Optimizing CNN Hyperparameters for Image Classification Using Evolutionary Algorithms: A Comparative Study (ACO, GA, PSO)

Shubhangi Jaiswal and Arpit Raj

Abstract:

This study investigates the use of evolutionary algorithms (EAs) to optimize the hyperparameters of a convolutional neural network (CNN) for multiclass image classification using the CIFAR-10 dataset. We explored three EA techniques, namely Ant Colony Optimization (ACO), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO), for searching the optimal hyperparameters. Results demonstrate that the EAs improved the model's performance by tuning the CNN hyperparameters, with PSO achieving the best results. This work highlights the potential benefits of applying EAs to hyperparameter optimization in CNNs for image classification tasks.

1. Introduction:

Convolutional Neural Networks (CNNs) have gained widespread popularity in recent years due to their exceptional performance in image classification tasks[2]. A key factor in a CNN's success is the choice of hyperparameters, which significantly impact the model's training and overall performance. Manual tuning of hyperparameters can be time-consuming and might not yield the best results. Traditional search methods like grid search and random search may also be computationally expensive.

Evolutionary Algorithms (EAs) provide an alternative method for optimizing hyperparameters. EAs are nature-inspired optimization techniques that emulate biological evolution. In this study, we focus on optimizing the training hyperparameters of a CNN to build a multiclass image classification model using the CIFAR-10 dataset[1]. We explore the use of Ant Colony Optimization (ACO)[3], Genetic Algorithm (GA), and Particle Swarm Optimization (PSO) to search for optimal hyperparameters and compare their performance in terms of test accuracy and ROC-AUC score.

2. Problem & Data Description:

The CIFAR-10 dataset comprises 60,000 32x32 color images, with 10 classes and 6,000 images per class. The dataset is divided into 50,000 training images and 10,000 test images. Our goal is to optimize the training hyperparameters of a CNN to maximize classification performance on the CIFAR-10 dataset. The performance of the model is measured using test accuracy and ROC-AUC score.

3. Approach:

We employed three EAs (ACO, GA, and PSO) to search for the optimal hyperparameters within a predefined search space. The search space included the number of filters for two convolutional layers (num_filters1 and num_filters2), the number of dense units (dense_units), and the learning rate (learning_rate).

The bounds for the hyperparameter search space were set to the same values for all algorithms. The bounds were defined as follows:

- num_filters1: [16, 128]
- num_filters2: [16, 128]
- dense_units: [128, 1024]
- learning_rate: [1e-5, 1e-2]

3.1 Ant Colony Optimization (ACO)

ACO is a nature-inspired optimization technique based on the foraging behavior of ants. The algorithm simulates the way ants communicate by depositing pheromones on their paths to guide other ants to food sources. In our study, we experimented with various parameters such as colony size, the number of iterations, alpha, beta, rho, q, and randomness factor to optimize the CNN hyperparameters.

Table 1.1: ACO Hyperparameters				
colony_size: 20; n_iterations: 20; num_params: 4; randomness_factor: 0.5				
Run	alpha	beta	rho	q
1	1.5	4	0.8	0.3
2	1	4.5	0.8	0.5

3.2 Genetic Algorithm (GA)

The Genetic Algorithm is inspired by the process of natural selection and genetics. In GA, candidate solutions are encoded as chromosomes, and their fitness is evaluated by an objective function. We used a population size of 10, a crossover probability of 0.5, and a mutation probability of 0.2. We also applied Gaussian mutation with parameters mu, sigma, and indpb.

Table 1.2: GA Hyperparameters		
Population size: 10; Generations: 20		
Crossover probability	Mutation probability	Gaussian mutation parameters (mu, sigma, indpb)
0.5	0.2	(0, 0.2, 0.1)

3.3 Particle Swarm Optimization (PSO)

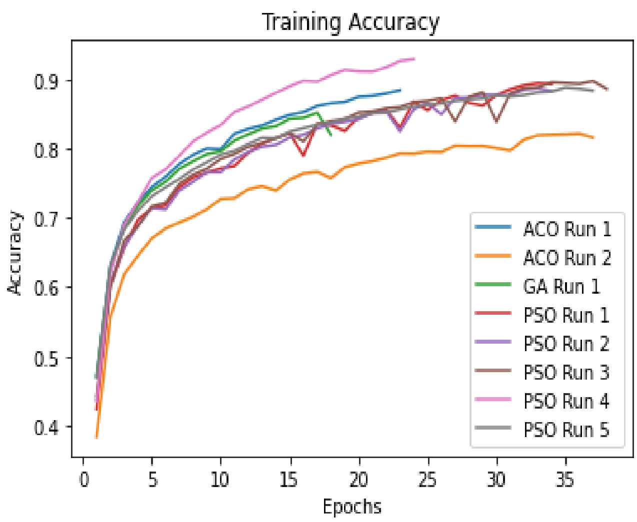
PSO is an optimization technique inspired by the flocking behavior of birds and fish. In PSO, particles representing potential solutions move in the search space to find the global optimum. We tested various combinations of cognitive (c1), social (c2), and inertia (w) parameters, along with the number of particles and iterations.

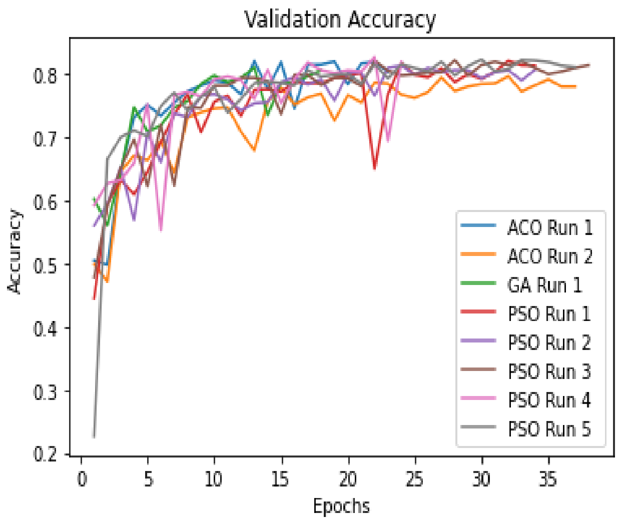
Table 1.3: PSO Hyperparameters	
n_particles: 20; optimiser_iters: 20	
Run	options
1	{'c1': 1.5, 'c2': 1.5, 'w': 0.9}
2	{'c1': 1.5, 'c2': 1.5, 'w': 0.9}
3	{'c1': 1.0, 'c2': 1.9, 'w': 0.8}
4	{'c1': 1.2, 'c2': 1.9, 'w': 0.7}
5	{'c1': 1.7, 'c2': 1.7, 'w': 0.92}

4. Results

Our results indicate that all three EAs improved the CNN's performance on the CIFAR-10 dataset. The best test accuracy and ROC-AUC score were achieved by the PSO algorithm with specific parameter configurations, followed by ACO and GA. The results obtained for each EA with different parameter configurations are shown in the table provided below:

Algorithm	Run No.	Test Accuracy	ROC-AUC Score
ACO	1	0.8122	0.9808
ACO	2	0.7938	0.9763
GA	1	0.8053	0.9799
PSO	1	0.8132	0.9806
PSO	2	0.8118	0.9807
PSO	3	0.8153	0.9815
PSO	4	0.7972	0.9781
PSO	5	0.8147	0.9809





4.1 ACO Results

ACO demonstrated a range of test accuracies and ROC-AUC scores depending on the selected parameters. The best results were achieved with a colony size of 20, 20 iterations, num_params set to 4, alpha at 1.5, beta at 4, rho at 0.8, q at 0.3, and a randomness factor of 0.5.

4.2 GA Results

For the Genetic Algorithm, the best results were obtained using a population size of 10, 20 generations, a crossover probability of 0.5, a mutation probability of 0.2, and Gaussian mutation parameters mu at 0, sigma at 0.2, and indpb at 0.1.

4.3 PSO Results

PSO provided the best overall results with a range of parameter configurations. The highest test accuracy and ROC-AUC score were achieved using options {'c1': 1.0, 'c2': 1.9, 'w': 0.8}, 20 particles, and 20 optimization iterations.

5. Discussion:

Our study demonstrates the potential of using EAs for optimizing CNN hyperparameters in multiclass image classification tasks. By searching for optimal hyperparameters within a predefined search space, we achieved improved test accuracy and ROC-AUC scores compared to default or randomly-selected hyperparameters. The choice of parameters for each EA was based on prior research and experimentation, aiming to strike a balance between exploration and exploitation.

Among the three EAs, PSO consistently outperformed ACO and GA in terms of test accuracy and ROC-AUC score. This may be attributed to the nature of PSO, which is known for its fast convergence and ability to adapt to changes in the search space. However, it is important to note that the performance of each EA is highly dependent on the choice of parameters, and there might be room for further improvement by fine-tuning these parameters.

In conclusion, our work highlights the importance of hyperparameter optimization in CNNs and shows that EAs can be an effective tool for this task. Further research may explore additional EA variants or investigate more advanced hyperparameter search spaces to further improve the performance of CNNs on image classification tasks. Additionally, future studies could focus on the application of EAs in optimizing the architecture of CNNs, as well as their potential use in other machine learning domains.

References:

[1] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Technical report, University of Toronto, 2009.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, May 2015.

[3] M. Dorigo and T. Stützle, "Ant colony optimization: Overview and recent advances," in Handbook of Metaheuristics, 3rd ed., Springer, Cham, 2019, pp. 311-351.

[4] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, Dec. 2014.