Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 136-151 e-ISSN: 2656-8632
RESEARCH ARTICLE OPEN ACCESS

Enhancing Deep Learning Model Using Whale Optimization Algorithm on Brain Tumor MRI

Winarno[1,2], and Agus Harjoko[2]

[1] Departement of Informatics, Universitas Sebelas Maret, Surakarta, Indonesia
[2] Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia

Corresponding author: Winarno (e-mail: win@staff.uns.ac.id), **Author Email**: Agus Harjoko (e-mail: aharjoko@ugm.ac.id)

Abstract: The increasing prevalence of brain cancer has emerged as a significant global health issue, with brain neoplasms, particularly gliomas, presenting considerable diagnostic and therapeutic obstacles. The timely and precise identification of such tumors is crucial for improving patient outcomes. This investigation explores the advancement of Convolutional Neural Networks (CNNs) for detecting brain tumors using MRI data, incorporating the Whale Optimization Algorithm (WOA) for the automated tuning of hyperparameters. Moreover, two callbacks, ReduceLROnPlateau and early stopping, were utilized to augment training efficacy and model resilience. The proposed model exhibited exceptional performance across all tumor categories. Specifically, the precision, recall, and F1-scores for Glioma were recorded as 0.997, 0.980, and 0.988, respectively; for meningioma, as 0.983, 0.986, and 0.984; for no tumors, as 0.998, 0.998, and 0.998; and for pituitary, as 0.997, 0.997, and 0.997. The mean performance metrics attained were 0.994 for precision, 0.990 for recall, and 0.992 for F1-score. The overall accuracy of the model was determined to be 0.991. Notably, incorporating callbacks within the CNN architecture improved accuracy to 0.994. Furthermore, when synergized with the WOA, the CNN-WOA model achieved a maximum accuracy of 0.996. This advancement highlights the effectiveness of integrating adaptive learning methodologies with metaheuristic optimization techniques. The findings suggest that the model sustains high classification accuracy across diverse tumor types and exhibits stability and robustness throughout training. The amalgamation of callbacks and the Whale Optimization Algorithm significantly bolster CNN performance in classifying brain tumors. These advancements contribute to the development of more reliable diagnostic instruments in medical imaging.

Keywords: **CNN; Brain Tumor; Callback; Grad-CAM; Metaheuristic; WOA**.

I. Introduction

The increasing incidence of brain cancer has become a significant public health concern globally. Recent findings highlight a growing trend in diagnosing cerebral neoplasms, which may result from various factors, including innovations in imaging methods and enhanced awareness among healthcare practitioners and patients [1]. The World Health Organization has emphasized that early detection is crucial for improving patient outcomes, as it allows for timely interventions that can significantly enhance the survival rate [2]. Brain tumors, particularly gliomas, present unique challenges due to their complex nature and the critical functions of the brain [3]. In early detection, technological advancements, especially those related to deep learning and artificial intelligence, have shown significant promise in improving diagnostic accuracy. Deep learning techniques, such as Convolutional Neural Networks (CNNs), have been employed to analyze medical imaging data, enabling the identification of tumors at earlier stages [4]. These methods can handle large volumes of data swiftly and accurately, which is vital given the urgency associated with cancer treatment [5]. For instance, integrating deep learning with magnetic resonance imaging (MRI) has been highlighted as an efficient approach for detecting brain tumors, offering improved precision [6]. Moreover, the development of applications that utilize deep learning for early brain cancer detection is gaining traction. These applications aim to provide healthcare professionals with tools that enhance diagnostic capabilities and facilitate better patient management [7].

Convolutional Neural Networks (CNNs) have become a cornerstone in medical imaging, particularly for classifying brain tumors from Magnetic Resonance Imaging (MRI) datasets. Recent advancements in deep learning techniques have significantly improved diagnostic accuracy and efficiency, making CNNs a vital tool in neuro-oncology. This article synthesizes findings from various studies published in the last three years, highlighting the effectiveness of CNNs in brain tumor classification, along with the reported accuracy and F1 scores.

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

Significant advancements in cancer research are evident. Among these developments, various CNN architectures have been employed in brain tumor classification, demonstrating unique strengths. For instance, Mahjoubi reported that their proposed CNN networks achieved an accuracy of 95.44% and an F1 score of 95.36% in classifying brain tumors from MRI images [8]. Similarly, Raza developed a hybrid deep learning model that combined CNNs with traditional machine learning techniques, achieving an accuracy of 94.5% [9]. Transfer learning has also been pivotal; for example, [10] utilized a transfer learning model based on AlexNet, achieving an impressive accuracy of 99.62%. Another study, Inception architecture by [11] reports that their proposed CNN networks achieved an accuracy of 0.993. Furthermore, the FL+VGG16 [12] study has achieved an accuracy of 0.98. There is also research using EfficientnetV2S that has achieved an accuracy of 0.989 [13]. The last one is research using a 2DCNN ensemble conducted by [14] and Xception conducted by [15], which achieved accuracies of 0.9647 and 0.940, respectively. The above state-of-the-art research still has room for improvement.

Data augmentation methodologies have played an essential role in enhancing the resilience of Convolutional Neural Network (CNN) architectures. Rasool [16] provided evidence that their integrative deep learning framework, which synergistically incorporated GoogleNet for feature extraction alongside Support Vector Machine (SVM) classifiers, attained a classification accuracy of 97.5%. This methodology highlights the crucial need to augment training datasets to mitigate the risk of overfitting and improve the model's generalizability. Furthermore, the investigation analyzed a range of convolutional neural network (CNN) methodologies and concluded that proficient data augmentation practices could lead to substantial improvements in classification efficacy [17], [18]. To enhance the efficacy of deep learning, a crucial step involves appropriately utilizing hyperparameters. Hyperparameters are variables that remain fixed during the training of a machine learning model, and their selection significantly influences the model's accuracy and performance. As noted by Shi et al., the effectiveness of deep learning methodologies heavily relies on the correct hyperparameter configurations, which the user establishes before the model training process begins [19]. In numerous instances, inadequate hyperparameter choices can result in suboptimal model performance, underscoring the significance of a strategic approach in this customization [20].

Researchers can employ two strategies to determine the best hyperparameters: manually performing repeated experiments or automatically using optimization algorithms. Some commonly utilized algorithms are Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Grey Wolf Optimization (GWO), and Whale Optimization Algorithm (WOA) [21] [22]. Among these options, WOA was selected for its benefits in locating optimal solutions through a method inspired by the hunting techniques of whales. WOA has demonstrated its effectiveness across various optimization tasks in deep learning, as noted by Sanmorino et al [23]. It was observed that this optimization technique can enhance model performance with greater efficiency [23].

WOA operates by exploring a broad search landscape to identify optimal hyperparameter values, and this algorithm has the advantage of avoiding local minima often encountered by traditional optimization methods. The research conducted by Ilemobayo et al. highlights how different techniques for hyperparameter tuning, including WOA, can minimize the time and resources necessary to attain high accuracy in a model [24]. Furthermore, WOA offers scalability and sustainability advantages when training becomes more intricate. This is elaborated upon in the study by Rao and Jaganathan [22].

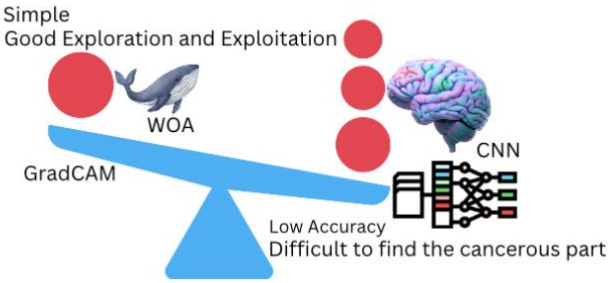


Fig. 1. Research Gap Visualization

Thus, employing a suitable optimization algorithm, such as WOA, enhances the efficiency of the hyperparameter search process and plays a crucial role in achieving superior performance in deep learning models. This is particularly important, as effective hyperparameter selection can significantly enhance the accuracy of the developed model. Based on the provided explanation, this research aims to construct a deep learning model that integrates automatic hyperparameter search using WOA on brain cancer MRI datasets. This research is expected to yield a model that outperforms previous studies conducted in this area.

This article presents a scenario illustrating the outcomes of a vanilla CNN-based deep learning model, derived from parameters obtained using the default hyperparameters. The second scenario employs callbacks with predetermined hyperparameters, whereas the third uses hyperparameters derived from the outcomes of the WOA search. To implement WOA,

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

an objective function will be established to achieve optimal accuracy.

Despite achieving optimal outcomes, deep learning models often remain enigmatic regarding their functionality and comprehensibility, primarily due to their reliance on black-box mechanisms for feature extraction [25], [26], [27]. This leads to another problem, namely the difficulty of revealing these mysterious parts. To explain this, explainable AI can be employed so that consumers are not perplexed about how the deep learning model generates results, and the tumor's location in the brain can be correctly identified. To address this, Grad-CAM can help deep learning models explain where the brain cancer is in an image [28]. The capability above is of paramount importance within the realm of medical applications, as healthcare professionals require the ability to validate and comprehend the underlying reasoning associated with model predictions, thereby fostering trust and promoting the integration of artificial intelligence tools into clinical practice [25], [29].

As previously explained, this study aims to answer the following research question: RQ1: How can a deep learning model be developed with improved accuracy compared to current research? RQ2: How can brain tumor area segmentation be visualized from a deep learning model?

Based on the explanation provided, Fig. 1 illustrates the research gap identified in this study. Research contributes by constructing a deep learning model automatically built to determine the best hyperparameters using the whale optimization technique. This model defines the hyperparameter value based on the dataset used. Additionally, the best results can be achieved by segmenting the model findings using Grad-CAM. This research is structured as follows: introduction, explanation of the methods used, including dataset preparation, data preprocessing, models employed, ablation experiments conducted on prepared scenarios, evaluation, and conclusion.

II. Method

The approach utilized in this research can be outlined in the sequence shown in Fig. 2. This investigation commences with the acquisition of datasets. The datasets were procured from the reference pages of prior research studies. Subsequently, the downloaded dataset undergoes a rigorous data preprocessing phase, which encompasses validating the dataset size, augmenting the dataset, and organizing appropriate folders for each class within the dataset. Following the data preprocessing stage, the next step involves constructing the Convolutional Neural Network (CNN) architecture and developing a basic model.

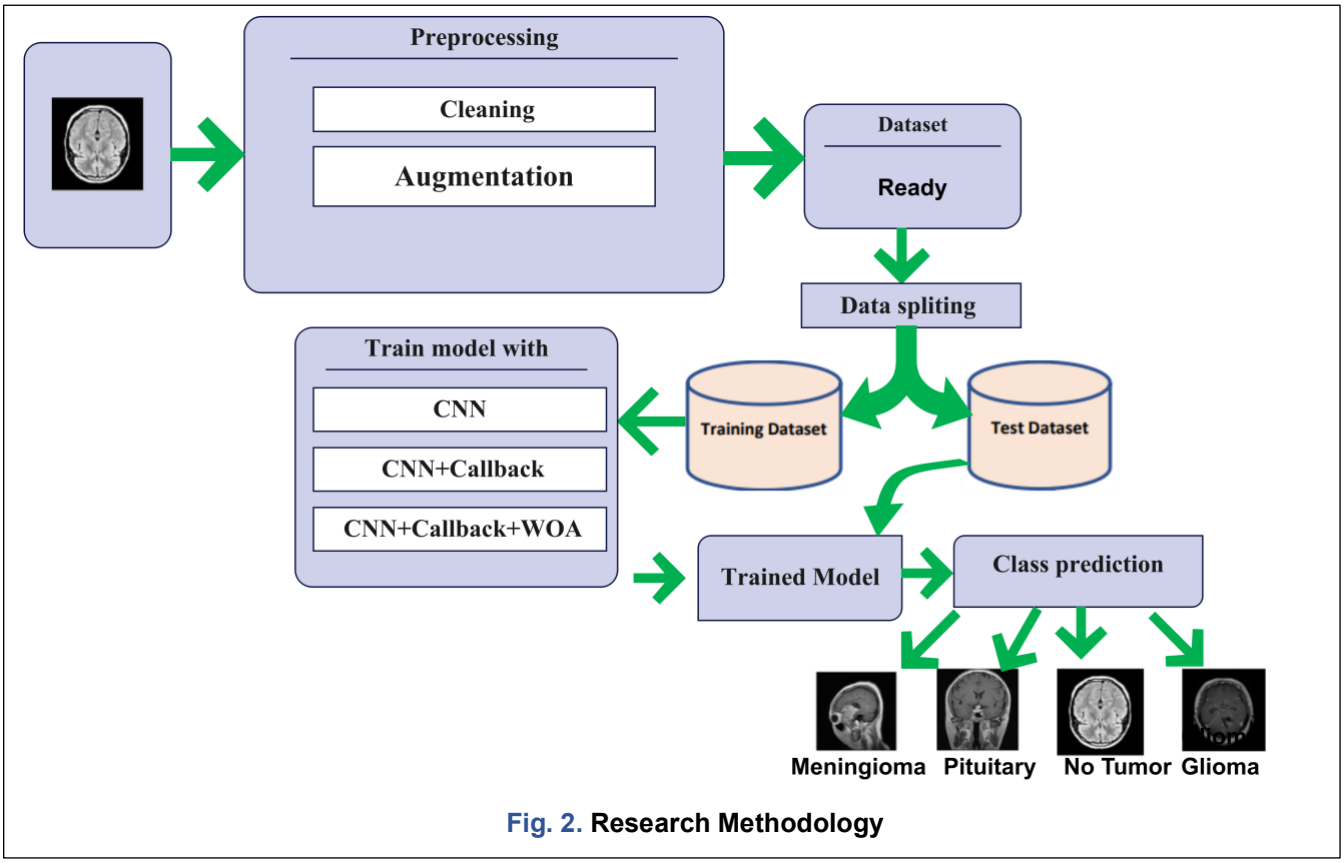


Fig. 2. Research Methodology

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

In addition to the foundational model, a supplementary model is devised by incorporating callbacks into the CNN framework. Upon completing the foundational model development, the code is further refined by integrating automated hyperparameter optimization utilizing the Whale Optimization Algorithm (WOA). After implementing this optimization, the process progresses to the training phase, which continues until the optimal model is derived from the results of the WOA search. The final step involves extensive evaluation of the constructed model utilizing testing data. This sequence of processes is illustrated in Fig. 2.

A. Dataset

This study focuses on preprocessing three pivotal datasets: Figshare [30], SARTAJ, and BR35H. These establish a comprehensive framework for training, validating, and testing models to identify brain tumors within MRI imaging. This dataset can be accessed on the Kaggle platform [31]. This dataset has been processed, and some files contain incomplete or inaccurate data. The total number of datasets is 6438, divided into two parts, namely, 5127 training datasets and 1311 testing datasets. This dataset is then used for the modeling process. Some examples of brain tumor images are shown in Fig. 3.

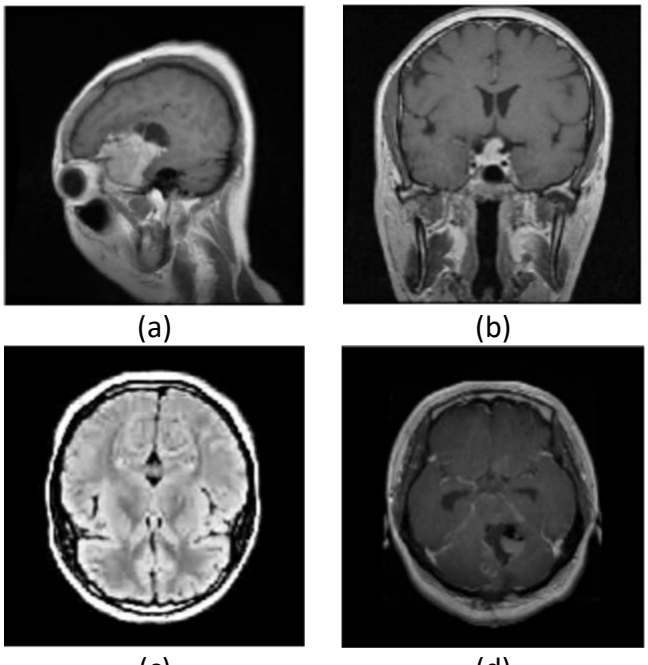


Fig. 3. The Brain Tumor Dataset a) Meningioma b) Pituitary c) No Tumor d) Glioma

B. Data Preprocessing

The pre-processing of datasets for detecting brain tumors constitutes a vital phase in guaranteeing the efficacy of machine learning models, especially those that leverage advanced deep learning methodologies. This dataset is then augmented with parameters such as those in Table 1. This technique allows the model to learn to recognize objects from both orientations, which is especially useful in situations where the orientation of an object is variable. Research has indicated that such transformations can substantially enhance the model's capacity for generalization; random rotation can yield considerable advancements in model accuracy as it exposes the network to a broader spectrum of object orientations and spatial configurations [32], [33] [34], [35].

Moreover, varying contrast throughout the training process can significantly aid in refining the model's ability to extract relevant features, thereby mitigating the risk of overfitting [33], [34], [36]. Random zoom techniques can also effectively reduce model overfitting and improve performance on previously unseen data [37], [38]. In contrast, random translation instructs the model to maintain robustness against variations in object placement, which frequently occur in real-world scenarios [38], [39].

Table 1. Dataset Augmentation Parameter

Action	Value
random flip	horizontal
random rotation	0.02; fill mode=constant
random contrast	0.1
random zoom	height factor=0.01, width factor=0.05
random translation	height factor=0.0015, width factor=0.0015, fill mode=constant

C. Modelling

Convolutional Neural Networks (CNNs) have emerged as a formidable tool within deep learning, particularly in image processing applications. The structural design of these networks is intended to autonomously and adaptively acquire spatial hierarchies of characteristics from input images, rendering them especially effective for applications such as medical image analysis and object recognition [40]. The input to a Convolutional Neural Network (CNN) is conventionally represented as a multi-dimensional array that encapsulates an image, with each pixel value signifying a particular intensity level. This input undergoes processing through a succession of convolutional layers, wherein the fundamental operation is the convolution itself. Convolution entails the systematic traversal of a filter (or kernel) across the input image, generating feature maps that accentuate distinct patterns or features present in the image [41]. The selection of kernel dimensions is paramount; small kernels can capture intricate details, whereas larger kernels can delineate more extensive features [42]. For instance, a typical

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

kernel size is 3x3, which strikes a balance between detail and computational efficiency.

Strides denote the number of pixels traversed by the filter across the image throughout the convolution process. A stride of one indicates that the filter progresses one pixel at a time. In contrast, a stride of two omits every other pixel, thereby diminishing the dimensions of the resultant feature map [41]. Padding represents a crucial concept in image processing, involving the addition of extra pixels surrounding the input image to control the spatial dimensions of the resulting output. Widely utilized padding methodologies encompass 'valid' padding (which entails no extra padding) and 'same' padding (where the dimensions of the production are congruent with those of the input). Lecun[43] defines that the core of CNN lies in the convolution operation, which can be expressed as Eq. (1) [44]

$$y_{i,j}^{(k)} = \sum_{m=1}^{M} \sum_{n=1}^{N} x_{i+m,j+n} \cdot w_{m,n}^{(k)} + b^{(k)}, \quad (1)$$

where $x_{i,j}$ denotes the input pixel intensity at location i,j, $w_{m,n}^{(k)}$ represents the convolution kernel weights of *the* k-th filter, $b^{(k)}$is the corresponding bias term, and $y_{i,j}^{(k)}$ is the resulting feature map. This formulation allows the network to learn spatially localized patterns critical for identifying tumor regions.

Activation functions are of paramount significance in the structural design of neural networks, as they facilitate the introduction of non-linearity, thereby empowering the models to discern intricate patterns within datasets. The Rectified Linear Unit (ReLU) is highly regarded among the various activation functions. It has established itself as a fundamental component that is beneficial for Convolutional Neural Networks (CNNs) due to their efficacy in alleviating challenges such as the vanishing gradient problem [45]. The mathematical formulation of ReLU permits it to yield a value of zero for negative inputs while enabling positive inputs to traverse linearly, thus maintaining the inherent characteristics of the data.

The widespread adoption of the Rectified Linear Unit (ReLU) activation function can be ascribed to its inherent simplicity and efficacy in training deep neural networks. Empirical studies show that the application of ReLU accelerates convergence during the training process compared to conventional activation functions such as sigmoid or hyperbolic tangent (tanh), mainly due to its ability to overcome the saturation dilemma, where gradients approach insignificant values for extreme input values [46], [47]. For instance, a research investigation highlighted that convolutional neural networks utilizing the Rectified Linear Unit (ReLU) activation function demonstrate superior efficacy and proficiently capture nonlinear characteristics, which is critical for applications such as image classification [48], [49]. Recent progressions in activation functions, encompassing variations of ReLU such as Leaky ReLU. Parametric ReLU (PReLU) has been scrutinized to mitigate its shortcomings, including the 'dying ReLU' phenomenon, wherein neurons may become dormant and thus fail to partake in the learning [50], [51]. Notwithstanding these apprehensions, ReLU remains a dominant feature in contemporary deep learning architectures, highlighting its integral position in the evolution of neural network functionalities [47]. Consequently, grasping the advantages and limitations of various activation functions is essential for formulating effective deep learning architectures, particularly as novel variants continue to emerge with prospective enhancements in performance attributes [52], [53].

The Rectified Linear Unit (ReLU) is recognized as one of the most prevalent activation functions utilized in Convolutional Neural Networks (CNNs), as delineated in Eq. (2) [54]

$$Relu(x) = \max(0, x), \quad (2)$$

where x is the input to the activation layer. This research uses Adam as an optimizer. The Adam optimizer is widely regarded as an advantageous algorithm for training convolutional neural networks (CNNs), primarily due to its inherent adaptive learning capabilities. This optimizer integrates the beneficial features of two prominent variants of stochastic gradient descent, specifically AdaGrad and RMSProp. This thereby facilitates practical training even in high-dimensional parameter spaces [12], [40]. Adam's Algorithm is shown as Algorithm 1.

Algorithm 1. Adam:Stochastic Optimization Algoritm
Input: Objective function $J(\theta)$; initial parameters θ_0; learning rate η; exponential decay rates β_1, β_2; numerical stability constant ε maximum iterations T.
Output: Optimized parameters θ_T
Initialize parameters:
θ_0 (initial parameters),
$m_0 = 0$ (first moment),
$v_0 = 0$ (second moment),
timestep t = 0.
The hyperparameters are: learning rate η, exponential decay rates β_1, β_2 (typically β_1=0.9, β_2=0.999), and ε for numerical stability.
for t = 1 **to** T **do**
    Compute gradient: gt ← ∇θ J(θt-1)
    Update biased first moment: mt ← β1mt-1 + (1 − β1)gt

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

```
8       Update biased second moment: vt
        ← β2vt-1 + (1 − β2)g2
9       Compute bias-corrected first
        moment: ^mt ← mt/(1−β1^t)
10      Compute bias-corrected second
        moment: ^vt ← vt/(1−β2^t)
11      Update parameters: θt ← θt-1 −
        (η · ^mt/(√(^vt+ε)))
12  end for
13  return θT
```

To reduce dimensionality while preserving essential information, CNNs typically apply pooling layers. In the case of max pooling, the operation selects the maximum value within a local neighborhood Ω, formulated as Eq. (3) [55]

$$P_{i,j} = \max_{(m,n)\in\Omega} X_{i+m,\,j+n} \quad (3)$$

where $x_{i,j}$ denotes the input pixel intensity at location i,j. Following several convolution and pooling stages, the network transitions into fully connected layers where linear combinations of previous activations are computed, written as Eq. (1). For multiclass prediction tasks, the final layer commonly employs the softmax function, which normalizes the logits into a probability distribution over K classes as formulated in Eq. (4) [56]

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad (4)$$

where the index $j \in \{1,2,\ldots,K\}$ iterates over all class logits to ensure probabilistic normalization. $K \in N$ is the total number of classes, and the variable z_i denotes the logit corresponding to class i. The resulting output satisfies $\sigma(z_i) \in [0,1]$. To facilitate the learning process, Convolutional Neural Networks (CNNs) are trained using appropriate loss functions, with the cross-entropy loss function being the most commonly chosen option. This loss function is denoted $\mathcal{L}$, measuring the difference between the predicted distribution $\hat{y}$ and the actual target distribution y as shown in Eq. (5) [57]

$$\mathcal{L} = -\sum_{i=1}^{K} y_i \log(\hat{y}_i). \quad (5)$$

variable $K \in N$ denotes the total number of classes in the classification problem. The vector $y \in R^K$ represents the ground-truth label encoded in one-hot format, where $y_i = 1$ if the sample belongs to the class i and $y_i = 0$ otherwise. The predicted probability vector is given by $\hat{y} \in [0,1]^K$, obtained from the softmax function applied to the final logit outputs of the network. Each scalar $\hat{y}_i$ denotes the model's estimated probability that the input sample belongs to class i. The preliminary modeling conducted in this investigation employed hyperparameters, as delineated in Table 2. This model shall be designated the vanilla model, representing the baseline model for this research endeavor. The performance and convergence behavior of a CNN are significantly influenced by the choice of hyperparameters that govern the optimization process.

In this study, the network is trained using a configuration comprising 50 epochs, a batch size of 32, an initial learning rate of 0.001, and the Adam optimization algorithm with a momentum coefficient $\beta_1 = 0.85$, and $\beta_2 = 0.9925$. Each of these settings makes a distinct contribution to the learning dynamics of the model.

Table 2. Default Hyperparameter

Hyperparameter	Values
epoch	50
batch size	32
learning rate	0.001
optimizer	Adam
β_1	0.85
β_1	0.9925

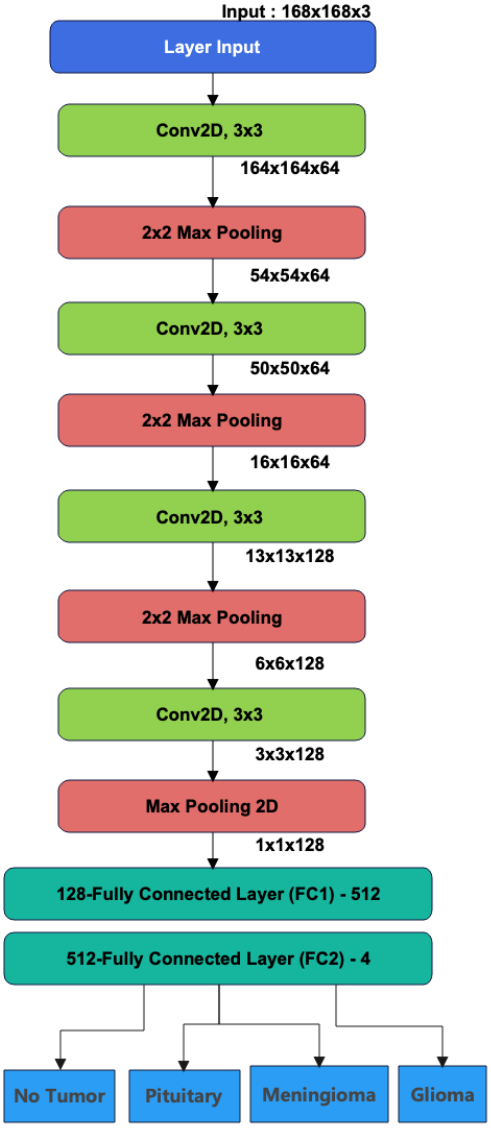


Fig. 4. Brain Cancer CNN Architecture

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

Training 50 epochs provides the optimizer with multiple opportunities to refine the parameter space, allowing the network to assimilate the discriminative patterns embedded within the data gradually. The batch size of 32 dictates how many samples are processed before a parameter update is performed. The learning rate of 0.001 represents the initial step size used in navigating the loss landscape. The Adam optimizer is employed due to its ability to adapt step sizes for individual parameters by maintaining moment estimates of both gradients and squared gradients.

The constructed Convolutional Neural Network (CNN) model features an architectural design similar to that depicted in Fig. 4. An early stopping mechanism and an adaptive learning rate strategy have been implemented within this architectural framework, specifically using the ReduceLROnPlateau function.

D. Whale Optimization Algorithm

The Whale Optimization Algorithm (WOA) was first proposed by Mirjalili and Lewis in 2016. This algorithm emulates the bubble-net foraging strategy of humpback whales, whereby the cetaceans encircle a congregation of fish and generate bubbles to trap them. The methodology employed in this algorithm is population-based, wherein each whale represents a prospective solution, and it methodically adjusts its position to move closer to the optimal solution identified. The principal phases of the algorithm encompass encircling the prey, executing bubble-net attacks, and conducting exploratory activities [58].

The Whale Optimization Algorithm (WOA) represents a sophisticated metaheuristic optimization technique that draws inspiration from the foraging behaviors exhibited by humpback whales, particularly their distinctive "bubble-net" hunting strategy. The underlying principle of WOA is to replicate the hunting dynamics of whales through three fundamental phases: the circling of prey, the spiral bubble-net approach, and a stochastic search to facilitate further exploration.

In this study, the Whale Optimization Algorithm (WOA) is employed to identify the optimal set of hyperparameters for a Convolutional Neural Network (CNN) trained on brain tumor classification tasks. The primary objective is to minimize the Cross-Entropy validation loss, thereby improving the network's generalization capability. Each whale in the population represents a four-dimensional hyperparameter vector, as shown in Eq. (7) [59]

$$X_i^t = [lr \quad factor \quad min\_lr \quad batch]_i^{t^\top} \in R^4 \tag{7}$$

where lr is the learning rate, $factor$ denotes the learning rate reduction factor on the plateau, min_lr is the minimum allowed learning rate, and $batch$ represents the minibatch size. WOA iteratively refines these parameters by exploring and exploiting the search space. For a multiclass brain tumor classification problem consisting of C tumor classes, each input image n yields a predicted probability vector $\widehat{y^n} \in R^C$ and a one-hot encoded ground truth vector $y^n \in R^C$. WOA performs optimization to minimize the loss function. The loss function used in this study is cross-entropy. The Cross-Entropy (CE) loss is defined as Eq. (8) [59]

$$\mathcal{L}_{\mathcal{CE}} = -\frac{1}{B}\sum_{n=1}^{B}\sum_{c=1}^{C} y_c^n \log \widehat{y_c^n}, \tag{8}$$

where B is the batch size. The fitness value of each whale i at iteration t is obtained by training the CNN using the hyperparameter vector $\mathrm{X_i^t}$ and computing Eq. (9) [60]

$$f_i^t = \mathcal{L}_{\mathcal{CE},val}(X_i^t). \tag{9}$$

The objective function can be shown by Eq. (10) [61]

$$X^* = arg\min_{X \in R^4} f\,(X). \tag{10}$$

A population of N whales are initialized randomly within the predefined hyperparameter bounds. WOA uses a linearly decreasing parameter $\mathrm{a(t)}$ to balance exploration and exploitation, where a(t) is shown as Eq. (11) [62]

$$a(t) = 2 - 2\frac{t}{T_{max}}, \tag{11}$$

where T_{max} notes the maximum number of iterations. For each whale, random vectors are generated as Eq. (12) [62]

$$r = [r_1, r_2, \ldots, r_4]^\top, \tag{12}$$

where component r_j denotes the j-th element of the random vector, corresponding to the j-th dimension of the hyperparameter search space. Each component of the random vector is sampled from a continuous uniform distribution on the interval [0,1], expressed concisely as shown in Eq. (13) [62]

$$r_j \sim \mathrm{U}(0,1), \tag{13}$$

where the notation ~ indicates that the random variable r_j is drawn according to a specified uniform probability distribution. The uniform distribution $U(0,1)$ assigns equal probability density to all values within the interval [0,1], and its probability density function is defined as Eq. (14) [63]

$$f_{U(0,1)}(x)\begin{cases}1, & 0 \le x \le 1\\ 0, & others\end{cases}. \tag{14}$$

Sampling each component r_j independently $U(0,1)$ ensures unbiased stochastic perturbations across all hyperparameter dimensions. These random values are subsequently used to compute the WOA coefficient vectors, Eq. 15 and Eq. 16 [64]

$$A = 2a(t)r - a(t), \tag{15}$$

$$C = 2r. \tag{16}$$

In the exploitation phase, when $|A| < 1$ the whale updates its position by moving toward the best-known hyperparameter vector, $\boldsymbol{X}^*(t)$ as shown in Eq. (17) [64]

$$\boldsymbol{D} = \left|\boldsymbol{C} \odot \boldsymbol{X}^*(t) - \boldsymbol{X_i^{(t)}}\right|, \tag{17}$$

$$\boldsymbol{X_i^{(t+1)}} = \boldsymbol{X}^*(t) - A \odot \boldsymbol{D}, \tag{18}$$

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

where $\odot$ denotes the Hadamard product, $i \in \{1, \ldots, N\}$ is the whale index and k is the index of a candidate spiral trajectory. $X_i^{(t)} \in R^d$ is the position of the whale i at iteration t. Furthermore $X^*(t) \in R^d$ is the best-known solution at iteration t. $D, D' \in R^d$ is the distance vectors.

WOA incorporates an additional exploitation mechanism modeled as whale spiral hunting behavior. With probability $p \geq 0.5$, the whale updates its position following a logarithmic spiral, which is modeled by Eq. (19) and Eq. (20) [64]

$$\boldsymbol{w^{(k)}}(t+1) = \boldsymbol{D'} \odot e^{bl} \cos(2\pi l) + \boldsymbol{w^*}(t), \quad (19)$$

where

$$\boldsymbol{D'} = \left| \boldsymbol{w^*}(t) - \boldsymbol{w^{(k)}}(t) \right|, \quad (20)$$

with $\mathrm{b} > 0$ is a spiral constant, and $l \in [-1,1]$ is a random number and $\boldsymbol{w^{(k)}}(t) \in R^d$ is the spiral candidate position.

E. Gradient-weighted Class Activation Mapping (Grad-CAM)

Gradient-weighted Class Activation Mapping (Grad-CAM) is a sophisticated visualization methodology that enhances the interpretability of convolutional neural networks (CNNs) by highlighting the regions within input images that have the most significant impact on the model's predictions. This technique produces a rudimentary localization map that accentuates substantial areas within the image, thereby facilitating researchers and clinicians in comprehending the cognitive processes underlying the decision-making of deep learning models [65]. Grad-CAM functions by employing the gradients associated with the target class that propagate into the terminal convolutional layer, thereby facilitating the identification of the prominent features that influence the model's output [66]. Grad-CAM involves computing the gradient of the class score with respect to the feature maps of the final convolutional layer. The key mathematical outputs can be expressed as Eq. (21) [67]

$$\alpha = \frac{\partial y^k}{\partial A_{ij}}, \quad (21)$$

where y^k is the score for the class k, and A_{ij} represents the activations of the feature map [29], [30]. The gradients are subsequently processed through global average pooling. This technique transforms these gradients into a weight vector [70]. The mathematical equation for this step is presented in Eq. (22) [67]

$$\alpha^{\mathrm{k}} = \frac{1}{Z} \sum i \sum j \frac{\partial y^k}{\partial Aij} \quad (22)$$

where Z is the number of pixels in the feature map A_{ij}, giving a normalization of the accumulated gradients for output stability. This vector delineates the significance of each feature map. This methodology accentuates areas that play a significant role in attaining the anticipated class score. The mathematical representation employed in Grad-CAM can be articulated as Eq. (23) [67]

$$L^k_{Grad-CAM} = RELU(\sum_k \alpha_k^c A^k) \quad (23)$$

where the RELU function is represented as shown in Eq. (2) and $L^k_{Grad-CAM} \in \mathbb{R}$ $L^k_{Grad-CAM}$ represents the class localization map, derived from the feature maps of the last convolutional layer [71], [72].

The primary role of Grad-CAM is to furnish visual explanations for the predictions generated by convolutional neural networks (CNNs). This aspect is especially vital in medical imaging, where comprehending the underlying rationale of a model's determinations can profoundly influence clinical outcomes. For example, in brain tumor identification, Grad-CAM can facilitate the localization of specific areas within MRI scans that the model deems indicative of tumor existence, thereby aiding radiologists in diagnostic evaluation. The F1-score integrates these two metrics, offering a harmonized measure that can be especially enlightening in scenarios characterized by class imbalance [73].

The aggregated insights from these metrics facilitate a more comprehensive understanding of model efficacy and empower more informed decision-making in advancing and implementing deep learning frameworks. For example, a research investigation demonstrated that disparate machine learning models exhibited varying degrees of performance in forecasting health outcomes, while using these metrics to benchmark performance proficiently [74]. Accuracy serves as a broad indicator of a model's efficacy; however, it may present a distorted representation when the dataset exhibits an imbalance [63]. To ascertain the value of accuracy, one may employ Eq. (24) [63] as delineated below

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP}. \quad (24)$$

This TP is for true positive, TN for true negative, TP for false positive, and FN FN for false negative. This metric measures the proportion of correct predictions from the total predictions made. To ascertain the ratio, where the accuracy of optimistic predictions is relative to the total number of optimistic predictions executed. The equation is shown as Eq. 25 [75]

$$Precision = \frac{TP}{TP+FP}. \quad (25)$$

Precision holds significant importance in scenarios where the repercussions of false positives are substantial, particularly in medical diagnostics [64]. The next metric is recall. The concept of recall is crucial, and identifying all positive examples, such as disease identification, is critical [75]. This metric assesses the proportion of accurate optimistic predictions about the overall count of actual positive occurrences. The formula is shown as Eq. 26 [75]:

$$Recall = \frac{TP}{TP+FN}. \quad (26)$$

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

The next metric is the F1 score. The F1-score represents the harmonic mean of precision and recall, striking a balance between the two metrics. The formula is shown as Eq. 27 [75]:

$$F1 - Score = \frac{2*precission*recall}{precission+recall} \quad (27)$$

III. Result

The results of this research commence with the preprocessing phase, which is precisely the augmentation technique. Overall, the combination of augmentations successfully increased data diversification without causing artifacts that damaged the image structure. Visual observation of the augmentation samples revealed that the main pattern characteristics were maintained, while the model acquired additional variation that facilitated the generalization process during training. The training process is carried out using three scenarios, and each scenario will be evaluated. The first scenario is using default parameters without adding callbacks. The second scenario is achieved by adding callbacks, specifically ReduceLROnPlateau and ModelCheckpoint. The last one uses callbacks and hyperparameter search results from WOA. The three results will be presented below.

Default parameters without adding callbacks: In this scenario, the accuracy of the training and validation results is obtained, as shown in Fig. 5. The loss function for this scenario is illustrated in Fig. 5b.

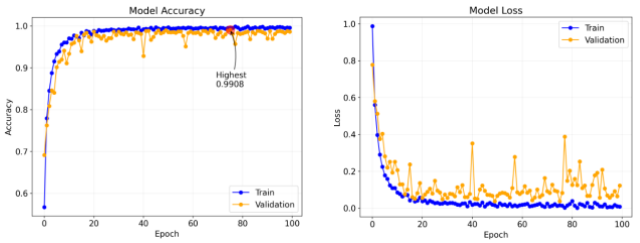


Fig. 5. Performance Vanilla Model (a) Accuracy Vanilla Model b) Loss Vanilla Model

Based on the calculations in Table 4, it can be inferred that the employed classification model exhibited commendable performance across all categories. The model achieved an accuracy rate of 99.14%, signifying that most predictions were accurate. Additionally, the Precision and Recall metrics for each category demonstrated remarkably high values, with the Notumor and Pituitary classes achieving nearly flawless precision and recall metrics. The F1-score, which indicates the equilibrium between precision and recall, also reflected outstanding values across the categories, with the Notumor class recording the highest F1-score of 0.998. Despite some misclassifications within the Glioma and Meningioma categories, these figures remained comparatively high, indicating that the model was proficient in distinguishing between the various tumor classes. In summary, the model exhibited an exceptional capability in tumor classification, characterized by exceedingly low error rates.

Table 4. Model Evaluation Result

Class	Precision	Recall	F1-Score
Glioma	0.997	0.98	0.988
Meningioma	0.983	0.986	0.984
Notumor	0.998	0.998	0.998
Pituitary	0.997	0.997	0.997
Accuracy			0.9914
Avg	0.994	0.99	0.992

Furthermore, the callback results are obtained by adding them, as shown in Fig. 6. Accuracy is illustrated in Fig. 6(a), and the loss function for this scenario is presented in Fig. 6(b).

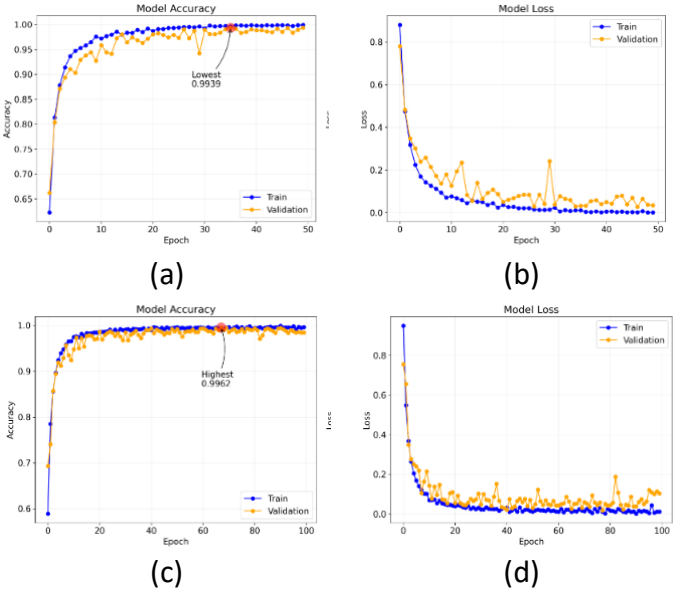
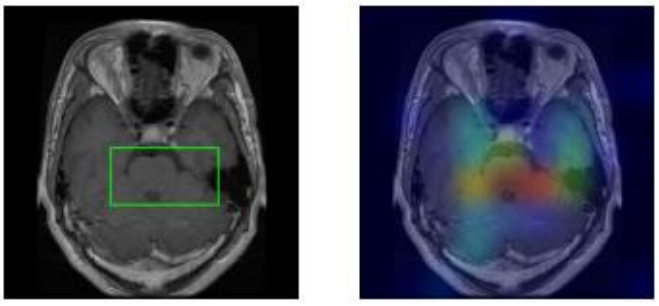


Fig. 6. Performance Proposed Model (a) Accuracy Vanilla Model with Callback (b) Loss Vanilla Model with Callback (c) Accuracy Vanilla Model with Callback and WOA (d) Loss Vanilla Model with Callback and WOA

This second scenario resulted in an increase in accuracy to 0.9936. This provides an improvement to the resulting model. As depicted in Fig. 6a, the model demonstrates convergence, improving accuracy compared to the preceding scenario. Next, a model with WOA is constructed to optimize performance. The

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

training results are shown in Fig. 9c. The accuracy obtained in this process is 0.9962. The Grad-CAM is illustrated in Fig. 7 to highlight its salient features. From Fig. 7, the area of the tumor on the brain MRI image can be identified. This can make it easier to understand and diagnose brain tumors.



(a) (b)

Fig. 7. Comparison Ground Truth and Grad-CAM

0.980, signifying commendable performance, albeit marginally trailing the CNN frameworks. 2DCNN+ensemble and Xception, while remaining valuable, exhibited comparatively lower accuracies of 0.9647 and 0.940, respectively, suggesting that ensemble methodologies and specific pre-trained networks may encounter limitations within the present context. The results of this study are compared with those of previous research. This comparison is illustrated in Table 5. Table 5 shows that the proposed method outperforms the earlier research, achieving an accuracy of 0.9962. The results of this study are compared with those of previous research. This comparison is illustrated in Table 5. The CNN model, devoid of Callback mechanisms (our methodology), attained an accuracy of 0.991, surpassing several prominent architectures, such as Xception and 2DCNN+ensemble, when evaluated on the integrated dataset. Table 5 highlights the proposed

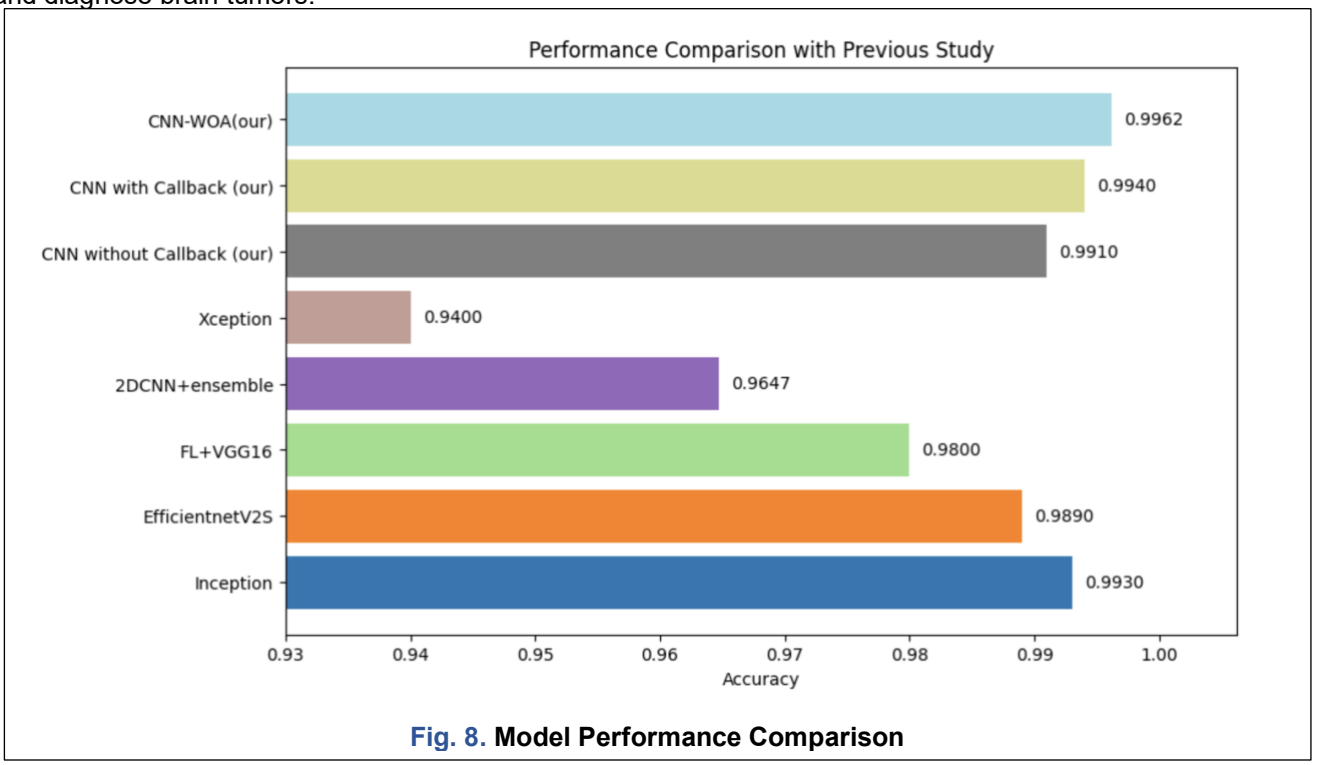


Fig. 8. Model Performance Comparison

IV. Discussion

From Fig. 8, it can be seen that the models derived from antecedent investigations, encompassing Inception, EfficientNetV2S, FL+VGG16, 2DCNN+ensemble, and Xception, all yielded significant outcomes; however, none attained the efficacy exhibited by the proposed models. For instance, Inception and EfficientNetV2S demonstrated commendable performance, achieving accuracies of 0.993 and 0.989, respectively. These architectures exemplify the effectiveness of pre-trained models in addressing intricate tasks. FL+VGG16 realized an accuracy of

method in bold, which achieves the best accuracy among the previous techniques. The CNN model incorporating callback mechanisms (ReduceLRonPlateau) demonstrated even superior accuracy, achieving a score of 0.994, which signifies the efficacy of employing callbacks to optimize the training process. The model that exhibited the highest performance in this investigation was CNN-WOA, which realized an accuracy of 0.9962, representing the pinnacle among all methods applied to the combined dataset. This observation implies that the Whale

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

Optimization Algorithm (WOA) enhances classification accuracy.

Table 5. Performance Comparison

Method	Dataset	Acc
Inception [11]	BR35H	0.993
EfficientNetV2S [13]	BR35H	0.989
FL+VGG16 [12]	Combine	0.980
2DCNN+ ensemble [14]	Combine	0.9647
Xception[15]	BR35H	0.940
CNN without Callback (our)	Combine	0.991
CNN with Callback (our)	Combine	0.994
CNN-WOA (our)	**Combine**	**0.9962**

This study set out to investigate whether a carefully tuned convolutional neural network (CNN) could achieve higher accuracy than existing approaches by leveraging training callbacks and meta-heuristic hyperparameter optimization. The comparative results in Table 5 show that the proposed CNN variants are competitive with, and in several cases superior to, previously reported models, including Inception [11], EfficientNetV2S [13], FL+VGG16 [12], 2DCNN+ensemble [14], and Xception [15].

When focusing on methods evaluated on the combined dataset, the baseline CNN without callbacks achieved an accuracy of 0.991. Introducing training callbacks in the form of ReduceLROnPlateau and ModelCheckpoint increased the accuracy to 0.994, corresponding to an absolute gain of 0.30 percentage points (difference callback vs. no-callback: 0.994 – 0.991 = 0.003, i.e., 0.3%). This suggests that relatively simple training strategies, such as adaptive learning-rate scheduling and model checkpointing, already yield a measurable performance improvement, facilitating the model's convergence to a better local optimum.

The integration of the Whale Optimization Algorithm (WOA) for hyperparameter tuning further boosted the accuracy of the CNN to 0.9962 on the same combined dataset. Compared to the CNN without callbacks, this represents an absolute improvement of 0.52 percentage points (difference CNN-WOA vs. CNN without callback: 0.9962 – 0.991 = 0.0052 ≈ 0.52%), and compared to the CNN with callbacks, the improvement is 0.22 percentage points (difference CNN-WOA vs. CNN with callback: 0.9962 – 0.994 = 0.0022 ≈ 0.22%). Although these numerical differences are below 1%, they occur in an already high-accuracy regime (>0.99), where small gains are typically difficult to achieve and can still be meaningful for safety-critical applications such as brain tumor analysis.

A broader comparison with other methods on the combined dataset reinforces this conclusion. The FL+VGG16[12] approach reports an accuracy of 0.980, whereas 2DCNN+ensemble achieves 0.9647. The proposed CNN variants improve upon FL+VGG16 by 1.10, 1.40, and 1.62 percentage points for the CNN without callbacks (0.991 – 0.980 = 0.011), CNN with callbacks (0.994 – 0.980 = 0.014), and CNN-WOA (0.9962 – 0.980 = 0.0162), respectively. Relative to 2DCNN+ensemble [14], the gains are even larger: 2.63, 2.93, and 3.15 percentage points for the same three CNN variants. These differences suggest that both the architectural design of the CNN and the proposed optimization strategy contribute to systematically higher performance on the combined dataset.

It is also informative to compare the proposed models with architectures evaluated on the BR35H dataset only. Inception achieves an accuracy of 0.993, EfficientNetV2S [13] achieves 0.989, and Xception [15] yields 0.940. While these results are reported on a single dataset, the proposed CNN-WOA on the more heterogeneous combined dataset still attains a slightly higher accuracy of 0.9962. The difference between CNN-WOA and Inception is approximately 0.32 percentage points (0.9962 – 0.993 = 0.0032), and CNN with callbacks exceeds Inception by about 0.10 percentage points (0.994 – 0.993 = 0.001). Although cross-dataset comparisons should be interpreted with caution, these numbers indicate that the proposed approach remains competitive even when contrasted with state-of-the-art deep architectures. According to prior research, deep learning methodologies can yield favorable outcomes contingent upon the optimal selection of hyperparameters.

Overall, the numerical analysis shows a consistent trend: (i) adding callbacks yields a modest but clear improvement over the plain CNN, and (ii) incorporating WOA-based hyperparameter tuning produces the highest accuracy among all evaluated methods. These findings support the conclusion that systematic hyperparameter optimization, combined with appropriate training strategies, can lead to measurable accuracy gains beyond what is achievable by relying solely on architectural complexity. Fig. 7 is the Grad-CAM visualization, highlighting the areas that contributed most to the model's decision. Warmer colors (red, orange, and yellow) indicate regions with a larger influence on classification, while cooler colors (blue and purple) represent less significant areas.

V. Conclusion

This investigation aimed to enhance the efficacy and interpretability of deep learning algorithms used for brain tumor classification by integrating training callbacks and meta-heuristic hyperparameter optimization techniques. The findings reveal that

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

integrating ReduceLROnPlateau and ModelCheckpoint callbacks resulted in an increase in the CNN accuracy to 0.994. Furthermore, applying hyperparameter tuning via the Whale Optimization Algorithm (WOA) further elevated the accuracy to 0.996. Moreover, supplementary analyses using Grad-CAM yielded more accurate visual representations that distinguished between tumor-affected and unaffected cerebral regions, thereby bolstering the interpretability of the model. The results also indicate that the implementation of meta-heuristic optimization significantly prolongs training durations, thereby imposing practical limitations on extensive experimental undertakings. Subsequent investigations should explore more rapid or adaptive optimization methodologies, such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), Grey Wolf Optimization (GWO), or hybrid approaches, and assess the proposed methodology on larger and more heterogeneous multi-institutional datasets to enhance both robustness and generalizability.

Acknowledgment

Sincere gratitude is extended to LPPM UNS for its significant financial support and assistance with this research, as outlined in the 2025 budget. The monetary aid rendered has been pivotal in facilitating the successful implementation of the project. Without the confidence and investment from LPPM UNS, this research initiative would not have been feasible. The support has not only aided in the progression of the study but has also played a significant role in enhancing and evolving the academic community. We express our gratitude for the ongoing dedication to promoting research excellence.

Funding

This research was funded by the Universitas Sebelas Maret RKAT Fiscal Year 2024 through the MITRASMART Scheme.

Author Contribution

Winarno conceptualized and designed the study, conducted data collection, and participated in data analysis and interpretation. Agus Suhajoko provided critical feedback on the manuscript.

Declarations

Ethical Approval

The research guide reviewed and ethically approved this manuscript for publication in the Journal.

Consent for Publication Participants.

Not applicable.

Competing Interests

The authors declarehas no conflict of interest related to this publication

References

[1] A. Roberts, M. Hu, and M. Hajizadeh, “Income and Education Inequalities in Brain and Central Nervous System Cancer Incidence in Canada: Trends over Two Decades,” *J Cancer Prev*, vol. 26, no. 2, pp. 110–117, June 2021, doi: 10.15430/JCP.2021.26.2.110.

[2] N. A. Larimi *et al.*, “An investigation of efficient nursing interventions in early diagnosis of cancer: A systematic review and meta-analysis,” *Journal of Family Medicine and Primary Care*, vol. 10, no. 8, pp. 2964–2968, Aug. 2021, doi: 10.4103/jfmpc.jfmpc_2148_20.

[3] P. S. Chaitanya and S. K. Satpathy, “A Multilevel De-Noising Approach for Precision Edge-Based Fragmentation in MRI Brain Tumor Segmentation,” *TS*, vol. 40, no. 4, pp. 1715–1722, Aug. 2023, doi: 10.18280/ts.400440.

[4] R. Asad, S. U. Rehman, A. Imran, J. Li, A. Almuhaimeed, and A. Alzahrani, “Computer-Aided Early Melanoma Brain-Tumor Detection Using Deep-Learning Approach,” *Biomedicines*, vol. 11, no. 1, p. 184, Jan. 2023, doi: 10.3390/biomedicines11010184.

[5] M. Praveena and M. K. Rao, “Brain Tumor Detection using Integrated Learning Process Detection (ILPD),” *IJACSA*, vol. 13, no. 10, 2022, doi: 10.14569/IJACSA.2022.0131018.

[6] Mrs. D. S. W. Dr. Selvarani Rangasamy, “Review On Deep Learning Approach For Brain Tumor Glioma Analysis,” *ITII*, vol. 9, no. 1, pp. 395–408, Mar. 2021, doi: 10.17762/itii.v9i1.144.

[7] K. Sailunaz, D. Bestepe, S. Alhajj, T. Özyer, J. Rokne, and R. Alhajj, “Brain tumor detection and segmentation: Interactive framework with a visual interface and feedback facility for dynamically improved accuracy and trust,” *PLoS ONE*, vol. 18, no. 4, p. e0284418, Apr. 2023, doi: 10.1371/journal.pone.0284418.

[8] M. A. Mahjoubi, S. Hamida, O. E. Gannour, B. Cherradi, A. E. Abbassi, and A. Raihani, “Improved Multiclass Brain Tumor Detection using Convolutional Neural Networks and Magnetic Resonance Imaging,” *IJACSA*, vol. 14, no. 3, 2023, doi: 10.14569/IJACSA.2023.0140346.

[9] A. Raza *et al.*, “A Hybrid Deep Learning-Based Approach for Brain Tumor Classification,” *Electronics*, vol. 11, no. 7, p. 1146, Apr. 2022, doi: 10.3390/electronics11071146.

[10] B. Badjie and E. Deniz Ülker, “A Deep Transfer Learning Based Architecture for Brain Tumor

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

Classification Using MR Images," *ITC*, vol. 51, no. 2, pp. 332–344, June 2022, doi: 10.5755/j01.itc.51.2.30835.

[11] D. Rastogi, P. Johri, V. Tiwari, and A. A. Elngar, "Multi-class classification of brain tumour magnetic resonance images using multi-branch network with inception block and five-fold cross validation deep learning framework," *Biomedical Signal Processing and Control*, vol. 88, p. 105602, Feb. 2024, doi: 10.1016/j.bspc.2023.105602.

[12] E. Albalawi *et al.*, "Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor," *BMC Med Imaging*, vol. 24, no. 1, p. 110, May 2024, doi: 10.1186/s12880-024-01261-0.

[13] M. M. E. Yurtsever, Y. Atay, B. Arslan, and S. Sagiroglu, "Development of brain tumor radiogenomic classification using GAN-based augmentation of MRI slices in the newly released gazi brains dataset," *BMC Med Inform Decis Mak*, vol. 24, no. 1, p. 285, Oct. 2024, doi: 10.1186/s12911-024-02699-6.

[14] S. Saeedi, S. Rezayi, H. Keshavarz, and S. R. Niakan Kalhori, "MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques," *BMC Med Inform Decis Mak*, vol. 23, no. 1, p. 16, Jan. 2023, doi: 10.1186/s12911-023-02114-6.

[15] S. Sundari.M, Y. Divya, K. Durga, V. Sukhavasi, M. D. Sugnana Rao, and M. S. Rani, "A Stable Method For Brain Tumor Prediction In Magnetic Resonance Images Using Fine-tuned XceptionNet," *IJCDS*, vol. 15, no. 1, pp. 67–79, Jan. 2024, doi: 10.12785/ijcds/150106.

[16] M. Rasool *et al.*, "A Hybrid Deep Learning Model for Brain Tumour Classification," *Entropy*, vol. 24, no. 6, p. 799, June 2022, doi: 10.3390/e24060799.

[17] Y. Xu, L. Huang, L. Zhang, L. Qian, and X. Yang, "Diffusion-Based Radio Signal Augmentation for Automatic Modulation Classification," *Electronics*, vol. 13, no. 11, p. 2063, May 2024, doi: 10.3390/electronics13112063.

[18] C. Zhang, G. Sheng, J. Su, and L. Duan, "Color fundus photograph-based diabetic retinopathy grading via label relaxed collaborative learning on deep features and radiomics features," *Front. Cell Dev. Biol.*, vol. 12, p. 1513971, Jan. 2025, doi: 10.3389/fcell.2024.1513971.

[19] C. Shi, A. K. Chiu, and H. Xu, "Evaluating Designs for Hyperparameter Tuning in Deep Neural Networks," *The New England Journal of Statistics in Data Science*, pp. 334–341, 2023, doi: 10.51387/23-NEJSDS26.

[20] S. S. Sulaiman, I. Nadher, and S. M. Hameed, "Credit Card Fraud Detection Using Improved Deep Learning Models," *CMC*, vol. 78, no. 1, pp. 1049–1069, 2024, doi: 10.32604/cmc.2023.046051.

[21] I. A. Dewi and M. A. Rizqullah, "Sentiment Analysis on Twitter Using Deep Belief Network Optimized with Particle Swarm Optimization," *E3S Web Conf.*, vol. 484, p. 02001, 2024, doi: 10.1051/e3sconf/202448402001.

[22] L. P. Swaminatha Rao and S. Jaganathan, "Adaptive Bayesian contextual hyperband: A novel hyperparameter optimization approach," *IJ-AI*, vol. 13, no. 1, p. 775, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp775-785.

[23] A. Sanmorino, L. Marnisah, and H. D. Kesuma, "Detection of DDoS Attacks using Fine-Tuned Multi-Layer Perceptron Models," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 5, pp. 16444–16449, Oct. 2024, doi: 10.48084/etasr.8362.

[24] J. A Ilemobayo *et al.*, "Hyperparameter Tuning in Machine Learning: A Comprehensive Review," *J. Eng. Res. Rep.*, vol. 26, no. 6, pp. 388–395, June 2024, doi: 10.9734/jerr/2024/v26i61188.

[25] H. Qu *et al.*, "BD-StableNet: a deep stable learning model with an automatic lesion area detection function for predicting malignancy in BI-RADS category 3–4A lesions," *Phys. Med. Biol.*, vol. 69, no. 24, p. 245002, Dec. 2024, doi: 10.1088/1361-6560/ad953e.

[26] N. Bai and I. Joe, "Deep Learning Methods With the Improved Attention for Explainable Image Recognition," *IEEE Access*, vol. 12, pp. 70559–70567, 2024, doi: 10.1109/ACCESS.2024.3397323.

[27] Dr. S. Singh, Dr. D. Pratap Singh, and Mr. K. Chandra, "Enhancing Transparency and Interpretability in Deep Learning Models: A Comprehensive Study on Explainable AI Techniques," *IJSREM*, vol. 08, no. 02, pp. 1–13, Feb. 2024, doi: 10.55041/IJSREM28675.

[28] F. Lin *et al.*, "Postoperative One Year Prediction for Patients with Cervical Spinal Cord Injury Based on Deep Learning and Radiomics," Oct. 08, 2024. doi: 10.21203/rs.3.rs-4848654/v1.

[29] S. Huang *et al.*, "Deep learning model to predict lupus nephritis renal flare based on dynamic multivariable time-series data," *BMJ Open*, vol. 14, no. 3, p. e071821, Mar. 2024, doi: 10.1136/bmjopen-2023-071821.

[30] J. Cheng, "brain tumor dataset." figshare, p. 879509079 Bytes, 2017. doi: 10.6084/M9.FIGSHARE.1512427.V5.

[31] Msoud Nickparvar, "Brain Tumor MRI Dataset." Kaggle. doi: 10.34740/KAGGLE/DSV/2645886.

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

[32] P. Hu, Y. Gao, Y. Zhang, and K. Sun, "Ultrasound image-based deep learning to differentiate tubal-ovarian abscess from ovarian endometriosis cyst," *Front. Physiol.*, vol. 14, p. 1101810, Feb. 2023, doi: 10.3389/fphys.2023.1101810.

[33] W. Sapitri, Y. N. Kunang, I. Z. Yadi, and M. Mahmud, "The Impact of Data Augmentation Techniques on the Recognition of Script Images in Deep Learning Models," *join*, vol. 8, no. 2, pp. 169–176, Dec. 2023, doi: 10.15575/join.v8i2.1073.

[34] K. Meethongjan, V. T. Hoang, and T. Surinwarangkoon, "Data augmentation by combining feature selection and color features for image classification," *IJECE*, vol. 12, no. 6, p. 6172, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6172-6177.

[35] X. Luo, Q. Nie, Y. Wang, and Z. Zhao, "Data augmentation techniques based on deep learning for Chinese paintings," in *Third International Conference on Computer Graphics, Image, and Virtualization (ICCGIV 2023)*, Y. Wang and A. J. Moshayedi, Eds., Nanjing, China: SPIE, Nov. 2023, p. 43. doi: 10.1117/12.3008104.

[36] R. N. D. Sousa and R. N. D. Sousa, "Development of a Convolutional Neural Network Architecture for Classifying Foliar Diseases in Plants," in *Anais do XVI Encontro Unificado de Computação do Piauí (ENUCOMPI 2023)*, Brasil: Sociedade Brasileira de Computação, Oct. 2023, pp. 1–8. doi: 10.5753/enucompi.2023.26610.

[37] A. Jokic, L. Djokic, M. Petrovic, and Z. Miljkovic, "Data augmentation methods for semantic segmentation-based mobile robot perception system," *Serb J Electr Eng*, vol. 19, no. 3, pp. 291–302, 2022, doi: 10.2298/SJEE2203291J.

[38] A. Nabilah, R. Sigit, A. Fariza, and M. Madyono, "Human Bone Age Estimation of Carpal Bone X-Ray Using Residual Network with Batch Normalization Classification," *JOIV : Int. J. Inform. Visualization*, vol. 7, no. 1, p. 105, Jan. 2023, doi: 10.30630/joiv.7.1.1024.

[39] S. Nesteruk *et al.*, "XtremeAugment: Getting More From Your Data Through Combination of Image Collection and Image Augmentation," *IEEE Access*, vol. 10, pp. 24010–24028, 2022, doi: 10.1109/ACCESS.2022.3154709.

[40] R. Nair, S. Vishwakarma, M. Soni, T. Patel, and S. Joshi, "Detection of COVID-19 cases through X-ray images using hybrid deep neural network," *WJE*, vol. 19, no. 1, pp. 33–39, Feb. 2022, doi: 10.1108/WJE-10-2020-0529.

[41] H. R. Mohammed and Z. M., "Detection and Recognition of Moving Video Objects: Kalman Filtering with Deep Learning," *IJACSA*, vol. 12, no. 1, 2021, doi: 10.14569/IJACSA.2021.0120118.

[42] C.-Y. Yang and H.-M. Lee, "Effects of the Hyperparameters on CNNs for MDD Classification Using Resting-State EEG," *Electronics*, vol. 13, no. 1, p. 186, Dec. 2023, doi: 10.3390/electronics13010186.

[43] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.

[44] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar: Association for Computational Linguistics, 2014, pp. 1746–1751. doi: 10.3115/v1/D14-1181.

[45] M.-E. Mickael *et al.*, "Using Copy Number Variation Data and Neural Networks to Predict Cancer Metastasis Origin Achieves High Area under the Curve Value with a Trade-Off in Precision," *CIMB*, vol. 46, no. 8, pp. 8301–8319, Aug. 2024, doi: 10.3390/cimb46080490.

[46] P. Purwono, A. Ma'arif, W. Rahmaniar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. U. Haq, "Understanding of Convolutional Neural Network (CNN): A Review," *IJRCS*, vol. 2, no. 4, pp. 739–748, Jan. 2023, doi: 10.31763/ijrcs.v2i4.888.

[47] S.-Y. Hwang and J.-J. Kim, "A Universal Activation Function for Deep Learning," *Computers, Materials & Continua*, vol. 75, no. 2, pp. 3553–3569, 2023, doi: 10.32604/cmc.2023.037028.

[48] A. Muis, E. M. Zamzami, and E. B. Nababan, "Convolutional Neural Network Activation Function Performance on Image Recognition of The Batak Script," *SinkrOn*, vol. 9, no. 1, pp. 182–195, Jan. 2024, doi: 10.33395/sinkron.v9i1.13192.

[49] D. Florek and M. Miłosz, "Comparison of an effectiveness of artificial neural networks for various activation functions," *J. Comput. Sci. Inst.*, vol. 26, pp. 7–12, Mar. 2023, doi: 10.35784/jcsi.3069.

[50] G. Madhu, S. Kautish, K. A. Alnowibet, H. M. Zawbaa, and A. W. Mohamed, "NIPUNA: A Novel Optimizer Activation Function for Deep Neural Networks," *Axioms*, vol. 12, no. 3, p. 246, Feb. 2023, doi: 10.3390/axioms12030246.

[51] Y. S. Kim, H. Pham, and I. H. Chang, "Deep-Learning Software Reliability Model Using SRGM as Activation Function," *Applied Sciences*, vol. 13, no. 19, p. 10836, Sept. 2023, doi: 10.3390/app131910836.

[52] N. Nimra, J. U. Rahman, and D. Lu, "Modified scaled exponential linear unit," *Math. Syst. Sci.*,

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

vol. 2, no. 2, Oct. 2024, doi: 10.54517/mss.v2i2.2870.
[53] H. Liang *et al.*, "Artificial Neurons Based on a Threshold Switching Memristor with Ultralow Threshold Voltage," *ACS Appl. Electron. Mater.*, vol. 7, no. 7, pp. 3019–3029, Apr. 2025, doi: 10.1021/acsaelm.5c00188.
[54] M Mesran, Sitti Rachmawati Yahya, Fifto Nugroho, and Agus Perdana Windarto, "Investigating the Impact of ReLU and Sigmoid Activation Functions on Animal Classification Using CNN Models," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 8, no. 1, pp. 111–118, Feb. 2024, doi: 10.29207/resti.v8i1.5367.
[55] C. Gulcehre, K. Cho, R. Pascanu, and Y. Bengio, "Learned-Norm Pooling for Deep Feedforward and Recurrent Neural Networks," in *Machine Learning and Knowledge Discovery in Databases*, vol. 8724, T. Calders, F. Esposito, E. Hüllermeier, and R. Meo, Eds., in Lecture Notes in Computer Science, vol. 8724. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 530–546. doi: 10.1007/978-3-662-44848-9_34.
[56] H. Afzaal *et al.*, "Detection of a Potato Disease (Early Blight) Using Artificial Intelligence," *Remote Sensing*, vol. 13, no. 3, p. 411, Jan. 2021, doi: 10.3390/rs13030411.
[57] P. K. Yadav *et al.*, "Citrus disease classification with convolution neural network generated features and machine learning classifiers on hyperspectral image data," in *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping VIII*, C. Bauer and J. A. Thomasson, Eds., Orlando, United States: SPIE, June 2023, p. 5. doi: 10.1117/12.2665768.
[58] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, May 2016, doi: 10.1016/j.advengsoft.2016.01.008.
[59] O. N. Oyelade and A. E. Ezugwu, "Characterization of abnormalities in breast cancer images using nature-inspired metaheuristic optimized convolutional neural networks model," *Concurrency and Computation*, vol. 34, no. 4, p. e6629, Feb. 2022, doi: 10.1002/cpe.6629.
[60] A. Bahaa, A. Sayed, L. Elfangary, and H. Fahmy, "A novel hybrid optimization enabled robust CNN algorithm for an IoT network intrusion detection approach," *PLoS ONE*, vol. 17, no. 12, p. e0278493, Dec. 2022, doi: 10.1371/journal.pone.0278493.
[61] R. M. Lewis and V. Torczon, "Pattern Search Algorithms for Bound Constrained Minimization," *SIAM J. Optim.*, vol. 9, no. 4, pp. 1082–1099, Jan. 1999, doi: 10.1137/S1052623496300507.
[62] R. Murugan, T. Goel, S. Mirjalili, and D. K. Chakrabartty, "WOANet: Whale optimized deep neural network for the classification of COVID-19 from radiography images," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 4, pp. 1702–1718, Oct. 2021, doi: 10.1016/j.bbe.2021.10.004.
[63] N. M. Ashraf, R. R. Mostafa, R. H. Sakr, and M. Z. Rashad, "Optimizing hyperparameters of deep reinforcement learning for autonomous driving based on whale optimization algorithm," *PLoS ONE*, vol. 16, no. 6, p. e0252754, June 2021, doi: 10.1371/journal.pone.0252754.
[64] A. Brodzicki, M. Piekarski, and J. Jaworek-Korjakowska, "The Whale Optimization Algorithm Approach for Deep Neural Networks," *Sensors*, vol. 21, no. 23, p. 8003, Nov. 2021, doi: 10.3390/s21238003.
[65] J. Kim, C. M. Park, S. Y. Kim, and A. Cho, "Convolutional neural network-based classification of cervical intraepithelial neoplasias using colposcopic image segmentation for acetowhite epithelium," *Sci Rep*, vol. 12, no. 1, p. 17228, Oct. 2022, doi: 10.1038/s41598-022-21692-5.
[66] H. Qi *et al.*, "Rice seed vigor detection based on near-infrared hyperspectral imaging and deep transfer learning," *Front. Plant Sci.*, vol. 14, p. 1283921, Oct. 2023, doi: 10.3389/fpls.2023.1283921.
[67] F. M. Talaat, S. A. Gamel, R. M. El-Balka, M. Shehata, and H. ZainEldin, "Grad-CAM Enabled Breast Cancer Classification with a 3D Inception-ResNet V2: Empowering Radiologists with Explainable Insights," *Cancers*, vol. 16, no. 21, p. 3668, Oct. 2024, doi: 10.3390/cancers16213668.
[68] K. Nakajo *et al.*, "Anatomical classification of pharyngeal and laryngeal endoscopic images using artificial intelligence," *Head & Neck*, vol. 45, no. 6, pp. 1549–1557, June 2023, doi: 10.1002/hed.27370.
[69] V. Jahmunah, E. Y. K. Ng, R.-S. Tan, S. L. Oh, and U. R. Acharya, "Explainable detection of myocardial infarction using deep learning models with Grad-CAM technique on ECG signals," *Computers in Biology and Medicine*, vol. 146, p. 105550, July 2022, doi: 10.1016/j.compbiomed.2022.105550.
[70] J. P. Cruz-Bastida, E. Pearson, and H. Al-Hallaq, "Toward understanding deep learning classification of anatomic sites: lessons from the development of a CBCT projection classifier," *J. Med. Imag.*, vol. 9, no. 04, July 2022, doi: 10.1117/1.JMI.9.4.045002.
[71] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941

Gradient-Based Localization," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice: IEEE, Oct. 2017, pp. 618–626. doi: 10.1109/ICCV.2017.74.

[72] M. Giavina-Bianchi, W. G. Vitor, V. Fornasiero De Paiva, A. L. Okita, R. M. Sousa, and B. Machado, "Explainability agreement between dermatologists and five visual explanations techniques in deep neural networks for melanoma AI classification," *Front. Med.*, vol. 10, p. 1241484, Aug. 2023, doi: 10.3389/fmed.2023.1241484.

[73] S. Bomrah *et al.*, "A scoping review of machine learning for sepsis prediction- feature engineering strategies and model performance: a step towards explainability," *Crit Care*, vol. 28, no. 1, p. 180, May 2024, doi: 10.1186/s13054-024-04948-6.

[74] N. A. Brennan, W. Shamp, E. Maynes, X. Cheng, and M. A. Bullimore, "Influence of age and race on axial elongation in myopic children: A systematic review and meta-regression," *Optom Vis Sci*, vol. 101, no. 8, pp. 497–507, Aug. 2024, doi: 10.1097/OPX.0000000000002176.

[75] A. Bahaa, A. Sayed, L. Elfangary, and H. Fahmy, "A novel hybrid optimization enabled robust CNN algorithm for an IoT network intrusion detection approach," *PLoS ONE*, vol. 17, no. 12, p. e0278493, Dec. 2022, doi: 10.1371/journal.pone.0278493.

Author Biography

Winarno has been a researcher at the Faculty of Information Technology and Data Science, Universitas Sebelas Maret Surakarta. He received a Bachelor's Degree in Mathematics and an M.Eng in Electrical Engineering from Universitas Gadjah Mada, Yogyakarta, Indonesia. He work in the Departement of Informatics, Universitas Sebelas Maret Surakarta. His academic pursuits encompass machine learning, deep learning, distributed systems, cybersecurity, and federated learning. He is pursuing doctoral studies within the Department of Computer Science and Electronics at FMIPA Gadjah Mada University. Furthermore, he frequently engages in collaborative endeavors with multiple domestic and international entities to advance software development with various organizations, including a partnership with Rapixus, Inc., located in Taiwan, focusing on cybersecurity initiatives. For matters of communication, he may be contacted at win@staff.uns.ac.id.

Agus Harjoko is a Professor at the Department of Computer Science and Electronics Universitas Gadjah Mada Yogyakarta. received the bachelor's degree in electronics and instrumentation from Universitas Gadjah Mada, Yogyakarta, Indonesia, in 1986, and the M.Sc. and Ph.D. degrees in computer science from the University of New Brunswick, Fredericton, NB, Canada, in 1990 and 1996, respectively. Studi beliau dibiayai oleh Bank Dunia. His research interests include computer vision, pattern recognition, instrumentation, and sensor networks. An accomplished scholar and educator, he devoted more than 35 years to Universitas Gadjah Mada (UGM), functioning as a lecturer from March 1987 until December 2022. During this extensive tenure, he contributed significantly to advancing the disciplines of electronics and computer science through pedagogical efforts, research endeavors, and academic leadership. From January 2016 to January 2021, he was Chair of the Department of Computer Science and Electronics, where he spearheaded departmental initiatives and promoted scholarly development. Thereafter, he was appointed Head of the Electronics and Instrumentation Laboratory from April 2021 to January 2022, followed by an additional term commencing in January 2022. To establish communication with him, kindly direct electronic correspondence to aharjoko@ugm.ac.id.

Manuscript received 25 May 2025; Revised 12 September 2025; Accepted 8 November 2025; Available online 16 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.941