

## Utilization of Metaheuristic Algorithms in Hyperparameter Tuning Deep Learning for Image Classification: Systematic Literature Review

Wirya Wonggo\*, Joko Siswantoro

Universitas Surabaya, Indonesia

Email: wonggowirya@gmail.com\*, joko\_siswantoro@staff.ubaya.ac.id

---

### Keywords

image classification; deep learning; hyperparameter tuning; metaheuristics; systematic literature review.

---

### ABSTRACT

This study aims to systematically examine the application of metaheuristic algorithms in hyperparameter tuning for deep learning-based image classification models. Selecting an appropriate hyperparameter configuration is a crucial factor that determines model performance, however, the hyperparameter selection process is often time consuming and inefficient when performed manually or through conventional approaches such as Grid Search and Random Search. Therefore, metaheuristic algorithms have emerged as a promising alternative, as they are capable of adaptively exploring and exploiting the search space in complex optimization problems. This research employed a Systematic Literature Review (SLR) approach following the PRISMA protocol, supported by the Zotero reference management tool. The PRISMA protocol ensures a transparent and structured process for identification, selection, and data extraction. A total of 41 scientific articles that met the inclusion criteria were analyzed based on publication year, application domain, classification models used, metaheuristic algorithms applied, datasets employed, and the performance evaluation results. The findings show that the application of metaheuristic algorithms significantly improves the performance of image classification tasks. Convolutional Neural Networks (CNNs) and their variants are the most widely used architectures, while common datasets such as MNIST and CIFAR-10 are frequently employed in experiments. This study provides an overview of current research trends, demonstrates the effectiveness of metaheuristic based hyperparameter optimization, and highlights opportunities for developing more efficient optimization strategies. The results are expected to serve as a valuable reference for future research in image classification using metaheuristic optimization within deep learning frameworks.

---

### INTRODUCTION

Image classification is one of the main problems in the field of computer vision and artificial intelligence (Artificial Intelligence). Its ability to extract, interpret, and categorize information, especially visual information, makes image classification a very important technology in various application domains, ranging from medical, security, agriculture, manufacturing industries, to environmental monitoring (Alzubaidi et al., 2021; Shah, 2024). In recent years, rapid advances in deep learning models have driven significant improvements in image classification accuracy and capabilities. These deep learning models are able to learn features automatically, resulting in performance that exceeds traditional classification methods.

However, the success of deep learning models is highly dependent on the right hyperparameter configuration. Hyperparameters such as learning rate, number of neurons, dropout, batch size, number of layers, and optimizer algorithms used, are important factors that determine the stability of the training process, generalization ability, convergence speed of the model, and model performance (Purnomo, Tad Gonsalves, et al., 2024). Inappropriate selection of hyperparameters can cause the model to overfit, underfit, and take a very long training process, and even fail to achieve convergence.

In practice, the selection of hyperparameters is still done manually with a trial-and-error approach. This approach requires large, very long computation, and does not guarantee optimal results, especially in large-scale models with many parameters that are interrelated with each other (Purnomo, Tad Gonsalves, et al., 2024). There are two methods that are most often used in hyperparameter selection, namely Grid Search and Random Search, but both conventional methods also have limitations in terms of search time in a very large multidimensional parameter space (Purnomo, Gonsalves, et al., 2024). Grid Search has a very high computational complexity, because Grid Search will explore all combinations of hyperparameters, while Random Search is more efficient because it does not explore all hyperparameters, but Random Search does not guarantee that it can provide the best combination of hyperparameters as a result.

In addition to using Grid Search and Random Search, hyperparameter optimization can be done using methodological optimization, which is optimization with traditional mathematical calculations. Examples of traditional mathematical optimization are Linear Programming (LP), Non-Linear Programming (NLP), Mixed-Integer-Programming, and others (Han et al., 2024). But optimization using traditional mathematical calculations also has disadvantages, namely dependence on mathematical models which if the objective function used and its limits do not meet the requirements of the first and second differentiation, it will cause the model to become non-convergent, in addition to this optimization depends on the starting point of the calculation, if the starting point is wrong, the results given will be low.

To overcome these challenges, metaheuristic algorithms are starting to be widely used as alternative solutions in the optimization of hyperparameter deep learning models (Benaissa et al., 2024; Raiaan et al., 2024). Metaheuristics is a nature-inspired optimization technique that relies on heuristic principles to explore solution spaces efficiently, flexibly, and adaptively, even in non-linear problems. This algorithm is designed to explore solution spaces more efficiently through a combination of local exploration and adaptive global exploration. Some of the metaheuristic algorithms that are widely adopted include Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Gray Wolf Optimizer (GWO), as well as various new algorithm variants such as Jellyfish Optimization (JFO), Henry Gas Solubility Optimization (HGSO), Harmony Search (HS), and Egyptian Vulture Optimization (EVO) (Alnabhan et al., 2022; Purnomo, Gonsalves, et al., 2024; Ribeiro, 2022; Saber et al., 2024; Yang, 2024). Various studies show that using metaheuristic can significantly improve the performance of deep learning by finding more optimal hyperparameter configurations compared to traditional methods as well as Grid Search and Random Search.

The novelty of this study is sixfold. First, it applies the PRISMA protocol to ensure a transparent, reproducible, and bias-free systematic literature review process, with clear documentation of inclusion and exclusion criteria at each stage (identification, screening,

\

eligibility, inclusion). Second, it analyzes 41 high-quality articles published between 2015-2025, extracting data across six dimensions: publication year trends, application domains, deep learning models, metaheuristic algorithms, datasets, and classification accuracy. Third, it provides quantitative synthesis of average accuracy per major metaheuristic algorithm, identifying top performers such as Hybrid Multi-Strategy Aquila Optimization (99.98%), Henry Gas Solubility Optimization (99.93%), and Grey Wolf Optimizer (99.80%). Fourth, it maps the distribution of application domains, revealing that medical image classification is the most dominant domain (34% of studies), followed by general/mixed applications (24%) and hyperspectral imaging (12%). Fifth, it identifies research gaps including under-explored metaheuristic algorithms and limited studies on computational cost analysis. Sixth, it provides evidence-based recommendations for future research directions, including hybrid approaches combining metaheuristics with Bayesian optimization and reinforcement learning.

In addition to being efficient, metaheuristic algorithms are also considered to be able to handle the problems of global exploration and local exploitation dynamically. These algorithms show the ability to adapt to the high complexity of the hyperparameter space of deep learning models. Metaheuristics is also model-agnostic, meaning that metaheuristics can be applied to various types of deep learning architectures without requiring significant modifications to the structure of the model created.

Although there have been many studies that have explored the use of metaheuristics for hyperparameter tuning in deep learning-based image classification, there are still few systematic studies that comprehensively map trends, models, and metaheuristic optimization. Therefore, this study conducted a Systematic Literature Review (SLR) to investigate the trend of using metaheuristic algorithms in deep learning hyperparameter optimization in image classification tasks. This SLR approach refers to the PRISMA protocol to ensure a transparent, systematic, and bias-free article selection process. This study aims to identify research gaps, classify the models and algorithms used, and provide an overview of the latest research developments to answer the main questions:

"How to improve the performance of deep learning-based image classification by optimizing hyperparameters using metaheuristic algorithms?"

## **METHOD**

This chapter explained the approach that used in the process of compiling a systematic literature review regarding the application of metaheuristic algorithms for hyperparameter tuning in deep learning models for image classification. The research method is designed to ensure that the process of identification, selection, evaluation, and synthesis of literature is carried out in a structured, transparent, and bias-free manner. The Systematic Literature Review (SLR) method will be adopted in this study, which refers to the guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA). Through this methodology, all stages of research start from the beginning, namely the formulation of research questions, article search strategies, selection criteria, quality assessment of articles, to the process of data extraction and synthesis.

### **Research Design**

This study used the Systematic Literature Review (SLR) method to identify, evaluate, and synthesize the scientific literature related to the use of metaheuristic algorithms in

optimizing hyperparameters for image classification using a deep learning model. This approach was chosen because it is able to present evidence systematically and thoroughly, and avoid bias in the literature selection. This study uses the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guide in conducting the review process.

### **Research Stages**

The research is carried out through several main stage processes.

1. Formulation of Research Questions, Determining the focus of the research to be carried out is the study of the use of metaheuristic algorithms in hyperparameters.
2. Article Search Strategy, conducting an article search using keywords that are relevant to the research being conducted.
3. Study Selection Using PRISMA, Conducting a screening process through the stages of identification, screening, eligibility, and inclusion.
4. Data Extraction, Collecting important data from selected articles according to the variables of the research.
5. Data Synthesis, conducting descriptive and comparative analysis to identify patterns, trends, and gaps from each study.

This stage is carried out to ensure that the SLR process is carried out methodologically and in accordance with the standards of scientific research.

### **Article Search Strategies**

The articles were collected from several databases, the databases used were Web of Science, Semantic Scholar, Scopus, and Science of Direct. The keywords used to query each database along with the number of each article obtained in each database can be seen in Table 1. Because Science Direct has a limited number of boolean connectors, which is eight, so it must be broken down in doing the query. From the four dataset sources, 1902 articles were obtained.

In the process, the articles used are articles published in 2015 and above. This study uses Zotero's reference manager in storing article metadata, referencing, tagging based on inclusion and exclusion criteria, and others. The total number of articles obtained from the results of the article search process from all databases is 1,902 articles.

### **Article Selection Process (Flow Prisma)**

The selection stage is carried out based on the PRISMA flow. The results of the screening using PRISMA, can be seen in Figure 1.

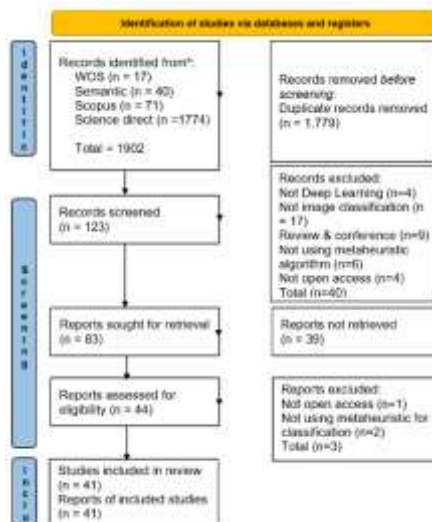


Figure 1. PRISMA Results

At the screening stage, out of a total of 1,902 articles, 1,779 were excluded due to duplication, resulting in 123 articles. A total of 123 articles entered the screening stage. In addition, data screening was also carried out, screening was carried out by including and excluding them. The inclusion and exclusion criteria used in the study are those contained in subchapter 3.4. The results of the excluded data are based on not using deep learning (4 articles), not image classification (17 articles), review & conference (9 articles), not using a metaheuristic algorithm (6 articles), not open access (5 articles), so that the total number of articles excluded in the screening process is 41 articles. A total of 82 articles (123-41) were tried to get the article file. A total of 39 articles were not successfully obtained (because they were not open access or not found), so the number of articles remaining was 41 articles. 1 article obtained was not a file from a journal article, 2 articles did not use metaheuristic for image classification. In the end, as many as 41 articles were used in the systematic review process.

Table 1. Database Query Article

Database	Query	Number of Article
Science Direct	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning” OR “Hyper-parameter Tuning”) AND (“Optimization” OR “Fine-tuning” OR “Improvement” OR “Algorithm”) (“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning”) AND (“Optimization”) (“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning”) AND (“Fine-tuning”) (“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning”) AND (“Improvement”) (“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning”) AND (“Algorithm”) (“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyper-parameter Tuning”) AND (“Optimization”)	1174

	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyper-parameter Tuning”) AND (“Fine-tuning”)	
	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyper-parameter Tuning”) AND (“Improvement”)	
	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyper-parameter Tuning”) AND (“Algorithm”)	
<b>Scopus</b>	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning” OR “Hyper-parameter Tuning”) AND (“Optimization” OR “Fine-tuning” OR “Improvement” OR “Algorithm”)	71
<b>Semantic Scholar</b>	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning” OR “Hyper-parameter Tuning”) AND (“Optimization” OR “Fine-tuning” OR “Improvement” OR “Algorithm”)	40
<b>Web of Science</b>	(“Metaheuristic” OR “Meta-heuristic”) AND (“Image Classification” OR “Object Classification” OR “Image Recognition”) AND (“Hyperparameter” OR “Hyper-parameter” OR “Hyperparameter Tuning” OR “Hyper-parameter Tuning”) AND (“Optimization” OR “Fine-tuning” OR “Improvement” OR “Algorithm”)	17

### Data Extraction

From the 41 selected articles, each article will be extracted based on the following variables.

1. Titles and authors
2. Year of publication
3. Application
4. Classification model used (classification model only, not feature extractor)
5. Metaheuristic algorithms used
6. Optimized hyperparameters and their ranges
7. Datasets used
8. Evaluation results (Accuracy, Precision, Recall, F1-score)
9. Average result of the accuracy of each metaheuristic algorithm

These data are used as the basis for analysis and synthesis of findings in the results and discussion sections.

### Data Synthesis

Data from 41 articles were analyzed descriptively and comparatively to identify trends in deep learning models, metaheuristic algorithm distribution, dataset mapping, performance improvement, and. Synthesis was carried out using narrative approaches and graphical visualization to facilitate interpretation.

## RESULT AND DISCUSSION

After a literature selection process with the PRISMA method and reference management using Zotero, a total of 41 articles were obtained that were relevant to the topic of image classification using deep learning which was optimized using metaheuristic algorithms. The findings will be discussed in this section.

## 1. Distribution of the Year of Publication

Based on the results of the study, research on the application of metaheuristic algorithms in the hyperparameter tuning process for deep learning-based image classification shows a significant increasing trend in the last five years. The results of article distribution per year can be seen in Figure 2. The distribution of publications shows that before 2018, the number of studies examining the integration of metaheuristics and deep learning in image classification was relatively limited. However, from 2020 to 2024 there has been a consistent increase in the number of publications, reflecting the increasing interest and relevance of this topic in the field of artificial intelligence.

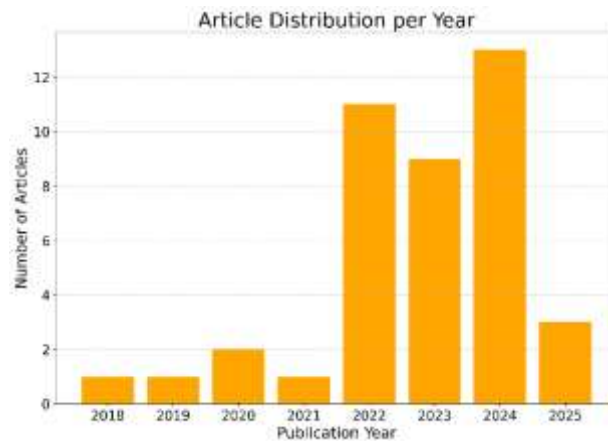


Figure 2. Article Distribution per Year

The increase in the number of publications is influenced by several main factors. First, rapid developments in deep learning algorithms, especially in convolutional neural network (CNN) architectures, have opened up new opportunities in image processing in various domains, including medical, industrial, agricultural, and facial recognition (Alzubaidi et al., 2021). Second, the limitations of traditional hyperparameter tuning methods such as grid search and random search in handling complex parameter spaces have prompted researchers to explore more efficient metaheuristic optimization approaches (Purnomo, Gonsalves, et al., 2024). In addition, the increasing availability of public datasets and advances in computing hardware (GPU and TPU) have also accelerated the process of experimenting and testing metaheuristic algorithms in various studies.

Overall, the increasing trend in the number of publications in the last five years indicates that the use of metaheuristic algorithms for hyperparameter tuning in image classification using deep learning is a growing and relevant research area, both academically and in its application.

## 2. Application Domain Distribution

Application domains are categorized based on the purpose of the article and the dataset used in the research of the article. For the results of the distribution of application domains, see Figure 3. The results of the systematic literature review analysis show that the application of metaheuristic algorithm-based hyperparameter tuning in image classification using deep learning has been widely applied to various application domains. The distribution of application domains analyzed in this study reflects the increasingly widespread diversification

of the use of deep learning, in line with the flexibility and ability of deep learning models in image processing from various fields.

In general, the medical image classification domain is the most dominant application area, including CT scan image classification (Alamgeer, 2022; Banu, 2025; Purnomo, Tad Gonsalves, et al., 2024; Rengasamy & Nadar, 2024; Zhu et al., 2023), MRI (Alnabhan et al., 2022; Baghdadi, 2022; Bose & Garg, 2024; R.Aishwarya et al., 2023; Rath, 2021; Xu & Mohammadi, 2024; Zhu et al., 2023), X-ray (Banu, 2025; Duhayyim et al., 2022), to eye imaging (Abu et al., 2023; Thanarajan et al., 2023). This is particularly relevant given the need for high accuracy in medical image-based diagnoses, where misclassification can have a significant impact on clinical decision-making. The use of metaheuristics in hyperparameter tuning for medical image classification allows for the optimization of the configuration of CNN models as well as other hybrid models in order to capture critical features more effectively (Al Duhayyim et al., 2022).

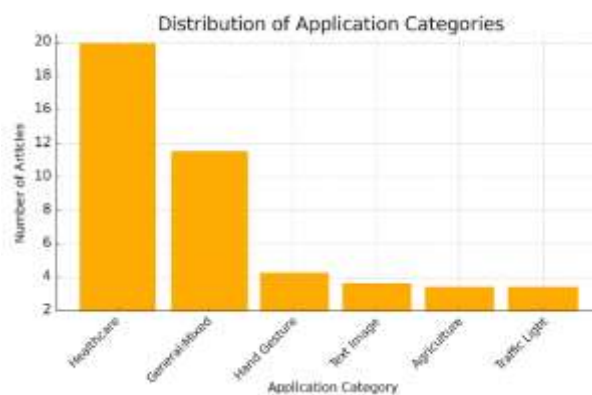


Figure 3. Application Category Distribution

In addition to the health domain, several studies have also utilized a similar approach to letter or digit recognition applications (Bacanin, 2020; Munsarif, 2024). Agricultural applications, namely land use (Aljebreen, 2024) and crop classification (Ahmed, 2023). Hand gesture applications, namely hand gesture recognition (M. H. Zafar et al., 2023) and sign language recognition (Alrowais, 2022). Sports applications are sports image classification (Zhou, 2024). Traffic light applications, namely stop sign classification (Loussaief & Abdelkrim, 2018). Hyperspectral applications, namely hyperspectral image classification (Alahmari, 2023; Munishamaiaha et al., 2023; Wu et al., 2024). In addition, there are also studies that are categorized as general/mixed applications where this is a category that contains articles that use general datasets and the research uses several datasets with different domains (Abdulsaed et al., 2024; Ahmad, 2020; Assiri, 2025; Fischetti & Stringher, 2019; Franco-Gaona et al., 2025; Purnomo, Gonsalves, et al., 2024; Ribeiro, 2022; Wei et al., 2024; Xin et al., 2024; A. Zafar et al., 2022).

This distribution shows that the development of metaheuristic algorithms for hyperparameter tuning not only contributes to improved classification accuracy, but also enables the adoption of deep learning models on a wide range of complex, heterogeneous, and high-precision application domains.

### 3. Distribution Metaheuristic Algorithm

The calculation of the metaheuristic algorithm distribution was carried out by looking at the main algorithm used in the article, the comparison algorithm was not included in this distribution. For the results of the metaheuristic algorithm distribution, it can be seen in Figure 4. The extraction results show that the diversity of metaheuristic algorithms is very wide used in the hyperparameter tuning process of image classification using deep learning. Some studies use more than one algorithm at a time, both for performance comparison needs and for the development of hybrid optimization.

Some of the algorithms that are the main algorithms in the main development of the study are, Henry Gas Solubility Algorithm (HGS) as the main algorithm in hyperparameter tuning in human gesture recognition applications based on sEMG signals based on Stacked CNN (Alrowais, 2022; M. H. Zafar et al., 2023). Genetic Algorithm (GA) remains one of the most widely used as the main algorithm (Loussaief & Abdelkrim, 2018; Purnomo, Gonsalves, et al., 2024). Harmony Search (HS) as the main algorithm in the optimization of the Xception architecture for medical classification (Banu, 2025) and in general data classification (Purnomo, Gonsalves, et al., 2024). Exploration and Exploitation Enhanced Firefly Algorithm (E3-FA) is applied for the classification of brain cancer (Bacanin, 2020). The Adaptive Habitat Biogeography-Based Optimizer (AHBBO) and the Adaptive Fick's Law Algorithm (AFLA) have also emerged as relatively newly developed main algorithms (Wu et al., 2024; Xin et al., 2024).

This distribution shows the diversity of innovations in the development of metaheuristic algorithms. Metaheuristic development innovations for deep learning hyperparameter tuning do not only focus on classical algorithms, but are growing with the advent of hybrid, adaptive, and bio-inspired algorithms. Most recent studies tend to develop hybrid algorithms that combine various optimization mechanisms to balance global exploration capabilities and local exploitation, given that the complexity of hyperparameter configuration in deep learning models is increasing as model architectures develop.

### 4. Deep Learning Model Distribution

The results of the systematic literature review analysis show that the Convolutional Neural Network (CNN) architecture remains the most dominant deep learning model used in image classification studies with the application of hyperparameter tuning based on metaheuristic algorithms. Of the total 41 articles reviewed, 14 articles use CNN. The results of the distribution of the 10 most widely used deep learning models can be seen in Figure 5.

CNN is the top choice due to its ability to effectively extract spatial features from image data, both for high-resolution and low-resolution data (Alzubaidi et al., 2021; Shah, 2024). For example, in the CNN study it is used in the form of Stacked CNN for the classification of sEMG signals in human-robot interactions (M. H. Zafar et al., 2023). Similarly, the CNN model is also applied in various benchmark tests on datasets such as CIFAR-10, Fashion-MNIST, MNIST, SVHN, GTSRB, and Tiny ImageNet in the development of the Adaptive Habitat Biogeography-Based Optimizer (AHBBO) (Xin et al., 2024).

Another model that is also found quite frequently is Xception, which integrates Harmony Search (HS) in hyperparameter tuning for medical image classification (Banu, 2025). In addition, there is also the use of Extreme Learning Machine (ELM) in several cancer

classification studies (Rengasamy & Nadar, 2024) as well as the application of hybrid models such as SqueezeNet + ANFIS (Adaptive Neuro-Fuzzy Inference System) for EEG motor imagery processing (Yang, 2024).

The use of these various deep learning models shows that the hyperparameter tuning process using metaheuristic algorithms is not limited to one type of architecture, but can be applied flexibly to various types of models, both conventional CNNs, lightweight models, and hybrid models, depending on the needs of the application and the complexity of the data used.



Figure 4. Metaheuristic Algorithm Distribution

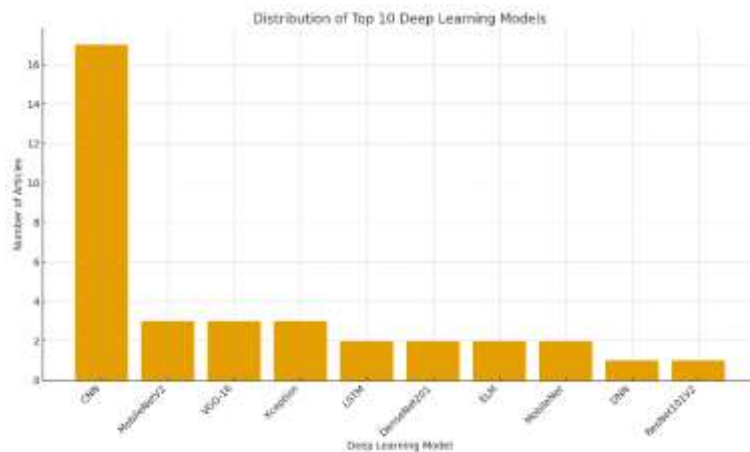


Figure 5. Distribution of the Top 10 DL Models

## 5. Dataset Distribution

The calculation of the distribution of the dataset was carried out by counting each dataset used from each article. The results of the distribution of the 10 most frequently used datasets can be seen in Figure 6. In this systematic literature review study, the analysis of the distribution of the dataset showed that the use of datasets in metaheuristic-based hyperparameter tuning testing varied greatly, depending on the domain of the intended application.

Common datasets such as MNIST, CIFAR-10, MNIST-Fashion, and CIFAR-100 are still the most widely used datasets. MNIST has 8 articles, CIFAR-10 has 6 articles, and MNIST-Fashion has 4 articles. CIFAR-100, Indian

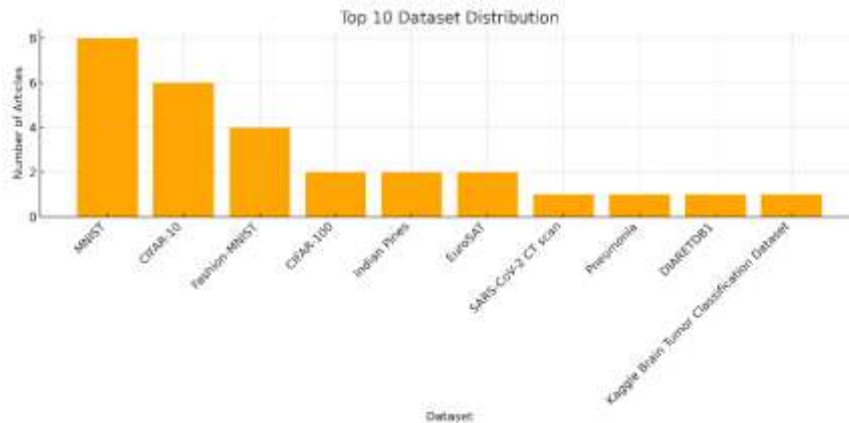


Figure 6. Distribution of the Top 10 Datasets

Pines, EuroSAT as many as 2 articles and the rest of the datasets are 1 article all. In the health domain, various medical datasets are widely used, such as SARS-CoV-2 CT scan and Brain Tumor MRI in the MEEDNets study optimized with Evolutionary Algorithm (Zhu et al., 2023).

BreakHis dataset for histopathological image-based breast cancer classification optimized with Manta Ray Foraging Optimization (MRFO) (Woźniak et al., 2022). LIDC-IDRI and Kaggle Lung Cancer dataset for Ant Lion Optimization (ALO) optimized lung cancer detection (Rengasamy & Nadar, 2024). Glioma dataset and Brain MRI dataset on Spotted Hyena Optimizer (SHO) optimized brain tumor classification (Rath, 2021). ISIC 2017 and ISIC 2020 for skin cancer classification based on Hybrid Salp Swarm Algorithm (HSSA) (Obayya et al., 2023).

In hand gesture applications, unique datasets such as sEMG sensor signal dataset are used (M. H. Zafar et al., 2023). American Sign Language (ASL) Alphabet Dataset from Kaggle (Alrowais, 2022).

For the sports domain, the sports imagery dataset collected from YouTube videos was used for Deep Neural Network (DNN) testing optimized with Novel Tuna Swarm Optimization (NTSO) (Zhou, 2024).

In the hyperspectral domain, the study used Indian Pines and Pavia University hyperspectral datasets for the classification of satellite imagery using the Adaptive Fick's Law Algorithm (AFLA) (Wu et al., 2024). The Indian Pines and Salinas dataset used CNN with Enhanced Elephant Herding Optimization (EEHO) (Munishamaiaha et al., 2023).

In the agriculture domain, datasets such as EuroSAT (Aljebreen, 2024) and UAV-based Food Crop Dataset (Ahmed, 2023) are used. In the text image classification domain, the MNIST dataset (Bacanin, 2020; Munsarif, 2024) is used. In the traffic light domain, the Caltech-256 dataset (Loussaief & Abdelkrim, 2018) is used.

As for the general/mixed domain, the MNIST, CIFAR-10, CIFAR-100, Fashion-MNIST, SVHN, EMNIST, GTSRB, Tiny ImageNet, Rice Image, Marmot Tables, Marmot Equations, Coronary Stenosis (Antczak), Stenosis608, Pneumonia, and EuroSAT datasets were used (Abdulsaed et al., 2024; Ahmad, 2020; Assiri, 2025; Fischetti & Stringher, 2019; Franco-Gaona et al., 2025; Purnomo, Gonsalves, et al., 2024; Ribeiro, 2022; Wei et al., 2024; Xin et al., 2024; A. Zafar et al., 2022).

## 6. Average Accuracy Per Major Metaheuristic

Based on the results of systematic literature reviews of studies that utilize metaheuristic algorithms in deep learning hyperparameter tuning for image classification, the results of accuracy vary widely, but in general show high performance compared to manual tuning or conventional optimization methods. The use of metaheuristics is able to find the best hyperparameters, so that deep learning models are able to achieve optimal configurations with very competitive accuracy levels on various types of datasets. From all the articles analyzed, the results of the classification accuracy show a fairly high performance in almost all metaheuristic combinations with deep learning models. A summary of the average accuracy per main algorithm can be seen in Table 3.

**Table 3. Average Accuracy of Each Metaheuristic Algorithm**

<i>Study ID</i>	<i>Algorithm Name</i>			<i>Accuracy (%)</i>	<i>Study ID</i>	<i>Algorithm Name</i>			<i>Accuracy (%)</i>
S1	Evolutionary Algorithm			99.1	S21	Modified Particle Swarm Optimization (LDRW-PSO)			<b>99.25</b>
S2	Novel	Tuna	Swarm	97.777	S22	Enhanced Elephant Herding Optimization (EEHO)			<b>99.73</b>
S3	Henry	Gas	Solubility	99.3	S23	Genetic Algorithm (GA)			<b>98.94</b>
S4	Non-dominated Genetic Algorithm-II (NSGA-II)		Sorted	91.5	S24	Modified Mating Probability-based Water Strider Algorithm (MMP-WSA)			<b>95.51</b>
S5	Jellyfish Optimization (JFO)			94.2	S25	Boltzmann Univariate Marginal Distribution Algorithm (BUMDA)			<b>98.086</b>
S6	Contracted Fox Optimization Algorithm (CFOA)			97.32	S26	Simulated Annealing (SA)			<b>85.44</b>
S7	Adaptive Biogeography-Based Optimizer (AHBBO)	Habitat		94.86	S27	Whale Optimization Algorithm (WOA)			<b>99.2</b>
S8	Adaptive Algorithm (AFLA)	Fick's	Law	98.95	S28	Manta-Ray Foraging Optimization (MRFO)			<b>99</b>
S9	Chaotic Algorithm (CSSA)	Sparrow	Search	98.61	S29	Manta Ray Foraging Optimization (MRFO)			<b>98.37</b>
S10	QIESOA Optimization)	(Egret	Swarm	92.85	S30	Gorilla Troops Optimizer (GTO)			<b>96.45</b>
S11	Harmony Search (HS)			97.18	S31	Exploitation Enhanced Tree Growth Algorithm (EE-TGA)			<b>99.33</b>
S12	Chaotic Optimization (CBO)	Butterfly		99.29	S31	Exploration and Exploitation Enhanced Firefly Algorithm (E3-FA)			<b>99.29</b>
S13	Grey Wolf Optimizer (GWO)			99.80	S32	Golden Jackal Optimization (GJO)			<b>97.39</b>
S13	Hunger Games Search (HGS)			99.60	S33	Henry Gas Solubility Optimization (HGSO)			<b>99.93</b>
S14	Single-Iteration Optimization (SIO)			77.88	S34	Egyptian Vulture Optimization (EVO)			<b>96.12</b>
S15	Ant Lion Optimization (ALO)			99.4	S35	River Formation Dynamics Algorithm (RFDA)			<b>98.12</b>
S16	Spotted	Hyena	Optimizer	98.14	S36	Moth Swarm Optimization (MSO)			<b>99.33</b>
	(SHO)								

<b>S17</b>	Modified Controlling Parameters-based African Vultures Optimization Algorithm (MCP-AVOA)	99.36	S37	Hybrid Multi-Strategy Aquila Optimization (HMAO)	<b>99.98</b>
<b>S18</b>	Ant Colony Optimization (ACO)	99.7	S38	Satin Bowerbird Optimization (SBO)	<b>97.43</b>
<b>S18</b>	Genetic Algorithm (GA)	97.7	S39	Binary Crow Search Algorithm (BCSA)	<b>84.83</b>
<b>S18</b>	Harmony Search (HS)	89.9	S40	Sailfish Optimization (SFO)	<b>98.47</b>
<b>S19</b>	Ebola Optimization Search Algorithm (EOSA)	91.5	S41	Modified Salp Swarm Algorithm (MSSA)	<b>96.84</b>
<b>S20</b>	Hybrid Salp Swarm Algorithm (HSSA)	94.95			

Most studies report classification accuracy above 90%, with some even reaching accuracy levels above 99%. Some algorithms show very prominent accuracy contributions, such as:

- a. Exploitation Enhanced Tree Growth Algorithm (EE-TGA) and Exploration and Exploitation Enhanced Firefly Algorithm (E3-FA) which are able to produce high accuracy in image classification based on EE-TGA and E3FA-CNN, which are 99.33% and 99.29%, respectively (Bacanin, 2020).
- b. Henry Gas Solubility Algorithm (HGS) and Harmony Search (HS) applied in several studies showed consistently high accuracy stability in medical classification tasks, signal imagery, and general datasets, namely 99.3%, 99.93%, 97.18%, and 89.9% (Alrowais, 2022; Banu, 2025; Purnomo, Gonsalves, et al., 2024; M. H. Zafar et al., 2023).
- c. Genetic Algorithm (GA) also contributes to competitive accuracy performance in hyperparameter tuning benchmarking studies (Loussaief & Abdelkrim, 2018; Purnomo, Gonsalves, et al., 2024).
- d. Hybrid Multi-Strategy Aquila Optimization (HMAO) provides the highest accuracy of 99.98%, followed by Henry Gas Solubility Optimization (HGSO) which also provides very high accuracy of 99.93% (Alahmari, 2023; Alrowais, 2022).

Overall, the overall average accuracy of the entire study reached 96.55%. The high average accuracy showed that the application of metaheuristics in hyperparameter tuning contributed significantly to improving the performance of deep learning-based image classification.

## CONCLUSION

The results of this systematic literature review show that the application of metaheuristic algorithms in hyperparameter tuning deep learning models for image classification has developed very rapidly in recent years. Various algorithms have been used to overcome the problem of exploration and exploitation of hyperparameter spaces that are very complex in deep learning models. Metaheuristic algorithms have continued to experience an increase in implementation in publications in the last five years and the largest domain of application is in the health domain. The Henry Gas Solubility Algorithm (HGS), Genetic Algorithm (GA), Non-dominated Sorted Genetic Algorithm-II (NSGA-II), Harmony Search (HS), as well as new algorithms such as the Exploration and Exploitation Enhanced Firefly Algorithm (E3-FA), Adaptive Habitat Biogeography-Based Optimizer (AHBBO), and Adaptive Fick's Law

Algorithm (AFLA) show an important role as the main metaheuristic algorithm in the development of hyperparameter tuning. The data distribution shows that some algorithms appear more than once in various studies, either as single algorithms or as part of multi-benchmark studies, namely the Henry Gas Solubility Algorithm (HGS), Genetic Algorithm (GA), and Harmony Search (HS). A careful validation process confirms that most modern hybrid algorithms are increasingly adopted in image classification applications, especially in the medical domain, hyperspectral imaging, gesture recognition, and image classification in general. Thus, the application of metaheuristics in hyperparameter tuning of deep learning models has been proven to be able to significantly improve the accuracy performance of image classification, and is projected to continue to develop as new innovations emerge in the field of metaheuristic algorithm optimization. Further research can combine metaheuristics with other methods or metaheuristics, to create new algorithms, in addition to being able to conduct more in-depth research, such as analyzing research time consumption and resource needs. Further research can develop a hyperparameter tuning approach with a more adaptive approach by combining metaheuristics and other methods such as Bayesian Optimization or Reinforcement Learning as well as with other metaheuristic methods to create a more efficient hybrid model. In addition, research can be conducted using metaheuristic algorithms that are still very rare or even never used in research, especially in the case of image classification, to see if the approach can provide significant improvements or not. Future research is also recommended to be able to add more in-depth analysis, especially in terms of the time consumption of research and resource requirements from the use of such approaches.

## REFERENCE

- Abdulsaed, E. H., Alabbas, M., & Khudeyer, R. S. (2024). Optimizing the Architecture of Convolutional Neural Networks Using Modified Salp Swarm Algorithm. *Dalam Journal of Al-Qadisiyah for Computer Science and Mathematics*. <https://doi.org/10.29304/jqcs.2024.16.11450>
- Abu, M., Zahri, N., Amir, A., Ismail, M. I., Yaakub, A., Fukumoto, F., & Suzuki, Y. (2023). Analysis of the Effectiveness of Metaheuristic Methods on Bayesian Optimization in the Classification of Visual Field Defects. *Dalam Diagnostics (Vol. 13)*. <https://doi.org/10.3390/diagnostics13111946>
- Ahmad, M. (2020). Image classification based on automatic neural architecture search using binary crow search algorithm. *Dalam IEEE Access (Vol. 8, hlm. 189891–189912)*. <https://doi.org/10.1109/ACCESS.2020.3031599>
- Ahmed, M. A. (2023). Satin Bowerbird Optimization with Convolutional LSTM for Food Crop Classification on UAV Imagery. *Dalam IEEE Access (Vol. 11, hlm. 41075–41083)*. <https://doi.org/10.1109/ACCESS.2023.3269806>
- Al Duhayyim, M., A. Malibari, A., Obayya, M., K. Nour, M., S. Salama, A., I. Eldesouki, M., Sarwar Zamani, A., & Rizwanullah, M. (2022). Metaheuristic with Deep Learning Enabled Biomedical Bone Age Assessment and Classification Model. *Computers, Materials & Continua, 73(3)*, 5473–5489. <https://doi.org/10.32604/cmc.2022.031976>
- Alahmari, S. (2023). Hybrid Multi-Strategy Aquila Optimization with Deep Learning Driven Crop Type Classification on Hyperspectral Images. *Dalam Computer Systems Science*

- and Engineering (Vol. 47, Nomor 1, hlm. 376–391).  
<https://doi.org/10.32604/csse.2023.036362>
- Alamgeer, M. (2022). Deep Learning Enabled Computer Aided Diagnosis Model for Lung Cancer using Biomedical CT Images. Dalam *Computers, Materials and Continua* (Vol. 73, Nomor 1, hlm. 1437–1448). <https://doi.org/10.32604/cmc.2022.027896>
- Aljebreen, M. (2024). Land Use and Land Cover Classification Using River Formation Dynamics Algorithm With Deep Learning on Remote Sensing Images. Dalam *IEEE Access* (Vol. 12, hlm. 11147–11156). <https://doi.org/10.1109/ACCESS.2023.3349285>
- Alnabhan, M., Habboush, A. K., Al-Haija, Q. A., Mohanty, A. K., Pattnaik, S., & Pattanayak, B. K. (2022). Hyper-Tuned CNN Using EVO Technique for Efficient Biomedical Image Classification. Dalam *Mobile Information Systems* (Vol. 2022). <https://doi.org/10.1155/2022/2123662>
- Alrowais, F. (2022). Sign Language Recognition and Classification Model to Enhance Quality of Disabled People. Dalam *Computers, Materials and Continua* (Vol. 73, Nomor 2, hlm. 3419–3432). <https://doi.org/10.32604/cmc.2022.029438>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Assiri, M. (2025). Synergy of Internet of Things and Software Engineering Approach for Enhanced Copy–Move Image Forgery Detection Model. Dalam *Electronics*. <https://doi.org/10.3390/electronics14040692>
- Bacanin, N. (2020). Optimizing convolutional neural network hyperparameters by enhanced swarm intelligence metaheuristics. Dalam *Algorithms* (Vol. 13, Nomor 3). <https://doi.org/10.3390/a13030067>
- Baghdadi, N. A. (2022). A3 C-TL-GTO: Alzheimer Automatic Accurate Classification Using Transfer Learning and Artificial Gorilla Troops Optimizer. Dalam *Sensors* (Vol. 22, Nomor 11). <https://doi.org/10.3390/s22114250>
- Banu, S. (2025). Harmony Search-Based Xception Model For Enhanced Lung Cancer Prediction. Dalam *Mitteilungen Klosterneuburg*. <https://doi.org/10.61586/qbeg9>
- Benaissa, B., Kobayashi, M., Al Ali, M., Khatir, T., & El Amine Elaissaoui Elmeliani, M. (2024). Metaheuristic optimization algorithms: An overview. *HCMCOU Journal of Science – Advances in Computational Structures*, 33–61. <https://doi.org/10.46223/HCMCOUJS.acs.en.14.1.47.2024>
- Bose, A., & Garg, R. (2024). Optimized CNN Using Manta-Ray Foraging Optimization for Brain Tumour Detection. *Procedia Computer Science*, 235, 2187–2195. <https://doi.org/10.1016/j.procs.2024.04.207>
- Duhayyim, M. A., Malibari, A. A., Obayya, M., Nour, M. K., Salama, A. S., Eldesouki, M. I., Zamani, A. S., & Rizwanullah, M. (2022). Metaheuristic with Deep Learning Enabled Biomedical Bone Age Assessment and Classification Model. *Computers, Materials and Continua*, 73(3), 5473–5489. <https://doi.org/10.32604/cmc.2022.031976>
- Fischetti, M., & Stringher, M. (2019). Embedded hyper-parameter tuning by Simulated Annealing. Dalam *ArXiv*. <https://doi.org/10.48550/arXiv.1906.01504>

- Franco-Gaona, E., Avila-Garcia, M., & Cruz-Aceves, I. (2025). Automatic Neural Architecture Search Based on an Estimation of Distribution Algorithm for Binary Classification of Image Databases. *Dalam Mathematics*. <https://doi.org/10.3390/math13040605>
- Han, M., Du, Z., Yuen, K. F., Zhu, H., Li, Y., & Yuan, Q. (2024). Walrus optimizer: A novel nature-inspired metaheuristic algorithm. *Expert Systems with Applications*, 239, 122413. <https://doi.org/10.1016/j.eswa.2023.122413>
- Loussaief, S., & Abdelkrim, A. (2018). Convolutional Neural Network Hyper-Parameters Optimization based on Genetic Algorithms. *Dalam International Journal of Advanced Computer Science and Applications* (Vol. 9). <https://doi.org/10.14569/IJACSA.2018.091031>
- Munishamaiaha, K., Kannan, S. K., Venkatesan, D., Jasinski, M., Novak, F., Goño, R., & Leonowicz, Z. (2023). Hyperspectral Image Classification with Deep CNN Using an Enhanced Elephant Herding Optimization for Updating Hyper-Parameters. *Dalam Electronics*. <https://doi.org/10.3390/electronics12051157>
- Munsarif, M. (2024). Convolution neural network hyperparameter optimization using modified particle swarm optimization. *Dalam Bulletin of Electrical Engineering and Informatics* (Vol. 13, Nomor 2, hlm. 1268–1275). <https://doi.org/10.11591/eei.v13i2.6112>
- Obayya, M., Saeed, M. K., Alruwais, N., Alotaibi, S. S., Assiri, M., & Salama, A. S. (2023). Hybrid Metaheuristics With Deep Learning-Based Fusion Model for Biomedical Image Analysis. *Dalam IEEE Access* (Vol. 11, hlm. 117149–117158). <https://doi.org/10.1109/ACCESS.2023.3326369>
- Purnomo, H., Gonsalves, T., Mailoa, E., Santoso, F. J., & Pribadi, M. R. (2024). Metaheuristics Approach for Hyperparameter Tuning of Convolutional Neural Network. *Dalam Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*. <https://doi.org/10.29207/resti.v8i3.5730>
- Purnomo, H., Tad Gonsalves, Evangs Mailoa, Santoso, F. J., & Pribadi, M. R. (2024). Metaheuristics Approach for Hyperparameter Tuning of Convolutional Neural Network. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 8(3), 340–345. <https://doi.org/10.29207/resti.v8i3.5730>
- Raiaan, M. A. K., Sakib, S., Fahad, N. M., Mamun, A. A., Rahman, Md. A., Shatabda, S., & Mukta, Md. S. H. (2024). A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks. *Decision Analytics Journal*, 11, 100470. <https://doi.org/10.1016/j.dajour.2024.100470>
- R.Aishwarya, Ganesan, D. S., & Babu, D. T. R. (2023). MT-MRANBrainTumorNet: A Novel Brain Tumour and Severity Level Classification Framework using Meta-heuristic-assisted Transformer based Multiscale Residual Attention Network. *Dalam Tuijin Jishu/Journal of Propulsion Technology*. <https://doi.org/10.52783/tjjpt.v44.i3.1775>
- Rath, P. (2021). An empirical development of hyper-tuned CNN using spotted hyena optimizer for bio-medical image classification. *Dalam Journal of Natural Science, Biology and Medicine* (Vol. 12, Nomor 3, hlm. 300–306). [https://doi.org/10.4103/jnsbm.JNSBM\\_12\\_3\\_5](https://doi.org/10.4103/jnsbm.JNSBM_12_3_5)
- Rengasamy, V., & Nadar, M. T. (2024). Optimizing Lung Cancer Classification with Extreme Learning Machine and Ant Lion Optimization for Enhanced Early Detection. *Dalam*

- Traitement Du Signal (Vol. 41, Nomor 4, hlm. 2185–2193).  
<https://doi.org/10.18280/ts.410447>
- Ribeiro, L. C. F. (2022). Convolutional neural networks ensembles through single-iteration optimization. Dalam *Soft Computing* (Vol. 26, Nomor 8, hlm. 3871–3882).  
<https://doi.org/10.1007/s00500-022-06791-9>
- Saber, A., Elbedwehy, S., Awad, W. A., & Hassan, E. (2024). An optimized ensemble model based on meta-heuristic algorithms for effective detection and classification of breast tumors. Dalam *Neural Computing and Applications*. <https://doi.org/10.1007/s00521-024-10719-9>
- Shah, N. N. (2024). Recent Advances In Image Classification.
- Thanarajan, T., Alotaibi, Y., Rajendran, S., & Nagappan, K. (2023). Eye-Tracking Based Autism Spectrum Disorder Diagnosis Using Chaotic Butterfly Optimization with Deep Learning Model. *Computers, Materials and Continua*, 76(2), 1995–2013.  
<https://doi.org/10.32604/cmc.2023.039644>
- Wei, P., Shang, M., Zhou, J., & Shi, X. (2024). Efficient adaptive learning rate for convolutional neural network based on quadratic interpolation egret swarm optimization algorithm. *Heliyon*, 10(18), e37814. <https://doi.org/10.1016/j.heliyon.2024.e37814>
- Woźniak, M., Ijaz, M. F., Shankar, K., Dutta, A., Kumar, S., Joshi, G. P., & Doo, I. (2022). Chaotic Sparrow Search Algorithm with Deep Transfer Learning Enabled Breast Cancer Classification on Histopathological Images. Dalam *Cancers* (Vol. 14).  
<https://doi.org/10.3390/cancers14112770>
- Wu, T.-Y., Li, H., Kumari, S., & Chen, C.-M. (2024). A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification. *Computers, Materials and Continua*, 79(1), 19–46.  
<https://doi.org/10.32604/cmc.2024.048347>
- Xin, J., Khishe, M., Zeebaree, D. Q., Abualigah, L., & Ghazal, T. M. (2024). Adaptive habitat biogeography-based optimizer for optimizing deep CNN hyperparameters in image classification. *Heliyon*, 10(7), e28147. <https://doi.org/10.1016/j.heliyon.2024.e28147>
- Xu, L., & Mohammadi, M. (2024). Brain tumor diagnosis from MRI based on Mobilenetv2 optimized by contracted fox optimization algorithm. *Heliyon*, 10(1), e23866.  
<https://doi.org/10.1016/j.heliyon.2023.e23866>
- Yang, E. (2024). Optimal Fuzzy Logic Enabled EEG Motor Imagery Classification for Brain Computer Interface. Dalam *IEEE Access* (Vol. 12, hlm. 46002–46011).  
<https://doi.org/10.1109/ACCESS.2023.3346674>
- Zafar, A., Aamir, M., Nawari, N. M., Arshad, A., Riaz, S., Alruban, A., Dutta, A. K., Almutairi, B., & Almotairi, S. (2022). An Optimization Approach for Convolutional Neural Network Using Non-Dominated Sorted Genetic Algorithm-II. *Computers, Materials and Continua*, 74(3), 5641–5661. <https://doi.org/10.32604/cmc.2023.033733>
- Zafar, M. H., Langås, E. F., & Sanfilippo, F. (2023). Empowering human-robot interaction using sEMG sensor: Hybrid deep learning model for accurate hand gesture recognition. *Results in Engineering*, 20, 101639. <https://doi.org/10.1016/j.rineng.2023.101639>
- Zhou, Z. (2024). Improved sports image classification using deep neural network and novel tuna swarm optimization. Dalam *Scientific Reports* (Vol. 14, Nomor 1).  
<https://doi.org/10.1038/s41598-024-64826-7>

Zhu, H., Wang, W., Ulidowski, I., Zhou, Q., Wang, S., Chen, H., & Zhang, Y. (2023). MEEDNets: Medical Image Classification via Ensemble Bio-inspired Evolutionary DenseNets. *Knowledge-Based Systems*, 280, 111035. <https://doi.org/10.1016/j.knosys.2023.111035>