Metaheuristic Optimization in Cancer-Related Deep Learning Applications

Anh Tuan Tran
Kent State University
atran16@kent.edu

Jared Young
Kent State University
jared.young@artificialinstinct.net

Abstract—**This report investigates the application of metaheuristic optimization algorithms in cancer-related prediction tasks, with emphasis on both training-level and representation-level optimization in deep learning models. Four studies are examined: (1) Bat-optimized Convolutional Neural Networks (CNN) for lung cancer detection, (2) Whale Optimization Algorithm (WOA)-based weight optimization for breast cancer classification, (3) Firefly-based segmentation for brain tumor classification, and (4) Bat-optimized feature selection combined with Extreme Learning Machine (ELM) and transfer learning. The analysis highlights how metaheuristic algorithms can operate at different stages of the machine learning pipeline, including learning-rate tuning, weight optimization, segmentation threshold selection, and feature subset selection. Experimental findings reported in the literature demonstrate performance improvements over conventional gradient-based methods. This study provides a comparative technical analysis of these approaches, discusses their mathematical formulations, and evaluates their computational trade-offs and limitations.**

Keywords—**Metaheuristic optimization, Convolutional Neural Network (CNN), Bat Algorithm (BA), Whale Optimization Algorithm (WOA), Firefly Algorithm (FA), Extreme Learning Machine (ELM), Medical image analysis, Representation-level optimization**

I. INTRODUCTION

Deep learning has achieved remarkable success in medical image analysis, particularly in cancer detection and classification tasks. Most deep learning models rely on gradient-based optimization methods, such as backpropagation, to update network parameters. While effective, these methods may suffer from issues such as local minima, sensitivity to hyperparameter selection, and slow convergence in complex search spaces.

Metaheuristic optimization algorithms provide an alternative approach by performing global search without relying exclusively on gradient information. Algorithms such as the Bat Algorithm (BA) and Whale Optimization Algorithm (WOA) have demonstrated potential in optimizing neural network parameters, learning rates, and feature subsets.

In cancer-related prediction tasks, metaheuristics have been applied in two primary ways: (1) training-level optimization, where learning rates or network weights are directly optimized, and (2) representation-level optimization, where segmentation thresholds or feature subsets are optimized prior to classification. This report analyzes four representative studies that apply metaheuristic optimization to lung cancer, breast cancer, and brain tumor classification problems. The mathematical formulations, optimization mechanisms, and reported performance improvements are examined, along with their computational trade-offs and limitations.

II. APPLICATION AREAS STUDIED

In this report, we focus on four applications of metaheuristic algorithms in cancer-related problems. The first application investigates the use of the Whale Optimization Algorithm (WOA), one of the more recent metaheuristic techniques, to update network weights during the training process instead of relying solely on gradient-based methods. The second application explores the Bat Algorithm (BA), a well-known metaheuristic optimization technique, to determine the most suitable learning rate at a specific stage of training. After identifying the optimal learning rate, the model parameters are further optimized using the traditional backpropagation algorithm. This hybrid approach combines the global search capability of metaheuristics with the fine-tuning efficiency of gradient-based optimization.

A. Breast cancer detection based on optimized neural network using whale optimization algorithm

The Whale Optimization Algorithm (WOA) is a relatively recent nature-inspired metaheuristic optimization algorithm in 2016. It is inspired by the unique bubble-net hunting strategy of humpback whales, a cooperative feeding behavior in which whales create spiral-shaped bubbles around their prey before attacking. Mathematically, this behavior is modeled using mechanisms of encircling prey, spiral position updating, and random exploration. Like other population-based metaheuristics, WOA initializes a set of candidate solutions and iteratively updates them using stochastic equations that balance exploration and exploitation.

The study in [2] proposes a computer-aided diagnosis (CAD) system for breast cancer detection using mammogram images. Instead of training a Multilayer Perceptron (MLP) using conventional backpropagation, the authors replace gradient descent with the Whale Optimization Algorithm (WOA) for direct weight optimization.

The system consists of two primary components:

1) A feed-forward Multilayer Perceptron (MLP).
2) A population-based Whale Optimization Algorithm that searches for optimal weight configurations.

There are many crucial points in modeling the Whale Optimization Algorithm. Each Whale is modeled as a set of neural network weights:

$$X = [w_1, w_2, \dots, w_d] \tag{1}$$

where d denotes the total number of weights and biases in the network. Furtheremore, The mathematical model of the bubble-net system is defined as follows:

$$A = 2ar - a \tag{2}$$

$$C = 2r \tag{3}$$

$$X(t+1) = \begin{cases} X^*(t) - AD, & \text{if } P < 0.5 \\ D' e^{bl} \cos(2\pi t) + X^*(t), & \text{if } P \geq 0.5 \end{cases} \tag{4}$$

$$D' = |CX^*(t) - X(t)| \tag{5}$$

where P, r, and l are random variables such that $P, r, l \in [0, 1]$, a decreases linearly from 2 to 0 over the iterations, and b denotes the shape parameter of the logarithmic spiral. It is easy to observe from formula (4) that when the random value $P < 0.5$, the position $X(t+1)$ is updated based on X^*, which represents the best solution found so far. The parameter A models the behavior of encircling the prey: when $|A| > 1$, the whale explores the search space globally, whereas when $|A| < 1$, it fine-tunes the CNN weights locally. Also in formula (4), when the random value $P \geq 0.5$, the equation represents the spiral movement of the whale in the wild. Here, D' denotes the distance between the whale and the best solution found so far, e^{bl} is an exponential term controlling the spiral decay, and b is the logarithmic spiral coefficient, which is typically set to 1 in most cases. This mechanism directly influences how the weights of the CNN model are updated during optimization.

The application of the Whale Optimization Algorithm (WOA) for optimizing a CNN model, following the feed-forward process, is described below:

1) Generate an initial population of N whales (agents) representing the MLP weights and evaluate the fitness of each agent.
2) Update the position of each whale according to the fitness value, using bubble-net feeding and prey-encircling behaviors.
3) Apply other WOA parameters to each agent as defined by the algorithm.
4) Check whether the network has achieved a sufficiently low error or satisfies the stopping criteria.
5) If the stopping criteria are not met, return to step 2.
6) If the stopping criteria are satisfied, proceed to the next step.
7) End the algorithm.

The complexity of model is computed as O(MaxGen × NP × FitnessCost)

Although the approach of using the Whale Optimization Algorithm (WOA) for optimization after the feed-forward process demonstrates superior performance compared to traditional backpropagation, there are several important limitations that should be considered before applying it in practice.

If the globally best solution becomes trapped in a local minimum, the entire population may converge prematurely toward a suboptimal region of the search space. The logarithmic spiral movement in (4) is fixed and that formula seems not dynamic adjustment which may result in inefficient exploration of non-promising regions of the search space.

The computational complexity of WOA is generally much higher than gradient-based methods because it relies on a population of candidate solutions. In large-scale models such as deep CNNs with millions of parameters, applying WOA for weight optimization can significantly increase processing time, making it impractical for real-world applications.

B. Brain Tumor Classification Using Firefly-Based Segmentation

In this study, the Firefly Algorithm (FA) is applied at the representation level to optimize MRI segmentation prior to CNN classification [1]. Unlike weight optimization methods that modify network parameters directly, this approach improves classification indirectly by refining the quality of the input representation.

The preprocessing pipeline consists of MRI normalization, histogram equalization, and background removal. The segmentation stage is formulated as an optimization problem in which each firefly encodes a candidate threshold vector:

$$S = [t_1, t_2, \ldots, t_K] \tag{6}$$

where K denotes the number of segmentation thresholds. Each threshold partitions the MRI intensity space into regions corresponding to potential tissue types. Pixels are assigned to the cluster whose threshold minimizes intensity difference.

The segmentation quality is evaluated using the fitness function:

$$Fitness(Y) = \frac{D_w + E_y}{\alpha + D_b} \tag{7}$$

where:

- D_w denotes intra-cluster distance (compactness within a segmented region),
- D_b denotes inter-cluster distance (separation between tumor and healthy tissue),
- E_y denotes entropy within segmented regions,
- α is a weighting constant.

The optimization objective is to maximize tumor-to-background separation while minimizing intra-region variance and entropy. This ensures that the extracted region of interest (ROI) is compact, well-separated, and structurally coherent before being passed to the parallel CNN architecture.

Firefly movement follows canonical FA dynamics:

$$x_i^{t+1} = x_i^t + \beta e^{-\gamma r_{ij}^2}(x_j^t - x_i^t) + \alpha\epsilon \tag{8}$$

where β controls attractiveness, γ controls light absorption, and ϵ represents stochastic perturbation. Brighter fireflies (higher fitness) attract others, enabling convergence toward high-quality segmentation configurations.

By improving segmentation before classification, the CNN receives cleaner tumor boundaries, reducing noise-induced misclassification. Experimental results reported in [1] demonstrate that integrating Firefly-based segmentation improves classification accuracy compared to models without segmentation optimization.

However, this approach introduces computational overhead, as segmentation fitness must be evaluated for each firefly at every iteration. Furthermore, hyperparameter sensitivity (e.g., choice of K, α, β, and γ) can significantly influence convergence behavior.

C. Brain Tumor Classification Using Bat-Optimized Feature Selection and ELM

In contrast to segmentation-level optimization, the Bat-based approach operates at the feature-selection level [4]. After preprocessing and median filtering, Stationary Wavelet Transform (SWT) is applied to extract frequency-domain features (LL, LH, HL, HH sub-bands). These wavelet coefficients form a high-dimensional feature space.

Instead of using all extracted features, the Bat Algorithm is employed to identify an optimal subset that maximizes classification performance. Each bat encodes a candidate feature subset:

$$x_i = [f_1, f_2, \ldots, f_n] \tag{9}$$

where each element represents inclusion or exclusion of a feature.

The Bat Algorithm updates frequency, velocity, and position as follows:

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{10}$$

$$v_i^{(t)} = v_i^{(t-1)} + f_i(x_i^{(t)} - x^*) \tag{11}$$

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t)} \tag{12}$$

The fitness function is defined as the classification accuracy (or equivalently, minimization of error rate) obtained using an Extreme Learning Machine (ELM) classifier trained on the selected feature subset.

Unlike gradient-based neural networks, ELM assigns hidden-layer weights randomly and computes output weights analytically. This makes it computationally efficient, while the Bat Algorithm handles the combinatorial feature selection process.

This hybrid approach yields two advantages:

- Reduction of redundant or noisy wavelet features.
- Improved generalization by lowering model dimensionality.

Experimental results in [4] indicate that combining Bat-based feature selection with ELM and transfer learning improves classification accuracy compared to standalone CNN models.

Nevertheless, this method introduces additional optimization cost due to iterative fitness evaluation. Feature encoding design, dataset imbalance, and limited hyperparameter reporting also present challenges for reproducibility and scalability.

D. Lung Cancer Prediction Using a Bat-Inspired Metaheuristic Optimized Convolutional Neural Network

The Bat Algorithm (BA) is a nature-inspired metaheuristic optimization algorithm proposed in 2010. It is inspired by the echolocation behavior of microbats, which emit ultrasonic pulses and listen to the returning echoes to detect prey, obstacles, and distance. Mathematically, this behavior is modeled through frequency tuning, velocity and position updating, loudness adaptation, and pulse emission rate control. Similar to other population-based metaheuristics, BA initializes a set of candidate solutions (bats) and iteratively updates their positions using stochastic equations that balance global exploration and local exploitation in the search space.

The study in [3] proposes a hybrid BAT-optimized Convolutional Neural Network (CNN) for lung cancer detection. The workflow consists of three main stages: (1) FCM-based segmentation, (2) feature enhancement using transform methods, and (3) CNN classification optimized by the Bat Algorithm. The key contribution lies in applying the Bat Optimization Algorithm (BA) to optimize the CNN learning rate rather than manually tuning it or relying solely on gradient-based adjustment.

$$\omega_{i+1} = \omega_i - \eta \nabla E_i(\omega_i) \tag{13}$$

The formula (11) use how the weight are updated in propagation of training CNN model. And the role of Bat optimization algorithm is determining the most profitable value for learning rate η.

In generally, the main idea of this approach is apply the whole original Bat algorithm idea into adjusting learning rate so there are not thing changed in the Bat model. In Application, each bath presents for a candidate learning rate η. CNN is trained using that learning rate. Then classification error is computed. That error may be temporally seen as objective function and BAT searches for the learning rate that minimizes error

III. CONCLUSION

This report analyzed four cancer-related applications of metaheuristic optimization integrated with deep learning models. The examined approaches demonstrate that metaheuristics can operate at multiple levels of the machine learning pipeline, including training-level optimization (learning rate and weight tuning) and representation-level optimization (segmentation thresholds and feature subset selection).

Across the reviewed studies, metaheuristic algorithms were shown to enhance model performance by introducing global search capabilities that complement or replace traditional gradient-based optimization. In the breast cancer study, the Whale Optimization Algorithm (WOA) directly optimized neural network weights, replacing backpropagation as the primary training mechanism. In lung cancer detection, the Bat Algorithm (BA) was used to optimize the CNN learning rate, acting as a supervisory layer over gradient descent. In contrast, the brain tumor studies demonstrated that optimization need not occur at the weight level; instead, segmentation thresholds (via Firefly Algorithm) and feature subsets (via Bat-based selection combined with ELM) were optimized prior to classification.

This distinction highlights an important conceptual contribution: metaheuristics are not limited to parameter tuning within neural networks but can be deployed at different stages of the computational pipeline. Representation-level optimization improves the quality of input data before classification, while training-level optimization modifies how models converge during learning. Both approaches ultimately influence classification accuracy, but they do so through different mechanisms.

Despite the promising performance improvements reported in the literature, several limitations remain. Metaheuristic algorithms generally introduce higher computational complexity due to population-based search. Convergence stability depends heavily on hyperparameter configuration, and reproducibility can be challenging when parameter schedules are insufficiently documented. Additionally, many reported results are validated on specific datasets, and broader cross-institutional validation is required to assess generalizability in clinical environments.

Future research directions may include hybrid frameworks that combine metaheuristic global search with adaptive gradient-based refinement, dynamic hyperparameter scheduling, and integration with modern deep architectures such as transformer-based medical imaging models. Furthermore, improved benchmarking standards and reproducibility protocols are necessary to ensure that reported performance gains translate into practical medical decision-support systems.

In summary, metaheuristic optimization provides a flexible and powerful framework for enhancing deep learning models in cancer-related applications. Whether applied to learning rates, network weights, segmentation thresholds, or feature subsets, these algorithms function as adaptive supervisory mechanisms that modify either how models learn or what they learn from. As computational resources continue to improve, such hybrid optimization strategies may play an increasingly important role in medical AI systems.

REFERENCES

[1] C. Li, F. Zhang, Y. Du, and H. Li, “Classification of brain tumor types through MRIs using parallel CNNs and firefly optimization,” *Scientific Reports*, vol. 14, Art. no. 15057, 2024, doi: 10.1038/s41598-024-65714-w.

[2] H. Fang, H. Fan, S. Lin, Z. Qing, and F. R. Sheykhahmad, “Automatic breast cancer detection based on optimized neural network using whale optimization algorithm,” *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 425–438, 2021, doi: 10.1002/ima.22468.

[3] P. Priyadharshini and B. S. E. Zoraida, “Bat-inspired metaheuristic convolutional neural network algorithms for CAD-based lung cancer prediction,” *Journal of Applied Science and Engineering*, vol. 24, no. 1, pp. 65–71, 2021, doi: 10.6180/jase.202102_24(1).0008.

[4] G. R. Sreekanth, A. F. Alrasheedi, K. Venkatachalam, M. Abouhawwash, and S. S. Askar, “Extreme learning bat algorithm in brain tumor classification,” *Intelligent Automation & Soft Computing*, vol. 34, no. 1, pp. 249–265, 2022, doi: 10.32604/iasc.2022.024538.