Jurnal Teknokes
Homepage: teknokes.org; Vol. 19, No. 3, pp. 123-138, September 2026;
e-ISSN: 2407-8964
p-ISSN: 1907-7904

RESEARCH PAPER

OPEN ACCES

Improved Autism Spectrum Disorder Detection Using Euclidean-Distance Facial Landmark Features to Enhance Classification Accuracy and Stability with Cross-Validation

Syifa Anzella[1], Yudha Nurdin[1], Melinda Melinda[1], Junidar Junidar [2] Rosminazuin Ab Rahim[3]

[1] Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia
[2] Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh, Indonesia
[3] Department of Electrical & Computer Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Malaysia

ABSTRACT

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that affects communication skills, social interaction, and behavioral patterns. Early detection is essential for timely intervention; however, conventional diagnostic methods remain time-consuming and subjective, as they rely heavily on clinical observations and expert judgment. This limitation highlights the need for an automated and objective approach to support early ASD screening. This study aims to analyze the performance, stability, and generalization of ASD classification using geometric features extracted from distances between facial landmarks. By representing facial morphology in terms of quantitative spatial relationships, this approach provides a more interpretable alternative to raw image-based methods. This study contributes by proposing a geometric feature representation based on facial landmark distances, providing a comparative analysis between linear and nonlinear classifiers, and ensuring robust evaluation through cross-validation. The dataset consists of 2,032 facial images, evenly distributed between children with ASD and those with typical development. A total of 68 facial landmark points were detected and used to compute pairwise Euclidean distances as classification features. Two classification algorithms, Logistic Regression and Extra Trees Classifier, were evaluated using 5-fold cross-validation to ensure reliable and unbiased performance estimation. The results show that Logistic Regression achieved an average accuracy of 89.91%, precision of 91.04%, recall of 88.56%, and F1-score of 89.76%. Meanwhile, the Extra Trees Classifier outperformed the linear model, achieving an average accuracy of 91.88%, precision of 92.69%, recall of 90.89%, and F1-score of 91.77%. Overall, both models demonstrated stable and consistent performance across validation folds, with the Extra Trees Classifier showing superior ability to capture nonlinear patterns in the data. These findings indicate that geometric feature extraction based on facial landmark distances is effective for ASD detection and has strong potential to be developed as an objective, interpretable, and efficient early screening tool using children's facial images.

PAPER HISTORY

Received July 18, 2026
Revised August 15, 2026
Accepted September 5, 2026
Published September 29, 2026

KEYWORDS

Autism Spectrum Disorder;
Facial Landmark;
Euclidean Distance;
Logistic Regression;
Extra Trees Classifier;
Cross-validation

CONTACT:

syifa.an@mhs.usk.ac.id
yudha.nurdin@usk.ac.id
melinda@usk.ac.id
junidar678@mhs.usk.ac.id
rosmi@iium.edu.my

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects an individual's communication skills, social interactions, and behavioral patterns [1]. The prevalence of ASD is significant, with an estimated 1 in 160 children affected globally according to the World Health Organization (WHO). Early detection of autism is crucial because it allows parents, medical personnel, and educators to provide appropriate interventions and therapies, enabling optimal development for children [2]. Ideally, ASD can be reliably identified as early as 18–24 months; however, in practice, the average age of diagnosis still ranges between 3 and 5 years [3]. This diagnostic delay significantly impacts developmental outcomes. Research shows that children with ASD who receive medical treatment before the age of four tend to have higher IQs than children diagnosed later [4]. Furthermore, delays in diagnosis contribute to increased socio-economic burdens, including long-term therapy costs, special education needs, and reduced productivity of caregivers [5], [6]. Despite its importance, current autism diagnosis still relies heavily on clinical observation and interviews, which are time-consuming, dependent on the expertise of medical personnel, and

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

potentially subjective in the assessment process [7]. Therefore, a technology-based approach is needed to facilitate faster, more objective, consistent, and measurable detection. Along with technological advances, the fields of artificial intelligence (AI) and computer vision are increasingly being used to assist in ASD detection. One promising approach is facial image analysis, as several studies have found differences in facial morphology between individuals with ASD and those without ASD [8]. Some individuals with ASD have been reported to have a wider face shape, a different interocular distance, or a different position of the nose and mouth than in normal individuals [9]. These characteristics can be identified using facial landmark points, specific points that represent the main anatomical structures of the face, such as the eyes, nose, and mouth [10], [11].

Several previous studies have investigated ASD classification directly from facial images without incorporating an explicit feature extraction stage. For example, studies in [12] and [13] utilize convolutional neural networks (CNN) to process raw RGB facial images, enabling the model to automatically learn discriminative visual features from large-scale datasets. These approaches are effective in capturing complex patterns and variations in facial appearance associated with ASD. Similarly, the study in [14] employs more advanced architectures, such as vision transformers, which can model global relationships in facial images and further improve classification performance.

In addition, the work presented in [15] conducts a comparative analysis of various machine learning and deep learning algorithms applied to facial image data, demonstrating that data-driven approaches can achieve competitive results across different model types. A comprehensive review in [16] also highlights that deep learning-based methods currently dominate ASD detection research, particularly in studies involving facial image analysis. Furthermore, another study in [17] confirms that direct classification using facial images, without explicit feature engineering, remains widely adopted for its simple pipeline design and strong predictive performance. Despite these advantages, deep learning-based approaches exhibit several important limitations. First, they typically require substantial computational resources and large annotated datasets to achieve optimal performance [18]. Second, these models operate as “black-box” systems, making it difficult to interpret the decision-making process and identify which specific facial regions or structures contribute to the classification results [19]. This lack of interpretability limits their applicability in clinical settings, where transparency and explainability are essential.

On the other hand, feature-based approaches that utilize facial landmarks provide a more structured and interpretable alternative. By representing the face using a set of key anatomical points, these methods enable the extraction of geometric features, such as distances and

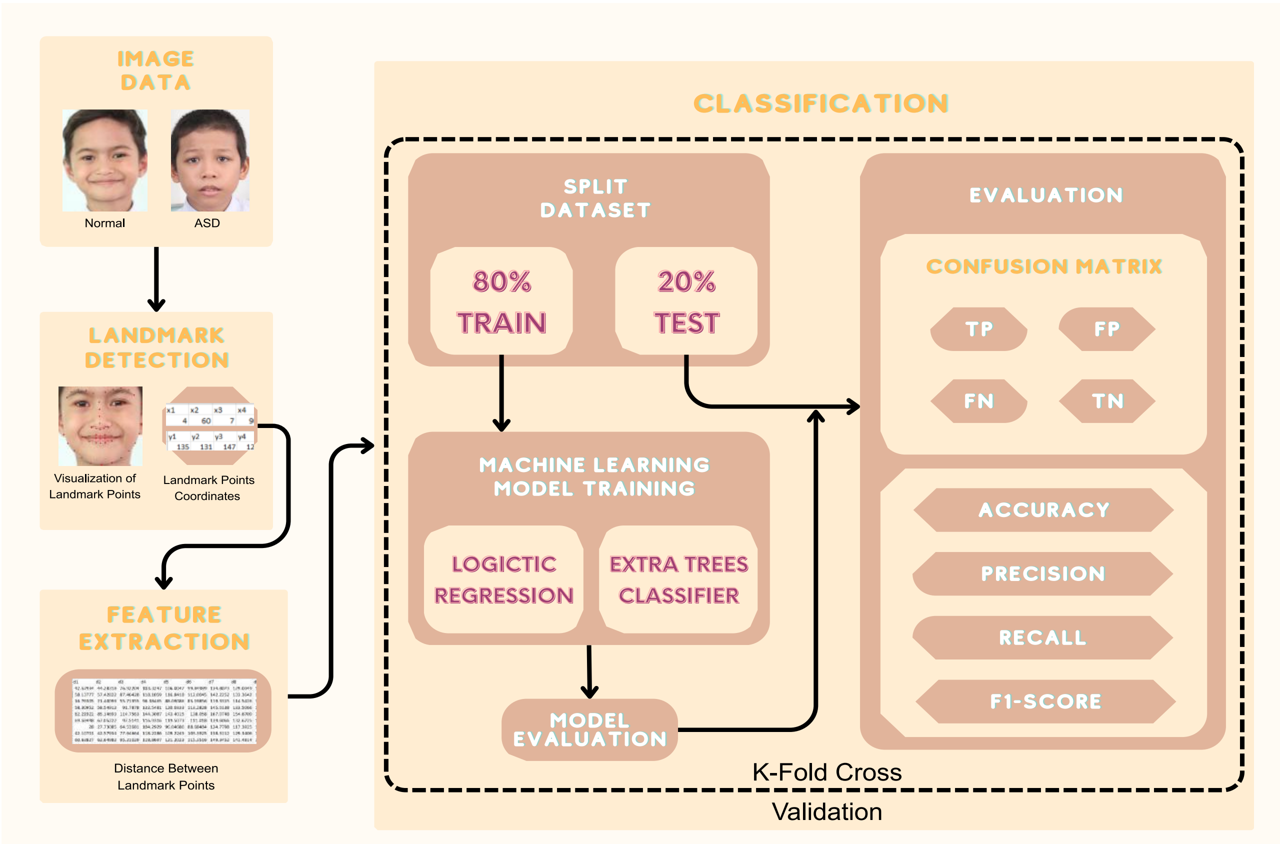

Fig. 1. Stages of research methodology for ASD classification based on child facial image analysis

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

spatial relationships among facial components. Such representations explicitly capture the face's proportional and morphological characteristics, which are particularly relevant to ASD analysis [20]. In addition, this approach reduces data dimensionality relative to raw image inputs and enables quantitative analysis of each feature, thereby enhancing model interpretability and supporting clinical insights. However, landmark-based approaches also have certain limitations. Since the extracted features are typically predefined and relatively simple, these methods may not fully capture complex nonlinear patterns present in facial data. As a result, their performance may be lower than that of deep learning models in highly complex classification tasks [21]. Therefore, there is a fundamental trade-off between interpretability and model complexity in ASD facial image analysis: highly accurate models often lack transparency, while more interpretable models may sacrifice some predictive performance.

Although previous studies have explored both image-based and landmark-based approaches, several limitations remain. First, many studies that use raw image classification do not explicitly capture geometric relationships among facial structures, which are essential for representing ASD-related morphological characteristics [22]. Second, existing landmark-based studies often focus primarily on feature extraction and do not provide a comprehensive comparison of different classification models, particularly between linear and nonlinear approaches [20]. Third, the stability and generalizability of classification performance are often not rigorously evaluated using robust validation techniques such as cross-validation, leading to potentially biased performance estimates. To address these gaps, this study applies a feature-extraction method that computes Euclidean distances between landmark points to represent facial proportions and structures numerically [20]. This representation converts two-dimensional coordinates into feature vectors that reflect the spatial relationships between facial points [23]. This approach not only reduces the dimensionality of the data compared to raw images but also increases the focus of the analysis on relevant geometric structures [24].

In the classification process, selecting the right algorithm is a crucial factor in achieving optimal analysis results [25]. Logistic Regression is a well-established linear classification technique that models the probability of a sample belonging to a particular class using the logistic (sigmoid) function, which maps real-valued inputs to a probability between 0 and 1. This approach assumes a linear decision boundary between classes, making it interpretable and well-suited for datasets with linearly separable features [26]. On the other hand, Extra Trees Classifier is an ensemble method that builds a large number of randomized decision trees and aggregates their predictions. It is particularly effective at capturing complex, nonlinear relationships within the data. Extra Trees work by selecting random splits at each tree node, which helps reduce variance and improve accuracy, especially in high-dimensional feature spaces [27]. By combining these two approaches, this study provides a balanced comparison between interpretable linear models and more flexible nonlinear ensemble methods.

To ensure that the evaluation results are stable and not influenced by specific data divisions, this study uses cross-validation. Through this approach, the dataset is divided into several folds, so that each part of the data is alternately used for training and testing. This strategy optimally utilizes all data in the evaluation process and provides a more representative estimate of the model's generalization to new data [28]. This study aims to analyze the performance of ASD facial image classification using a landmark-based geometric feature extraction approach and compare the performance of the Logistic Regression and Extra Trees Classifier algorithms with cross-validation-based evaluation. This study also emphasizes the importance of developing models that are not only accurate but also stable and adequately interpretable within an artificial intelligence-based diagnostic support system.

The main contributions of this research can be summarized into three aspects. First, this study proposes a geometric feature extraction method based on the Euclidean distance between facial landmarks to explicitly capture facial morphological characteristics associated with Autism Spectrum Disorder (ASD). This approach provides a compact and interpretable representation of facial structures that may serve as discriminative biomarkers for ASD classification. Second, this research presents a comprehensive comparative analysis of the performance of both linear and nonlinear ensemble machine learning algorithms using the extracted geometric facial features. The comparison aims to identify the most suitable classification model for distinguishing ASD and non-ASD subjects while highlighting the strengths and limitations of each algorithm. Third, a cross-validation evaluation framework is implemented to assess the robustness, stability, and generalizability of the proposed classification models. This evaluation strategy minimizes the risk of overfitting and provides a more reliable estimate of model performance when applied to unseen data..

The structure of this article is as follows. Section II describes the dataset and preprocessing steps, including landmark detection and feature extraction. Section III presents the results of the classification experiments and the performance evaluation of the proposed models. Section IV discusses the obtained findings and their relevance to ASD detection using facial image features. Section V concludes the paper and offers recommendations for future work.

II. MATERIALS AND METHOD

Fig. 1 shows the stages of the research methodology used to classify Autism Spectrum Disorder (ASD) in children from facial images. The research process begins with child facial images, divided into two categories: children with ASD and those with typical facial features. Each image then goes through a facial landmark-point

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

Jurnal Teknokes
Homepage: teknokes.org; **Vol. 19, No. 3, pp. 123-138, September 2026;**
e-ISSN: 2407-8964
p-ISSN: 1907-7904

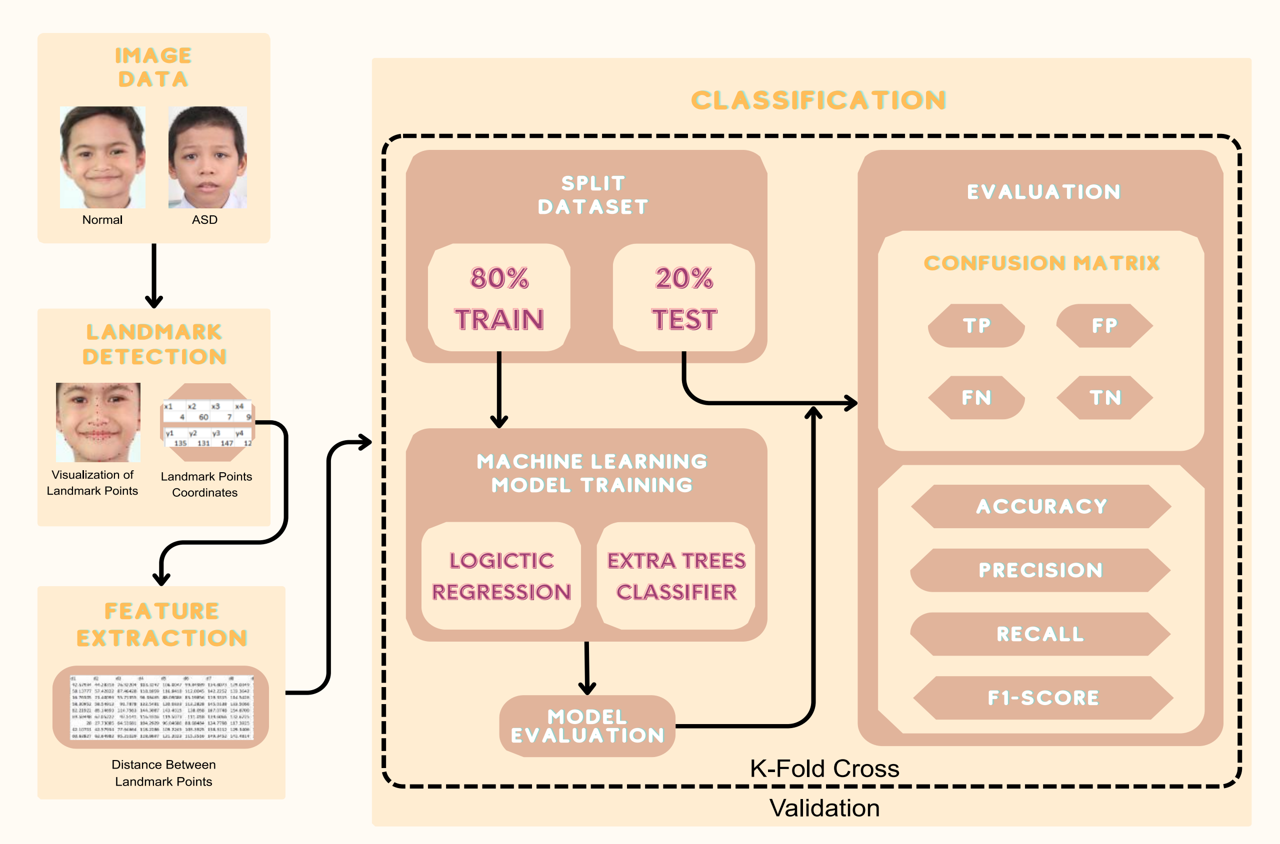

Fig. 1. Stages of research methodology for ASD classification based on child facial image analysis

marking stage to identify key facial points that represent the face's geometric structure. This process produces two-dimensional coordinates for each landmark point. The obtained landmark point coordinates are then used in the feature extraction stage. At this stage, the distance between landmark points is computed using the Euclidean distance formula to obtain a numerical representation that can serve as a feature for classification. The feature extraction results are compiled into a dataset, which is then split 80% for training and 20% for testing.

The training data is used to train a classification model using two machine learning algorithms: Logistic Regression and Extra Trees Classifier. After the model training process is complete, the next stage is model evaluation to assess the model's ability to learn patterns from the training data. Next, the model's performance was evaluated on data that had never been used in training. The model's predictions were evaluated using a confusion matrix and several metrics, including accuracy, precision, recall, and F1 Score. Furthermore, this study used k-fold cross-validation to assess model stability and consistency by repeatedly training and testing on different subsets of the dataset. Thus, the resulting evaluation can provide a more objective picture of the developed classification's performance.

A. Dataset

This study used a primary dataset of children's facial images collected directly from participants. The dataset comprised facial images from 60 children, including 30 diagnosed with ASD and 30 with typical development, aged 5 to 15 years. Data were collected from Special Education Schools (SLB) and public elementary schools in Banda Aceh. Each participant contributed approximately 20 facial images taken with different facial expressions, including neutral, smiling, sad, and angry, to increase variability and represent natural facial conditions. Initially, a total of 1,016 facial images were obtained, consisting of 508 images of children with ASD and 508 images of children with typical development. To increase dataset variability and improve model generalization, data augmentation was performed using a vertical flip technique, resulting in 1,016 images for each class, for a total of 2,032 images. This balanced composition was intended to minimize classification bias and ensure that the model could learn the characteristics of both classes equally.

Before the feature extraction and classification stages, the dataset was carefully selected and prepared. The image selection process was based on visual quality and feasibility for geometric facial analysis. Each image required a frontal face orientation, a relatively upright head position, and adequate lighting conditions to ensure clear detection of key facial anatomical structures, such as the eyes, nose, and mouth. Images with low resolution, excessive blur, or occlusion from external objects were excluded from the analysis to maintain consistency and improve the accuracy of facial landmark detection.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

The data collection process was conducted in a controlled environment following standard experimental protocols. The acquisition room measured 7 m × 7 m, with a specific imaging area of 2 m × 2 m. Environmental conditions were maintained consistently, with temperatures ranging from 25 to 27°C, humidity between 40% and 45%, and lighting provided by approximately 30-watt ceiling lights. The shooting sessions took place between 9:00 AM and 12:00 PM, and each participant underwent a 10- to 15-minute acclimatization period to reduce potential sensory discomfort and anxiety. The technical setup involved a Canon EOS M100 mounted on a tripod approximately 1 meter from the subject. A 17-inch LCD screen was used to display visual stimuli to elicit natural facial expressions, while a blackout curtain was installed to minimize interference from external lighting. This controlled setting was designed to reduce environmental noise and ensure consistent image quality, thus supporting reliable feature extraction. Finally, the collection and use of the dataset in this study were approved by the ethics committee under reference number 036/EA/FK/2025, which ensures compliance with ethical and legal standards governing the use of children's facial image data.

B. Facial Landmark Detection

Facial landmark detection is the process of identifying key facial points that correspond to anatomical structures such as the eyes, nose, lips, and jaw [29]. These points describe the shape and geometric proportions of the face numerically, enabling computational analysis [30]. This study used a 68-point facial landmark model commonly used in the Dlib algorithm. Each point has two-dimensional coordinates (x, y) indicating its position relative to the entire face. This structure can represent the shape, proportions, and variations in facial expressions in detail [31], [32]. At this stage, all images undergo face detection to ensure the analysis focuses solely on the facial region. A face detector algorithm is used to automatically locate the face in each image, thereby minimizing the influence of background noise. After face detection, the detected facial region is cropped and resized to 224 × 224 pixels to ensure uniform input dimensions across all samples.

Prior to landmark extraction, additional preprocessing steps are applied to improve data quality and consistency. Each image is converted into grayscale format to reduce computational complexity while preserving essential structural information. Furthermore, Intensity normalization was performed using min-max normalization, where pixel values were scaled from the original range of 0–255 to 0–1. This transformation ensures consistency across images and reduces the impact of lighting variations [33]. To further enhance image quality, a Gaussian filter is applied to reduce noise and minor variations that may interfere with accurate landmark detection [34]. These preprocessing steps are essential to ensure that the extracted landmark points are stable and consistent across different images. Following preprocessing, the next stage involves detecting and marking 68 facial landmark points on each image. These landmarks correspond to key anatomical regions, including the eyes, eyebrows, nose, mouth, and jawline, and are represented as (x, y) coordinate pairs. The resulting landmark annotations provide a structured representation of facial geometry that serves as the basis for subsequent feature extraction. The visualization of landmark points is stored in .jpg format for documentation purposes, while the coordinate data are saved in .csv

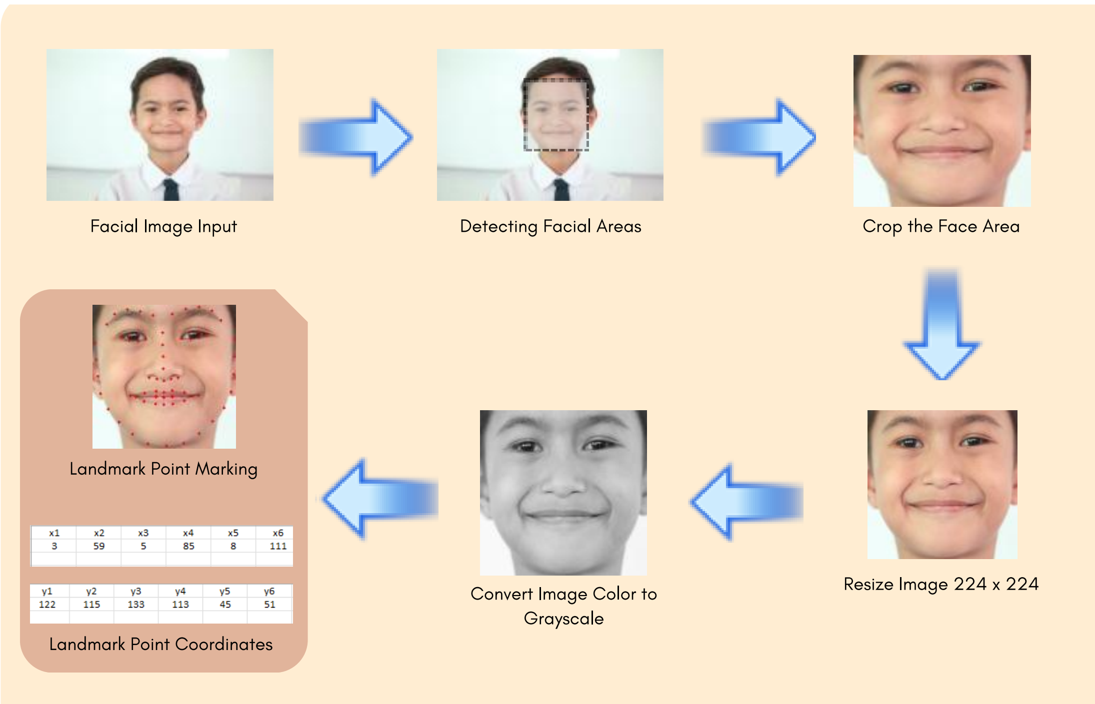

Fig. 2. Stages of the Detection and Marking Process of 68 Facial Landmark Points

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

format for further analysis and classification. Fig. 2 illustrates the overall process of facial landmark detection, including the distribution of the 68 landmark points across the facial region.

C. Feature Extraction

Feature extraction is a crucial stage in image processing that transforms visual data into a structured numerical representation while preserving essential information [35]. In this study, feature extraction was

performed using a geometric approach based on the spatial relationships among facial landmark points obtained from a 68-point landmark model [36]. This approach relies on the assumption that facial morphology can be represented through geometric distances between key anatomical points.

The extraction process is based on the Euclidean distance, which measures the straight-line distance between two points in a two-dimensional space. This concept is derived from the Pythagorean theorem, in which the distance is the hypotenuse of a right triangle formed by the horizontal and vertical differences between two points. The Euclidean distance is defined in Eq. (1) [36], as follows:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

Where (x_1, y_1) and (x_2, y_2) denote the coordinates of two facial landmark points. In the context of facial image analysis, this distance quantifies the geometric relationships among facial components, such as inter-eye distance, eye-to-nose distance, and nose-to-mouth distance [24]. These measurements capture structural characteristics of the face that are relevant for distinguishing between ASD and typically developing children.

All computed distances are then arranged into a feature vector representing the overall facial geometry. This geometric representation is relatively invariant to illumination, skin color, and texture variations, making it suitable for robust facial analysis [37]. Additionally, Euclidean distance has low computational complexity, allowing efficient processing of large datasets [36]. The resulting numerical representation is then used as input to the classification stage to quantitatively identify patterns of facial structural differences between children with ASD and typically developing children.

D. Logistic Regression

Logistic Regression is a probabilistic classification method that models the relationship between input features and a binary output using a logistic (sigmoid) function [38]. Unlike linear regression, which predicts continuous values, Logistic Regression estimates the probability that a given input belongs to a particular class [39]. The mathematical formulation is given in Eq. (2):

$$P(y = 1 \mid x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \quad (2)$$

This formulation is derived from the logit (log-odds) function, which maps a linear combination of input features to a probability between 0 and 1. The classification decision is determined by applying a threshold, typically 0.5.

In this study, the input features consist of geometric distances between facial landmark points that represent the face's structural proportions. Logistic Regression models the relationship between these geometric features and class labels by constructing a linear decision boundary in the feature space. Each coefficient β_1 reflects the contribution of a specific facial distance feature to the classification outcome. This allows the model to identify which facial proportions, such as inter-eye distance or the relative position between facial components, are most influential in distinguishing between children with ASD and typically developing children.

The model training process involves estimating parameters that maximize the likelihood or minimize the loss function, generally using optimization methods such as Gradient Descent [39]. The advantages of Logistic Regression lie in its ease of implementation, high model interpretability, and good performance on data with linear relationships between features [40]. However, this algorithm has limitations in modeling complex non-linear relationships, so its performance can decrease on facial image datasets with non-linear geometric patterns [41].

E. Extra Trees Classifier

The Extra Trees Classifier (Extremely Randomized Trees) is an ensemble learning algorithm that constructs multiple decision trees and combines their predictions to improve classification performance [42]. This method is based on the principle of variance reduction through randomization, where multiple weak learners are aggregated to produce a more robust and stable model.Each decision tree is trained on the entire dataset, with randomness introduced in the selection of feature subsets and split thresholds. Unlike traditional decision trees or Random Forest, which determine optimal split points based on impurity measures, the Extra Trees Classifier selects split points randomly. This strategy increases model diversity and reduces the risk of overfitting [43]. The final prediction is obtained through majority voting [44], as expressed in Eq. (3):

$$\hat{y} = mode(h_1(x), h_2(x), \ldots h_n(x)) \quad (3)$$

Where $h_i(x)$represents the prediction of the i-th decision tree. In this study, the input features are derived from Euclidean distances between facial landmarks, which encode the geometric structure of the face. In the context of facial image analysis, relationships between facial features are often complex and non-linear due to variations in facial proportions, expressions, and individual anatomical differences. The Extra Trees Classifier is particularly well-suited for this scenario, as it can model non-linear interactions between features and capture subtle variations in facial morphology. This enables the model to better distinguish patterns associated with ASD compared to linear approaches [27].

F. K-Fold Cross-validation

K-Fold Cross-Validation is an evaluation method that measures model performance more objectively and

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

Jurnal Teknokes
Homepage: teknokes.org; Vol. 19, No. 3, pp. 123-138, September 2026;
e-ISSN: 2407-8964
p-ISSN: 1907-7904

consistently, especially on datasets with a limited sample size [45]. This method aims to reduce bias from unrepresentative training and test data allocation and minimize the risk of overfitting to a particular data allocation scenario [46].

In K-Fold Cross Validation, the dataset is divided into k subsets (folds) of relatively equal size. The training and testing processes are performed in k iteration [47]. At each iteration to-i, one fold is used as the test data (D_i), while the other $k - 1$ folds folds are used as training data. After all iterations are complete, the model performance value is calculated as the average of all test results [48], [49], [50]. Mathematically, this average value is expressed in Eq. (4) [51], as follows:

$$\mathrm{CVscore} = \frac{1}{k}\sum_{i=1}^{k} S_i \tag{4}$$

Where S_i is the evaluation metric value (e.g., accuracy, precision, recall, or F1-score) at iteration to-i, dan k is the number of folds used. With this approach, each data point is alternately used for training and testing data, yielding a more comprehensive and representative evaluation of the dataset's distribution [51]. In this study, K-fold cross-validation was used to evaluate the performance of a classification model based on geometric facial features. This approach yields more reliable performance estimates and reflects the model's ability to generalize to data not used in training.

G. Evaluation Matrix

Classification model performance evaluation is a crucial step in research to assess the model's ability to predict on the test data. In this study, model performance was evaluated using several metrics commonly used in classification problems, namely accuracy, precision, recall, and F1 Score. These metrics are calculated from values in a confusion matrix, a table that compares the actual class labels with the model's predicted results [52].

The confusion matrix consists of four main components, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), as shown in Table 1**.** True Positive (TP) indicates the number of data correctly predicted as positive. True Negative (TN) indicates the number of data correctly predicted as negative. False Positive (FP) is the number of data points that are actually in the negative class but are predicted as positive by the model. Meanwhile, False Negative (FN) denotes the number of data points that are actually in the positive class but are predicted as negative [53].

Table 1. Confusion Matrix

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Based on the values in the confusion matrix, several evaluation metrics can be calculated to assess the model's performance more comprehensively. Accuracy is used to measure the proportion of correct predictions against all tested data [54]. The accuracy can be calculated as shown in Eq. (5).

$$\text{Accuracy} = \frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN})} \tag{5}$$

In addition, precision measures the model's accuracy in predicting a particular class, namely, how many positive predictions actually belong to the positive class [54]. The precision calculation is given by Eq. (6).

$$\text{Precision} = \frac{(\mathrm{TP})}{(\mathrm{TP}+\mathrm{FP})} \tag{6}$$

Next, recall measures the model's ability to detect all data that is actually included in the positive class. The recall value shows how much positive data the model successfully identified [54]. The recall equation is shown in Eq. (7).

$$\text{Recall} = \frac{(\mathrm{TP})}{(\mathrm{TP}+\mathrm{FN})} \tag{7}$$

To balance precision and recall, the F1-score is used, the harmonic mean of precision and recall. The F1-score value provides a more balanced picture of model performance, especially in conditions of unbalanced data distribution [54]. The F1-Score is given by Eq. (8).

$$\text{F1}-\text{Score} = 2 \times \frac{(\mathrm{Precision} \times \mathrm{Recall})}{(\mathrm{Precision} + \mathrm{Recall})} \tag{8}$$

The use of these multiple evaluation metrics aims to provide a more comprehensive assessment of the classification model's performance. By considering various evaluation metrics, the model's performance in classifying the data in this study can be analyzed more objectively.

III. RESULTS

The overall research workflow is presented in Fig. 1, which shows the sequential process of ASD classification, starting with data acquisition, preprocessing, facial landmark detection, feature extraction, classification, and performance evaluation. The results of facial landmark annotation are shown in Fig. 2, where 68 predefined landmark points are successfully detected on each facial image, representing key facial regions such as the eyes, nose, mouth, and jawline. These landmarks form the basis for subsequent feature extraction.

A. Feature Extraction Results: Distances Between Landmark Points

Feature extraction is performed after 68 landmark points are marked on each facial image. Each landmark point is represented by a two-dimensional coordinate that describes its geometric position on the image plane. Based on these coordinates, the Euclidean distance between pairs of landmark points is calculated, as stated in Eq. (1), to obtain a numerical representation of facial structure. Through this process, each facial image generates a set of distance values that represent the spatial relationships among facial features, such as the distances between the eyes, the eyes and the nose, and the nose and the mouth. These distance values are then arranged into a feature vector that comprehensively

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

describes the face's geometric characteristics.

In this study, the feature extraction process generated 67 distance features from each facial image. From a dataset of 2,032 images, a 2,032 × 67 feature matrix was obtained and used as input to the classification stage. This numerical representation allows for quantitative analysis of facial morphological patterns and is relatively independent of variations in lighting, skin color, or facial surface texture.

B. Model Evaluation Results Using 5-Fold Cross Validation

The performance evaluation of the Logistic Regression and Extra Trees Classifier models was conducted using 5-fold cross-validation. In this approach, the dataset was divided into five subsets, with each subset used once as test data and the remaining subsets used for training. This process ensures that all data is evaluated, providing a more robust estimate of model performance. To provide a concise yet comprehensive evaluation, the performance results are summarized as the mean and standard deviation (mean ± SD) for each metric across all folds, as presented in Table 2.

Table 2. Comparison of Average Values of K-Fold Cross Validation

Metric	Logistic Regression	Extra Trees Classifier
Accuracy (%)	89.91 ± 1	91.88 ± 2
Precision (%)	91.04 ± 1	92.69 ± 1
Recall (%)	88.56 ± 3	90.89 ± 4
F1-Score (%)	89.76 ± 1	91.77 ± 2

Based on the results shown in Table 2, the Logistic Regression model produced an average accuracy of 89.91 ± 1, precision of 91.04 ± 1, recall of 88.56 ± 3, and an F1-score of 89.76 ± 1. Meanwhile, the Extra Trees Classifier model achieved an average accuracy of 91.88 ± 2, precision of 92.69 ± 1, recall of 90.89 ± 4, and F1 Score of 91.77 ± 2. The standard deviation (±) value represents the level of variation or dispersion of the evaluation results obtained at each fold during the K-Fold Cross Validation process. A relatively small standard deviation indicates that model performance tends to be consistent across data splits, while a larger standard deviation indicates greater performance variation between folds.

C. Confusion Matrix Analysis

In the Logistic Regression model, the relatively small standard deviations across most metrics indicate consistent performance during validation. Meanwhile, the Extra Trees Classifier model also shows standard deviations that remain within a relatively low range, though slightly higher for some metrics, particularly recall. This indicates that both models exhibited good performance stability during the evaluation process using K-Fold Cross Validation, with variation in results remaining within acceptable limits across folds.

In the confusion matrix representation, the y-axis represents the true label, while the x-axis represents the predicted class label. The class label is defined as 1 for ASD and 0 for Non-ASD. Each value in the matrix represents the number of samples (n) for each classification result. Fig. 3 shows the confusion matrix for the Logistic Regression model. Based on the confusion matrix for Fold 1, the Logistic Regression model correctly classified 723 ASD samples as ASD (True Positive) and 743 Non-ASD samples as Non-ASD (True Negative). However, the model incorrectly classified 90 ASD samples as Non-ASD (False Negative) and 70 Non-ASD samples as ASD (False Positive).

In Fold 2, the model obtained 724 True Positive predictions and 747 True Negative predictions. Meanwhile, 89 ASD samples were incorrectly classified as Non-ASD (False Negative), and 66 Non-ASD samples were incorrectly classified as ASD (False Positive). In Fold 3, the model correctly identified 733 ASD samples (True Positive) and 750 Non-ASD samples (True Negative). However, there were 80 False-negative cases and 63 False-positive cases, indicating a relatively lower number of misclassifications compared to the previous fold. In Fold 4, the model generated 731 True Positive predictions and 742 True Negative predictions. The total number of misclassifications was 82 False-negative cases and 71 False-positive cases, indicating a balanced distribution. In Fold 5, the Logistic Regression model correctly classified 730 ASD samples as ASD (True Positive) and 741 Non-ASD samples as Non-ASD (True Negative). Meanwhile, 83 ASD samples were incorrectly classified as Non-ASD (False Negative), and 872 Non-ASD samples were incorrectly classified as ASD (False Positive).

Overall, across all folds, the number of True Positive and True Negative predictions consistently exceeds the number of misclassifications (False Positive and False Negative), indicating that the Logistic Regression model is generally effective at distinguishing between ASD and Non-ASD classes. However, variations in the numbers of False Positive and False Negative cases are observed across folds, suggesting that the model's performance is influenced by differences in data distributions. In particular, certain folds exhibit higher misclassification rates, reflecting the limitation of Logistic Regression in capturing complex patterns in facial geometric features. Despite these variations, the model demonstrates relatively stable performance with a dominant proportion of correct classifications across all folds.

Fig. 4 shows the confusion matrix for the logistic Extra Trees Classifier. Based on the confusion matrix for Fold 1, the Extra Trees Classifier correctly classified 799 ASD samples as ASD (True Positive) and 804 Non-ASD samples as Non-ASD (True Negative). Meanwhile, 14 ASD samples were incorrectly classified as Non-ASD (False Negative), and 9 Non-ASD samples were incorrectly classified as ASD (False Positive). In Fold 2, the model obtained 787 True Positive predictions and 798 True Negative predictions. Meanwhile, 26 ASD samples were misclassified as Non-ASD, and 15 Non-ASD

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

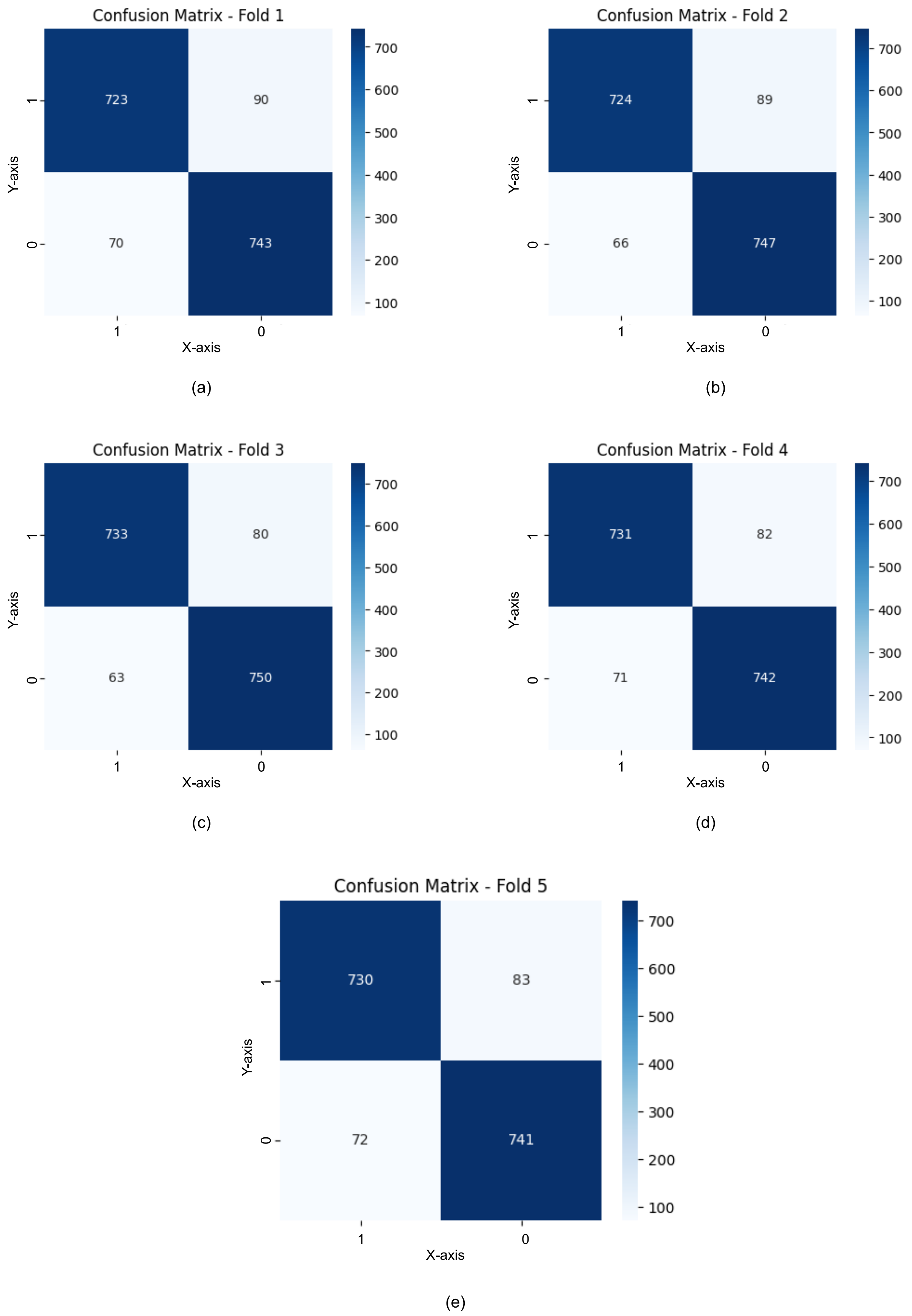

Fig. 3. Confusion Matrix for Each Fold in 5-Fold Cross-Validation Logistic Regression Model.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

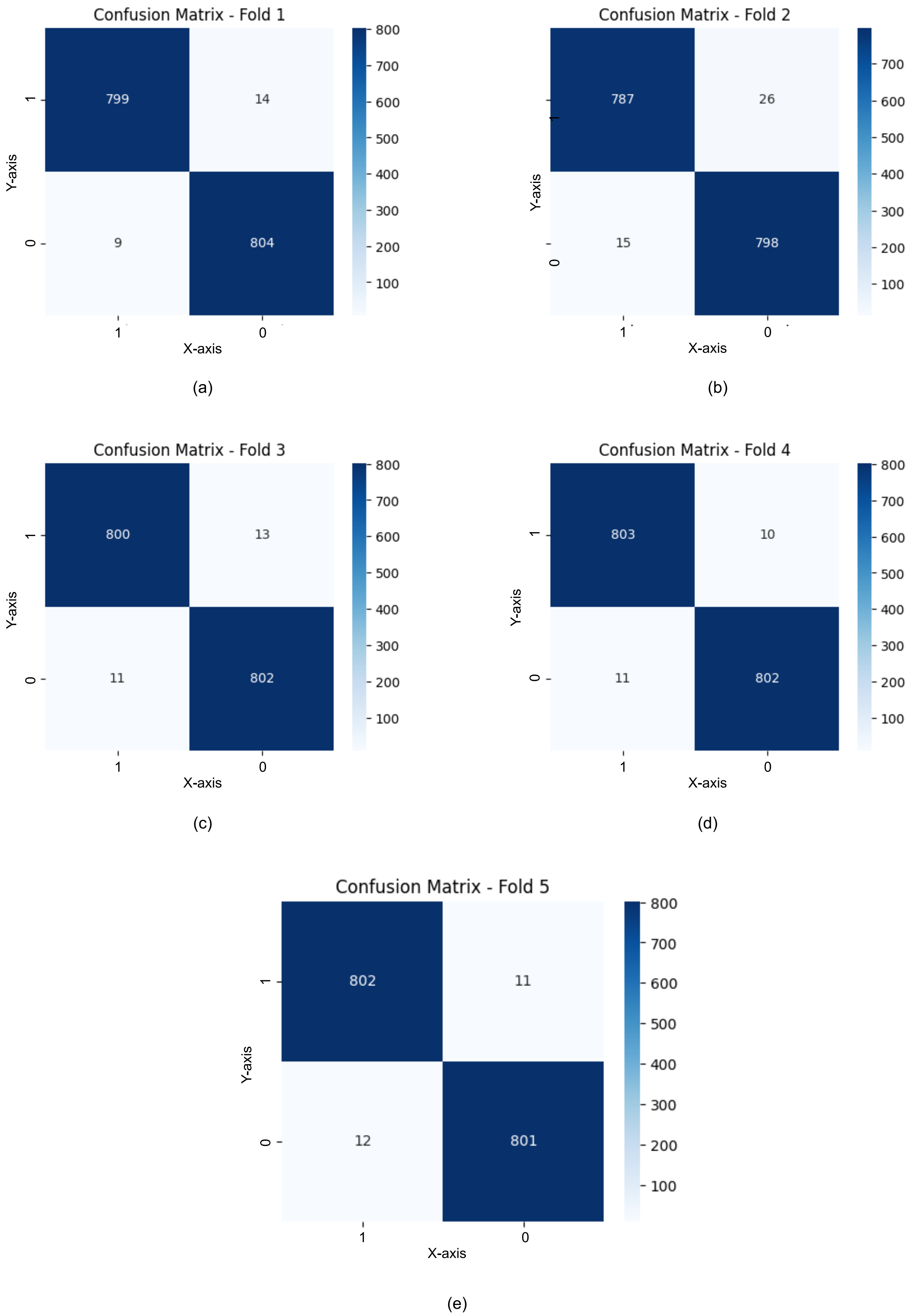

Fig. 4. Confusion Matrix for Each Fold in 5-Fold Cross-Validation Extra Trees Classifier Model.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

samples were incorrectly classified as ASD. In Fold 3, the model correctly identified 800 ASD samples (True Positive) and 802 Non-ASD samples (True Negative), with 13 False Negatives and 11 False Positives. In Fold 4, the model generated 803 True Positive predictions and 802 True Negative predictions, while the number of misclassifications consisted of 10 False Negative and 11 False Positive cases. In Fold 5, the Extra Trees Classifier correctly classified 802 ASD samples as ASD (True Positive) and 801 Non-ASD samples as Non-ASD (True Negative). Meanwhile, 11 ASD samples were misclassified as Non-ASD, and 12 Non-ASD samples were incorrectly classified as ASD.

Overall across all folds, the Extra Trees Classifier consistently produces a significantly higher number of correct classifications (True Positive and True Negative) than misclassifications (False Positive and False Negative). The number of False Negative and False Positive cases remains relatively low and stable across all folds, indicating that the model is highly effective in distinguishing between ASD and Non-ASD classes. This consistency reflects the model's strong generalization ability and robustness to variations in data distribution. Furthermore, the low misclassification rates suggest that the Extra Trees Classifier can capture complex non-linear relationships among facial landmark-based features, resulting in more accurate and stable classification performance than Logistic Regression.

IV. DISCUSSION

Based on the evaluation results, both models demonstrated a strong ability to classify children's facial images as ASD or non-ASD, with consistent performance across 5-fold cross-validation. However, a clear performance difference is observed between the two models. The Extra Trees Classifier achieved a higher accuracy (91.88%) than Logistic Regression (89.91%) and showed more stable results across folds. This difference can be explained by the nature of the extracted features. The Euclidean distance between facial landmarks reflects inherently nonlinear spatial relationships due to variations in facial morphology. Logistic Regression assumes linear relationships between features and class labels, which limits its ability to capture complex geometric patterns. In contrast, the Extra Trees Classifier can model nonlinear interactions through multiple randomized decision trees, allowing it to better capture combinations of facial distances and morphological variations associated with ASD. This explains the lower misclassification rates and more consistent performance observed in the confusion matrix results.

Compared with previous studies, the proposed approach demonstrates competitive performance despite using a simpler, more interpretable feature representation. Studies [12] reported 92% accuracy with ResNet50, while [13] achieved 87.9% with EfficientNetB0. Similarly, [15] achieved 88% with VGG16, and [14] reported 77% with a Vision Transformer. Higher accuracies were achieved in [16] (99% using MobileNet) and [17] (98% using VGG19), which rely on deep learning models applied directly to raw facial images. In comparison, the Extra Trees model in this study achieved an accuracy of 91.88%, indicating that geometric feature-based approaches can achieve comparable performance while offering better interpretability and lower computational complexity.

Despite these promising results, several limitations should be noted. The dataset used in this study was collected as primary data, but it may still have limitations in terms of subject diversity, including age range and imaging conditions. The use of only frontal facial images restricts the model's ability to generalize to variations in head pose or orientation. In addition, the method's accuracy depends on the reliability of facial landmark detection, which can be affected by image quality, lighting conditions, and occlusions. The dataset size is also relatively limited compared to large-scale datasets used in deep learning studies, and no external validation using an independent dataset was performed. These findings indicate that geometric facial features derived from landmark distances can effectively capture structural characteristics relevant to ASD classification. The proposed approach has practical potential for early screening applications, as it can be implemented in lightweight systems such as mobile or web-based tools. Such systems can assist healthcare practitioners, educators, and parents in identifying early signs of ASD and supporting further clinical assessment.

V. CONCLUSION

This study aims to evaluate the effectiveness of geometric features derived from facial landmark distances for Autism Spectrum Disorder (ASD) classification, compare the performance of linear and nonlinear classification models, and assess the stability and generalization capability of the models using cross-validation. The proposed method utilizes 68 facial landmark points, from which geometric features are generated by computing Euclidean distances between landmark pairs to numerically represent facial morphological characteristics.

The experimental results, obtained using 5-fold cross-validation, demonstrate that both models achieved strong and stable performance. The Logistic Regression model achieved an average accuracy of 89.91%, precision of 91.04%, recall of 88.56%, and F1-score of 89.76%. Meanwhile, the Extra Trees Classifier achieved higher performance, with an average accuracy of 91.88%, precision of 92.69%, recall of 90.89%, and F1-score of 91.77%. These results indicate that the nonlinear model is more effective in capturing complex relationships among facial features, while the linear model provides competitive performance with higher interpretability.

In conclusion, geometric features based on Euclidean distances between facial landmarks are effective for distinguishing ASD and non-ASD facial characteristics, achieving a classification performance of up to 91.88%

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

accuracy. In addition, cross-validation confirms that the proposed models yield stable, generalizable results. This approach also offers advantages in terms of interpretability and computational efficiency compared to deep learning-based methods, making it suitable for early screening applications.

For future research, several improvements can be considered. First, expanding the dataset with more diverse subjects in terms of age, ethnicity, and imaging conditions may improve model generalization. Second, evaluation on external or independent datasets is necessary to further validate the robustness of the model. Third, the proposed method can be developed into a real-time application, such as a mobile or web-based screening system. Fourth, integrating additional modalities, such as behavioral or physiological data, may enhance classification performance. Finally, further validation through clinical studies is required to assess the reliability and applicability of the method in real-world medical settings.

REFERENCES

[1] L. Qin, H. Wang, W. Ning, M. Cui, and Q. Wang, "New advances in the diagnosis and treatment of autism spectrum disorders," *European Journal of Medical Research*, pp. 1–11, 2024, doi: 10.1186/s40001-024-01916-2.

[2] J. Ostrowski, U. Religioni, B. Gellert, J. Sytnik-Czetwertyński, and J. Pinkas, "Autism Spectrum Disorders: Etiology, Epidemiology, and Challenges for Public Health," *Medical Science Monitor*, vol. 30, pp. 1–6, 2024, doi: 10.12659/MSM.944161.

[3] D. Control, "Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network , 11 Sites , United States , 2020," *Morbidity and Mortality Weekly Report*, vol. 72, no. 2, 2023, doi: 10.15585/mmwr.ss7202a1.

[4] S. P. Collins *et al.*, "A Study on Autism Detection from Facial Images," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 11, no. XI, pp. 167–186, 2023, doi: https://doi.org/10.22214/ijraset.2023.56908.

[5] H. M. Dubrock, H. D. Germack, M. Gauthier, L. Jörg, and L. Ambika, "Economic Burden of Delayed Diagnosis in Patients with Pulmonary Arterial Hypertension (PAH)," *PharmacoEconomics - Open*, vol. 8, no. 1, pp. 133–146, 2024, doi: 10.1007/s41669-023-00453-8.

[6] L. Berbel, A. María, A. Joan, C. Marta, and A. Annika, "Association Between Diagnostic Delay and Economic and Clinical Burden in Axial Spondyloarthritis : A Multicentre Retrospective Observational Study," *Rheumatology and Therapy*, vol. 12, no. 2, pp. 255–266, 2025, doi: 10.1007/s40744-024-00742-6.

[7] P. McCarty and R. E. Frye, "Early Detection and Diagnosis of Autism Spectrum Disorder: Why Is It So Difficult?," *Seminars in Pediatric Neurology*, vol. 35, 2020, doi: 10.1016/j.spen.2020.100831.

[8] M. A. Mahmood, L. Jamel, N. Alturki, and M. A. Tawfeek, "Leveraging artificial intelligence for diagnosis of children autism through facial expressions," *Scientific Reports*, vol. 15, no. 1, pp. 1–20, 2025, doi: 10.1038/s41598-025-96014-6.

[9] G. Quatrosi *et al.*, "Cranio-Facial Characteristics in Autism Spectrum Disorder: A Scoping Review," *Journal of Clinical Medicine*, vol. 13, no. 3, 2024, doi: 10.3390/jcm13030729.

[10] M. S. ALAM, M. M. RASHID, A. R. FAIZABADI, and H. F. M. ZAKI, "Power of Alignment: Exploring the Effect of Face Alignment on Asd Diagnosis Using Facial Images," *IIUM Engineering Journal*, vol. 25, no. 1, pp. 317–327, 2024, doi: 10.31436/iiumej.v25i1.2838.

[11] K. K. Mujeeb Rahman and M. Monica Subashini, "Identification of Autism in Children Using Static Facial Features and Deep Neural Networks," *Brain Sci*, vol. 12, 2022, doi: 10.3390/brainsci12010094.

[12] M. Faheem, I. Ahmad, J. Rashid, N. A. Khan, and R. Amin, "Autism spectrum disorder detection using facial images : A performance comparison of pretrained convolutional neural networks," *Healthcare Technology Letters published*, vol. 11, pp. 227–239, 2024, doi: 10.1049/htl2.12073.

[13] P. Reddy and J. Andrew, "Diagnosis of Autism in Children Using Deep Learning Techniques by Analyzing Facial Features †," vol. 59, no. 1, p. 198, 2024, doi: https://doi.org/10.3390/engproc2023059198.

[14] M. F. Almufareh, S. Tehsin, and M. Humayun, "Facial Classification for Autism Spectrum Disorder," vol. 3, pp. 1–8, 2024, doi: 10.57197/JDR-2024-0025.

[15] T. Farhat, S. Akram, H. S. Alsagri, Z. Ali, A. Ahmad, and A. Jaffar, "Facial Image-Based Autism Detection : A Comparative Study of Deep Neural Network Classifiers," *Computers, Materials & Continua*, vol. 78, no. 1, pp. 105–126, 2024, doi: 10.32604/cmc.2023.045022.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

[16] F. Mohammadi, “Artificial Intelligence in Autism Spectrum Disorder Diagnosis : A Scoping Review of Face , Voice , and Text Analysis Methods,” *Health Science Reports*, vol. 8, no. 11, 2025, doi: 10.1002/hsr2.71476.

[17] M. Melinda, M. Oktiana, Y. Nurdin, I. Pujiati, and M. Irhamsyah, “Performance of ShuffleNet and VGG-19 Architectural Classification Models for Face Recognition in Autistic Children,” *International Journal on Advance Science Engineering Information Technology*, vol. 13, no. 2, pp. 674–680, 2023.

[18] S. Yoonesi, R. A. Azar, M. A. Bafrani, S. Yaghmayee, and H. Shahavand, “Facial expression deep learning algorithms in the detection of neurological disorders : a systematic review and meta - analysis,” *BioMedical Engineering OnLine*, vol. 24, 2025, doi: 10.1186/s12938-025-01396-3.

[19] Q. Teng, Z. Liu, Y. Song, K. Han, and Y. Lu, “A survey on the interpretability of deep learning in medical diagnosis,” *Multimedia Systems*, vol. 28, pp. 2335–2355, 2022, doi: 10.1007/s00530-022-00960-4.

[20] B. Banire, D. Al Thani, M. Qaraqe, and B. Mansoor, “Face-Based Attention Recognition Model for Children with Autism Spectrum Disorder,” *Journal of Healthcare Informatics Research*, vol. 5, no. 4, pp. 420–445, 2021, doi: 10.1007/s41666-021-00101-y.

[21] B. Al-baker, X. Ju, P. Mossey, and A. Ayoub, “The accuracy of automated facial landmarking - a comparative study between Cliniface software and patch- based Convoluted Neural Network algorithm,” *European Journal of Orthodontics*, vol. 47, 2025, doi: 10.1093/ejo/cjaf009.

[22] K. Rezaee, T. S. Javanmardi, and A. M. Heydari, “A deep learning framework for classifying autism spectrum disorder from children ’ s facial images using a multi-scale ViT architecture and edge computing,” *Exploration of Medicine*, vol. 6, 2025, doi: 10.37349/emed.2025.1001376.

[23] G. Sanil, K. Prakash, S. Prabhu, V. C. Nayak, and S. Sengupta, “2D-3D Facial Image Analysis for Identification of Facial Features Using Machine Learning Algorithms With Hyper-Parameter Optimization for Forensics Applications,” *IEEE Access*, vol. 11, no. August, pp. 82521–82538, 2023, doi: 10.1109/ACCESS.2023.3298443.

[24] Y. An, J. Lee, E. Bak, and S. Pan, “Deep Facial Emotion Recognition Using Local Features Based on Facial Landmarks for Security System,” *Computers, Materianls and Continua*, vol. 76, p. 2, 2023, doi: 10.32604/cmc.2023.039460.

[25] Q. Y. Shambour, N. Qandeel, Y. Alraba’nah, A. Abumariam, and M. K. Shambour, “Artificial Intelligence Techniques for Early Autism Detection in Toddlers: A Comparative Analysis,” *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 1754–1764, 2024, doi: 10.47738/jads.v5i4.353.

[26] L. Kampa, K. Yamini, A. Basavaraju, and K. Anoop, “A Stack Based Ensemble Learning Method for Diagnosing Autism Spectrum Disorder,” *Mathematical Statistician and Engineering Applications*, vol. 71, no. 3, pp. 237–251, 2022, doi: https://doi.org/10.17762/msea.v71i3.162.

[27] B. Twala and E. Molloy, “On effectively predicting autism spectrum disorder therapy using an ensemble of classifiers,” *Scientific Reports*, vol. 13, no. 1, pp. 1–18, 2023, doi: 10.1038/s41598-023-46379-3.

[28] T. J. Bradshaw, Zachary Huemann, Junjie Hu, and Arman Rahmim, “A Guide to Cross-Validation for Artificial Intelligence in,” *Radiology: Artificial Intelligence*, vol. 5, no. 12, 2023, doi: 10.1148/ryai.220232.

[29] K. Takahashi, Y. Shimamura, C. Tachiki, Y. Nishii, and M. Hagiwara, “Cephalometric landmark detection without X-rays combining coordinate regression and heatmap regression,” *Scientific Reports*, vol. 13, no. 1, pp. 1–13, 2023, doi: 10.1038/s41598-023-46919-x.

[30] T. Terada, R. Kimura, and Y. Chen, “3D facial landmark detection based on differential cylindrical projection and multi-task learning,” *Communications in Information and Systems*, vol. 20, no. 4, pp. 443–459, 2020, doi: 10.4310/CIS.2020.v20.n4.a3.

[31] R. S. Raj, D. Pratiba, and R. P. Kumar, “Facial Expression Recognition using Facial Landmarks: A novel approach,” *Advances in Science, Technology and Engineering Systems*, vol. 5, no. 5, pp. 24–28, 2020, doi: 10.25046/aj050504.

[32] W. Yu-ching, J. Jhen-yu, and T. Lan, “The Accuracy of Facial Landmark Models in Different Facial Expressions,” *Journal of Clinical & Experimental Dermatology Research*, vol. 15, pp. 3–6, 2024, doi: 10.35841/2155-9554.24.15.674.Citation.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

[33] G. Gao, C. Chen, K. Xu, K. Liu, and A. Mashhadi, “Automatic face detection based on bidirectional recurrent neural network optimized by improved Ebola optimization search algorithm,” *Scientific Reports*, vol. 14, pp. 1–19, 2024, doi: 10.1038/s41598-024-79067-x 1.

[34] R. H. Al-abboodi, “Facial Expression Recognition Based on GSO Enhanced Deep Learning in IOT Environment,” *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 3, pp. 445–459, 2024, doi: 10.22266/ijies2024.0630.35.

[35] M. A. Khan, A. Power, P. Corcoran, and C. Thorpe, “An Evaluation of Features Extracted from Facial Images in the Context of Binary Age Classification,” *Learning and Analytics in Intelligent Systems*, vol. 48, pp. 99–112, 2025, doi: 10.1007/978-3-031-88649-2_11.

[36] R. Hidayat, M. O. B. Wibowo, B. Y. Satria, and A. Winursito, “Implementation of Face Recognition Using Geometric Features Extraction,” *Jurnal Ilmiah Kursor*, vol. 11, no. 2, p. 83, 2022, doi: 10.21107/kursor.v11i2.284.

[37] S. Li, J. Wang, L. Tian, J. Wang, and Y. Huang, “A fine-grained human facial key feature extraction and fusion method for emotion recognition,” *Scientific Reports*, vol. 15, no. 1, pp. 1–29, 2025, doi: 10.1038/s41598-025-90440-2.

[38] J. J. Levy, J. J. Levy, J. J. Levy, and A. J. O’Malley, “Don’t dismiss logistic regression: The case for sensible extraction of interactions in the era of machine learning,” *BMC Medical Research Methodology*, vol. 20, no. 1, pp. 1–15, 2020, doi: 10.1186/s12874-020-01046-3.

[39] M. Anshori, M. S. Haris, and A. Wahyudi, “Logistic Regression’s Effectiveness in Feature Selection with Information Gain in Predicting Heart Failure Patients,” *Journal of Enhanced Studies in Informatics and Computer Applications*, vol. 1, no. 2, pp. 35–39, 2024, doi: 10.47794/jesica.v1i2.8.

[40] O. Old, “Benefits and challenges of using logistic regression to assess neuropsychological performance validity: evidence from a simulation study,” *Clinical Neuropsychology*, vol. 37, no. 1, pp. 34–59, 2023, doi: 10.1007/978-3-658-38618-4_5.

[41] E. Dumitrescu, S. Hué, C. Hurlin, and S. Tokpavi, “Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects,” *European Journal of Operational Research*, vol. 297, no. 3, pp. 1178–1192, 2022, doi: 10.1016/j.ejor.2021.06.053.

[42] T. E. Mathew, “an Optimized Extremely Randomized Tree Model for Breast Cancer Classification,” *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 16, pp. 5234–5246, 2022.

[43] F. G. Arage, Z. B. Tadese, E. A. Taye, T. K. Tsegaw, T. G. Abate, and E. A. Alemu, “Cervical cancer screening uptake and its associated factor in Sub-Sharan Africa: a machine learning approach,” *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, 2025, doi: 10.1186/s12911-025-03039-y.

[44] Y. I. Sulistya and M. Istighosah, “Obesity Prediction with Machine Learning Models Comparing Various Algorithm Performances,” *International Journal of Artificial Intelligence in Medical Issues*, vol. 3, no. 1, pp. 1–13, 2025, doi: 10.56705/ijaimi.v3i1.181.

[45] J. Li, “Asymptotics of K-Fold Cross Validation,” *Journal of Artificial Intelligence Research*, vol. 78, pp. 491–526, 2023, doi: https://doi.org/10.1613/jair.1.13974.

[46] V. Singh, M. Pencina, A. J. Einstein, J. X. Liang, D. S. Berman, and P. Slomka, “Impact of train / test sample regimen on performance estimate stability of machine learning in cardiovascular imaging,” *Scientific Reports*, vol. 11, pp. 1–8, 2021, doi: 10.1038/s41598-021-93651-5.

[47] J. Qiu, “An Analysis of Model Evaluation with Cross-Validation : Techniques , Applications , and Recent Advances,” vol. 0, pp. 69–72, 2024, doi: 10.54254/2754-1169/99/2024OX0213.

[48] V. Teodorescu and L. O. Bras, “Assessing the Validity of k-Fold Cross-Validation for Model Selection : Evidence from Bankruptcy Prediction Using Random Forest and XGBoost,” vol. 13, p. 127, 2025, doi: 10.3390/computation13050127.

[49] J. White and S. D. Power, “k-Fold Cross-Validation Can Significantly Over-Estimate True Classification Accuracy in Common EEG-Based Passive BCI Experimental Designs : An Empirical Investigation,” vol. 23, p. 6077, 2023, doi: 10.3390/s23136077.

[50] Y. Wang, M. Khodadadzadeh, and R. Zurita-milla, “Spatial + : A new cross-validation method to evaluate geospatial machine learning models,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 121, p. 103364, 2023, doi: 10.1016/j.jag.2023.103364.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

e-ISSN: 2407-8964
p-ISSN: 1907-7904

[51] T. Abedin, H. Xu, and S. Uddin, "The impact of K selection in K - fold cross-validation on bias and variance in supervised learning models," vol. 16, pp. 1–13, 2026, doi: 10.1038/s41598-026-37247-x 1.

[52] J. C. Obi, "A Comparative Study of Several Classification Metrics and Their Performances on Data," *World Journal of Advanced Engineering Technology and Sciences*, vol. 8, no. 1, pp. 308–314, 2023, doi: 10.30574/wjaets.2023.8.1.0054.

[53] D. Chicco and G. Jurman, "An Invitation to Greater Use of Matthews Correlation Coef fi cient in Robotics and Arti fi cial Intelligence," *Frontiers in Robotics and AI*, vol. 9, pp. 1–4, 2022, doi: 10.3389/frobt.2022.876814.

[54] G. M. F. Id, "Challenges in the real world use of classification accuracy metrics : From recall and precision to the Matthews correlation coefficient," *PLoS ONE*, vol. 18, no. 10, pp. 1–27, 2023, doi: 10.1371/journal.pone.0291908.

AUTHOR BIOGRAPHY

Syifa Anzella was born in Meulaboh on June 22, 2005. She is currently a student in the Electrical and Computer Engineering Program at Syiah Kuala University, where she began her undergraduate studies in 2022. Throughout her academic journey, she has developed a strong interest in multimedia technology, particularly in digital image processing, pattern recognition, and computer-based system development. She has been actively involved in several academic projects in machine learning and image analysis, focusing on applying computational methods to real-world problems. Her research interests also extend to artificial intelligence and data-driven systems. This interest has motivated her to continually explore innovative approaches to using visual data to support technological advancement. She can be contacted via email at: syifa.an@mhs.usk.ac.id.

Yudha Nurdin was born in Bangkok, Thailand, on October 1st, 1979. He received his B.Eng. degree in Electrical Engineering in 2005 and M.Eng. in Electrical Engineering in 2009, both from Bandung Institute of Technology, Indonesia. His research focuses on the challenges and solutions of Artificial Intelligence and Machine Learning models, particularly in cybersecurity and practical real-world applications. Concurrent with his academic activities, he serves as a lecturer and researcher in the Department of Electrical and Computer Engineering, Faculty of Engineering, University of Syiah Kuala, Indonesia. His research interests include cybersecurity, cloud computing, multi-agent systems, and machine learning. He has been actively involved in various research projects, academic publications, and collaborative studies related to intelligent systems and secure computing. He can be contacted by email at: yudha.nurdin@usk.ac.id.

Melinda was born in Bireuen, Aceh, on June 10, 1979. She received a B.Eng degree from the Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh, in 2002. She completed her master's degree in the Faculty of Electrical Department, University of Southampton, United Kingdom, with a concentration in field study of Radio Frequency Communication Systems in 2009. She has already completed her Doctoral degree at the Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, in February 2018. She has been with the Department of Electrical Engineering, Faculty of Engineering, Universitas Syiah Kuala since 2002. She is also a member of IEEE. Her research interests include multimedia signal processing and fluctuation processing.
She can be contacted by email at melinda@usk.ac.id.

J. Junidar received the B.Sc. degree in Mathematics from the Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Banda Aceh, Indonesia, in 2001. Since 2006, she has been a lecturer with the Department of Informatics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University. She obtained her M.Sc. degree in Computer Science from the Faculty of Computer Science, Universitas Sumatera Utara, Indonesia, in 2013. Her research interests include computer science and soft computing, with a particular focus on deep learning and data-driven methods. She has been actively involved in teaching, supervising student research, and participating in various academic and research collaborations. She can be contacted by email at: junidar678@mhs.usk.ac.id
.

Rosminazuin Ab Rahim received her BEng (Hons) in Electrical & Electronics Engineering from University Science Malaysia. She commenced her professional journey as a quality control engineer at Delphi Packard Electric (M) Sdn Bhd, Kuala Terengganu, from October 1999 to November 2001. In January 2003, she assumed the position of Assistant Lecturer at the Department of Electrical & Computer Engineering, Kulliyyah of Engineering. In 2004, she earned a Master of Science (Microelectronics), followed by a Doctor of Philosophy (Microengineering and Nanoelectronics) in

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163

2012, both from University Kebangsaan Malaysia. Her research interests include electrochemical sensors, MEMS, and energy harvesters. She currently serves as an Associate Professor in the Department of Electrical & Computer Engineering, Kulliyyah of Engineering at the International Islamic University Malaysia.

She can be contacted by email at: rosmi@iium.edu.my.

Corresponding author: Yudha Nurdin, yudha.nurdin@usk.ac.id, Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Darussalam, Banda Aceh, 23111, Indonesia
DOI: https://doi.org/10.35882/jteknokes.v19i3.163