

Integrative Hybrid Metaheuristic Algorithm for Hyperparameter Optimisation in Pre-Trained Convolutional Neural Network Models (I-HAHO)

Nazleeni Samiha Haron, Jafreezal Jaafar, Izzatdin Abdul Aziz, Mohd Hilmi Hasan, Muhammad Hamza Azam
Department of Computing, Universiti Teknologi Petronas, Perak, Malaysia

Abstract—Hyperparameter optimisation (HPO) remains a fundamental challenge in deep learning, especially for pre-trained convolutional neural networks (CNNs). While pre-trained models reduce the computational burden of training from scratch, their effectiveness depends heavily on tuning parameters such as learning rate, batch size, dropout, weight decay, and optimizer type. The search space of hyperparameters is large, nonlinear, and highly dataset-dependent, making traditional techniques like grid search, random search, and Bayesian optimisation insufficient. This paper introduces I-HAHO, an Integrative Hybrid Metaheuristic Algorithm that combines Artificial Bee Colony (ABC) for global exploration and Harris Hawks Optimisation (HHO) for local exploitation. A diversity-based phase-switching mechanism dynamically regulates exploration and exploitation, allowing the optimiser to adapt its search behaviour to varying landscape conditions. Experiments on CIFAR-10, CIFAR-100, SVHN, and TinyImageNet with three CNN architectures (VGG16, ResNet50, EfficientNet-B0) demonstrate up to 6.9% accuracy improvements. I-HAHO enhances adaptability, scalability, and robustness for hyperparameter tuning.

Keywords—Hyperparameter Optimisation (HPO); Convolutional Neural Networks (CNNs); Artificial Bee Colony (ABC); Harris Hawks Optimisation (HHO); Hybrid Metaheuristic Algorithm

I. INTRODUCTION

Convolutional Neural Networks (CNNs) have become indispensable in modern computer vision, enabling breakthroughs in tasks such as classification, object detection, and medical imaging. Pre-trained models like VGG16, ResNet50, and EfficientNet-B0, trained on large datasets such as ImageNet, provide transferable feature representations that accelerate training and improve performance, especially when target datasets are small [1], [2].

Despite their success, the performance of pre-trained CNNs is highly sensitive to hyperparameter settings such as learning rate, batch size, dropout rate, and optimizer type. Selecting these parameters is challenging because the search space is high-dimensional, non-convex, and dataset-dependent, making manual tuning and traditional search strategies inefficient. Techniques such as grid search and random search require extensive computational effort, while Bayesian optimisation often struggles with scalability and mixed-variable types [3], [4].

To address these challenges, metaheuristic algorithms have been widely adopted for hyperparameter optimisation [5], [6]. Their ability to balance exploration (global search for promising

regions) and exploitation (local refinement near good solutions) makes them well-suited for navigating complex search spaces [7], [8]. Among these, the Artificial Bee Colony (ABC) and Harris Hawks Optimisation (HHO) algorithms have shown promise in deep learning optimisation tasks [9], [10]. However, both algorithms have limitations: ABC often lacks strong exploitation, while HHO may converge prematurely in multimodal landscapes [11], [12].

This motivates the development of hybrid approaches that combine complementary strengths. In this study, we propose I-HAHO (Integrative Hybrid ABC-HHO), which unifies ABC's exploration capability with HHO's exploitative refinement. A dynamic phase-switching mechanism guided by population diversity regulates the balance between exploration and exploitation, allowing the optimiser to adapt its search behaviour to varying landscape conditions, preventing premature convergence and preserving diversity [13], [14].

The main contributions of this paper can be summarized as follows. First, we propose a novel hybrid metaheuristic algorithm, I-HAHO, which integrates the global exploration strength of ABC with the local exploitation capability of HHO to achieve robust hyperparameter optimisation for pre-trained CNNs. Second, we introduce a diversity-driven dynamic phase-switching strategy that adaptively balances exploration and exploitation during the optimisation process, preventing premature convergence and preserving solution diversity. Third, we provide an extensive evaluation of I-HAHO across multiple benchmark datasets, including CIFAR-10, CIFAR-100, SVHN, and Tiny ImageNet, applied to three widely used CNN architectures: VGG16, ResNet50, and EfficientNet-B0. Finally, we demonstrate that I-HAHO achieves substantial accuracy improvements of up to 6.9% compared with grid search, random search, Bayesian optimisation, and standalone metaheuristics, highlighting its effectiveness, scalability, and generalizability for deep learning optimisation tasks.

The remainder of this manuscript is organized as follows. Section II presents the theoretical background of pre-trained CNN architectures and the hyperparameter optimisation problem, followed by a review of recent metaheuristic and hybrid approaches in the Section III. Section IV introduces the proposed I-HAHO algorithm, detailing its design, modifications to ABC and HHO, and the diversity-driven phase-switching mechanism. Section V describes the experimental setup, including datasets, CNN models, hyperparameter search spaces, and preprocessing strategies. It also presents and analyses the

results, comparing I-HAHO with conventional and metaheuristic baselines in terms of accuracy and convergence. It discusses the implications, strengths, limitations, and potential extensions of the proposed approach. Finally, Section VI concludes the paper by highlighting key findings and suggesting directions for future research.

II. BACKGROUND OF STUDY

This section provides a discussion of relevant theoretical foundations and an introduction to the hyperparameter optimization problem.

A. Pre-Trained Models

Deep learning has revolutionized computer vision, with Convolutional Neural Networks (CNNs) emerging as the foundation of state-of-the-art models across a wide range of applications, including medical imaging, object recognition, and autonomous driving [15]. CNNs learn hierarchical feature representations directly from raw data, eliminating the need for handcrafted feature extraction and enabling scalable solutions for high-dimensional inputs such as images.

However, training CNNs from scratch is computationally expensive and requires large labelled datasets. To address these limitations, pre-trained CNNs such as VGG16, ResNet50, and EfficientNet-B0 are commonly adopted in transfer learning. These models, initially trained on large datasets like ImageNet, serve as general-purpose feature extractors and can be fine-tuned for domain-specific tasks.

1) *VGG16*: Introduced by [16], is characterized by its simplicity, uniform 3×3 convolution filters, and deep sequential layers. While powerful in feature extraction, its large parameter size ($\sim 138\text{M}$) makes it computationally demanding. Fig. 1 depicts the VGG16 architecture.

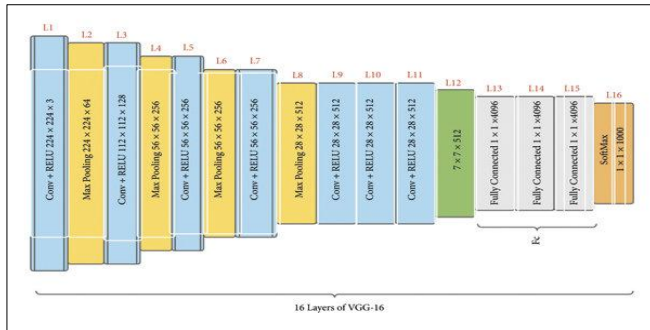


Fig. 1. VGG16 architecture [16].

2) *ResNet50*: Developed by [17], incorporates residual connections that enable the effective training of deeper architectures by mitigating vanishing gradient problems. This design has significantly improved CNN scalability and accuracy in challenging tasks [18]. Fig. 2 depicts the ResNet50 architecture.

3) *EfficientNet-B0*: Proposed by [20], introduces a compound scaling strategy that jointly scales depth, width, and resolution. It achieves high accuracy with a lightweight parameter count ($\sim 5.3\text{M}$), making it suitable for mobile and

embedded applications [21] (Şafak & Barışçıl, 2024). Fig. 3 depicts the EfficientNet-B0 architecture.

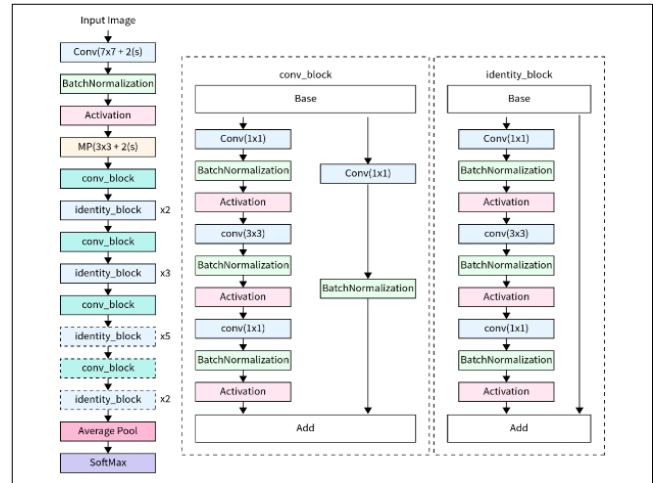


Fig. 2. ResNet50 architecture [19].

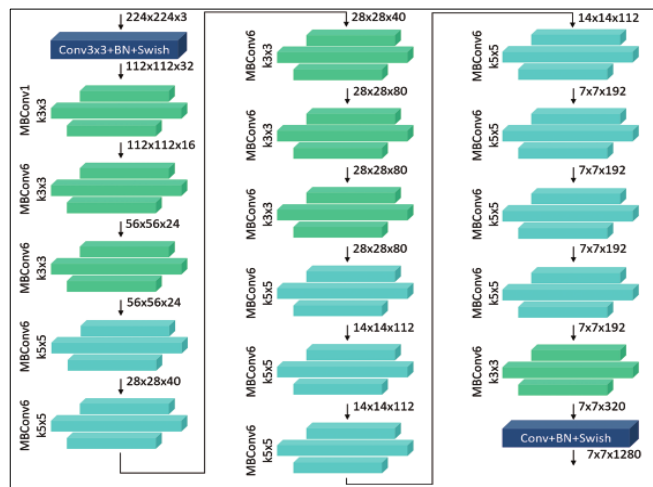


Fig. 3. EfficientNet-B0 architecture [20].

4) *Comparison among pre-trained CNN models or architectures*: To validate the proposed HPO strategies and algorithms in real-world scenarios, three pre-trained CNN architectures (VGG16, ResNet50, and EfficientNet-B0) are selected for experimental evaluation. These models represent a range of design philosophies and computational complexities, allowing for comprehensive testing against the three core performance objectives: accuracy, adaptability, and computational efficiency. Table I summarises the characteristics of the chosen pre-trained CNN models.

B. Hyperparameter Optimisation (HPO)

The primary goal of HPO is to automate the hyperparameter tuning process to ensure optimal performance of ML models on specific tasks. The motivation for employing HPO lies in its ability to address challenges posed by manual tuning, such as non-linear hyperparameter interactions and the computational burden of evaluating large search spaces.

TABLE I. CHARACTERISTICS OF CHOSEN PRE-TRAINED CNN MODELS

Criteria	VGG16	ResNet50	EfficientNet-B0
Architecture Depth	16 layers (13 convolutional + 3 fully connected)	50 layers with residual connections	82 layers (scalable depth, width, and resolution)
Key Innovation	Sequential architecture with uniform 3×3 convolutions	Residual connections to solve vanishing gradient problem	Compound scaling method to balance network dimensions
Model Size	~138 million parameters	~25.5 million parameters	~5.3 million parameters
Computational Complexity	High due to depth and fully connected layers (~15.4 GFLOPs)	Moderate (~3.9 GFLOPs)	Low (~0.39 GFLOPs)
Real-World Applications	Medical imaging, video classification, and object detection	Autonomous vehicles, semantic segmentation, and face recognition	Mobile vision, IoT devices, and scalable vision applications

In the context of convolutional neural networks (CNNs), hyperparameters play a pivotal role in shaping the architecture, learning dynamics, and generalization ability. These settings, determined prior to training, influence feature extraction, overfitting prevention, and convergence behaviour. CNN hyperparameters can generally be categorized into three main groups [22], [23].

- Hyperparameters of convolutional layers (i.e., number of convolutional layers, number of kernels, size of kernels, stride, padding, dilation rate, activation functions, and filter initialization methods).
- Hyperparameters of fully connected layers (i.e., dropout rate, connectivity pattern, number of neurons, activation functions, weight regularization, and initialization methods).
- General Hyperparameters (i.e., batch size, learning rate, learning rate decay, optimizer type, number of epochs, momentum, and gradient clipping).

1) *HPO problem formulation*: HPO is the process of identifying the optimal hyperparameters for a machine learning model to maximize its performance on a given task. This is formulated as an optimisation problem where the objective is to find the hyperparameter configuration that minimizes a loss function or maximizes an evaluation [24], [25].

a) *General formulation*

Let,

$\lambda = \lambda_1, \lambda_2, \dots, \lambda_n$: The hyperparameter vector, where each λ_i represents a specific hyperparameter.

$M(\lambda)$: The machine learning model trained with hyperparameter configuration λ .

$L(M(\lambda), D)$: The loss function of the model on dataset D . $D = \{D_{train}, D_{val}\}$ where D_{train} and D_{val} are the training and validation datasets, respectively.

The goal of HPO is to find the optimal hyperparameter configuration λ^* such that:

$$\lambda^* = \underset{\lambda \in \Lambda}{\operatorname{argmin}} L(M(\lambda), D_{val})$$

where Λ is the search space of hyperparameters.

2) *Search space*: In the context of HPO for pre-trained CNNs, the search space defines the range and type of hyperparameters that are optimized to improve model performance. This space encompasses various types of parameters that differ in structure and optimization complexity. The hyperparameter search space Λ is typically composed of categorical, discrete and continuous parameters.

Each category introduces different challenges for optimisation algorithms. Categorical parameters often require non-gradient-based search techniques, discrete parameters increase combinatorial complexity, and continuous parameters demand fine-grained exploration.

C. *Metaheuristic Algorithms*

Metaheuristics are population-based or single-solution optimisation methods that incorporate stochastic search strategies to efficiently explore large, complex landscapes. They are particularly suited to HPO, as they do not rely on gradient information and can handle mixed-variable and non-differentiable objective functions [26], [27].

Metaheuristics achieve success by balancing:

- Exploration: discovering new regions of the search space.
- Exploitation: refining promising candidate solutions [7], [8].
- Several algorithms have been applied to CNN HPO:
- Genetic Algorithms (GA): Inspired by natural evolution, used for optimising filter sizes and learning rates [28].
- Particle Swarm Optimisation (PSO): Mimics social behaviour of birds, effective in refining CNN configurations [29].
- Whale Optimisation Algorithm (WOA): Employs simulated hunting behaviour for global exploration [30].
- Harmony Search and Variants: Used for fine-tuning CNNs through adaptive exploration-exploitation trade-offs [31].

Among these, two algorithms stand out for their complementary strengths:

- Artificial Bee Colony (ABC): A swarm-based algorithm inspired by bee foraging behaviour [9]. ABC excels at exploration by maintaining population diversity through scouts and onlookers but suffers from weak local refinement in later iterations [11].
- Harris Hawks Optimisation (HHO): A predator-prey inspired method with dynamic exploration and exploitation phases [10]. HHO is effective in local

search but is prone to premature convergence in highly multimodal spaces [12].

Individual metaheuristic algorithms often underperform in the complex hyperparameter landscape of pre-trained CNNs, leading to stagnation, poor convergence, or lack of diversity [13]. To address these challenges, hybrid metaheuristics have emerged, aiming to combine complementary strengths while mitigating weaknesses [14]. Recent studies, such as PSO-GWO [32] and HHOS [33] highlight the effectiveness of hybridisation for deep learning optimisation.

Nevertheless, the integration of ABC and HHO for CNN hyperparameter tuning remains underexplored. This motivates the present study, which introduces I-HAHO, an integrative hybrid approach that leverages ABC's exploratory power and HHO's exploitative refinement under a diversity-based switching strategy.

III. RELATED WORKS

This section presents reviews on recent studies that apply hybrid metaheuristic algorithms to HPO in pre-trained CNN models. Recent studies have explored hybrid metaheuristic algorithms for HPO in pre-trained CNN models, yet consistent challenges remain across adaptability, accuracy, and computational efficiency dimensions.

Ulutaş et al. [34] proposed a hybrid PSO-GWO approach for tuning hyperparameters within ensemble models for diabetes detection, achieving a peak accuracy of 98.10% using Random Forest classifiers. Their findings indicate that the hybrid method outperforms standalone Particle Swarm Optimization (PSO) and Grey Wolf Optimization (GWO), offering improved robustness and generalisation across classifiers.

Yan et al. [35] addressed key challenges in CNN hyperparameter tuning, including mixed-variable encoding, high computational cost, and convergence stability. They introduced GPPSO, a hybrid algorithm combining Gaussian

Processes with PSO. The method incorporates a novel encoding scheme for mixed-variable hyperparameters and a surrogate model to reduce evaluation overhead. Experiments on CIFAR-10 and CIFAR-100 confirmed GPPSO's superior efficiency and accuracy, with practical relevance for applications such as metal fracture diagnosis.

Kumar and Kondaveeti [33] applied transfer learning with a Hybrid Hyperparameter Optimization Scheme (HHOS), integrating manual domain knowledge and random search strategies. Using pre-trained ImageNet models with additional trainable layers, they demonstrated that fine-tuning EfficientNetB0 achieved 99.12% accuracy, outperforming ResNet18 (93.24%). These results highlight the value of hybrid HPO in ecological classification tasks.

Akkuş et al. [36] explored HPO for CNN-based image classification, specifically diabetic retinopathy detection. Their study compared the Adolescent Identity Search Algorithm (AISA) with Bayesian Optimization (BO) across pre-trained models including AlexNet, MobileNetV2, ResNet18, and GoogLeNet. AISA-based tuning yielded higher accuracy with fewer iterations, demonstrating its efficiency in medical imaging contexts.

Kiliçarslan [32] investigated hybrid optimisation for cardiovascular disease prediction using 1D and 2D VGG-16 architectures. The study compared PSO, Cat Swarm Optimization (CSO), and hybrid PSO-GWO, showing that the hybrid consistently achieved the highest classification accuracy across medical and MNIST datasets.

Atteia et al. [37] further emphasised the importance of HPO in CNN-based diagnostic systems, applying optimisation techniques to diabetic maculopathy detection using optical coherence tomography and fundus retinography. Their work reinforces the critical role of HPO in enhancing model performance for clinical applications.

The summary of the related works is presented in Table II.

TABLE II. SUMMARY OF RELATED WORKS

Authors	Hybrid Method	Domain	Models / Datasets	Key Findings
Ulutaş et al. [34]	PSO + GWO	Diabetes detection (Ensemble Learning)	Random Forest, other classifiers	Achieved 98.10% accuracy; hybrid PSO-GWO outperformed standalone PSO and GWO, improving robustness and model performance.
Yan et al. [35]	GP + PSO (GPPSO)	CNN HPO, Metal fracture diagnosis	CNNs on CIFAR-10, CIFAR-100	Proposed mixed-variable encoding + surrogate model; improved efficiency, convergence, and accuracy compared to traditional methods.
Kumar & Kondaveeti [33]	Hybrid Hyperparameter Optimisation Scheme (HHOS) (manual + random search)	Ecological application (transfer learning)	EfficientNetB0, ResNet18 (pre-trained on ImageNet)	Fine-tuned EfficientNetB0 achieved 99.12%, ResNet18 achieved 93.24%; hybrid scheme improved accuracy and adaptability.
Akkuş et al. [36]	AISA + Bayesian Optimisation	Medical imaging (Diabetic Retinopathy detection)	AlexNet, MobileNetV2, ResNet18, GoogLeNet; Kaggle DR dataset	Hybrid AISA-based optimisation achieved higher accuracy with fewer iterations than BO; more efficient for CNN HPO in medical imaging.

IV. MATERIALS AND METHODS

In this section, I-HAHO is introduced as our approach to improve the accuracy of the HPO for pre-trained CNN models.

A. Dataset

This study utilizes four benchmark datasets: CIFAR-10, CIFAR-100, SVHN, and TinyImageNet to evaluate the performance of the proposed hybrid metaheuristic HPO algorithms. These datasets were selected due to their widespread

adoption in deep learning literature and their varying levels of complexity, class granularity, and domain diversity, which support robust validation of adaptability and generalisation.

CIFAR-10 and CIFAR-100 were obtained from the official dataset repository maintained by the Canadian Institute for Advanced Research (CIFAR). Each dataset comprises 60,000 32×32 colour images. CIFAR-10 consists of 10 mutually exclusive classes, whereas CIFAR-100 is structured with 100 fine-grained categories across 20 superclasses. Both datasets were split into 50,000 training and 10,000 testing images. During preprocessing, images were normalized using dataset-specific mean and standard deviation values, converted to float tensors, and standardized to improve gradient stability during model training. No class rebalancing was performed, preserving original distribution characteristics for realism.

Street View House Numbers (SVHN), sourced from the official Stanford University repository, was employed in its cropped digits format, ensuring consistent spatial resolution with other datasets. Comprising 73,257 training and 26,032 test images, SVHN represents a real-world digit classification task under non-ideal lighting and alignment conditions. Raw .mat files were parsed and converted to tensor format. Images were resized to 32×32 pixels and normalized to zero mean and unit variance. Dataset labels were integer-encoded and subsequently transformed to one-hot format for compatibility with CNN classifiers. It has a real-world and noisy data characteristic which makes it a challenging test case.

TinyImageNet, a scaled-down variant of the ILSVRC ImageNet dataset, was downloaded from the official ImageNet challenge portal. It includes 200 object categories, with 100,000 training samples (500 images per class), 10,000 validation samples (50 per class), and 10,000 test samples. Unlike CIFAR and SVHN, the input resolution for TinyImageNet was preserved at 64×64 pixels to maintain semantic richness. Due to its relatively small sample size per class and high inter-class similarity, extensive preprocessing was applied, including data augmentation techniques such as random horizontal flipping, random cropping with padding, and brightness/contrast jittering to prevent overfitting and improve generalization under low-data regimes. It represents real-world scenarios and tests the scalability and efficiency of optimisation algorithms in complex tasks.

A summary of the datasets is provided in Table III, highlighting their complexity based on factors such as image resolution, number of classes, and dataset size.

TABLE III. SUMMARY OF DATASETS

Dataset	Number of Classes	Total Images	Image Resolution	Complexity
CIFAR-10	10	60,000	32×32	Moderate
CIFAR-100	100	60,000	32×32	High
TinyImageNet	1,000	1.2 million	Varies	Very High
SVHN	10	600,000+	32×32	High (Real-world)

B. Proposed I-HAHO Algorithm

The pseudocode for the I-HAHO is presented in the Algorithm 1.

Algorithm 1: I-HAHO

Input:
Population Size N , Maximum iterations T , Initial diversity threshold D_0 ,
Hyperparameters for ABC and HHO, Objective function $F(x)$
Output:
Best solution x_{best} with optimized hyperparameters.

Initialise the population X using Latin Hypercube Sampling
Set iteration counter $t = 0$
Calculate the initial diversity threshold $D_{thresh}(t) = D_0$
Evaluate the fitness $F(x)$ for each solution $x_i \in X$
While $t < T$:
 Calculate population diversity $D(t)$
 Update the diversity threshold $D_{thresh}(t)$
 Dynamic Phase-Switching Strategy:
 if $D(t) > D_{thresh}$ (High Diversity):
 Perform Modified ABC Exploration
 else $D(t) < D_{thresh}$
 Perform Modified HHO Exploitation
 Evaluate the $F(x)$ of the updated solutions
 Update the global best solution x_{best} based on the fitness
 Calculate the change in fitness $\Delta F = |F_{best}(t) - F_{best}(t-1)|$.
 Convergence Check:
 If $\Delta F < \text{threshold}$ or if $D(t) > D_{min}$
 Increment $t = t + 1$
End While
Return the best solution x_{best}

Fig. 4 illustrates the detailed operational flow of the proposed Integrative Hybrid Metaheuristic Algorithm (I-HAHO), which unifies exploration and exploitation through a diversity-based phase-switching mechanism. At each iteration, the population diversity $D(t)$ is computed and compared with a dynamic threshold $D_{thresh}(t)$. When $D(t) > D_{thresh}$, indicating sufficient population diversity, the algorithm prioritises Modified ABC exploration to expand the search space. This phase adaptively updates candidate positions, reinitialises stagnated solutions using Distance-Aware Novelty Reinitialisation (DANR), and invokes Chaos-Guided Candidate Recovery (CGCR) when diversity falls below acceptable limits, ensuring sustained global exploration. Conversely, when $D(t) < D_{thresh}$, the process transitions to Modified HHO exploitation, which refines promising regions through Refined Surround Prey, Escaping Energy, and Success-History-Based Adaptation (SHEA) to intensify local search and accelerate convergence.

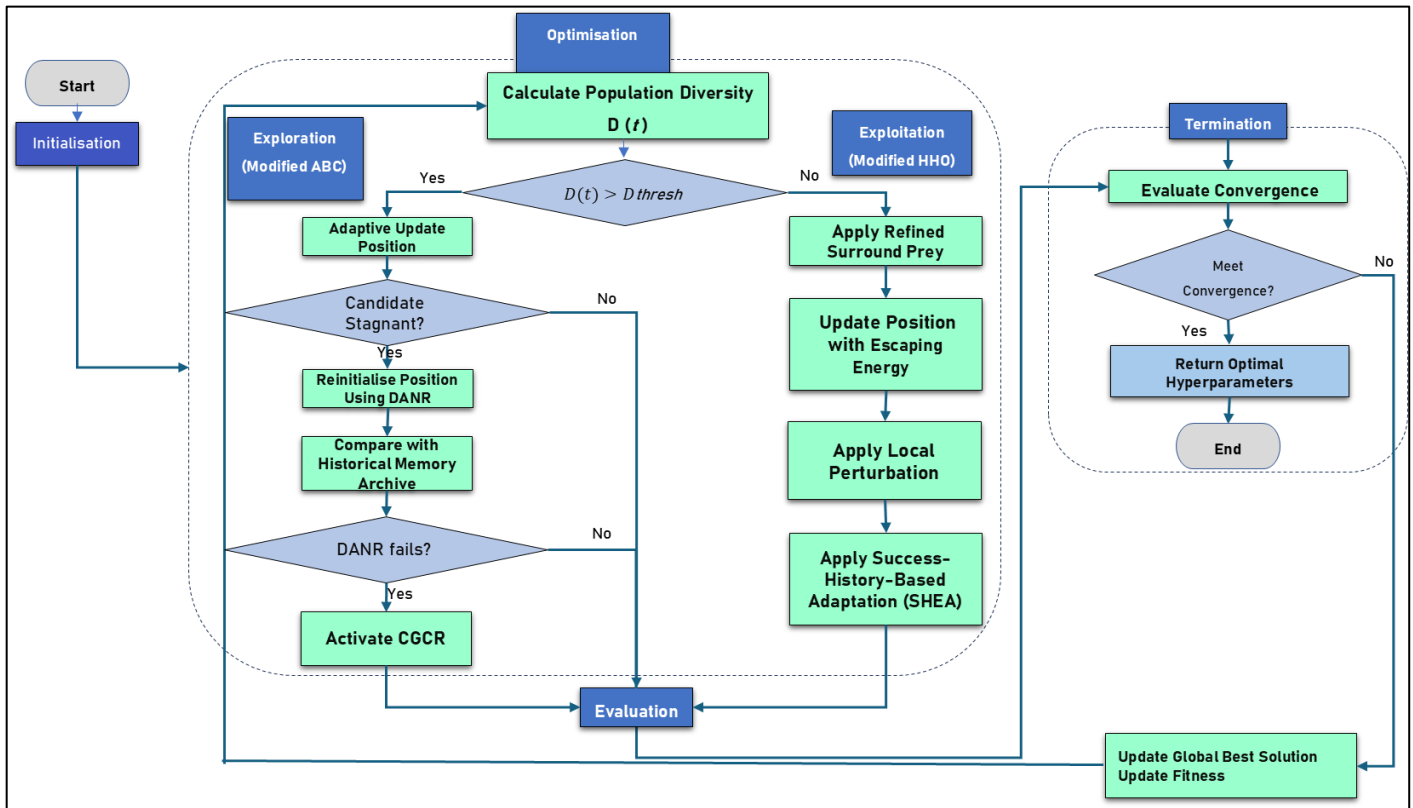


Fig. 4. Detailed algorithmic flowchart for I-HAHO with key strategies.

Unlike previous hybrid metaheuristics that rely on static schedules, random alternation, or fixed probability-based phase switching, the proposed diversity-driven framework continuously adapts based on the evolutionary state of the population. This allows I-HAHO to intelligently balance exploration and exploitation according to real-time feedback from the search dynamics, rather than predetermined iteration counts. The integration of DANR, CGCR, and SHEA further enhances the system's self-adaptive behaviour. DANR promotes novelty and prevents stagnation by introducing diversity-sensitive perturbations, CGCR leverages chaotic patterns to recover lost diversity while preserving promising regions, and SHEA enables adaptive parameter fine-tuning guided by historical success rates. Collectively, these mechanisms fill the gap left by earlier hybrid strategies that lacked dynamic responsiveness, allowing I-HAHO to navigate complex, multimodal hyperparameter landscapes more effectively and maintain balanced convergence performance across datasets and CNN architectures.

C. Performance Evaluation Measures

The evaluation measures employed in this research were selected to comprehensively assess the performance of the proposed I-HAHO algorithm.

1) *Classification accuracy*: Classification accuracy is the primary metric to evaluate the performance of HPO in pre-

trained CNN models. It measures the proportion of correctly classified instances relative to the total instances.

$$\text{Accuracy}(\theta) = \frac{\text{Correctly Classified Instances}(\theta)}{\text{Total Instances}} \quad (1)$$

A higher classification accuracy indicates better optimisation of hyperparameters and superior adaptability performance of the model.

V. RESULTS AND DISCUSSION

A. Overall Comparative Performance

The experiment results for I-HAHO comparison with baseline HPO algorithms and canonical metaheuristic algorithms are presented in Table IV. The results demonstrate the highest accuracy achieved by different HPO algorithms across four pre-trained CNN models on various benchmark datasets. The accuracy results are presented in Table IV.

The proposed I-HAHO algorithm consistently outperformed all benchmarked HPO algorithms across the evaluated pre-trained CNN models and datasets. On the CIFAR-10 dataset, I-HAHO achieved superior validation accuracies of 93.194% (VGG16), 94.503% (ResNet50), and 93.008% (EfficientNet-B0), outperforming traditional and baseline search-based algorithms (GS, RS and BO). This consistent improvement illustrates I-HAHO's ability to effectively navigate the hyperparameter space through adaptive hybridization of global and local search strategies.

TABLE IV. ACCURACY RESULTS FOR ALL HPO ALGORITHMS AND I-HAHO

Model	Dataset	GS (%)	RS (%)	BO (%)	Vanilla ABC (%)	Vanilla HHO (%)	I-HAHO (%)
VGG16	CIFAR-10	88.2	88.6	89.2	87.5	87.7	93.1
	SVHN	95.3	95.6	96.2	94.7	95.0	96.9
	CIFAR-100	67.1	68.04	68.5	66.0	66.5	78.5
	Tiny ImageNet	57.8	58.3	58.9	57.0	57.4	65.4
ResNet50	CIFAR-10	90.5	90.9	91.3	89.0	89.3	94.5
	SVHN	93.0	93.7	94.1	92.5	92.9	95.8
	CIFAR-100	68.9	69.7	70.4	67.6	68.3	76.0
	Tiny ImageNet	61.2	61.9	62.5	60.2	60.9	69.5
EfficientNet-B0	CIFAR-10	89.0	89.7	90.1	87.5	87.9	93.0
	SVHN	87.3	87.9	88.4	86.0	86.5	91.0
	CIFAR-100	68.0	68.9	69.5	66.6	67.3	75.5
	Tiny ImageNet	61.8	62.4	62.9	59.6	60.3	68.0

On the more complex CIFAR-100 dataset, which includes 100 classes with fewer samples per class, I-HAHO achieved 78.507% (VGG16), 76.011% (ResNet50), and 75.512% (EfficientNet-B0), confirming its adaptability in high-dimensional, class-imbalanced learning tasks. For the SVHN dataset, characterized by large-scale digit recognition with significant intra-class variance and noise, I-HAHO achieved 96.903% (VGG16), 95.803% (ResNet50), and 91.007% (EfficientNet-B0). These results demonstrate I-HAHO's ability to maintain high accuracy and generalization performance even under challenging, real-world-like scenarios with substantial class overlap and varying input conditions.

On TinyImageNet, a particularly challenging dataset with 200 categories and lower resolution images, I-HAHO again demonstrated superior adaptability, achieving 65.497% (VGG16), 69.502% (ResNet50), and 68.013% (EfficientNet-B0). Notably, the EfficientNet-B0 model exhibited strong performance due to its compound scaling, which supports effective feature extraction even in low-resolution, high-category contexts. These outcomes suggest that I-HAHO not only improves training outcomes for well-known benchmarks but also scales effectively to more complex domains with less structured data distributions.

Collectively, these results are aligned with recent studies highlighting the limitations of single-strategy optimizers in deep learning [3], [37], [38] and emphasize the benefit of hybrid metaheuristics for pre-trained CNN models HPO tasks.

VI. CONCLUSION

In conclusion, the proposed Integrative Hybrid Artificial Bee Colony–Harris Hawks Optimization (I-HAHO) algorithm improves hyperparameter optimization (HPO) for pre-trained convolutional neural networks (CNNs) through an adaptive hybrid structure. The framework integrates a modified Artificial Bee Colony (ABC) algorithm for global exploration with a modified Harris Hawks Optimization (HHO) algorithm for local exploitation. This integration provides a balanced search process that minimizes premature convergence and sustains population diversity, which conventional approaches such as grid search, random search, and Bayesian optimization fail to achieve. The

adaptive phase-switching mechanism regulates exploration and exploitation dynamically according to population diversity and fitness variance, allowing the optimiser to adjust its search intensity based on the topological characteristics of the loss landscape. This adaptivity enables I-HAHO to sustain population diversity during early iterations and progressively intensify local refinement as convergence nears, resulting in consistent and efficient optimisation in high-dimensional, non-convex search spaces. Experimental results on CIFAR10, CIFAR100, SVHN, and TinyImageNet datasets across VGG16, ResNet50, and EfficientNetB0 architectures demonstrate up to 6.9% improvement in classification accuracy compared to baseline optimizers. These results validate IHAHO as a scalable and generalizable hyperparameter optimization framework that improves search efficiency, convergence stability, and model adaptability across diverse deep learning tasks. The diversity-based phase-switching mechanism enables dynamic control over exploration and exploitation, supporting robust performance in high-dimensional, nonconvex spaces.

REFERENCES

- [1] K. Wagh, K. S. Suresh, M. Sayed, and S. Shahane, "Exploring transfer learning in image analysis using feature extraction with pre-trained models," *ICTACT Journal on Image & Video Processing*, vol. 14, no. 3, 2024.
- [2] R. N. Wanjiku, L. Nderu, and M. Kimwele, "Improved transfer learning using textural features conflation and dynamically fine-tuned layers," *PeerJ Computer Science*, vol. 9, p. e1601, 2023.
- [3] S. Iqbal, A. Qureshi, A. Ullah, J. Li, and T. Mahmood, "Improving the robustness and quality of biomedical CNN models through adaptive hyperparameter tuning," *Applied Sciences*, vol. 12, no. 22, p. 11870, 2022.
- [4] F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., *Automated Machine Learning: Methods, Systems, Challenges*. Cham: Springer, 2019.
- [5] V. Silva, "Optimizing transfer learning and fine-tuning hyperparameters in image classification problems with firefly algorithm," in *Proc. Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2024)*, pp. 484–495, 2024. doi: 10.5753/eniac.2024.245032.
- [6] S. B. Şüyun, M. Yurdakul, Ş. Taşdemir, and S. Biliş, "Triple-stream deep feature selection with metaheuristic optimization and machine learning for multi-stage hypertensive retinopathy diagnosis," *Applied Sciences*, vol. 15, no. 12, p. 6485, 2025.
- [7] A. Gaspar, D. Oliva, E. Cuevas, D. Zaldivar, M. Pérez, and G. Pajares, "Hyperparameter optimization in a convolutional neural network using metaheuristic algorithms," in *Metaheuristics in Machine Learning*:

- Theory and Applications*. Cham: Springer International Publishing, pp. 37–59, 2021.
- [8] R. Narayanan and N. Ganesh, “A comprehensive review of metaheuristics for hyperparameter optimization in machine learning,” in *Metaheuristics for Machine Learning: Algorithms and Applications*, pp. 37–72, 2024.
- [9] D. Karaboga and B. Basturk, “Artificial bee colony algorithm,” *Journal of Global Optimization*, vol. 39, no. 3, pp. 459–471, 2007.
- [10] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: Algorithm and applications,” *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019.
- [11] B. Akay and D. Karaboga, “A modified artificial bee colony algorithm for real-parameter optimization,” *Information Sciences*, vol. 192, pp. 120–142, 2012.
- [12] T. Yang, J. Fang, C. Jia, Z. Liu, and Y. Liu, “An improved Harris hawks optimization algorithm based on chaotic sequence and opposite elite learning mechanism,” *PLoS One*, vol. 18, no. 2, p. e0281636, 2023.
- [13] S. Sharma, P. Singh, and P. Kaur, “Metaheuristic optimization algorithms for hyperparameter tuning in deep learning: A comprehensive review and future directions,” *Artificial Intelligence Review*, vol. 55, pp. 4895–4953, 2022. doi: 10.1007/s10462-021-10032-7.
- [14] A. Seyyedabbasi, “A hybrid multi-strategy optimization metaheuristic algorithm for multi-level thresholding color image segmentation,” *Applied Sciences*, vol. 15, no. 13, p. 7255, 2025.
- [15] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [16] K. Simonyan and A. Zisserman, “Very deep convnets for large-scale image recognition,” *Computing Research Repository*, 2014.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning,” *Image Recognition*, vol. 7, no. 4, pp. 327–336, 2015.
- [18] L. Mahin Sheikh, A. Shaikh, A. Sandupatla, R. Pudale, A. Bakare, and M. Chavan, “Classification of simple CNN model and ResNet50,” *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, 2024. doi: 10.22214/ijraset.2024.60677.
- [19] Q. Ji, J. Huang, W. He, and Y. Sun, “Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images,” *Algorithms*, vol. 12, no. 3, p. 51, 2019.
- [20] M. Tan and Q. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proc. International Conference on Machine Learning (ICML)*, pp. 6105–6114, May 2019.
- [21] E. Şafak and N. Barışçı, “Detection of fake face images using lightweight convolutional neural networks with stacking ensemble learning method,” *PeerJ Computer Science*, vol. 10, p. e2103, 2024.
- [22] M. A. K. Raiaan, S. Sakib, N. M. Fahad, A. Al Mamun, M. A. Rahman, S. Shatabda, and M. S. H. Mukta, “A systematic review of hyperparameter optimization techniques in convolutional neural networks,” *Decision Analytics Journal*, vol. 11, p. 100470, 2024.
- [23] E. Tuba, N. Baćanin, I. Strumberger, M. Tuba, and R. Jovanović, “Convolutional neural networks hyperparameters tuning,” in *Artificial Intelligence: Theory and Applications*. Cham: Springer, pp. 65–84, 2021. doi: 10.1007/978-3-030-72711-6_4.
- [24] M. Feurer and F. Hutter, “Hyperparameter optimization,” *Automated Machine Learning*, vol. 3, no. 2, pp. 33–50, 2019.
- [25] A. Morales-Hernández, I. Van Nieuwenhuyse, and S. Rojas Gonzalez, “A survey on multi-objective hyperparameter optimization algorithms for machine learning,” *Artificial Intelligence Review*, vol. 56, pp. 8043–8093, 2023. doi: 10.1007/s10462-022-10359-2.
- [26] F. Peres and M. Castelli, “Combinatorial optimization problems and metaheuristics: Review, challenges, design, and development,” *Applied Sciences*, vol. 11, no. 14, p. 6449, 2021.
- [27] S. Adamu, H. Alhussian, N. Aziz, S. J. Abdulkadir, A. Alwadin, A. A. Imam, A. Garba, and Y. Saidu, “Optimizing hyperparameters for improved melanoma classification using metaheuristic algorithm,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 10, pp. 531–540, 2023. doi: 10.14569/IJACSA.2023.0141062.
- [28] M. R. Apriyadi and D. P. Rini, “Hyperparameter optimization of support vector regression algorithm using metaheuristic algorithm for student performance prediction,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 2, 2023.
- [29] N. Bacanin, T. Bezdan, E. Tuba, I. Štrumberger, and M. Tuba, “Optimizing convolutional neural network hyperparameters by enhanced swarm intelligence metaheuristics,” *Algorithms*, vol. 13, no. 3, p. 67, 2020. doi: 10.3390/a13030067.
- [30] A. Brodzicki, “The whale optimization algorithm approach for deep learning hyperparameter tuning,” *Sensors*, vol. 21, no. 23, p. 8003, 2021. doi: 10.3390/s21238003.
- [31] D. Liu, H. Ouyang, S. Li, and C. Zhang, “Hyperparameters optimization of convolutional neural network based on local autonomous competition harmony search algorithm,” *Journal of Computational Design and Engineering*, vol. 10, no. 4, pp. 1280–1297, 2023.
- [32] S. Kiliçarslan, “PSO+GWO: A hybrid particle swarm optimization and grey wolf optimization based algorithm for fine-tuning hyper-parameters of convolutional neural networks for cardiovascular disease detection,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 1, pp. 87–97, 2023.
- [33] S. V. S. Kumar and H. K. Kondaveeti, “Bird species recognition using transfer learning with a hybrid hyperparameter optimization scheme (HHOS),” *Ecological Informatics*, 2024. doi: 10.1016/j.ecoinf.2024.102510.
- [34] H. Ulutas, R. B. Günay, and M. E. Sahin, “Detecting diabetes in an ensemble model using a unique PSO-GWO hybrid approach to hyperparameter optimization,” *Neural Computing and Applications*, vol. 36, no. 29, pp. 18313–18341, 2024.
- [35] H. Yan, C. Zhong, Y. Wu, L. Zhang, and W. Lu, “A hybrid-model optimization algorithm based on the Gaussian process and particle swarm optimization for mixed-variable CNN hyperparameter automatic search,” *Frontiers of Information Technology & Electronic Engineering*, vol. 24, no. 11, pp. 1557–1573, 2023.
- [36] E. Akkuş, U. Bal, F. Ö. Koçoğlu, and S. Beyhan, “Hyperparameter optimization of pre-trained convolutional neural networks using adolescent identity search algorithm,” *Neural Computing and Applications*, vol. 36, no. 4, pp. 1523–1537, 2024.
- [37] G. Atteia, N. Samee, E. El-kenawy, and A. Ibrahim, “CNN-hyperparameter optimization for diabetic maculopathy diagnosis in optical coherence tomography and fundus retinography,” *Mathematics*, vol. 10, no. 18, p. 3274, 2022.
- [38] A. Amou, H. Costa, and R. Ferreira, “A novel MRI diagnosis method for brain tumor classification based on CNN and Bayesian optimization,” *Healthcare*, vol. 10, no. 3, p. 494, 2022.