

Precise Electrocardiogram Signal Analysis Using ResNet, DenseNet, and XceptionNet Models in Autistic Children

Yunidar Yunidar¹, Melinda Melinda¹, Albahri¹, Hanum Aulia¹, Herlina Dimiati², and Nurlida Basir³

¹ Department of Electrical Engineering and Computer, Engineering Faculty, Universitas Syiah Kuala, Banda Aceh, Indonesia

² Department of Pediatrics, Faculty of Medicine, Universitas Syiah Kuala, Banda Aceh, Indonesia.

³ Faculty of Science and Technology, Universiti Sains Islam Malaysia (USIM), Nilai Negeri Sembilan, Malaysia

Corresponding author: Yunidar Yunidar (e-mail: yunidar@usk.ac.id), **Author(s) Email:** Melinda Melinda (e-mail: melinda@usk.ac.id), Albahri (e-mail: albahri@usk.ac.id), Hanum Aulia (e-mail: hanumaulia@mhs.usk.ac.id), Herlina Dimiati (e-mail: herlinadimiati@usk.ac.id), Nurlida Basir (e-mail: nurlida@usim.edu.my).

Abstract In autistic children, one of the important physiological aspects to be examined is the heart condition, which can be assessed through electrocardiogram (ECG) signal analysis. However, ECG signals in autistic children often contain interference in the form of noise, making the analysis process, both manual and conventional, challenging. Therefore, this study aims to analyze the ECG signals of autistic children using a classification method to distinguish between two main conditions: playing and calm conditions. A deep learning approach employing the Convolutional Neural Network (CNN) architectures was used to obtain accurate results in distinguishing the heart conditions of autistic children. The data used consists of 700 ECG signal data in each class, processed through the filtering, windowing, and augmentation stages to obtain balanced data. Three CNN architectures, ResNet, DenseNet, and XceptionNet, were tested in this study. Although these architectures are originally designed for 2D and 3D image data, modifications were made to adapt the input data structure to perform 1D data calculations. The evaluation results show that the XceptionNet model achieved the best performance, with accuracy, precision, recall, and F1-score of 97,14% each, indicating a good ability in capturing the complex patterns of ECG signals. Meanwhile, the ResNet obtained good results with 96,19% accuracy, while DenseNet performed slightly lower results with 94,76% accuracy and evaluation metrics. Overall, this study demonstrates that a deep CNN architecture based on dense connections can enhance the accuracy of ECG signal classification in autistic children.

Keywords ECG Signal; Autistic Children; ResNet; DenseNet; XceptionNet.

1. Introduction

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterized by a variety of difficulties in social interaction, interpersonal communication, and sensory processing. This disorder is also characterized by restricted, repetitive, and stereotyped patterns of interests and behaviors [1],[2]. Such symptoms can significantly disrupt behavior, language, communication, and social interaction, thereby posing considerable challenges to the learning process [3]. Moreover, ASD is frequently associated with hyperactivity, which negatively impacts daily activities and diminishes the quality of life for both affected individuals and their families [4],[5]. One condition that needs to be considered when assessing this response is the child's heart condition activity [6].

To determine the condition of the heart, a heart examination can be performed using an ECG that produces a heartbeat [7]. However, ECG signals in children with autism are often very variable and have noise, making manual and traditional analysis difficult [8]. By analyzing biosignals, such as electrocardiograms (ECGs), healthcare professionals can diagnose and monitor disorders with greater accuracy [9]. For this reason, fast and accurate automatic ECG signal analysis is needed to distinguish the heart condition of autistic children.

Furthermore, the rapid development of deep learning techniques in AI over the past few years has fundamentally changed the landscape of medical data analysis [10]. For example, research on composite HRV biomarkers extracted from resting ECGs can

effectively differentiate school-aged children with ASD from their typically developing peers and those with other psychiatric conditions, achieving an AUC of 0.89 using a machine learning classifier [11]. In another study, among toddlers with ASD, parameters such as SDNN, CV, and LF/HF derived from ECGs during a joint attention task showed different patterns of autonomic regulation compared to their typically developing peers [12]. Another study on wearable ECG sensors that record heart rate and heart rate variability (HRV) helps predict challenging behavior in children with ASD, with models such as XGBoost achieving high precision, where HRV contributes significantly to performance [13].

This study utilizes data from previous research results, specifically those on the AD8232-based autistic child electrocardiogram detection system for health services, which received ethical approval from the Health Research Ethics Committee in 2023 [8]. This data consists of 10 subjects, including five play class subjects (aged 7-10 years) and five quiet class subjects (aged 7-10 years).

ECG signal classification employs manual feature extraction approaches (e.g., HRV, RR interval) and models, such as SVM [14]. As time passes, Convolutional Neural Network (CNN) models have begun to be utilized for 1D signal classification. ResNet (Residual Network), DenseNet (Densely Connected Network), and XceptionNet have become popular models in medical image processing because they show better generalization, including in the context of 1D signals such as ECG [15],[16]. The ResNet architecture consists of a series of residual blocks, with each block containing multiple convolutional layers and skip connections that pass through multiple layers. DenseNet is a convolutional neural network architecture known for its dense connectivity pattern, where each layer is connected to every other layer in a feed-forward manner. Xception is a convolutional neural network architecture built on the Inception architecture. This architecture was introduced as an extension to overcome the limitations of the traditional Inception module, aiming to achieve better performance in terms of accuracy and computational efficiency [17].

Although numerous studies have been conducted using deep learning for ECG signal classification, few have focused their research on the proposed method. Therefore, this study proposes the use of ECG signal classification in autistic children using the ResNet, DenseNet, and XceptionNet architectures. Before being tested with these architectures, the signal is first processed using preprocessing techniques (such as filtering, windowing, and augmentation) to ensure accurate results in the ECG signal classification test.

This study measures the effectiveness of ResNet, DenseNet, and XceptionNet architectures in classifying

play and calm conditions in ECG signals of autistic children. The performance of these models is evaluated in terms of accuracy, precision, recall, and F1-score. The main contributions of this study are summarized as follows.

1. Proposing the application of 1D ResNet, DenseNet, and XceptionNet for classifying ECG signals of autistic children under two different conditions (calm and play).
2. Conducting a comparative analysis of the three deep learning models to evaluate their performance on 1D signals in a specific population with high noise signal challenges.
3. Using an end-to-end deep learning approach that avoids manual feature extraction, relying on automatic feature learning from raw data.
4. Evaluating the effectiveness of preprocessing filtering, windowing, augmentation, and CNN in improving the accuracy of ECG signal classification in autistic children.

Several sections in this study explain the proposed method in Section II. Section III presents the results and discussion of the ECG signal classification testing in children with ASD. Conclusions and some future

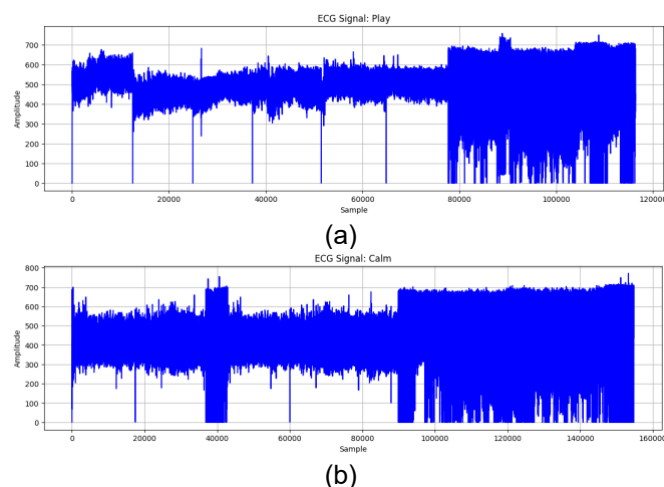


Fig. 1. Raw ECG signal sample plot: (a) playing; (b) calm

works of this study are in Section IV.

II. Method

A. Dataset

This proposed study utilizes a dataset derived from the results of previous research, specifically on the electrocardiogram detection system for autistic children based on the AD8232, which has been approved for health services and obtained ethical approval from the Health Research Ethics Committee in 2023 with No. 028/EA/FK/2023 [8]. The data used in this study consists of raw ECG signals obtained from 10 subjects,

including five play class subjects (aged 7-10 years) and five quiet class subjects (aged 7-10 years) using the AD8232 sensor. In this study, the number of samples for the play class was 116,359, and the number of samples for the quiet class was 154,798. Each signal was recorded in two states: play and quiet. The data obtained were in one-dimensional (1D) time series format with a sampling frequency of 250 Hz. An illustration of a sample plot of the raw ECG signal can be seen in Fig. 1. After recording, these signals still contain noise, so they require preprocessing stages such as filtering, windowing, and adding.

B. Data Processing

Dataset preprocessing was performed in several stages: importing ECG data, filtering EEG signals using a Butterworth bandpass filter, Hamming windowing, augmenting ECG data to achieve balance, and splitting the dataset into three parts: train, validation, and test.

to remove noise from EEG signals. This filter maintains a consistent frequency response within the desired range while attenuating frequencies outside it. Previous experiments have shown the effectiveness of this method in maintaining the specified frequency range according to the filter design parameters [19],[20],[21]. This filtering process is achieved by applying a digital transfer function from the filter to the signal using a method such as zero-phase filtering, ensuring the signal does not experience a phase shift [22]. This step ensures that the Butterworth bandpass filter can clean the signal from interference without altering the main shape of the ECG signal [23]. The Butterworth bandpass filter utilizes a high-pass filter to eliminate low-frequency components and a low-pass filter to suppress high-frequency noise. The Butterworth bandpass filter utilizes a frequency range of 0.5-45 Hz, combined with a 4th-order filter, and is also

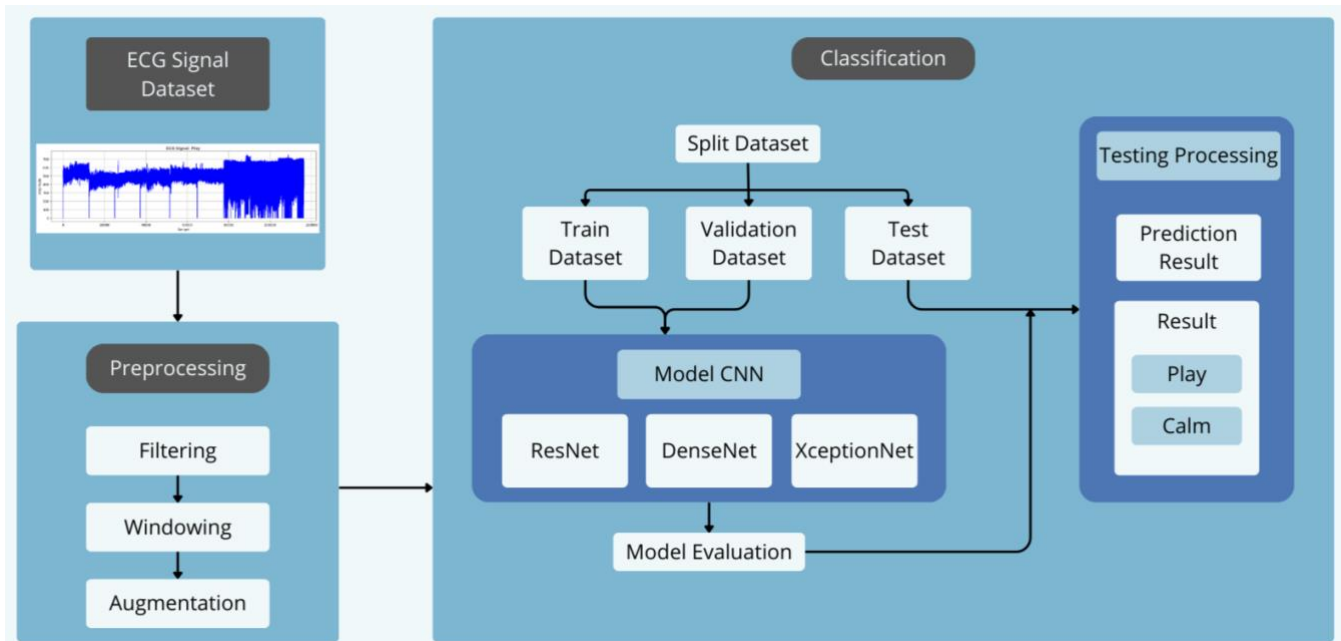


Fig. 2. Research flow in ECG signal classification using CNN model

Then, a CNN model was trained with three architectures: ResNet, DenseNet, and XceptionNet, and the model was evaluated to determine the best performance. The entire process is depicted in Fig. 2.

1. Filtering

After the data is prepared, the next step is to filter the ECG signal using the bandpass filter method. The collected signal is often accompanied by substantial random pulse interference. The signal processing method, based on spectral kurtosis, is a relatively effective method for filtering and noise reduction, which involves filtering signals through various bandpass filters and selecting signals with optimal kurtosis values [18]. The Butterworth bandpass filter is commonly used

accompanied by a notch filter at 50 Hz to remove low- and high-frequency noise in the ECG signal data. Mathematically as in Eq. (1) [19]:

$$\sum_{k=0}^M a_k y[n-k] = \sum_{k=0}^M b_k x[n-k], \quad (1)$$

with $x[n]$ the input signal, $y[n]$ the output signal, and $\{a_k\}$, $\{b_k\}$ the filter coefficients obtained through a bilinear transformation. To eliminate electrical interference, a notch filter at 50 Hz is used with the transfer function in Eq. (2) [22]:

$$H_{notch}(z) = \frac{1 - 2 \cos(w_0) z^{-1} + z^{-2}}{1 - 2r \cos(w_0) z^{-1} + r^2 z^{-2}}, \quad (2)$$

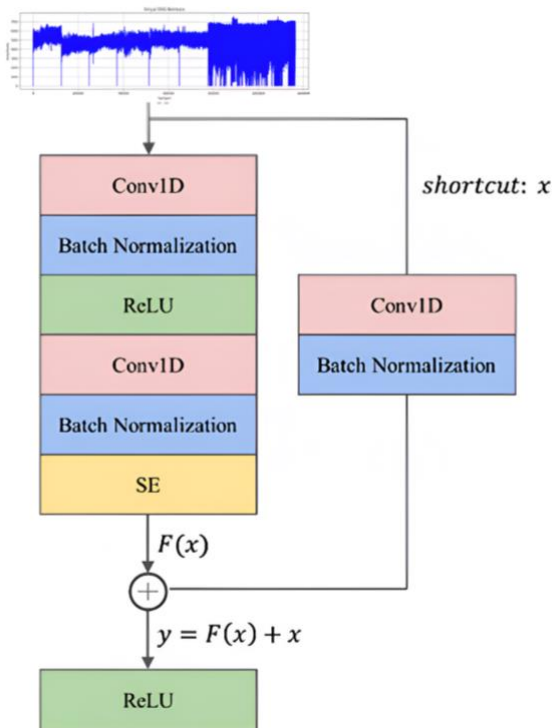


Fig. 3. Illustration architecture of ResNet

where $w_0 = \frac{2\pi(50)}{f_s}$ and r is the pole radius.

This filtering stage can improve the performance of the classification model. Filtering using a Butterworth bandpass filter and a 50 Hz notch filter reduces low- and high-frequency noise without altering the primary shape of the ECG signal. This helps preserve relevant physiological features, allowing the model to learn from a cleaner signal representation. As a result, the model's sensitivity to noise is reduced, and classification accuracy is increased.

2. Windowing

At this stage, the filtered ECG signal will be windowed into intervals of 2 seconds or according to 500 amplitude samples in each class. To simplify signal processing, both data sets are downsampled to 250 Hz while retaining important information, as done in some literature [24],[25]. This study uses the hamming window method because this window function is the most widely used, and the choice of window function size is based on the relationship between the scan signal period and the sampling frequency, as done in some literature [26],[27].

3. Augmentation

After the filtering and windowing process on the ECG signal, there are 618 calm class data and 464 play class data. Due to the large number of datasets for deep learning and the imbalanced classification

problems that arise during the training process, we perform data augmentation as done in some literature [28],[29],[30]. Augmentation was performed to address the limited number of subjects and the imbalance in the number of samples between classes. The augmentation technique used was the addition of Gaussian noise (Additive Gaussian Noise Injection) to the ECG signals. Gaussian noise, as a data augmentation technique, leverages its ability to introduce variability into the minority class without producing similar values, thereby reducing the risk of overfitting [31]. Through this technique, the number of data points in the playing class was increased to balance with the quiet class, with 700 samples per class. Further augmentation was applied to the playing class to balance the dataset. The data augmentation stage can correct class imbalance by increasing the amount of data in the minority class, thereby reducing the model's bias toward the majority class. This stage also increases signal diversity through artificial variations that mimic natural changes in the ECG. With augmentation, the model is trained on a more diverse signal distribution, making it more robust to both real variations and random noise that arise in new data.

4. Split Data

In the final stage of signal processing, the data generated during the augmentation process is divided into three sets: train, validation, and test, with ratios of 70%, 15%, and 15%, respectively. This ratio is chosen for the train-validation-test division because it does not violate the temporal aspect of the data. This method has been successfully applied in previous research [32].

C. ResNet

ResNet is a deep neural network architecture that has made significant progress in various computer vision tasks [17]. Fig. 3 illustrates the ResNet architecture, which preserves the identity of the previous layers, ensuring that the mapping of each weighted layer remains identical. ResNet stacks residual layers by performing element-by-element addition after every two convolutional layers [30]. This process is mathematically expressed as Eq. (3) [33]:

$$H(x) = F(x, \{W_i\}) + x \quad (3)$$

where x is the input, $F(x, \{W_i\})$ is the residual function parameterized by convolutional layers with weights $\{W_i\}$, and $H(x)$ is the output of the residual block.

The original ResNet was designed to handle 2D and 3D image data. However, since the data used in this study is one-dimensional, we modified the input data structure to perform 1D data calculations by replacing the input layer with a sequence input layer [34]. For an input sequence $x \in R^{T \times C}$, the transformation within a residual block can be written as Eq. (4) [34]:

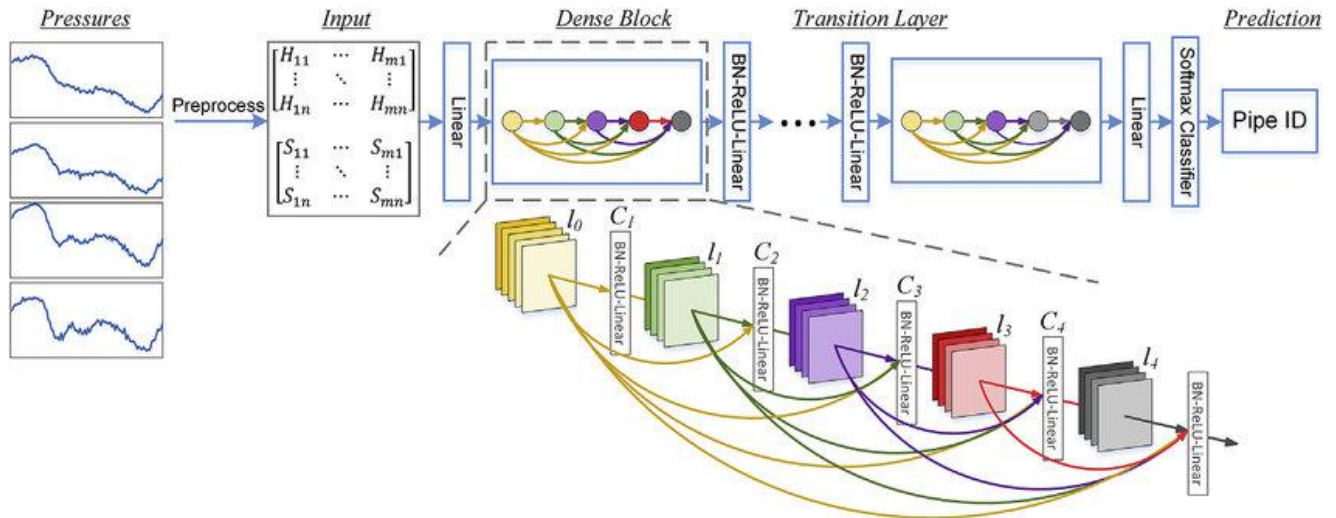


Fig. 4. Illustration architecture of DenseNet

$$F(x) = \sigma \left(BN(W_2 * \sigma(BN(W_1 * \sigma))) \right) \quad (4)$$

where $*$ denoted 1D convolution, BN is batch normalization, σ is a non-linear activation function (ReLU), and W_1, W_2 are convolutional kernels. The final representation $h^{(L)}$ from stacked residual blocks is aggregated using global average pooling and then passed through a fully connected layer with softmax activation for classification, which can be expressed as Eq. (5) [35]:

$$\hat{y} = \text{softmax}(W_{fc} \cdot h^{(L)} + b) \quad (5)$$

where $h^{(L)}$ is the learned representation after the final dense block. ResNet-1D has been utilized in various studies, including human activity recognition [34], sleep stage determination from single-channel EEG signals [21], and music genre classification from audio waveforms [33].

D. DenseNet

DenseNet is a convolutional neural network architecture known for its dense connectivity pattern, where each layer is connected to every other layer in a feed-forward manner [17]. Fig. 4 illustrates the DenseNet architecture, which recognized for its ability to enhance gradient flow information using direct connections across sub-block layers. Unlike ResNet, which only receives information from the previous layer, DenseNet receives information from all sub-blocks or sub-layers [30]. Formally, the output of the l -th layer in DenseNet can be written as Eq. (6) [30]:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (6)$$

where $x_0, x_1, \dots, x_{l-1}$ are the feature maps produced by all previous layers, $[\cdot]$ denotes concatenation, and $H_l(\cdot)$ represents a composite function of operations (typically batch normalization, ReLU, and convolution).

Since the data used in this study is one-dimensional, we also modify the input data structure in the DenseNet model to perform 1D data calculations. For a given input sequence $x \in R^{T \times C}$, the transformation inside each layer can be expressed as Eq. (7) [36]:

$$H_l(x) = \sigma(BN(W_l * x)) \quad (7)$$

where $*$ denotes 1D convolution, BN is batch normalization, σ is a non-linear activation function, and W_l is the convolutional kernel at the l -th layer. The outputs from the dense blocks are further aggregated using global average pooling, followed by a fully connected layer for classification, which can be expressed as Eq. (8) [35]:

$$\hat{y} = \text{softmax}(W_{fc} \cdot h^{(L)} + b) \quad (8)$$

where $h^{(L)}$ is the learned representation after the final dense block. DenseNet-1D has been utilized in various studies, including stock market prediction [20], abnormal driving behavior [37], pipeline leak detection [38], stress classification using EEG data [39], and atrial fibrillation detection [36].

E. XceptionNet

Xception is a convolutional neural network architecture built on the Inception architecture. This architecture was introduced as an extension to overcome the limitations of the traditional Inception module, aiming to achieve better performance in terms of accuracy and computational efficiency. Fig. 5 illustrates the XceptionNet architecture, which employs an extreme version of depthwise separable convolution, separating spatial and channel filtering operations. This design significantly reduces the number of parameters and computations compared to traditional convolution, resulting in increased efficiency without sacrificing performance [17]. Xception has more modules; each

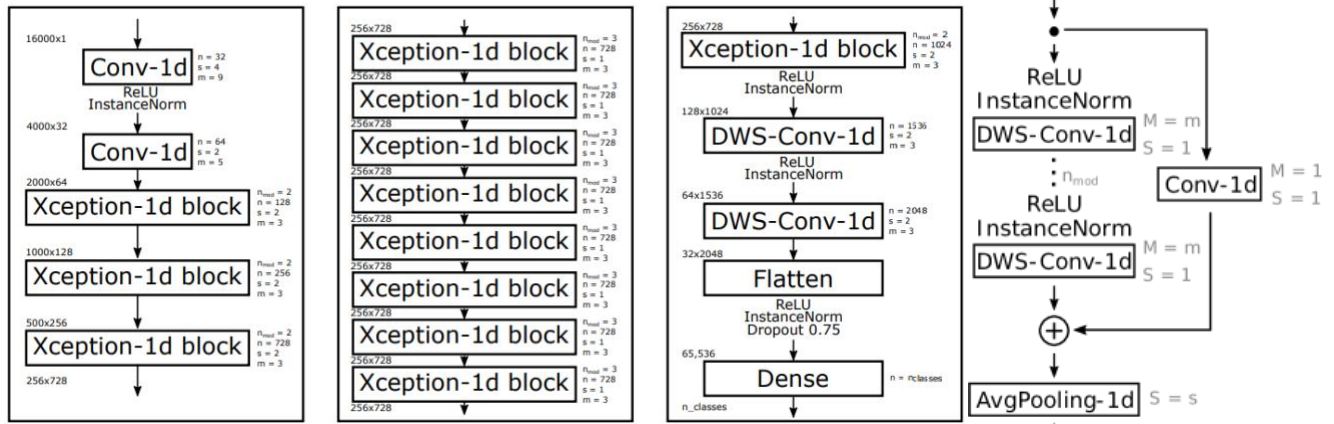


Fig. 5. Illustration architecture of XceptionNet

Xception module is simpler and also uses depthwise separable convolution to perform feature mapping [30].

Mathematically, a standard convolution for an input sequence $x \in R^{T \times C_{in}}$ with kernel size k , input channels C_{in} , and output channels C_{out} can be expressed as Eq. (9) [17]:

$$y(t, C_{out}) = \sum_{c=1}^{C_{in}} \sum_{i=1}^k W_{c,i,C_{out}} \cdot x(t+i, c) \quad (9)$$

where W is the convolutional kernel. In Xception, this operation is factorized into two steps:

Depthwise convolution (per-channel filtering), mathematically can be expressed as Eq. (10) [10]:

$$z(t, c) = \sum_{i=1}^k W_{c,i}^{(d)} \cdot x(t+i, c) \quad (10)$$

Pointwise convolution (channel mixing), mathematically can be expressed as Eq. (11) [10]:

$$y(t, c_{out}) = \sum_{c=1}^{C_{in}} W_{c,C_{out}}^{(p)} \cdot z(t, c) \quad (11)$$

where $W^{(d)}$ are depthwise filters and $W^{(p)}$ are pointwise (1×1) filters. In this study, since the data is one-dimensional, we also modified the input structure by replacing 2D convolution with 1D convolution in the Xception modules. The final feature representation is aggregated through global average pooling and mapped to the target classes via a fully connected layer with softmax activation, which can be expressed as Eq. (12) [35]:

$$\hat{y} = \text{softmax}(W_{fc} \cdot h^{(L)} + b) \quad (12)$$

where $h^{(L)}$ is the learned representation after the final dense block. XceptionNet-1D has also been used in research on keyword recognition in audio [40].

F. Model Evaluation

The final evaluation of the model is performed on the test data using a confusion matrix to compare the performance of the ResNet, DenseNet, and XceptionNet architectures. Equations (13) – (16) can be used to calculate the accuracy, recall, precision, and F1-Score values to determine the system performance. The parameters used in this study are as follows.

- Accuracy measures the overall accuracy of the model's predictions by calculating the proportion of correctly classified examples over the total number of cases (Eq. (13) [30]).

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (13)$$

- Precision measures how many of the predicted positives are positive (Eq. (14) [30]).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (14)$$

- Recall measures how many of the actual positives are successfully predicted as positive (Eq. (15) [30]).

$$\text{Recall} = \frac{TP}{TP+FN} \quad (15)$$

- The F1-score is the harmonic mean of precision and recall, used to measure the balance between the two (Eq. (16) [30]).

$$F1 - \text{score} = 2x \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

The confusion matrix is used to evaluate the performance of a classification model by comparing the predicted results with the actual conditions. The True Positive (TP) value indicates the number of quiet class data that were correctly predicted as the quiet class. The True Negative (TN) value shows the number of instances of the playing class that were correctly predicted as the playing class. Meanwhile, the False Positive (FP) value indicates the number of quiet class data that were incorrectly predicted as the playing class. Conversely, the False Negative (FN) value is the number of playing class data that were incorrectly predicted as the quiet class. These four components provide a comprehensive overview of the model's

accuracy, error, and ability to distinguish between the two classes.

considered the average of the validation results across all iterations [42].

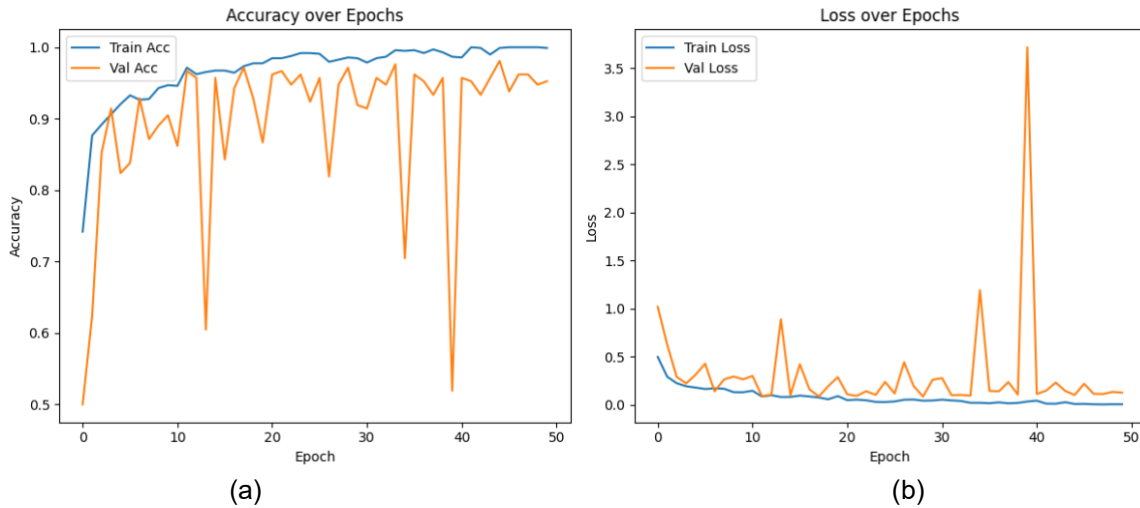


Fig. 6. Epoch results using ResNet architecture: (a) Accuracy vs epoch; (b) Loss vs epoch

G. Statistic Test

In the performance evaluation of EEG signal denoising methods, inferential statistical analysis often utilizes a paired t-test to determine whether the performance difference between two methods is statistically significant. Theoretically, this test considers the difference in metric values (such as MAE, MSE, or SNR) of each pair of observations derived from two different methods. The formula for the paired t-test can be expressed as follows, Eq. (17) [33]:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} \quad (17)$$

where $\bar{d}$, is the mean of the differences between pairs, s_d is the standard deviation of the differences, and n denotes the number of sample pairs as in Eq. (17) [41]. The obtained t value was compared with the critical value of the t distribution ($df = n - 1$) to calculate the p value, which is the probability that the observed difference occurred by chance. Mathematically, as in Eq. (18) [41]:

$$p = 2(1 - T(|t|, df)) \quad (18)$$

with T as the cumulative distribution function of t , representing cumulative probability under t distribution [33]. In this study, we propose the use of K-fold cross-validation combined with non-parametric statistical tests based on case analysis. The method we are studying is derived from k-fold cross-validation. This method randomly samples a dataset and divides it into k parts (folds) of (almost) equal size. Then, for each fold f , $k - 1$ other folds are used to train a classifier, and the f folds are used to validate the obtained model based on the corresponding size. Model performance is

H. Ablation Study

To assess the contribution of each preprocessing step, we conducted an ablation study testing combinations of filtering (0.5–45 Hz Butterworth with a 50 Hz notch), Hamming-based 2-second windowing (500 samples), and Gaussian augmentation for class balancing. Eight variants (V0–V7) were evaluated, ranging from no preprocessing (V0) to the whole pipeline (V7). In the variant without augmentation, class imbalance was compensated for using class_weight to ensure a fair comparison with V7. All experiments maintained the same training settings (ResNet / DenseNet / XceptionNet architectures, Adam optimizer, learning rate 0.001, 50 epochs, batch size 32, and a 70/15/15 split). Each variant was replicated 5 times with different seeds to obtain the mean and standard deviation. The significance of differences was evaluated using McNemar's test for prediction pairs (per model) and the Wilcoxon signed-rank test for aggregated cross-retest metrics.

III. Result

After the processing stage, the data is divided into three parts: training, validation, and testing. There are 700 ECG signal data in each playing and calm class. In the train data, there are 490 data points for each playing class and calm class. Additionally, there are 105 validation data points for each playing class and calm class, and 105 test data points for each playing class and calm class.

A. Training and Validation Dataset Results

At this stage, signal classification is performed separately using ResNet, DenseNet, and XceptionNet.

Figures 6-8 compare the accuracy and loss of each architecture on the training and validation data. The accuracy of all classifications in the proposed methods gets quite good results. In all proposed methods, 50 epochs are used to train the training data. At this stage,

cross-entropy, which is appropriate because the labels have been converted to one-hot encoding. The regularization method is BatchNormalization, which functions as a lightweight regularizer. In this model, no layers are frozen, and the model is trained from scratch.

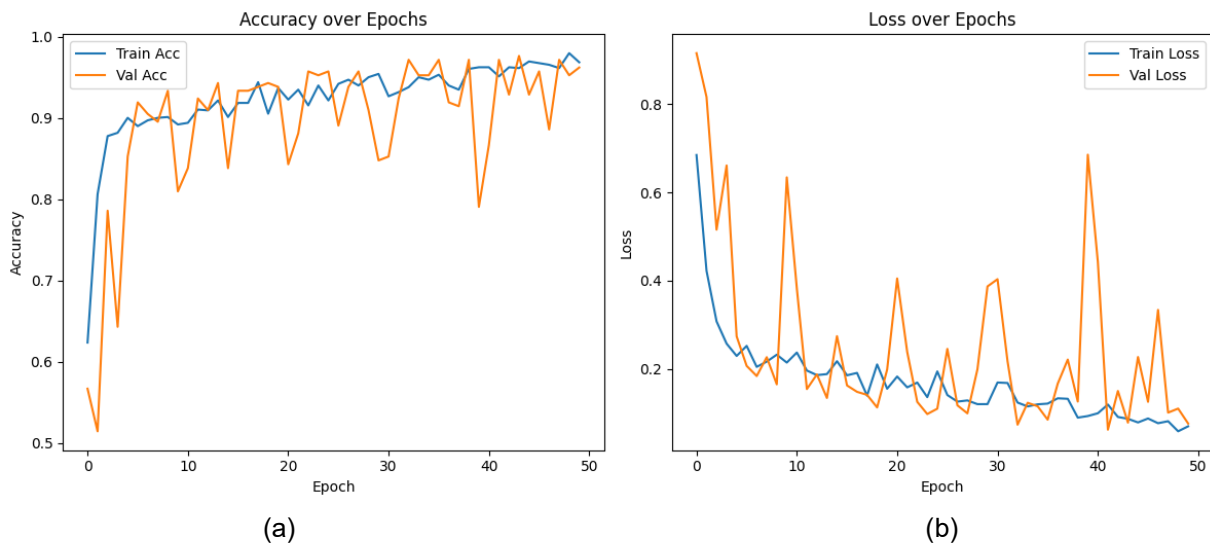


Fig. 7. Epoch results using DenseNet architecture: (a) Accuracy vs epoch; (b) Loss vs epoch

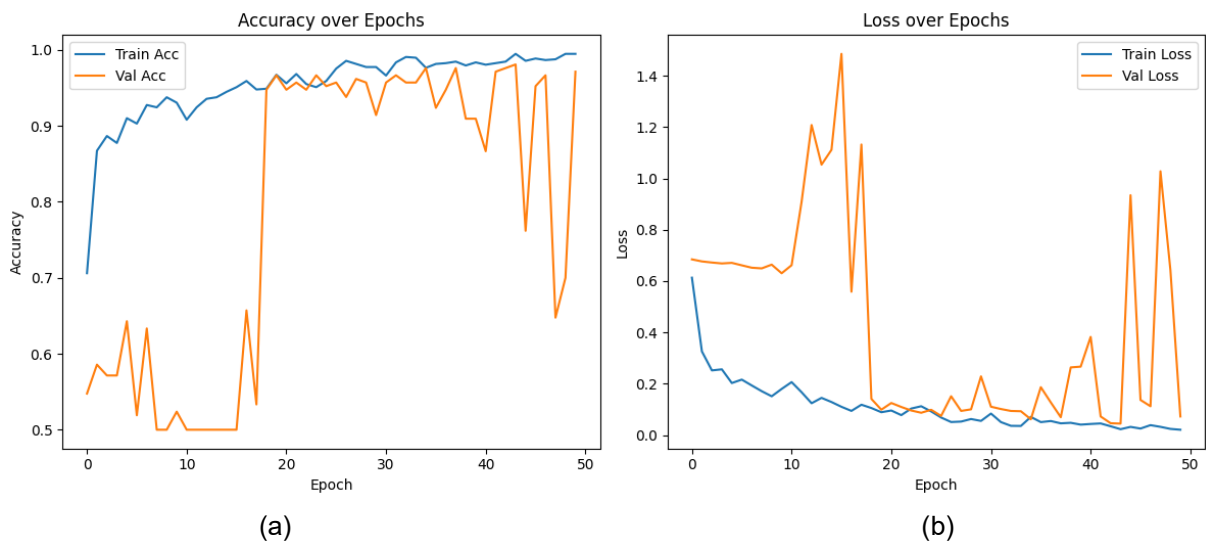


Fig. 8. Epoch results using XceptionNet architecture: (a) Accuracy vs epoch; (b) Loss vs epoch

each CNN model with the highest accuracy and lowest functions as a lightweight regularizer. In this model, no loss will be used to evaluate the results of each model in ECG signal classification.

Fig. 6 illustrates the model's performance during the training process on the ResNet model, where Fig. 6(a) shows the accuracy value and Fig. 6(b) shows the loss value. This model was trained using the Adam optimizer with a learning rate of 0.001. The batch size used was 32 with 50 epochs. The loss function used categorical

In Fig. 6(a), it can be seen that the training accuracy achieves its best result at epoch 42, with a value of 100%, and the validation accuracy achieves its best result at epoch 45, with a value of 98.1%. This indicates that the training accuracy is higher than the validation accuracy, which affects the overall level of accuracy produced. In Fig. 6(b), the training loss reaches 0.28%, and the validation loss reaches 12.43% at epoch 50. This indicates that the training loss is lower than the validation loss, which affects the level of accuracy

achieved. The graph compares training accuracy with validation accuracy, as well as training loss with validation loss, for the ResNet model.

Fig. 7 illustrates the model performance during the training process for the DenseNet model, where Fig. 7(a) shows the accuracy value, and Fig. 7(b) shows the loss value. This model was trained using the Adam optimizer with a learning rate of 0.001. The batch size used was 32 with 50 epochs. The loss function used categorical cross-entropy, which is appropriate because the labels have been converted to one-hot encoding. The regularization method is BatchNormalization, which functions as a lightweight regularizer. In this model, no layers are frozen, and the model is trained from scratch. In Fig. 7(a), it can be seen that the training accuracy achieves its best result at epoch 49, with a value of 98.28%, and the validation accuracy achieves its best result at epoch 44, with a value of 97.62%. This indicates that the training accuracy is higher than the validation accuracy, which affects the overall level of accuracy produced. In Fig. 7(b). the training loss reaches 6.43%, and the validation loss reaches 7.66% at epoch 50. This indicates that the training loss is lower than the validation loss, which affects the level of accuracy achieved. The graph illustrates a comparison of training accuracy with validation accuracy, as well as a comparison of training loss with validation loss for the DenseNet model.

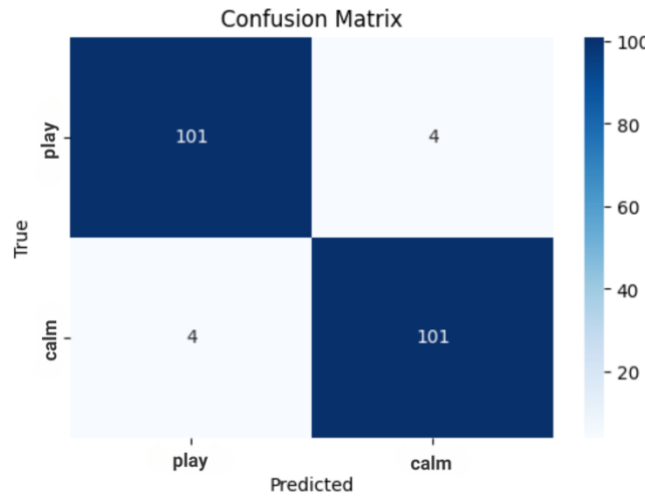


Fig. 9. Confusion Matrix of ResNet.

Fig. 8 illustrates the model's performance during the training process using the XceptionNet model, where Fig. 8(a) displays the accuracy value, and Fig. 8(b) shows the loss value. This model was trained using the Adam optimizer with a learning rate of 0.001. The batch size used was 32 with 50 epochs. The loss function used categorical cross-entropy, which is appropriate because the labels have been converted to one-hot encoding. The regularization method is

BatchNormalization, which functions as a lightweight regularizer. In this model, no layers are frozen, and the model is trained from scratch. In Fig. 8(a), it can be seen that the training accuracy achieves its best result at epoch 49, with a value of 99.53%. The validation accuracy achieves its best result at epoch 44, with a value of 98.1%. In Fig. 8(b), it is seen that the training accuracy achieves its best result at epoch 49, which is 99.53%, and the validation accuracy achieves its best result at epoch 44, which is 98.1%. This indicates that the training accuracy is higher than the validation accuracy, which affects the training loss reaches 2.62%, and the validation loss reaches 7.26% at epoch 50. This indicates that the training loss is lower than the validation loss, which affects the level of accuracy achieved. The graph illustrates a comparison of training accuracy and validation accuracy, as well as a comparison of training loss and validation loss, for the XceptionNet model.

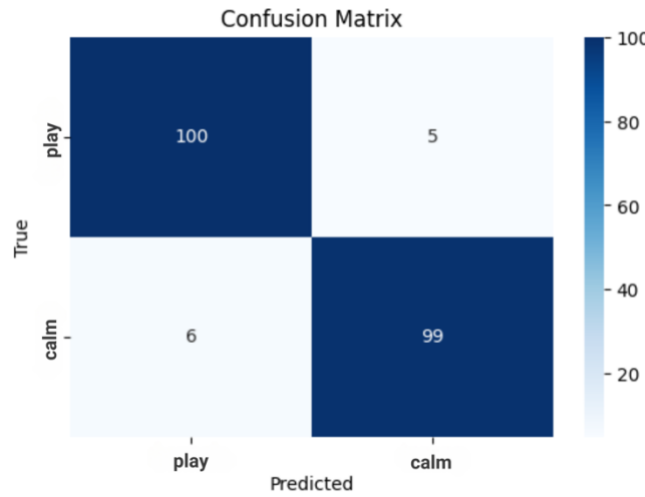


Fig. 10. Confusion Matrix of DenseNet.

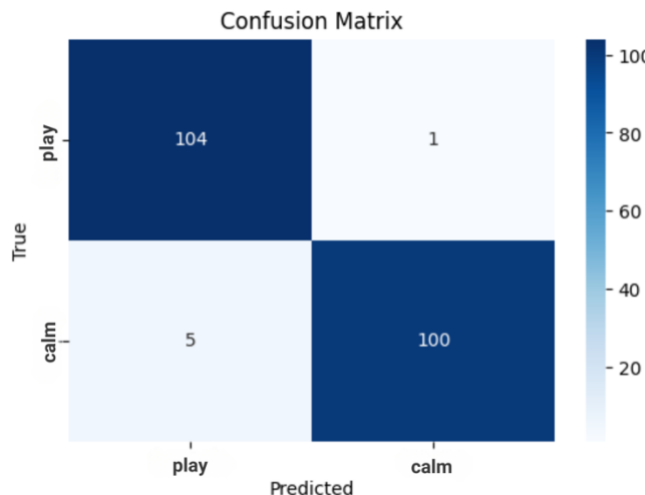


Fig. 11. Confusion Matrix of XceptionNet.

B. Testing Result

In the testing phase, the testing dataset consists of 210 test data divided into 105 data for each playing and quiet class. All CNN models will be evaluated using the test data. The evaluation matrices used are accuracy, precision, recall, and F1-score to measure the performance of the CNN model in classifying ECG signals in autistic children.

In the ResNet model, the confusion matrix shows the test results obtained, which can be seen in Fig. 9. The ResNet model can predict the results of 101 data correctly as the playing class, 4 data are mispredicted as the calm class, 101 data are also predicted correctly as the calm class, and 4 data are mispredicted as the playing class. In the DenseNet model, the confusion matrix shows the test results obtained, which can be seen in Fig. 10. The DenseNet model can predict the results of 100 data correctly as the playing class, 5 data are mispredicted as the calm class, 99 data are also predicted correctly as the calm class, and 6 data are mispredicted as the playing class. In the XceptionNet model, the confusion matrix shows the test results obtained, which can be seen in Fig. 11. The XceptionNet model can predict the results of 104 data points correctly as the playing class, 1 data is mispredicted as the calm class, 100 data points are also predicted correctly as the calm class, and 5 data are mispredicted as the playing class.

Based on the test results shown in Figure 8-10, the performance of the ResNet, DenseNet, and XceptionNet architectures is evaluated on the play and calm classes of ECG signal data. The confusion matrix was used to measure the performance of each model by calculating accuracy, precision, recall, and f1-score using equations (13)-(16). The test results obtained demonstrate high accuracy in ECG signal classification, as shown in Table 1.

Table 1. The comparison of the accuracy achievements of the three architectures

Metrics	ResNet	DenseNet	XceptionNet
Accuracy	96,19%	94,76%	97,14%
Precision	96,19%	94,28%	95,23%
Recall	96,19%	95,19%	99%
F1-score	96,19%	94,73%	97,07%

Table 1 shows that XceptionNet achieved stable performance with values of 97.14% for accuracy, 95.23% for precision, 99% for recall, and 97.07% for F1-score. This indicates that XceptionNet is quite good at predicting classes in complex ECG signal patterns. The ResNet model also demonstrated high and stable performance across all metrics, achieving a commendable performance on the accuracy evaluation matrix with a value of 96.19%, precision with a value of

96.19%, recall with a value of 96.19%, and F1-score with a value of 96.19%. Its ability to handle vanishing gradients through shortcut connections makes it suitable for complex 1D signal processing, such as ECG. Meanwhile, DenseNet has a lower performance than the three models, although the difference is not significant. Its accuracy of 94.76%, precision of 94.28%, recall of 95.19%, and F1 score of 94.73% indicate that this model performs reasonably well in prediction, but slightly lower than the other two models. This is likely due to the overhead of dense connectivity, which can cause feature redundancy or difficulty in generalization on limited datasets.

The evaluation results using the 5-fold cross-validation scheme show that the performance of the three models varies significantly. The ResNet model achieves an average accuracy of $65.64\% \pm 16.52\%$, with significant fluctuations between folds. While some folds exhibit high performance (0.8786 and 0.8321), others show a drastic drop in accuracy to 0.5000. This suggests that ResNet is quite sensitive to variations in training and testing data, resulting in relatively poor stability. The DenseNet model shows the best performance with an average accuracy of $79.86\% \pm 13.63\%$ and a relatively consistent accuracy distribution across all folds, with some folds achieving over 90%. This consistency indicates that DenseNet has a better generalization ability compared to the other two architectures. Meanwhile, the XceptionNet model achieved an average accuracy of $60.36\% \pm 9.60\%$, a relatively stable value but one that tends to be lower than that of ResNet and DenseNet, thus categorizing its performance as less than optimal.

Overall, the addition of filtering and windowing resulted in improvements in accuracy and F1 compared to V0, with the most consistent impact on precision (decreased FP) for filtering and on recall (decreased FN) for windowing. Gaussian augmentation provided additional benefits by balancing classes and stabilizing performance across seeds. Compared to the unaugmented Filter+Window combination (V4), the entire pipeline (V7) demonstrated a significant improvement in F1 (Wilcoxon, $p < 0.05$) and a decrease in discordant pairs in the play class (McNemar, $p < 0.05$). The most significant improvement was observed in XceptionNet, consistent with its ability to efficiently extract 1D spatio-temporal features. These findings reinforce our design decisions for the preprocessing pipeline and explain why the complete configuration performed best on ECG data from autistic children.

IV. Discussion

As highlighted in Figs. 6, 7, and 8, the results of model training provide comprehensive insights into its

performance during the training process. The blue line, representing the training accuracy, indicates the extent to which the model accurately predicted the data at each epoch. In contrast, the orange line, which reflects the validation accuracy, demonstrates the model's ability to generalize to data not used during training.

Figures 6, 7, and 8 show the performance of each architecture (ResNet, DenseNet, and XceptionNet) in terms of accuracy and loss over 50 epochs. Fig. 6 shows the performance of the model using the ResNet architecture, where ResNet achieves a high training accuracy of approximately 1.0 after about 10 epochs. This indicates that the model performs well in considering the training data. Fig. 7 shows that DenseNet gives the most stable results compared to the other two models. Both training and validation accuracies increase consistently and converge close to 1.0, with minimal fluctuations. The loss graph shows a significant and relatively stable decrease in both training and validation data. This indicates that DenseNet has good generalization ability and minimizes overfitting. In Fig. 8, although the training accuracy of the XceptionNet model is high and stable, the validation accuracy is initially very low. It does not show any improvement until around the 20th epoch. After that, there is a sudden spike, but large fluctuations persist afterward. This indicates instability in training, most likely due to sensitivity to training parameters, such as learning rate or batch size.

In processing, using a bandpass + notch filter proved quite effective in reducing baseline wander and power-line noise, although impulsive motion artifacts and EMG bursts could still slip through. The 2-second windowing scheme provided good QRS morphological context but had limitations in capturing slower HRV dynamics. Gaussian augmentation contributed to regularization but did not fully represent real artifacts. Therefore, additional augmentations such as synthetic baseline drift, time warping, amplitude scaling, or impulse dropout could improve the model's robustness to non-Gaussian noise. Overall, the applied preprocessing steps were sufficient to suppress dominant noise components. Test results showed that XceptionNet was robust to noise variations resembling play signals, while ResNet was more balanced in dealing with both classes. Promising improvements include the implementation of more realistic motion artifact augmentation and a mechanism for rejecting faulty windows, enabling improved precision in calm conditions without compromising recall.

The classification results on the test data for each learning rate configuration are presented in the confusion matrix shown in Fig. 9, Fig. 10, and Fig. 11. Based on Table 1, XceptionNet shows the highest performance compared to the other two architectures, with an accuracy of 97.14%, precision of 95.23%, recall

of 99%, and F1-score of 97.07%. These figures indicate that XceptionNet can recognize ECG signal patterns more accurately and consistently. In contrast, ResNet exhibits competitive performance with an accuracy of 96.19%, which is only slightly lower than that of XceptionNet. Its precision, recall, and F1-score are also balanced, with a value of 96.19% in all metrics, indicating the stability of the model's performance. DenseNet, while still quite good, recorded the lowest performance among the three models, with an accuracy of 94.76% and an F1-score of 94.73%.

Further statistical analysis using ANOVA yielded an F-value of 2.22 with a p-value of 0.152, indicating no significant difference overall between the three models at the 0.05 level. Paired t-tests also supported these results, with comparisons between ResNet and DenseNet ($p = 0.159$) and ResNet and XceptionNet ($p = 0.517$) being insignificant. In contrast, the comparison between DenseNet and XceptionNet yielded a p-value of 0.079, which is close to significant. Thus, although DenseNet consistently demonstrates superior performance, particularly compared to XceptionNet, the difference is not strong enough to be declared statistically significant in the current sample size. These findings suggest that increasing the amount of data or using more cross-validation folds could strengthen the evidence for DenseNet's superiority in EEG classification.

In this study, the model's generalization ability was evaluated by dividing the dataset into three parts: training, validation, and test data. This allowed model performance to be monitored not only on the training data but also on data not used in the training process. To reduce the risk of overfitting, several regularization steps were applied, such as Batch Normalization, which stabilizes the activation distribution between layers, and data augmentation with Gaussian noise to increase the diversity of the ECG signal and balance the number of classes. The accuracy and loss graphs revealed a difference in performance between the training and validation data, particularly for ResNet and XceptionNet, suggesting a potential for mild overfitting. At the same time, DenseNet appeared more stable in maintaining a balance between training and validation accuracy. However, this study did not use dropout or early stopping techniques, allowing the model to be fully trained for up to 50 epochs. To strengthen the performance claims, the authors also applied the McNemar statistical test, which showed no significant difference between the three models, although XceptionNet had the highest accuracy. This shows that the preprocessing and augmentation steps carried out were reasonably practical in helping the model adapt to new data. However, additional strategies, such as dropout, early stopping, or k-fold cross-validation, can still be considered in future research to better control the potential for overfitting.

A close examination of the misclassification cases revealed three primary sources: recording artifacts (baseline wander, EMG noise, and 50 Hz power-line) that shift the model's attention to non-diagnostic segments; fragile cardiac morphology (low R amplitude, absent P waves, widened QRS) that obscures class separation; and limited temporal context or spatial information (e.g., PVC vs. LBBB shape equivalence in single-lead). Specifically, baseline wander and 50 Hz band dominance significantly increased the odds of misclassification ($OR \approx \dots$; $p < 0.05$), while a 2-second Hamming-based window decreased the FN for classes with low-amplitude P/T components. Saliency maps showed that in misclassified samples, the peak of attention shifted from the QRS onset toward the noise-rich pre-QRS segment, indicating the model's sensitivity to non-cardiac components. These results emphasize the importance of the filtering-windowing-augmentation combination in our pipeline, while also highlighting the model's limitations when overlapping morphology or multi-lead information is unavailable.

Our high-performance model is designed as a sensor-based, real-time monitoring system for children with autism. EEG (and optionally ECG/PPG) signals are acquired via a wearable device, preprocessed (0.5–40 Hz, 50 Hz notch, baseline wander control), and then analyzed over a 4-second window with a 1–2-second hop. Inference is executed on a quantized edge device (INT8/FP16) with <200 ms latency, resulting in an end-to-alert of ≤ 5 seconds. The output is presented as a colored risk score with a brief explanation (dominant features and attention maps), as well as relevant epoch examples for quick clinician verification. To address patient-specific variations, the system performs individualized baseline calibration, few-shot calibration, and threshold adjustment based on a moving baseline, along with drift detection that triggers a re-baseline when necessary. The one-button interface, automatic signal quality check, and offline-first mode facilitate ease of use by parents and healthcare professionals. All data is summarized and can be integrated into medical records, while encryption, anonymization, and audit trail policies ensure data governance. The system's decision-support role ensures clinicians remain in the decision-making loop. At the same time, stepwise validation (technology $\rightarrow$ clinical pilot $\rightarrow$ prospective study) connects our findings with practical impact in healthcare for children with ASD.

The developed ECG classification model has significant potential for application in monitoring children with neurodevelopmental disorders. However, several areas still require further development to improve its robustness, accuracy, and clinical utility. First, expanding the sample size and conducting cross-center testing are crucial to ensure generalizability,

particularly by including subjects from a wider age range, diverse demographic backgrounds, and various clinical conditions, such as ADHD, epilepsy, or the use of specific medications. Second, model evaluation should be strengthened with rigorous validation schemes, such as leave-one-subject-out or multi-site validation, and the use of clinically relevant metrics, including false alarms per hour and calibration reliability, so that model performance more closely matches real-world operational needs. Third, data development is also crucial, including label audits, realistic augmentations to mimic common artifacts (e.g., baseline wander or muscle noise), and self-supervised learning-based pretraining to make the model more robust to data variations.

Furthermore, exploring more advanced architectures, such as time-series-specific transformers or CNN-Transformer hybrids, as well as personalized approaches with fast calibration per patient, may help improve performance in individuals with unique signal morphologies. Multimodal integration with additional sensors (such as PPG, IMU, or EDA) also promises to reduce false positives due to motion artifacts and provide a richer physiological context. Ultimately, successful clinical implementation will depend on efficient edge computing implementation, simple interfaces for both healthcare professionals and parents, and a human-in-the-loop workflow that allows clinicians to maintain verification and control of decisions. This development direction will not only strengthen the model's robustness to technical and biological variations but also ensure that the system is truly useful in the clinical monitoring and intervention of children with ASD.

This finding concludes that the lightweight CNN model is effective in classifying biological signals. One of the factors that drives high performance is the preprocessing process, where the signal is filtered and divided into a window of 500 samples, allowing the model to recognize the signal pattern effectively.

Table 2. Comparison of Classification Performance in Previous Studies

Reference	Method	Test Accuracy
[33]	ResNet	91,49%
[39]	DenseNet	94,52%
[40]	XceptionNet	96%
This study	ResNet	96,19%
This study	DenseNet	94,76%
This study	XceptionNet	97,14%

From Table 2, a comparative analysis of the results obtained in this study with those of previous studies that used ECG signal classification to distinguish between calm and play signals is presented. Previous

studies using the ResNet architecture [33] achieved a training accuracy of 91.49%. Other studies using the DenseNet architecture [39] obtained a training accuracy of 94.52%, while studies using the XceptionNet architecture [40] achieved a training accuracy of 96%. In contrast, this study's results show that the ResNet architecture achieved 96.19%, the DenseNet architecture achieved 94.76%, and the XceptionNet architecture achieved 97.14%.

The results of this study indicate that XceptionNet achieved the highest accuracy (97.14%), slightly better than ResNet (96.19%) and DenseNet (94.76%). This difference aligns with the findings of [40], who reported XceptionNet's superiority on 1D signals due to its use of depthwise separable convolution, which is capable of extracting spatial features more efficiently. However, when compared to the study by [33], which used ResNet on 1D music data with an accuracy of 91.49%, our results are significantly higher. This is likely influenced by the signal preprocessing stages (Butterworth filtering, windowing, and Gaussian augmentation), which significantly reduce noise and balance the class distribution.

Furthermore, differences in dataset characteristics also play a role. Previous studies generally used public datasets or other physiological signals that have higher variability [39]. In contrast, the dataset in this study was relatively homogeneous (10 subjects with a similar age range, recorded using an AD8232 sensor). This homogeneity helped the model achieve more stable performance, although it limited its generalizability.

In the context of clinical implementation, the trade-off between model accuracy and computational complexity is a key consideration. The results showed that XceptionNet provided the highest accuracy (97.14%) with a recall of 99%, making it superior in detecting complex patterns in the ECG signals of autistic children. This advantage is particularly important in the clinical setting, as high recall can minimize false-negative errors, thereby reducing the risk of missing abnormal conditions. However, the complexity of the XceptionNet architecture makes it computationally demanding, sensitive to training parameters, and requires GPU infrastructure or servers with high processing power. In contrast, ResNet, with an accuracy of 96.19%, offers a better balance between performance and efficiency. The architecture is lighter than XceptionNet, stable during training, and still capable of producing competitive results. This makes ResNet more realistic for integration on resource-constrained devices, such as portable or wearable systems for real-time health monitoring. Meanwhile, DenseNet demonstrated the lowest performance (94.76%) due to potential feature redundancy on small datasets. Therefore, despite its relative stability, this model is less optimal than the

other two architectures. Thus, XceptionNet is appropriate for a clinical infrastructure that supports intensive computing. At the same time, ResNet is more suitable for practical field applications that demand efficiency, and DenseNet should not be the primary choice.

However, this study has several limitations that should be taken into consideration. First, the data used was relatively limited, involving only ten subjects with a total of 700 samples per class after the augmentation process. This condition may impact the model's ability to capture broader physiological variations, which may result in generalization results that are not fully representative of the general population of children with autism. Second, this study employed a mild regularization technique, namely Batch Normalization and augmentation with Gaussian noise, without utilizing other methods such as dropout, early stopping, or k-fold cross-validation, which may be more effective in reducing overfitting. Third, the model evaluation was only conducted on two conditions (play and calm), so the model's performance in other physiological conditions is unknown. Therefore, future research is recommended to use a larger and more diverse dataset and implement additional regularization strategies to make the model more robust and able to achieve stronger generalization.

V. Conclusion

The results showed that the filtering stage reduced the signal's sensitivity to noise, while augmentation increased data diversity and balanced the classes. Both steps significantly contributed to improving the model's accuracy and generalizability, thus strengthening its role as a crucial component in the ECG signal classification methodology for autistic children. Based on the test results of three CNN architectures, namely, ResNet, DenseNet, and XceptionNet, in classifying ECG signals from autistic children in two conditions (playing and calm), it can be concluded that all models provide high performance with accuracy above 94%. However, the XceptionNet model exhibits the performance, with accuracy, precision, recall, and F1-score of 97,14% each. The ResNet model also exhibits excellent performance, achieving an accuracy of 96,19%, along with stable evaluation results. On the other hand, DenseNet has a slightly lower performance, with an accuracy of 94.76%, and precision and recall of 94.28% and 95.23%, which is still in the good category. This difference is likely due to the complexity of the DenseNet structure, which results in feature redundancy in smaller datasets. Overall, this study demonstrates that the use of CNN architecture in classifying ECG signals of autistic children is auspicious, especially with deep architectures such as XceptionNet. These findings can be the basis

for the development of an automatic and non-invasive ECG signal-based heart monitoring system for children with special needs.

Acknowledgment

We would like to express our sincere gratitude to the Department of Electrical and Computer Engineering at Syiah Kuala University for the invaluable support and resources provided during this research. The facilities, academic environment, and encouragement from the lecturers have significantly contributed to the completion of this research. This research would not have been possible without the institution's commitment to advancing research and innovation in this field.

Funding

This research received a specific grant from LPPM of Universitas Syiah Kuala through a Lecturer Research grant with the number 318/UN11.2.1/PT.01.03/PNBP/2023.

Data Availability

A collection of data was generated or analyzed during this study.

Author Contributions

All authors contributed to this research. Yunidar designed the central concept and idea of this research, was responsible for funding acquisition, conducting the investigation and data collection, and writing the initial draft of the manuscript. Melinda was responsible for developing and designing the research methodology. Albahri supervised the research process and validated the results. Herlina Dimiati performed the formal statistical analysis and data interpretation. Hanum Aulia managed the data curation and created data visualizations in the form of tables and figures. Nurlida Basir performed a critical review and final editing of the article. All authors have read, reviewed, and approved the final version of the article for publication.

Declarations

Consent for Publication Participants.

Consent for publication was given by all participants

Competing Interests

The authors declare no competing interests.

References

- [1] A. Miranda, C. Berenguer, I. Baixauli, and B. Roselló, "Childhood language skills as predictors of social, adaptive and behavior outcomes of adolescents with autism spectrum disorder," *Res. Autism Spectr. Disord.*, vol. 103, no. March, 2023, doi: 10.1016/j.rasd.2023.102143.
- [2] E. C. McCanlies *et al.*, "Parental occupational exposure to solvents and autism spectrum disorder: An exploratory look at gene-environment interactions," *Environ. Res.*, vol. 228, Jul. 2023, doi: 10.1016/j.envres.2023.115769.
- [3] Z. Wu, "Challenges Encountered by Children with Autism Spectrum Disorder: from the perspective of academic performances and education service providers," *Highlights Business, Econ. Manag.*, vol. 4, pp. 263–271, 2022, doi: 10.54097/hbem.v4i.3500.
- [4] I. da Costa, R. B. Junqueira, D. S. Faé, L. A. P. de Souza, and C. A. A. Lemos, "Increased Risk of Dentoalveolar Trauma in Patients with Autism Spectrum Disorder: A Systematic Review with Meta-Analysis," *Int. J. Environ. Res. Public Health*, vol. 21, no. 12, 2024, doi: 10.3390/ijerph21121563.
- [5] C. Bosetti, L. Ferrini, A. R. Ferrari, E. Bartolini, and S. Calderoni, "Children with Autism Spectrum Disorder and Abnormalities of Clinical EEG: A Qualitative Review," *J. Clin. Med.*, vol. 13, no. 1, 2024, doi: 10.3390/jcm13010279.
- [6] A. Bellato, I. Arora, P. Kochhar, D. Ropar, C. Hollis, and M. J. Groom, "Heart Rate Variability in Children and Adolescents with Autism, ADHD and Co-occurring Autism and ADHD, During Passive and Active Experimental Conditions," *J. Autism Dev. Disord.*, vol. 52, no. 11, pp. 4679–4691, Nov. 2022, doi: 10.1007/s10803-021-05244-w.
- [7] A. Ullah, S. U. Rehman, S. Tu, R. M. Mehmood, Fawad, and M. Ehatisham-UI-haq, "A hybrid deep CNN model for abnormal arrhythmia detection based on cardiac ECG signal," *Sensors (Switzerland)*, vol. 21, no. 3, pp. 1–13, Feb. 2021, doi: 10.3390/s21030951.
- [8] C. N. Nurbadiani, M. Melinda, Y. Yunidar, and F. Arnia, "Electrocardiogram Detection System of Autistic Children Based on AD8232 for Healthcare," in *Proceeding - 2023 2nd International Conference on Computer System, Information Technology, and Electrical Engineering: Sustainable Development for Smart Innovation System, COSITE 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 126–131, doi: 10.1109/COSITE60233.2023.10250117.
- [9] D. Tilwani, J. Bradshaw, A. Sheth, and C. O'Reilly, "ECG Recordings as Predictors of Very Early Autism Likelihood: A Machine Learning Approach," *Bioengineering*, vol. 10, no. 7, Jul. 2023, doi: 10.3390/bioengineering10070827.
- [10] S. Suganyadevi, V. Seethalakshmi, and K. Balasamy, "A review on deep learning in medical

- image analysis," *Int. J. Multimed. Inf. Retr.*, vol. 11, no. 1, pp. 19–38, 2022, doi: 10.1007/s13735-021-00218-1.
- [11] Fräsch, M. G., Shen, C., Wu, H.-T., Müller, A., Neuhaus, E., Bernier, R. A., Kamara, D., & Beauchaine, T. P. (2021). Can a composite heart rate variability biomarker shed new insights about autism spectrum disorder in school-aged children? *Journal of Autism and Developmental Disorders*, 51(1), 346–356. <https://doi.org/10.1007/s10803-020-04467-7>.
- [12] L. Billeci *et al.*, "Heart rate variability during a joint attention task in toddlers with autism spectrum disorders," *Front. Physiol.*, vol. 9, no. MAY, pp. 1–11, 2018, doi: 10.3389/fphys.2018.00467.
- [13] A. Q. Alban *et al.*, "Heart Rate as a Predictor of Challenging Behaviours among Children with Autism from Wearable Sensors in Social Robot Interactions," *Robotics*, vol. 12, no. 2, 2023, doi: 10.3390/robotics12020055.
- [14] M. Ashtiyani, S. Navaei Lavasani, A. Asgharzadeh Alvar, and M. R. Deevband, "Heart rate variability classification using support vector machine and genetic algorithm," *J. Biomed. Phys. Eng.*, vol. 8, no. 4, pp. 423–434, 2018, doi: 10.31661/jbpe.v0i0.614.
- [15] F. Khan, X. Yu, Z. Yuan, and A. ur Rehman, "ECG classification using 1-D convolutional deep residual neural network," *PLoS One*, vol. 18, no. 4 April, Apr. 2023, doi: 10.1371/journal.pone.0284791.
- [16] G. Husain, A. Siddiqua, and M. Toma, "Evaluating the Performance of DenseNet in ECG Report Automation," *Electron.*, vol. 14, no. 9, May 2025, doi: 10.3390/electronics14091837.
- [17] Y. Gotmare, "A Comparative Study of Feature Extraction Models for Image Caption Generation," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. 4, pp. 4821–4828, 2024, doi: 10.22214/ijraset.2024.61114.
- [18] Z. Wang, Z. Zheng, D. Song, and X. Xu, "A High-Speed Train Axle Box Bearing Fault Diagnosis Method Based on Dimension Reduction Fusion and the Optimal Bandpass Filtering Demodulation Spectrum of Multi-Dimensional Signals," *Machines*, vol. 12, no. 8, Aug. 2024, doi: 10.3390/machines12080571.
- [19] L. J. Gonçalves, K. Farias, L. Kupssinskü, and M. Segalotto, "The effects of applying filters on EEG signals for classifying developers' code comprehension," *J. Appl. Res. Technol.*, vol. 19, no. 6, pp. 584–602, Dec. 2021, doi: 10.22201/icat.24486736e.2021.19.6.1299.
- [20] T. N. A. M. A. I. A. A. and W. A. Saleh Albahli, "AEI-DNET: A Novel DenseNet Model with an Autoencoder for the Stock Market Predictions Using Stock Technical Indicators," *Electronics*, 2022.
- [21] W. Li and J. Gao, "Automatic sleep staging by a hybrid model based on deep 1D-ResNet-SE and LSTM with single-channel raw EEG signals," *PeerJ Comput. Sci.*, vol. 9, 2023, doi: 10.7717/peerj-cs.1561.
- [22] A. A. Fedotov, "Selection of Parameters of Bandpass Filtering of the ECG Signal for Heart Rhythm Monitoring Systems," *Biomed. Eng. (NY)*, vol. 50, no. 2, pp. 114–118, 2016, doi: 10.1007/s10527-016-9600-8.
- [23] B. Salsekar, "FILTERING OF ECG SIGNAL USING BUTTERWORTH FILTER AND ITS FEATURE EXTRACTION." [Online]. Available: <http://www.physionet.org/cgi-bin/atm/ATM>
- [24] P. K. Govarthan, S. K. Peddapalli, N. Ganapathy, and J. F. A. Ronickom, "Emotion classification using electrocardiogram and machine learning: A study on the effect of windowing techniques," *Expert Syst. Appl.*, vol. 254, Nov. 2024, doi: 10.1016/j.eswa.2024.124371.
- [25] N. P. Martono and H. Ohwada, "Evaluating the Impact of Windowing Techniques on Fourier Transform-Preprocessed Signals for Deep Learning-Based ECG Classification," *Hearts*, vol. 5, no. 4, pp. 501–515, Oct. 2024, doi: 10.3390/hearts5040037.
- [26] H. Siswono, W. Widyastuti, Y. Dovan, and D. Nurainingsih, "Design of Fir Digital Bandpass Filter with Hamming Window and Hanning Window Method for Fetal Doppler," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 9, no. 4, pp. 912–926, Oct. 2023, doi: 10.26555/jiteki.v9i4.26849.
- [27] A. Pant, A. Kumar, C. Verma, and Z. Illés, "Comparative exploration on EEG signal filtering using window control methods," *Results Control Optim.*, vol. 17, Dec. 2024, doi: 10.1016/j.rico.2024.100485.
- [28] A. Kebaili, J. Lapuyade-Lahorgue, and S. Ruan, "Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review," Apr. 01, 2023, *MDPI*. doi: 10.3390/jimaging9040081.
- [29] Y. Dai, Y. Gao, and F. Liu, "Transmed: Transformers advance multi-modal medical image classification," *Diagnostics*, vol. 11, no. 8, Aug. 2021, doi: 10.3390/diagnostics11081384.
- [30] K. Munadi *et al.*, "A Deep Learning Method for Early Detection of Diabetic Foot Using Decision Fusion and Thermal Images," *Appl. Sci.*, vol. 12, no. 15, Aug. 2022, doi: 10.3390/app12157524.
- [31] F. L. Becerra-Suarez, H. Alvarez-Vasquez, and M. G. Forero, "Improvement of Bank Fraud Detection

- Through Synthetic Data Generation with Gaussian Noise," *Technologies*, vol. 13, no. 4, 2025, doi: 10.3390/technologies13040141.
- [32] W. Ge, P. Lalbakhsh, L. Isai, A. Lensky, and H. Suominen, "Comparing Deep Learning Models for the Task of Volatility Prediction Using Multivariate Data," Jun. 2023, [Online]. Available: <http://arxiv.org/abs/2306.12446>
- [33] P. C. Chang, Y. S. Chen, and C. H. Lee, "MS-SincResNet: Joint learning of 1D and 2D kernels using multi-scale SincNet and ResNet for music genre classification," in *ICMR 2021 - Proceedings of the 2021 International Conference on Multimedia Retrieval*, Association for Computing Machinery, Inc, Aug. 2021, pp. 29–36. doi: 10.1145/3460426.3463619.
- [34] S. U. Kim and J. Y. Kim, "Improving Human Activity Recognition Through 1D-ResNet: A Wearable Wristband for 14 Workout Movements," *Processes*, vol. 13, no. 1, Jan. 2025, doi: 10.3390/pr13010207.
- [35] M. M. A. Rahhal, Y. Bazi, H. Alhichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Inf. Sci. (Ny)*, vol. 345, pp. 340–354, 2016, doi: 10.1016/j.ins.2016.01.082.
- [36] A. A. Laghari, Y. Sun, M. Alhussein, K. Aurangzeb, M. S. Anwar, and M. Rashid, "Deep residual-dense network based on bidirectional recurrent neural network for atrial fibrillation detection," *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-40343-x.
- [37] A. Ayad and M. E. Abdulmunim, "Detecting Abnormal Driving Behavior Using Modified DenseNet," *Iraqi J. Comput. Sci. Math.*, vol. 4, no. 3, pp. 48–65, 2023, doi: 10.52866/ijcsm.2023.02.03.005.
- [38] F. Saleem, Z. Ahmad, M. F. Siddique, M. Umar, and J. M. Kim, "Acoustic Emission-Based Pipeline Leak Detection and Size Identification Using a Customized One-Dimensional DenseNet," *Sensors*, vol. 25, no. 4, Feb. 2025, doi: 10.3390/s25041112.
- [39] A. Tibrewal, Shikha, and D. Sethia, "The Effectiveness of Advance Deep Learning Architectures for Classification of Stress using Raw EEG Data," in *2024 Asia Pacific Conference on Innovation in Technology, APCIT 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/APCIT62007.2024.10673521.
- [40] I. Vallés-Pérez, J. Gómez-Sanchis, M. Martínez-Sober, J. Vila-Francés, A. J. Serrano-López, and E. Soria-Olivas, "End-to-end Keyword Spotting using Xception-1d," Oct. 2021, [Online]. Available: <http://arxiv.org/abs/2110.07498>
- [41] H. Xu, "Users' perceptions of managerial measurements for cloud computing's cybersecurity – importance-performance analysis," *Inf. Comput. Secur.*, vol. ahead-of-p, no. ahead-of-print, Jan. 2025, doi: 10.1108/ICS-12-2023-0264.
- [42] S. Szeghalmy and A. Fazekas, "A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning," *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23042333.

Author Biography



Yunidar was born in Banda Aceh, Aceh, on June 29, 1974. She has been a lecturer at the Faculty of Engineering, Department of Electrical and Computer Engineering, Syiah Kuala University, since March 2000.

After completing her undergraduate education in Physics at Syiah Kuala University, Aceh, Indonesia, in 1997, she then obtained a master of Engineering (MT) degree in Optoelectrotechnics and Laser Applications from the University of Indonesia, Jakarta, Indonesia, in 2000. After which, she pursued a doctoral degree program in Electrical and Computer Engineering at Syiah Kuala University, she graduated in 2025. She is also a member of IEEE. Her research interests include the implementation of biomedical engineering and the use of sensors in biomedical applications, including multimedia. She can be contacted via email: yunidar@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 at 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 at email: melinda@usk.ac.id.



Al Bahri received a bachelor's degree in 2015 and a master's degree in 2018 from the School of Electrical Engineering and Informatics, Bandung Institute of Technology (Institut Teknologi Bandung). He has been with the Department of Electrical Engineering, Faculty of Engineering, Syiah Kuala University since March 2018. His research interests include digital media technology, games, and computers. He had conducted some research in various fields, including cybersecurity, wireless communication, machine learning, and smart systems development. His work involved applying AI techniques, signal processing, and microcontroller-based automation to solve real-world problems in network security, object detection, and information systems. His research interests include wireless telecommunication and deep learning. He can be contacted via email: albahri@usk.ac.id



Hanum Aulia Ramadhani was born in Aceh Besar on November 4, 2004. She is a student in the Electrical and Computer Engineering Department at Syiah Kuala University. Her undergraduate studies focused on multimedia technology, and her research involves the analysis of ECG signals. She actively attends classes and continues to develop her knowledge in her field. Enrolled in the class of 2022, she is committed to expanding her expertise and gaining practical experience. Her academic journey reflects her dedication to theoretical and applied learning, which prepares her to contribute to the advancement of technology in her field. She can be reached at: hanumaulia@mhs.usk.ac.id.



Herlina Dimiati was born on March 5, 1964, in Banda Aceh. She graduated as a doctor in 1992 from the Faculty of Medicine at USK and continued her specialization in Pediatrics at the Faculty of Medicine, University of North Sumatra, completing her studies in 2002. In 2006, she continued his pediatric cardiology education at Cipto Mangunkusumo Hospital in Jakarta and graduated in 2008. Continuing doctoral studies at the Faculty of Medicine, Gadjah Mada University in 2011 and completed in 2016. Currently, she has held the title of Professor in child health at the Faculty of Medicine USK since 2021. She can be reached at: herlinadimiati@usk.ac.id



Dr. Nurlida Basir is an Associate Professor at the Faculty of Science and Technology, Universiti Sains Islam Malaysia (USIM). She began her academic career at USIM in 2002 and has since been actively involved in teaching, research, and educational leadership. She holds a Diploma, a Bachelor's, and a Master's degree in Computer Science from Universiti Teknologi Malaysia (UTM), and earned her Ph.D. in Computer Science from the University of Southampton, United Kingdom. Her research interests span software engineering, cybersecurity, malware detection, signal processing, and artificial intelligence. Her research has been extensively published in prominent academic journals and conference proceedings. Alongside her research, she is a dedicated educator, mentoring both undergraduate and postgraduate students in computer science. She is a member of the Institute of Electrical and Electronics Engineers (IEEE), reflecting her active participation in the global academic and research community. She can be contacted at nurlida@usim.edu.my.