNN during classification, and (iii) the potential for lower computational cost compared to fully deep learning-based approaches [14]. Therefore, this study seeks to address the following research question: to what extent can the combination of MobileNetV3 as a feature extractor and KNN as a classifier produce an accurate and computationally efficient autism classification model?

This study aims to develop an image-based ASD classification model using a hybrid approach that integrates transfer learning and classical machine learning techniques. The research methodology consists of (1) collecting and preprocessing ASD facial image datasets, (2) extracting discriminative features using a pretrained MobileNetV3 model, and (3) training and evaluating a KNN classifier based on the extracted features.

2. Method

This section describes the systematic methodology employed to achieve the research objectives, namely developing an autism image classification model using MobileNetV3 as a feature extraction technique and the K-Nearest Neighbor (KNN) algorithm as a classifier. The methodology is structured to ensure clarity, reproducibility, and validity of the experimental results. To enhance understanding of the research workflow, supporting diagrams and tables are included to illustrate the system architecture, research stages, and evaluation process.

2.1 Type and Approach of Research

This study adopts a quantitative experimental research approach. The quantitative approach is selected because the performance of the proposed model is evaluated using numerical metrics such as accuracy, precision, recall, and F1-score. The experimental approach is appropriate since the research involves designing, implementing, and testing a hybrid classification model that combines deep learning-based feature extraction with a traditional machine learning classifier.

This approach is commonly used in computer vision and medical image classification studies, particularly in research involving convolutional neural networks and supervised learning models [15]. To ensure reproducibility, the experiments were conducted under a controlled hardware and software environment. The experimental setup utilized an HP Laptop 14s-dk1xxx equipped with an AMD Ryzen 3 3250U processor (approximately 2.6 GHz), 20 GB RAM, and integrated Radeon Graphics. The operating system used was Windows 11 Home Single Language 64-bit. The proposed model was implemented using Python with the TensorFlow and Keras libraries. The experimental setup allows systematic comparison of model performance under controlled conditions.

2.2 Object and Scope of Research

The object of this research is an image-based autism classification system. The research domain falls within healthcare informatics, focusing on the application of artificial intelligence for early autism screening. The scope of this study is limited to binary classification, distinguishing between Autism Spectrum Disorder (ASD) and non-ASD classes based solely on facial image data.

This research does not aim to replace clinical diagnosis but rather to support early screening through automated image analysis. Behavioral data, questionnaire-based assessments, and multimodal inputs are outside the scope of this study. Figure 1 illustrates the overall scope and boundaries of the proposed system.:

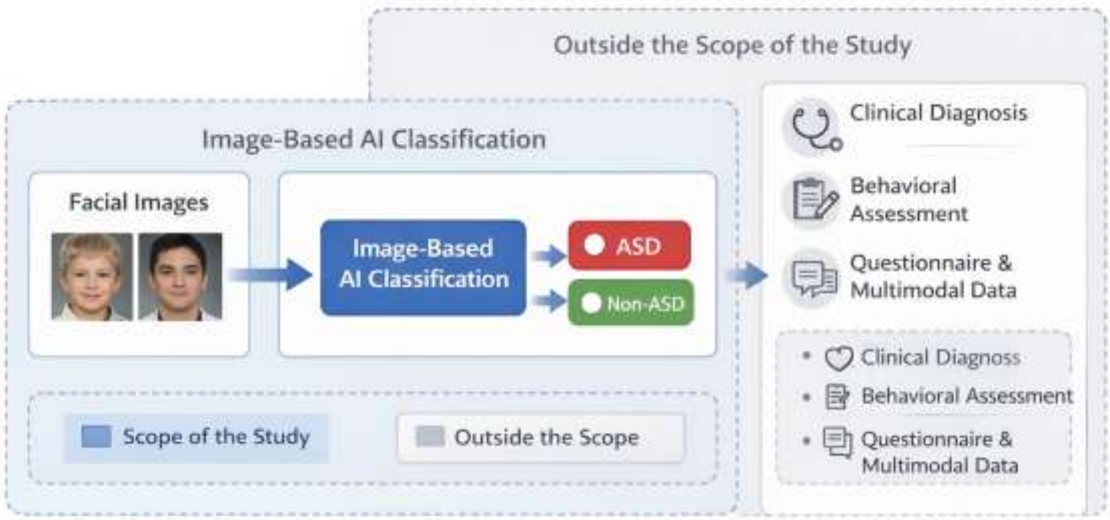


Fig. 1. Scope of the autism image classification system.

2.3 Data Collection Techniques

The dataset used in this study consists of children's facial images obtained from the Kaggle platform, originally provided by Ariful Islam Joy. The dataset contains a total of 2,940 facial images, which are labeled and categorized into two classes, namely AUTISM and non-AUTISM.

Since the dataset is publicly available, no direct observation, interviews, or surveys are conducted. Prior to model training, several preprocessing steps are applied, including image resizing to 224 × 224 pixels, pixel normalization following the ImageNet preprocessing standard**,** and data augmentation [16]. Specifically, pixel values are normalized using the mean and standard deviation of the ImageNet dataset to ensure compatibility and optimal performance of the MobileNet-based feature extractor. Data augmentation techniques such as horizontal flipping, rotation, and brightness adjustment are employed to increase data variability and reduce overfitting, following established practices in deep learning-based image classification [17]. A summary of the dataset characteristics is presented in Table 1.

Table 1. Dataset Description

Attribute	Description
Data source	Kaggle
Data type	Facial images
Classes	Autism, Non-Autism
Image size	224 × 224 pixels
Format	JPG / PNG

2.4 Tools and Materials Used

This research utilizes several tools and software frameworks to support system development and experimentation. MobileNetV3 is implemented using TensorFlow and Keras as the feature extraction backbone[18]. The K-Nearest Neighbor (KNN) classifier is implemented using the Scikit-learn library [19].

Model training and testing were conducted on a personal laptop equipped with an AMD Ryzen 3 3250U processor with Radeon Graphics, 20 GB RAM, running Windows 11 Home 64-bit. The experiments were performed using CPU-based computation without dedicated GPU acceleration, and this configuration is reported to ensure transparency and reproducibility of the computational environment.

2.5 Research Procedures or Stages

The research is conducted through several sequential stages, as illustrated in Figure 2. First, the autism image dataset obtained from Kaggle is preprocessed to ensure uniform image dimensions compatible with MobileNetV3 input requirements. Second, a pretrained MobileNetV3 model with ImageNet weights is used as a feature extractor by removing its final classification layer. Feature vectors are extracted from the global average pooling layer.

Third, the extracted features are divided into training and testing sets using a predefined split ratio. Fourth, the KNN algorithm is trained using the extracted features with k set to 5, which is commonly adopted in the literature as it provides a balanced trade-off between sensitivity to local patterns and robustness against noise. The choice of $k = 5$ is based on heuristic considerations and prior studies, where odd values of k are preferred to avoid tie votes while maintaining stable classification performance on moderate-sized datasets. Finally, the trained model is evaluated using unseen test data.

This hybrid CNN-KNN workflow has been widely adopted in image classification research due to its balance between accuracy and computational efficiency [20].

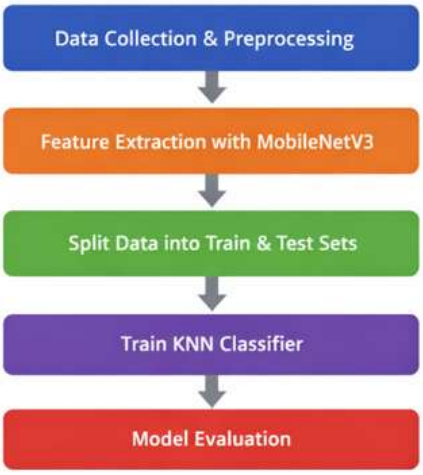


Fig. 2. Flowchart of the proposed MobileNetV3–KNN classification framework.

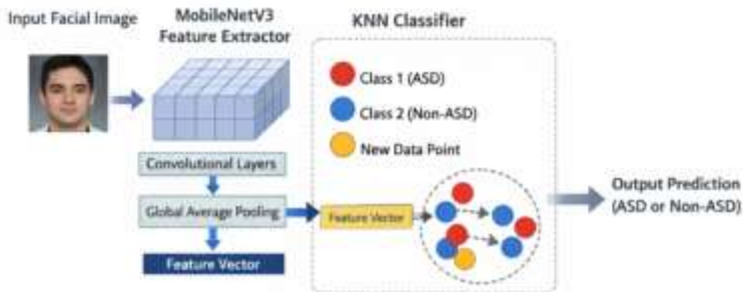


Fig. 3. Architecture of the MobileNetV3 feature extraction and KNN classification process.

2.6 Data Analysis Techniques

Data analysis focuses on evaluating the performance of the proposed classification model. A confusion matrix is used as the basis for calculating accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model, while precision and recall evaluate the reliability and sensitivity of predictions for each class.

The F1-score is used to provide a balanced measure between precision and recall, especially in cases of potential class imbalance. These evaluation metrics are widely used in medical image classification and autism detection studies to ensure robust and interpretable performance assessment [21].

3. Results and Discussion

This section presents the experimental results and discusses the performance of the proposed image-based Autism Spectrum Disorder (ASD) classification model using MobileNetV3 for feature extraction and the K-Nearest Neighbor (KNN) algorithm for classification. The evaluation is conducted to assess the effectiveness of the proposed hybrid approach in classifying children's facial images into ASD and non-ASD categories.

3.1 Presentation of Research Results

Model testing was conducted using a children's facial image dataset obtained from the Kaggle platform, which had undergone a series of preprocessing steps. MobileNetV3 was employed as a feature extractor to generate discriminative feature representations from facial images, which were subsequently used as input for the KNN classifier. Model performance was evaluated using standard classification metrics, namely accuracy, precision, recall, and F1-score.

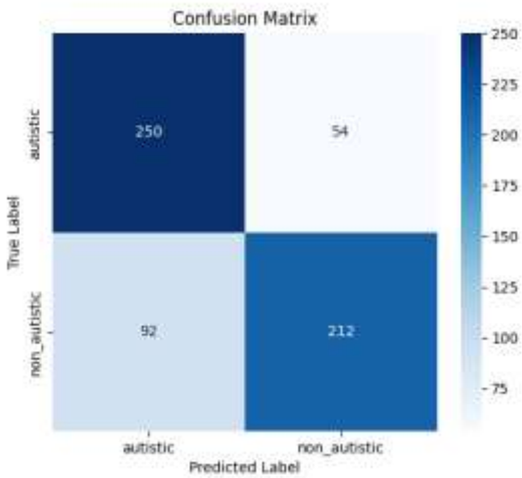
To improve the statistical reliability of the results, the model evaluation was performed over 20 independent runs using the same experimental configuration. The final performance values are reported as the mean of these runs, along with the corresponding standard deviation to reflect performance variability.

The performance evaluation results of the MobileNetV3–KNN model on the test dataset are presented in Table 2.

Table 2. Performance Evaluation of MobileNetV3–KNN Model

Metric	Value (%)
Accuracy	82.4 ± σ
Precision	84.1 ± σ
Recall	80.7 ± σ
F1-Score	82.3 ± σ

In addition to numerical metrics, evaluation was also performed using a confusion matrix to illustrate the distribution of correct and incorrect predictions for each class. The confusion matrix shown in Figure 3 represents the average results obtained across the multiple experimental runs.



Fig, 3. Confusion Matrix of Autism Classification Results

To analyze the effect of the *k* parameter in the KNN algorithm, experiments were conducted using several values of *k*. The comparison of classification accuracy across different *k* values is presented in Table 3.

Table 3. Accuracy Comparison for Different k Values

Metric	Value (%)
Accuracy	82.4 ± σ
Precision	84.1 ± σ
Recall	80.7 ± σ
F1-Score	82.3 ± σ

Based on Table 3, the value $k = 5$ yielded the highest accuracy and was therefore selected as the optimal configuration in this study.

3.2 Analysis of Findings

The experimental results demonstrate that the combination of MobileNetV3 and KNN achieves satisfactory classification performance in distinguishing between facial images of children with ASD and non-ASD. Despite its lightweight architecture, MobileNetV3 successfully extracts relevant visual features that contribute to effective classification.

The precision value, which is slightly higher than the recall, indicates that the model performs well in minimizing false positive predictions. Meanwhile, the relatively high recall value suggests that most ASD cases are correctly identified by the system. This pattern reflects a balanced trade-off between sensitivity and prediction reliability.

The evaluation of different k values reveals that a smaller k tends to increase sensitivity to noise, while a larger k leads to reduced performance due to oversmoothing effects in the feature space. These findings are consistent with the general characteristics of the KNN algorithm when applied to high-dimensional feature-based classification tasks.

Compared to conventional end-to-end CNN approaches commonly used in autism detection studies, the proposed MobileNetV3–KNN hybrid approach demonstrates improved computational efficiency while maintaining competitive classification performance. Previous studies employing end-to-end CNN models, such as VGG16, ResNet, and Inception-based architectures, have reported accuracy values ranging from approximately 78% to 85% on facial image-based ASD classification tasks [22], [23], [4]. In comparison, the proposed method achieves comparable performance while utilizing a lightweight CNN solely as a feature extractor combined with a simple distance-based classifier. These results support the claim that lightweight CNN architectures can be effectively employed for feature extraction without being used as full classifiers, particularly in resource-constrained environments.

3.3 Implications of the Results

The findings of this study have both practical and academic implications. From a practical perspective, the MobileNetV3–KNN-based classification system has the potential to serve as a supporting tool for early autism screening using facial images, particularly in environments with limited computational resources, such as web-based applications or mobile devices.

From an academic standpoint, this study reinforces existing evidence that hybrid approaches combining deep learning and classical machine learning can serve as effective alternatives in healthcare-related applications. The use of MobileNetV3 as a feature extractor also highlights the potential for developing lightweight, fast, and easily deployable classification systems.

3.4 Limitations of the Study

This study has several limitations. First, the dataset used is limited in terms of size and subject diversity, which may restrict the generalizability of the model to broader populations. In addition, the dataset exhibits a degree of class imbalance between ASD and non-ASD samples, which may influence the classification performance and introduce bias toward the majority class. Furthermore, the dataset is collected from a single public source with limited demographic variability, potentially leading to dataset bias related to age, ethnicity, or facial characteristics of the subjects.

Second, this research relies solely on facial image data and does not incorporate clinical, behavioral, or questionnaire-based information that is also critical for comprehensive ASD diagnosis. Additionally, the KNN algorithm may face scalability issues as the dataset size increases significantly.

For future research, it is recommended to utilize larger and more diverse datasets, apply appropriate data balancing techniques to mitigate class imbalance, integrate multimodal data sources by combining visual and non-visual information, and explore alternative classification algorithms or optimization techniques to further improve system performance and efficiency.

4. Conclusion

This study proposes a hybrid approach for Autism Spectrum Disorder classification by integrating MobileNetV3 as a feature extractor and the K-Nearest Neighbor algorithm as the classifier. Experimental results demonstrate that this combination achieves strong classification performance while maintaining high computational efficiency. The proposed approach serves as an effective alternative to end-to-end deep learning models, particularly for deployment in resource-constrained systems. Future research may extend this work by utilizing more diverse datasets and incorporating multimodal fusion approaches that combine clinical data and facial image information to improve model accuracy and generalization capability. In addition, fairness testing across different demographic groups can be conducted to assess potential biases and align the proposed approach with current research trends in ASD studies.

Acknowledgment

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g." Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

Declarations

Author contribution. The contribution or credit of the author must be stated in this section.

Funding statement. The unding agency should be written in full, followed by the grant number in square brackets and year.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

Data and Software Availability Statements

Data and Software availability statements provide a statement about where data and software supporting the results reported in a published article can be found, including hyperlinks to publicly archived datasets and software analyzed and generated during the study/experiments.

References

[1] A. Novianto dan M. D. Anasanti, "Autism Spectrum Disorder (ASD) Identification Using Feature-Based Machine Learning Classification Model," vol. 17, no. 3, hal. 259–270, 2023.

[2] R. K. Reghunathan, P. Nanjagoundan, dan P. Venkidusamy, "Machine Learning-Based Classification of Autism Spectrum Disorder across Age Groups †," 2024.

[3] F. M. Talaat, Z. H. Ali, R. R. Mostafa, dan N. El-rashidy, "Real-time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism children," *Soft Comput.*, vol. 28, no. 9, hal. 6695–6708, 2024, doi: 10.1007/s00500-023-09477-y.

[4] S. Srividhya dan D. Vidhya, "Detecting Autism Spectrum Disorder Using Machine Learning Techniques," vol. 12, no. 03.

[5] H. Li, Y. Gu, J. Han, Y. Sun, H. Lei, dan C. Li, "Faster R-CNN-MobileNetV3 Based Micro Expression Detection for Autism Spectrum Disorder," hal. 1–13, 2025.

[6] X. Wang, "Deep Learning Architectures for Alzheimer ' s Diagnosis," 2025.

[7] N. Zuzzaifa dan R. Rianto, "Convolutional Neural Network Untuk Perbandingan Optimizer Pada Citra Batang Pohon," *J. Sist. Cerdas*, vol. 6, no. 3, hal. 179–188, 2023, doi: 10.37396/jsc.v6i3.268.

[8] N. P. Novani, D. R. Salsabila, R. Aisuwarya, L. Arief, dan N. Afriyeni, "Sistem Pendeteksi Gejala Awal Tantrum Pada Anak Autisme Melalui Ekspresi Wajah Dengan Convolutional Neural Network," *JITCE (Journal Inf. Technol. Comput. Eng.*, vol. 5, no. 02, hal. 93–106, 2021, doi: 10.25077/jitce.5.02.93-106.2021.

[9] K. R. R. Wardani, H. Suryalim, V. J. L. Engel, dan H. Christian, "Analisis Pemilihan Optimizer dalam Arsitektur Convolution Neural Network VGG16 dan Inception untuk Sistem Pengenalan Wajah," *J. Edukasi dan Penelit. Inform.*, vol. 9, no. 2, hal. 186, 2023, doi: 10.26418/jp.v9i2.60432.

[10] Y. C. Oktaviani dan Y. Wahyuningsih, "Face Expression Recognizer Dengan Convolutional Neural Network Untuk Membantu Penderita Autisme Mengenali Ekspresi Wajah Seseorang," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3, 2023, doi: 10.23960/jitet.v11i3.3108.

[11] M. A. Hossain, S. Sakib, H. M. Abdullah, dan S. E. Arman, "Deep learning for mango leaf disease identification: A vision transformer perspective," *Heliyon*, vol. 10, no. 17, hal. e36361, 2024, doi: 10.1016/j.heliyon.2024.e36361.

[12] A. Howard *dkk.*, "Searching for mobileNetV3," *Proc. IEEE Int. Conf. Comput. Vis.*, hal. 1314–1324, 2019, doi: 10.1109/ICCV.2019.00140.

[13] S. F. D. Wardhana dan A. Nugroho, "Perbandingan Arsitektur MobileNetV2 dan MobileNetV3 Dalam Klasifikasi Jenis Jeruk," *J. Ilmu Komput. dan Bisnis*, vol. 16, no. 1, hal. 25–34, 2025, doi: 10.47927/jikb.v16i1.916.

[14] R. C. B. Rego, V. M. L. Silva, dan V. M. Fernandes, "Predicting Gender by First Name Using Character-level Machine Learning," 2021.

[15] G. Litjens *dkk.*, "A Survey on Deep Learning in Medical Imaage Analysis," no. 1995, 1998.

[16] P. Wahyu Setiyo Aji, R. Dijaya, dan F. Sains dan Teknologi, "KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen) Prediksi Penyakit Stroke Menggunakan Metode Random

Forest," *J. Penerapan Sist. Inf.*, vol. 4, no. 4, hal. 916–924, 2023.

[17] C. Shorten dan T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, 2019, doi: 10.1186/s40537-019-0197-0.

[18] R. R. Chevuri, "The Role of GPUs in Accelerating Machine Learning Workloads," vol. 11, no. 2, hal. 2676–2684, 2025.

[19] S. Alden dan B. N. Sari, "Implementasi Algoritma CNN Untuk Pemilahan Jenis Sampah Berbasis Android Dengan Metode CRISP-DM," *J. Inform.*, vol. 10, no. 1, hal. 62–71, 2023, doi: 10.31294/inf.v10i1.14985.

[20] R. D. Wicaksono dan G. F. Shidiq, "Comparison of KNN and CNN Algorithms for Gender Classification Based on Eye Images," vol. 11, no. 4, hal. 915–924, 2024, doi: 10.15294/sji.v11i4.13529.

[21] D. Mazumder dan T. Islam, "Franklin Open Evaluating the efficacy and site-specific performance of machine learning approaches : A comprehensive review of autism detection models," *Franklin Open*, vol. 11, no. April, hal. 100275, 2025, doi: 10.1016/j.fraope.2025.100275.

[22] E. Prihatini *dkk.*, "Facial Expression Detection of Autism Children Using ResNet-50 in Convolutional Neural Network Algorithm," vol. 8, no. 3, hal. 765–775, 2025.

[23] Kadek Ninda Nandita Putri, "Penerapan Fine-Tuning Resnet50 Dan Support Vector Machine (Svm) Untuk Klasifikasi Autism Specturm Disorder (Asd) Berbasis Citra Wajah," *2025*, hal. 5–6, 2025.