Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 11, No. 2, May, Pp. 273-284

Enhancing CNN performance for alzheimer's disease classification through genetic algorithm optimization

Wildan Arif Maulana[1], Zainul Abidin*[1], Rahmadwati[1]
Department of Electrical Engineering, Universitas Brawijaya, Indonesia[1]

Article Info
Keywords:
Alzheimer, Albumentation, CutMix, Genetic Algorithm, Optimization

Article history:
Received: October 21, 2025
Accepted: December 25, 2025
Published: May 01, 2026

Cite:
W. A. Maulana, Z. Abidin, and R. Rahmadwati, "Enhancing CNN Performance for Alzheimer's Disease Classification through Genetic Algorithm Optimization", *KINETIK*, vol. 11, no. 2, May. 2026. https://doi.org/10.22219/kinetik.v11i2.2543

*Corresponding author.
Zainul Abidin
E-mail address:
zainulabidin@ub.ac.id

Abstract
The rise in global life expectancy has contributed to a rapidly expanding elderly population and a corresponding increase in Alzheimer's disease cases, highlighting the need for more accurate and objective diagnostic methods. Although MRI is widely used for brain assessment, early-stage Alzheimer's detection remains challenging because structural differences between disease stages are often subtle and prone to subjective interpretation by clinicians. To address this limitation, this study proposes a custom Convolutional Neural Network (CNN) developed from scratch for classifying Alzheimer's disease using brain MRI images. Data diversity was enhanced through augmentation comparison strategies, including Albumentations, which achieved 84.8% accuracy; CutMix, which achieved 88.3% accuracy, and a combined Albumentations-CutMix approach, which enabled the base model to achieve 92.1% classification accuracy. Subsequently, a Genetic Algorithm (GA) was applied to optimize key hyperparameters, enabling efficient exploration of the solution space compared to manual tuning and improving model performance to 96.4% accuracy. The optimized model demonstrated improved stability and generalization across all classes, highlighting the capability of the proposed computational framework to function as a reliable tool for supporting the early detection of Alzheimer-related cognitive decline.

1. Introduction

Alzheimer's disease is a degenerative brain disorder that progressively worsens over time and ultimately leads to fatal outcomes by gradually damaging brain function, resulting in memory loss and cognitive impairment [1]. As life expectancy increases, the elderly population worldwide, including in Indonesia, continues to grow. This situation poses new challenges in the healthcare sector, particularly the increasing prevalence of progressive neurological disorders such as Alzheimer's disease. Currently, an estimated 50 million people worldwide are living with Alzheimer's disease, and this number is projected to more than double by 2050, with the number of patients tending to double every five years [2]. Early detection of Alzheimer's disease is critically important because treatment initiated at the earliest possible stage can help delay symptom progression while improving patients' overall quality of life. To support diagnosis, Magnetic Resonance Imaging (MRI) is widely used as a non-invasive imaging modality that employs strong magnetic fields to produce highly detailed anatomical images of the brain [3]. Through MRI analysis, clinicians can observe hippocampal atrophy, measure reductions in cortical volume, and identify degeneration of other brain structures that commonly serve as early indicators of neurodegenerative processes [4]. However, MRI interpretation still relies heavily on the expertise of medical personnel, increasing the risk of subjectivity and diagnostic errors [5]. Artificial intelligence technology, particularly deep learning, now offers a solution to overcome these limitations through automated medical image analysis.

On the other hand, Convolutional Neural Networks (CNNs) have become the leading and most commonly used architecture for medical image analysis, including MRI, CT scans, and X-ray imaging [6]. CNNs excel in recognizing spatial patterns and visual features, achieving up to 91.35% accuracy in classifying the severity of Alzheimer's disease [7]. However, one of the main challenges is the limited availability of training data, which can lead to overfitting. To address data limitations, data augmentation is used as the main strategy by increasing the diversity of training images through various transformations. One of the most widely adopted libraries due to its efficiency and flexibility is Albumentations [8]. Previous studies have shown that increasing the diversity of training images and reducing overfitting through Albumentations improved model accuracy to 82.14% [9]. In addition, CutMix effectively improves model performance by combining two images and their labels proportionally, thereby enriching training data diversity. CutMix has proven highly effective when applied to ResNet-50 for MRI tumor image classification, both in terms of classification accuracy and confidence calibration. This technique was reported to increase the Area Under the Receiver Operating Characteristic Curve (AUROC) by up to 14.1% [10].

To improve model performance, a Genetic Algorithm (GA) is employed as an optimization technique for hyperparameter tuning by mimicking natural evolutionary mechanisms to identify the optimal combination of learning rate, number of neurons, dropout rate, and batch size without requiring time-consuming manual searches [11]. In MobileNet-based models, GA optimization successfully increased training accuracy to 97% and improved the F1-score to 87% [12]. These results prove that GA effectively improves prediction accuracy and consistency.

Although previous studies have demonstrated performance improvements using CNNs, augmentation techniques, and GA-based optimization, several issues persist, including limited generalization capability, narrow augmentation exploration, and the lack of comprehensive hyperparameter tuning in some studies. Table 1 summarizes key related studies and their corresponding limitations.

Table 1. Comparison Related Studies

No	Researcher	Architecture	Augmentation	Weakness
1	Elnaghi [13]	VGG-16, ResNet-50, 2DCNN	Rotation, Flip, Zoom	Weak generalization
2	Kumar [14]	CNN, VGG-16, MobileNet, DenseNet	Rotation, Flip, Contrast	Narrow augmentation exploration
3	Kamardi [15]	CNN, VGG-16, DenseNet	Albumentation, Oversampling	Lack of hyperparameter optimization

The core issue investigated in this research concerns the absence of an integrated and optimized deep learning pipeline that combines advanced augmentation strategies with GA-based hyperparameter tuning for multi-class Alzheimer's MRI classification. To address this issue, the research focused on constructing a specialized convolutional neural architecture capable of distinguishing Alzheimer's severity levels based on neuroimaging scans, evaluating the effectiveness of Albumentations, CutMix, and their combined augmentation strategy, and applying GA optimization to enhance overall model performance. Unlike existing studies that examine augmentation or optimization techniques in isolation, this research integrates both components into a unified framework to maximize classification robustness and generalization capability. The contributions of this research include a systematic comparison of three augmentation approaches to evaluate their influence on data diversity and overfitting reduction, the design of an optimized CNN architecture tailored for multi-class Alzheimer's classification, and the integration of GA-driven hyperparameter tuning that achieves higher classification accuracy along with improved model stability. Collectively, these contributions establish a reliable and streamlined pipeline for enhancing early-stage Alzheimer's disease detection.

2. Research Method

The research stages for Alzheimer's disease severity classification using deep learning and GA optimization follow the workflow illustrated in Figure 1. The study begins with a literature review, followed by data preparation and augmentation design using Albumentations, CutMix, and a combined Albumentation-CutMix strategy. The augmented data are subsequently used for CNN-based classification optimized using a GA. Finally, the results are analyzed to evaluate the model's performance in Alzheimer's disease detection.

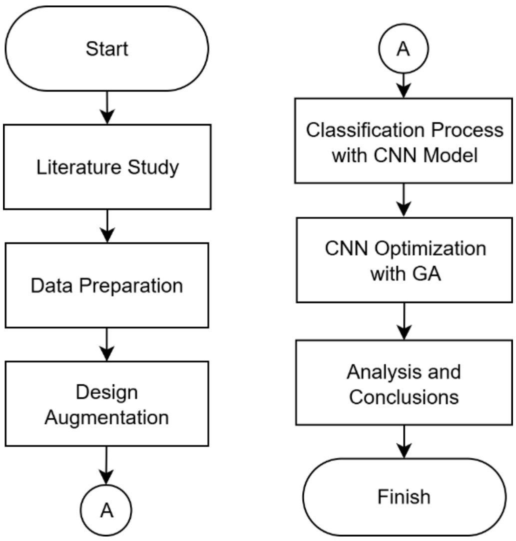


Figure 1. Research Stages

2.1 Dataset Preparation

This research utilized a dataset from Alzheimer's Disease Neuroimaging Initiative (ADNI), consisting of 6,400 axial-view brain MRI images with dimensions of 176 × 208 pixels. The dataset is categorized into four classes, representing different levels of Alzheimer's severity [16]. Specifically, the dataset comprises 3,200 Non-Demented images, 2,240 Very Mild Demented images, 896 Mild Demented images, and 64 Moderate Demented images.

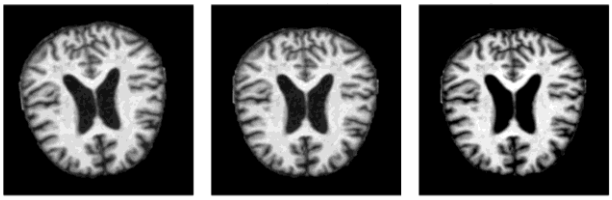
2.2 Data Splitting

For subsequent analysis, the dataset was divided into three independent subsets, with 70% allocated for training, 20% for validation, and 10% for testing. This partitioning strategy aims to ensure balanced model evaluation while minimizing the risk of overfitting during training [17]. Due to dataset imbalance, particularly the limited number of Moderate Demented samples, data augmentation and resampling techniques were applied to address the uneven class distribution and improve generalization capability.

The effectiveness of deep learning architectures strongly depends on well-distributed training data to achieve optimal performance and reduce overfitting. Since collecting additional MRI data was not feasible, resampling techniques, particularly oversampling, were employed to address class imbalance [18]. Oversampling was applied exclusively to the training set to preserve all majority-class samples while improving minority-class representation [19]. Although this approach effectively mitigates class imbalance, it may increase the risk of overfitting due to duplicated samples. To reduce this risk and improve model generalization, data augmentation techniques were implemented using the Albumentations library. Furthermore, this research investigates the use of Albumentations, CutMix, and their combined application to generate more diverse training samples and enhance classification performance.

2.3 Data Augmentation

This approach combines the two techniques to generate richer and more complex training data, enabling the model to learn greater pattern diversity compared to using a single technique alone. In this study, the CutMix technique is applied first, followed by Albumentations, as illustrated in Figure 2.

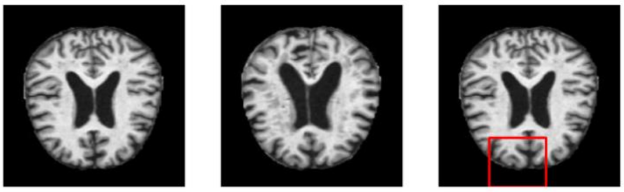


(a) (b) (c)

Figure 2. Albumentations

The Albumentations library provides a wide range of images transformation techniques, including both fundamental and advanced operations [9]. Figure 2 further illustrates these transformations, where (a) represents a rotation transformation of less than 10°, (b) shows a horizontal flip transformation, and (c) demonstrates brightness and contrast adjustments with a 20% probability randomly applied to the input images during training. This visualization highlights how different augmentation operations alter image appearance while preserving essential diagnostic features. These augmentation processes allow the data distribution to better reflect practical clinical variations, enabling the network to generalize more effectively and improve diagnostic discrimination across Alzheimer's severity categories.

Figure 3 illustrates an example of the CutMix technique applied to a medical images, where (a) represents the original base image, (b) shows the image patch extracted from another sample, and (c) depicts the final composite image generated through the CutMix process.



(a) (b) (c)

Figure 3. CutMix

The CutMix technique combines two images within the spatial domain by cutting a region from one image, overlaying it onto another image, and proportionally blending their corresponding labels based on the replaced region

[20]. This approach effectively enhances traning sample diversity by integrating spatial and semantic information from multiple images. By exposing the model to composite visual patterns, CutMix encourages the network to focus on more discriminative and localized features rather than relying on background correlations. Consequently, this technique improves model generalization and robustness, particularly in medical imaging tasks where data variability is often limited. In this study, the mixture proportion was limited to 20% to avoid compromising critical information in the MRI images.

2.4 Modelling CNN Classification

The next stage involves designing the proposed CNN model, which serves as the primary architecture in this study. The model is designed to develop a network structure capable of effectively extracting critical patterns from MRI data to support reliable classification of different stages of Alzheimer's disease. The CNN model employed in this study is a custom architecture developed entirely from scratch without relying on pre-trained models. This architecture was specifically constructed to align with the characteristics of the Alzheimer's MRI dataset, with the objective of efficiently extracting essential features while achieving optimal classification performance. The complete architecture of the proposed CNN framework is illustrated in Figure 4, showing the sequential arrangement of feature extraction and classification layers used to process MRI data.

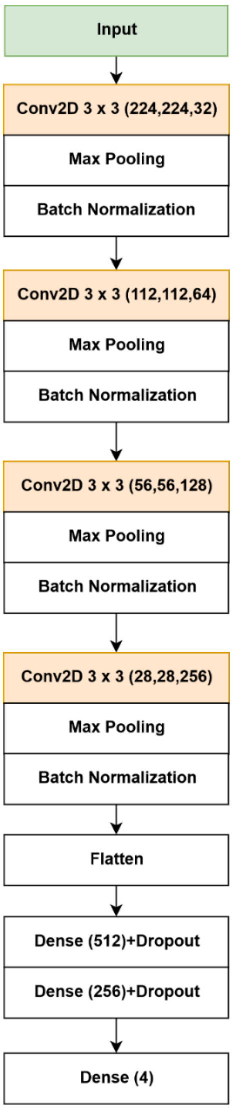


Figure 4. CNN Model

The proposed CNN framework incorporates a sequence of feature extraction, normalization, and regularization layers designed to mitigate overfitting during training. Each convolutional layer is intended to capture spatial patterns within the images, while the pooling layers are employed to reduce feature dimensionality without losing essential information. Following the feature extraction process, the resulting representations are passed through fully connected layers, which function as the classification component responsible for identifying and categorizing the progression stages of Alzheimer’s disease.

2.5 Model Evaluation

In the development of classification models based on deep learning techniques, particularly for medical data related to Alzheimer’s disease severity assessment, the evaluation stage serves as a critical component. The effectiveness of the classification models is generally assessed using standard performance metrics, including accuracy, precision, recall, and the F1-score, as defined in Equations 1-4.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - Score = 2x\frac{Precision\ x\ Recall}{Precision + Recall} \quad (4)$$

Table 2. Confusion Matrix

	Positive (P)	Negative (N)
Positive (P)	True Positive (TP)	False Positive (FN)
Negative (N)	False Positive (FP)	True Negative (TN)

The confusion matrix presented in Table 2 provides a comprehensive overview of the model’s prediction behavior by illustrating both correctly classified and misclassified test samples [21].

2.6 GA Optimization

The GA functions in CNN optimization by identifying optimal network structures, hyperparameters, or weight configurations. Inspired by the principle of natural selection, GA represents a population-based optimization approach that repeatedly performs selection, crossover, and mutation to evolve candidate solutions [22]. It effectively addresses complex optimization problems characterized by large search spaces that are difficult to solve using deterministic approaches. By mimicking evolutionary processes, GA efficiently explores the solution space and identifies optimal parameter combinations. Figure 5 illustrates the overall GA optimization workflow.

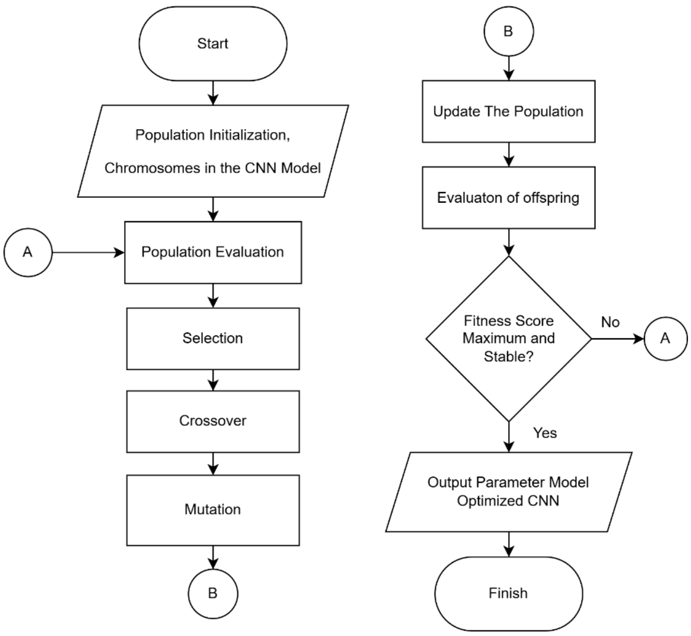


Figure 5. GA Optimization Process

In this study, GA is utilized to optimize key CNN parameters, including kernel size, number of filters, batch size, learning rate, and the number of neurons in hidden layers [23]. The optimization process begins with population initialization, where each chromosome represents a unique parameter combination. The population is then evaluated using a fitness function, such as validation accuracy, and the best-performing individuals are selected as parent candidates. Through crossover and mutation operations, new offspring are generated to explore improved solutions [24]. This iterative process continues until the fitness value converges, resulting in an optimized CNN configuration expected to achieve superior classification performance.

a. Genetic Algorithm Parameters

The GA parameters used for optimizing the CNN model hyperparameters are summarized in Table 3, which presents the configuration settings applied during the GA-based hyperparameter optimization process.

Table 3. GA Parameters

Parameters	Value
Population	10
Generation	5
Crossover Method	Two-point crossover

In this study, a population size of 10 individuals was used for each generation, with the evolutionary process executed over five generations to iteratively refine parameter combinations through selection, crossover, and mutation operations. The fitness function was defined based on validation accuracy, with the objective of maximizing model performance. A Two-Point Crossover method was employed to exchange genetic segments between individuals, thereby promoting population diversity and increasing the likelihood of obtaining optimal CNN hyperparameter configurations.

b. Population Initialization

The population initialization stage involves generating an initial set of candidate solutions (individuals) that form the first generation in the GA, as presented in Table 4. Each individual, or chromosome, in the initial population represents a unique CNN architecture configuration, including parameters such as kernel size, number of filters in each convolutional layer, batch size, learning rate, and the number of neurons in the fully connected layer [24].

Table 4. Population Initialization

Parameters	Value
Number of Filters	16, 32, 64, 128, 256, 512
Kernel Size	(3,3), (5,5), (7,7)
Batch Size	8, 16, 32, 64
Learning Rate	0.0001, 0.001, 0.01, 0.1
Hidden Layer	128, 256, 512, 1024

c. Chromosome Representation

In the GA, a chromosome serves as a fundamental element representing a single candidate solution in the form of a genetic vector. Each chromosome consists of a set of genes, where each gene encodes the value of a parameter to be optimized. In other words, a chromosome represents the encoded configuration of the model for which the optimal parameter values are sought [25]. In this study, each chromosome was designed to include all CNN parameters subject to optimization, namely the kernel size, number of filters in each convolutional layer, batch size, learning rate, and the number of neurons in the fully connected layer. The chromosome representation can be expressed by the vector formulation in Equation 5:

$$C = [k_1, f_1, k_2, f_2, k_3, f_3, k_4, f_4, b, \eta, h_{1,} h_{2,}] \tag{5}$$

Where:

$k_i \in \{(3,3), (5,5), (7,7)\}$ represents the kernel size in the i-th convolutional layer.
$f_i \in \{16, 32, 64, 128, 256, 512\}$ denotes the number of filters.
$b \in \{8, 16, 32, 64\}$ indicates the batch size.
$\eta \in \{0.0001, 0.001, 0.01, 0.1\}$ refers to the learning rate.
$h_i \in \{128, 256, 512, 1024\}$ represents the number of neurons in the hidden layer.

d. Fitness Function

Each individual was evaluated using a fitness function defined by the validation accuracy of the CNN trained with its corresponding parameter configuration. This CNN architecture follows the structure illustrated in Figure 4, where the network adopts a hierarchical feature extraction design followed by a classification layer tailored for multi-class prediction. Using the learning rate specified within the chromosome, the network employed the Adamax optimization algorithm, and the training process was conducted for five epochs using the designated batch size. The final validation accuracy was used as the fitness value, while any configuration errors occurring during training resulted in a fitness score of 0.0.

e. Selection

Selection is the process of determining the parent individuals for reproduction. During this stage, candidate solutions within the population are evaluated and ranked according to their fitness values. This study employed the tournament selection method, in which individuals with higher fitness values have greater probability of being selected to contribute genetic material to the next generation [26].

f. Crossover

During the evolutionary process, the two-point crossover technique was employed to generate new offspring from selected parent individuals. Two individuals from the current generation were selected as parents, each contributing part of their genetic composition to produce improved offspring for the next generation. The process involved defining two crossover points on the chromosome and exchanging the genetic segments between the parent chromosome within these boundaries. This mechanism enables the offspring to inherit a more favorable combinations, thereby increasing the overall fitness value in subsequent generations [27].

g. Mutation

Mutation involves performing stochastic modifications on specific genes within a chromosome to preserve population diversity and prevent premature convergence [28]. In this study, the mutation rate was set to 0.15, indicating that each gene had a 15% probability of undergoing modification. During the mutation process, several parameters, including batch size, number of filters, kernel size, learning rate, and hidden layer configurations were randomly altered to explore new parameter combinations that could potentially enhance model performance in subsequent generations.

h. Updating the Population

The population update process was carried out by replacing the previous population with the best-performing individual from the prior generation along with all newly generated individuals produced through crossover and mutation operations. Once the population was updated, each individual in the new generation was re-evaluated based on the validation accuracy obtained from CNN model training. This iterative process, beginning with evaluation and continuing through population renewal, was repeated until the predetermined number of generations was reached. Ultimately, the optimal hyperparameter configuration was obtained from the best-performing individual in the final generation.

3. Results and Discussion

In the initial stage, the model was trained using comparative augmentation strategies and deep learning architecture, as presented in Table 5. Three augmentation conditions were evaluated: Albumentations-only, CutMix-only, and a combined Albumentations+CutMix approach.

Table 5. Augmentation Comparison

Augmentations	Accuracy			
	CNN from scratch	MobileNet	ResNet	DenseNet
Albumentations	84.8%	90.2%	87.4%	83.4%
CutMix	88.3%	87.5%	84.7%	85.6%
Albumentations+CutMix	92.1%	86.3%	90.4%	85.1%

Table 5 presents a performance comparison of the three augmentation strategies: Albumentations, CutMix, and a hybrid Albumentations+CutMix approach, evaluated across four CNN architectures, namely a custom CNN trained from scratch, MobileNet, ResNet, and DenseNet. The results show that Albumentations alone improved accuracy consistently, achieving 84.8% on the custom CNN and up to 90.2% on MobileNet. CutMix provided slightly higher performance gains for the custom CNN with 88.3% accuracy, although its performance varied across the other architectures. The hybrid augmentation strategy demonstrated the most substantial improvement, particularly for the custom CNN, which achieved 92.1% accuracy, and for ResNet, which improved to 90.4%. These findings indicate that combining Albumentations and CutMix enhances data diversity more effectively than using either technique individually, enabling better generalization and more robust feature learning, especially in models trained from scratch.

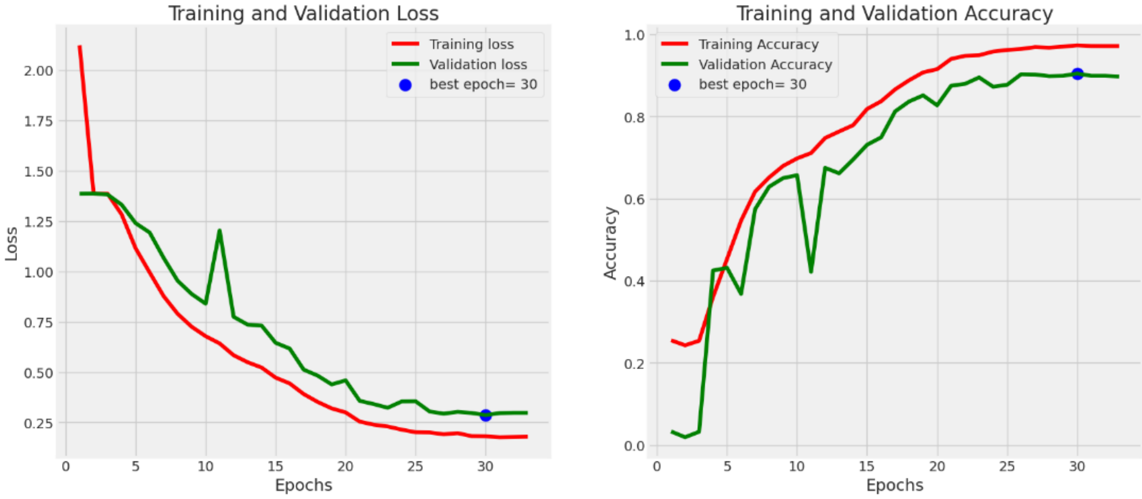


Figure 6. Hybrid Albumentations and CutMix

Following the training and validation processes, the combined augmentation strategy achieved the highest performance, resulting in a classification accuracy of 92.1%, as shown in Figure 6. The figure presents the training and validation loss curves as well as the training and validation accuracy over 35 epochs. The training loss showed a steady downward trend throughout the learning process, indicating that the model was successfully learned from the data, while the validation loss also exhibited a downward trend, demonstrating good generalization capability.

The fluctuations observed in the validation accuracy and loss curves were mainly caused by augmentation variability and the limited number of samples in certain classes, which are common in deep learning training using augmented data. Despite these fluctuations, the overall convergence trend indicates stable learning behavior and satisfactory model generalization. To further improve performance, a GA was applied for hyperparameter optimization using a CNN developed from scratch, allowing flexible adjustment of feature extraction and training parameters. The hybrid Albumentations+CutMix augmentation strategy was selected for GA optimization due to its superior classification accuracy compared to individual methods, thereby supporting an effective and high-performing optimization process.

3.1 Constructing the Final Model

After completing the optimization process using the GA, the next step involves constructing the parameter configuration identified as the most effective. The best chromosome obtained from the optimization process was decoded into a set of architectural parameters, as illustrated in Figure 4, which was then used to build a model identical in structure to the one employed during the fitness evaluation phase. The CNN architecture consists of four convolutional layers with an optimal combination of filters, kernel sizes, and batch size, followed by two hidden layers whose number of neurons was determined through the GA selection process. The entire model was then compiled using the optimal learning rate obtained from the optimization results. The GA optimization process required approximately 4 hours and 30 minutes to complete all generations, reflecting the computational cost associated with the evolutionary process. The results of the GA-based hyperparameter optimization are presented in Table 6, which summarizes the optimal parameters used in constructing the final model.

Table 6. Optimal Parameters

Parameter	Value
Number of Filters	16, 64, 64, 512
Kernel Size	(5,5), (7,7), (7,7), (7,7)
Batch Size	64
Learning Rate	0.0001
Hidden Layer	512, 1024

After obtaining the optimal parameters from the GA optimization results shown in Table 6, the model was retrained to evaluate its accuracy, precision, recall, and F1-score, in order to determine whether these performance metrics could be further improved. Figure 7 presents the training results after applying GA optimization.

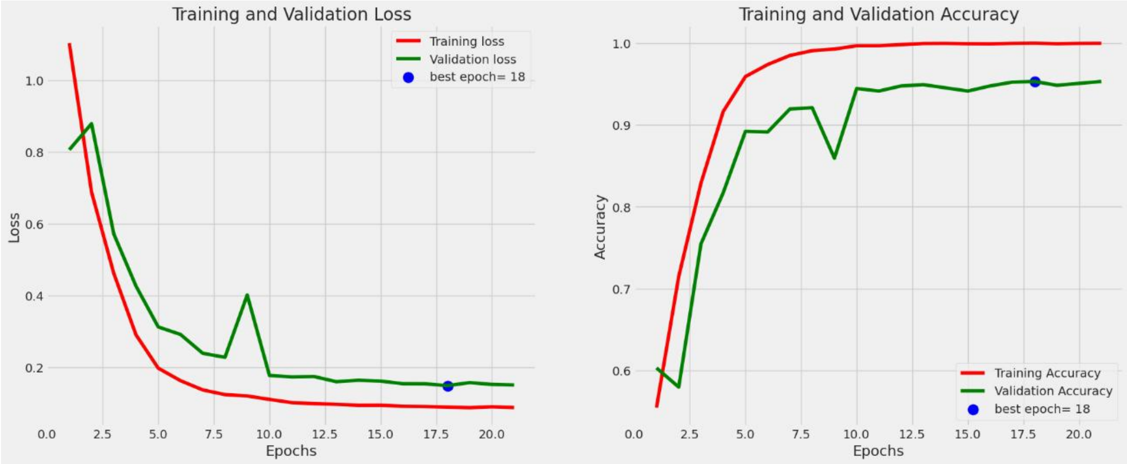
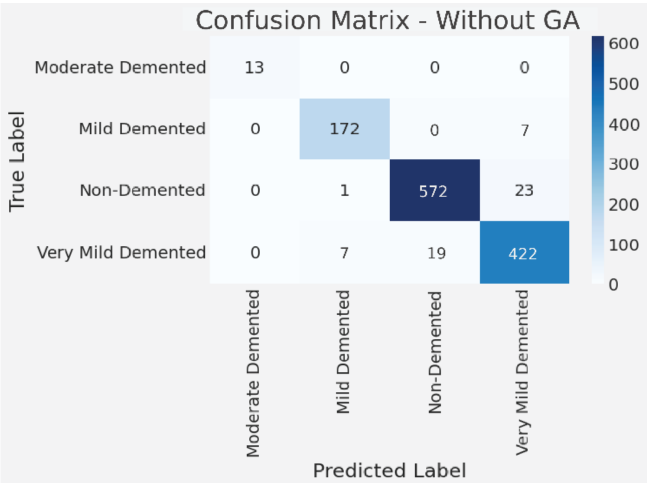


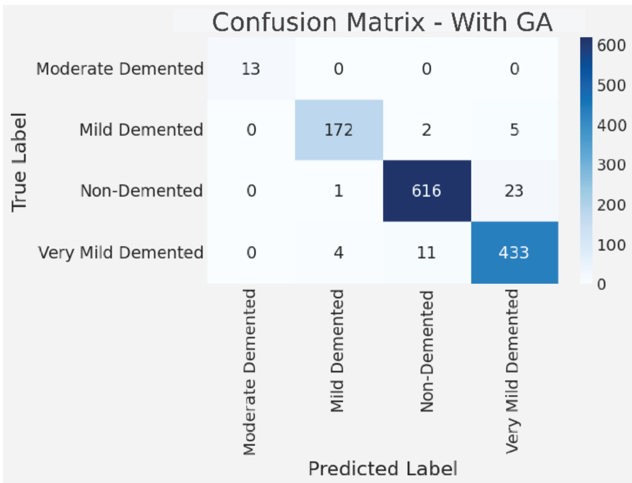
Figure 7. Optimization Results with GA

Both the training and validation loss curves exhibit a smooth and consistent decrease, indicating effective model optimization, with the optimal convergence point occurring around epoch 18. Meanwhile, the accuracy curves show significant improvement, reaching values close to 1.0, with the highest validation accuracy observed at epoch 18. These results demonstrate that the GA optimization effectively improved model stability and successfully identified an optimal hyperparameter configuration without observable signs of overfitting.

However, slight fluctuations can still be observed in the training curves, which may be attributed to dataset imbalance and the absence of a pre-trained model during training. Nevertheless, the overall training trend remains positive, indicating that the model successfully captured informative features throughout the learning process. These findings further suggest that the GA-optimized hyperparameters contributed to improved convergence behavior and enhanced generalization performance across the dataset. Despite these fluctuations, the training process eventually converged and achieved a classification accuracy of 96.4%. This result further demonstrates that the GA successfully identified an optimal parameter for the CNN model, thereby enhancing its overall performance and generalization capability. To evaluate the performance of the optimized model, a confusion matrix analysis was conducted to assess the model's classification capability. Figure 8 presents the comparative performance results of the CNN model: (a) without GA optimization and (b) with GA optimization, illustrating the distribution of correct and incorrect predictions across each Alzheimer's disease class.



(a) Confusion Matrix – Without GA



(b) Confusion Matrix – With GA

Figure 8. Comparative Result Model Performance

Figure 8 presents a comparative confusion matrix illustrating the performance of the CNN model before and after GA optimization. The baseline CNN, trained using the combined Albumentations+CutMix augmentation strategy, initially employed a fixed 3×3 convolutional kernel across all convolutional layers, which is commonly used due to its computational efficiency and local feature extraction capability. After GA optimization, the model demonstrates higher classification accuracy, correctly identifying 616 Non-Demented, 433 Very Mild Demented, 172 Mild Demented, and 13

Moderate Demented samples. Most predictions are concentrated along the diagonal of the confusion matrix, reflecting stronger discriminative capability across the four Alzheimer's disease stages.

Compared to the baseline model, the GA-optimized version exhibits a noticeable reduction in misclassification errors, particularly between the Non-Demented and Very Mild Demented classes, which previously showed substantial overlap. This improvement indicates that the GA effectively optimized critical hyperparameters, including kernel size, number of filters in each convolutional layer, batch size, learning rate, and the number of neurons in the fully connected layer. Notably, the GA replaces uniform 3×3 convolution kernels with larger 5×5 and 7×7 kernels in deeper layers, enabling the network to capture broader spatial context and more global structural patterns within MRI scans. The larger receptive field allows the model to distinguish subtle anatomical variations more effectively, leading to improved feature extraction and reduced class overlap.

Overall, integrating hybrid augmentation with GA-based hyperparameter optimization produced a CNN model with improved training stability and more consistent class-wise performance. The shift toward larger convolutional kernels demonstrates that GA-driven architectural adaptation effectively identifies suitable feature scales for complex neuroimaging patterns, resulting in smoother convergence and reduced performance variance.

4. Conclusion

The primary objective of this research was to develop a reliable model based on advanced neural network techniques for distinguishing Alzheimer's disease categories using MRI images through the integration of advanced augmentation techniques and optimization algorithms. The findings indicate that Albumentations alone achieved an accuracy of 84.8%, while CutMix alone achieved 88.3%, demonstrating that both methods independently improve classification performance. The combined augmentation strategy significantly enhanced the model's generalization capability and feature extraction performance, increasing classification accuracy to 92.1%. These results demonstrate that the hybrid augmentation approach effectively increases data variability and reduces overfitting, which is particularly important for limited medical imaging datasets. The application of GA for hyperparameter optimization, including kernel size, number of filters, batch size, learning rate, and neuron count, further improved the model's accuracy to 96.4%. This improvement highlights the ability of GA to efficiently explore a broad and complex search space beyond manual tuning, leading to more optimal and stable training configurations. These findings confirm the effectiveness of GA in identifying optimal configurations that improve model stability and convergence behavior.

Although the findings are encouraging, the research still faces several limitations, particularly the relatively small dataset size and the reliance solely on MRI data. Future work is recommended to expand the dataset and investigate alternative optimization approaches, such as Bayesian optimization or swarm-based algorithms. In adiition, the GA optimization process required significantly higher computational time, approximately 4–5 hours, compared to only 10 minutes for training without GA optimization. Therefore, future research should consider computational efficiency improvement strategies, such as reducing the hyperparameter search space, implementing early stopping mechanisms, or utilizing higher-performance GPUs.

References

[1] Y. S. Austin *et al.*, "Klasifikasi Penyakit Alzheimer Dari Scan Mri Otak Menggunakan Convnext," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 6, pp. 1223–1232, Dec. 2024. https://doi.org/10.25126/jtiik.2024118117

[2] A. Calderaro *et al.*, "The Neuroprotective Potentiality of Flavonoids on Alzheimer's Disease," Dec. 01, 2022, *MDPI*. https://doi.org/10.3390/ijms232314835

[3] P. Nasra and S. Gupta, "DenseNet-Based Approach for Early Detection and Classification of Alzheimer's Disease Using MRI Images," in *2025 International Conference on Pervasive Computational Technologies, ICPCT 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 170–175. https://doi.org/10.1109/ICPCT64145.2025.10940932

[4] Y. Zuo, X. Hao, M. Song, F. Qi, B. Qiu, and X. Wang, "Automated brain atrophy quantification and evaluation using spatial resolution enhancement," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/EMBC53108.2024.10782685

[5] A. Stabile *et al.*, "Factors Influencing Variability in the Performance of Multiparametric Magnetic Resonance Imaging in Detecting Clinically Significant Prostate Cancer: A Systematic Literature Review," *Eur Urol Oncol*, vol. 3, no. 2, pp. 145–167, Apr. 2020. https://doi.org/10.1016/j.euo.2020.02.005

[6] V. Marakala, G. V. Sriramakrishnan, G. Jakka, C. J. Shingadiya, H. P. Widiastuti, and G. G. R. Ortiz, "Use of Deep Learning Application in Medical Devices," in *4th International Conference on Inventive Research in Computing Applications, ICIRCA 2022 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 935–939. https://doi.org/10.1109/ICIRCA54612.2022.9985537

[7] G. Singh, K. Guleria, and S. Sharma, "A Deep Learning-based Convolutional Neural Network Model for Alzheimer's Disease Detection," in *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*, IEEE, Nov. 2024, pp. 1–5. https://doi.org/10.1109/IC3TES62412.2024.10877583

[8] C. Sinha, A. S. Raghuvanshi, and B. Acharya, "Enhancing Liver and Tumor Segmentation with Res50UNet using CDFL and Albumentations," in *2025 2nd International Conference on Circuits, Power and Intelligent Systems (CCPIS)*, IEEE, Sep. 2025, pp. 1–6. https://doi.org/10.1109/CCPIS65231.2025.11234192

[9] H. Bumpenje, Rahmadwati, and Z. Abidin, "Chest Cancer Classification from Chest CT-Scan Images using Deep Learning," in *2024 International Conference on Smart Computing, IoT and Machine Learning, SIML 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 49–55. https://doi.org/10.1109/SIML61815.2024.10578120

[10] A. Rao, J.-Y. Lee, and O. Aalami, “Studying the Impact of Augmentations on Medical Confidence Calibration,” Aug. 2023. https://doi.org/10.48550/arXiv.2308.11902
[11] O. Salih and K. J. Duffy, “Optimization Convolutional Neural Network for Automatic Skin Lesion Diagnosis Using a Genetic Algorithm,” *Applied Sciences (Switzerland)*, vol. 13, no. 5, Mar. 2023. https://doi.org/10.3390/app13053248
[12] K. V. Sridhar, V. K. Tiwari, R. Mounica, and K. Tejaswi, “Brain Tumor Classification Using Enhanced CNN and Optimization with Metaheuristic Algorithms,” in *2024 7th International Conference on Signal Processing and Information Security, ICSPIS 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/ICSPIS63676.2024.10812605
[13] L. M. Elnaghi and Y. M. Eltariny, “Evaluation of Deep Learning Models on Alzheimer' s MRI Dataset: AD-VGG 1 6, AD-Resnet50, and AD-2DCNN,” in *6th International Conference on Computing and Informatics, ICCI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 237–242. https://doi.org/10.1109/ICCI61671.2024.10485046
[14] M. S. Kumar, G. Charmi, Y. Chandana, D. V. Jaiyesh, and M. H. Kartheek, “Advanced Multimodal Deep Learning for Predicting Cognitive Decline in Alzheimer’s Disease,” in *2025 Fourth International Conference on Smart Technologies, Communication and Robotics (STCR)*, IEEE, May 2025, pp. 1–6. https://doi.org/10.1109/STCR62650.2025.11019800
[15] C. Kamardi ***et al.***, “Classification of Alzheimer’s Disease using Random Oversampling and Albumentations on Convolutional Neural Network,” in *2023 8th International Conference on Informatics and Computing, ICIC 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. https://doi.org/10.1109/ICIC60109.2023.10382106
[16] P. Kaushik and A. Singh, “Severity-Level Classification of Alzheimer’s Disease from MRI Scans using Convolutional Neural Networks,” in *2nd International Conference on Self Sustainable Artificial Intelligence Systems, ICSSAS 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 187–192. https://doi.org/10.1109/ICSSAS64001.2024.10760525
[17] R. Rahmadwati, A. Z. Imran, M. Aswin, And K. Ferdiana, “Identifikasi Penyakit Katarak berdasarkan Citra Fundus menggunakan Siamese Convolutional Neural Network,” *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 12, no. 4, p. 838, Dec. 2024. https://doi.org/10.26760/elkomika.v12i4.838
[18] N. Meenakshisundaram and Sajiv. G, “Evaluating Oversampling Strategies for Imbalanced Cervical Cancer Risk Prediction: A Comparative Analysis of SMOTE, Borderline-SMOTE and ADASYN,” in *2025 Third International Conference on Networks, Multimedia and Information Technology (NMITCON)*, IEEE, Aug. 2025, pp. 1–6. https://doi.org/10.1109/NMITCON65824.2025.11188220
[19] S. Rani, T. Ahmad, and S. Masood, “Handling Class Imbalance Problem using Oversampling Techniques for Breast Cancer Prediction,” in *2023 International Conference on Recent Advances in Electrical, Electronics and Digital Healthcare Technologies, REEDCON 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 693–698. https://doi.org/10.1109/REEDCON57544.2023.10150702
[20] R. Maheshwari, A. Sharma, J. P. Meena, and S. K. Nagar, “Optimized Deep Learning Architecture with EfficientNet and with MixUp & CutMix for Robust Plant Disease Classification,” in *2025 8th International Conference on Circuit, Power and Computing Technologies, ICCPCT 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 993–998. https://doi.org/10.1109/ICCPCT65132.2025.11176735
[21] K. Dinesh Kumar, K. J. Deepthi, S. Saravanakumar, S. Balamurugan, I. Govindharaj, and P. A. Reddeppa, “Early Melanoma Detection and Classification Using CNN and Confusion Matrix Analysis,” in *2024 International Conference on System, Computation, Automation and Networking, ICSCAN 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/ICSCAN62807.2024.10894452
[22] R. N. Pathapati, V. S. Anumala, V. S. C. G. Jupudi, and N. Pasam, “Improving Parkinson’s Disease Diagnosis: A Genetic Algorithm-Guided CNN Approach,” in *2nd International Conference on Signal Processing, Communication, Power and Embedded Systems, SCOPES 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/SCOPES64467.2024.10991308
[23] L. Gongalla and M. Bordoloi, “Optimized Deep Learning for Tea Leaf Age and Quality Classification using EGACNN and SHEDA-based Hyperparameter Tuning,” in *Proceedings of 5th International Conference on Pervasive Computing and Social Networking, ICPCSN 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 1072–1077. https://doi.org/10.1109/ICPCSN65854.2025.11036046
[24] K. Rajagopal, V. S. Kumari, S. Saraswathy, V. S. Kumar, S. Ponmaniraj, and A. Deepa, “Hybrid Deep Learning Models with Genetic Algorithm Optimization for Enhanced Kidney Tumor Detection,” in *5th International Conference on Electronics and Sustainable Communication Systems, ICESC 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 1525–1530. https://doi.org/10.1109/ICESC60852.2024.10689997
[25] K. V. Sridhar, V. K. Tiwari, R. Mounica, and K. Tejaswi, “Brain Tumor Classification Using Enhanced CNN and Optimization with Metaheuristic Algorithms,” in *2024 7th International Conference on Signal Processing and Information Security, ICSPIS 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. https://doi.org/10.1109/ICSPIS63676.2024.10812605
[26] H. M. Yumil, F. Sia, T. S. Fun, and L. P. Hung, “Optimized Convolutional Neural Network Using Genetic Algorithm for Music Genre Classification,” in *6th IEEE International Conference on Artificial Intelligence in Engineering and Technology, IICAIET 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 547–550. https://doi.org/10.1109/IICAIET62352.2024.10729917
[27] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, “Automatically Designing CNN Architectures Using the Genetic Algorithm for Image Classification,” *IEEE Trans Cybern*, vol. 50, no. 9, pp. 3840–3854, Sep. 2020. https://doi.org/10.1109/TCYB.2020.2983860
[28] F. Y. Santoso, E. Sediyono, and H. D. Purnomo, “Genetic Algorithm For Convolutional Neural Network Hyperparameter Tuning,” in *2023 3rd International Conference on Computing and Information Technology, ICCIT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 232–236. https://doi.org/10.1109/ICCIT58132.2023.10273928