Evaluation Report:

Performance Evaluation of Convolutional Neural Networks with Evolutionary Algorithm Optimized Hyperparameters: A CIFAR-10 Multiclass Image Classification Analysis

Table of Contents:

Introduction:

In this evaluation report, we present a comprehensive analysis of the results obtained from multiple runs of Evolutionary Algorithm (EA) optimization techniques applied to a multiclass image classification task using Convolutional Neural Networks (CNNs) on the CIFAR-10 dataset. The EAs considered in this study include Ant Colony Optimization (ACO), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO). Each run consists of a specific set of hyperparameters for the respective algorithm. The main purpose of this report is to evaluate the performance of the CNN with different hyperparameters obtained from the optimization algorithms, with a focus on key metrics such as test accuracy, ROC-AUC score, and classification report.

The report is structured as follows: For each run, we present the algorithm used along with its specific hyperparameters, followed by the results, which include test accuracy, ROC-AUC score, and a detailed classification report that covers precision, recall, F1-score, and support for each class. Additionally, the epoch vs accuracy plot for training and validation accuracy is provided to visualize the learning progress of the CNN for each set of hyperparameters. This comprehensive evaluation aims to shed light on the effectiveness of each EA in optimizing the CNN's training hyperparameters and provide insights into the potential improvements that can be achieved through optimization.
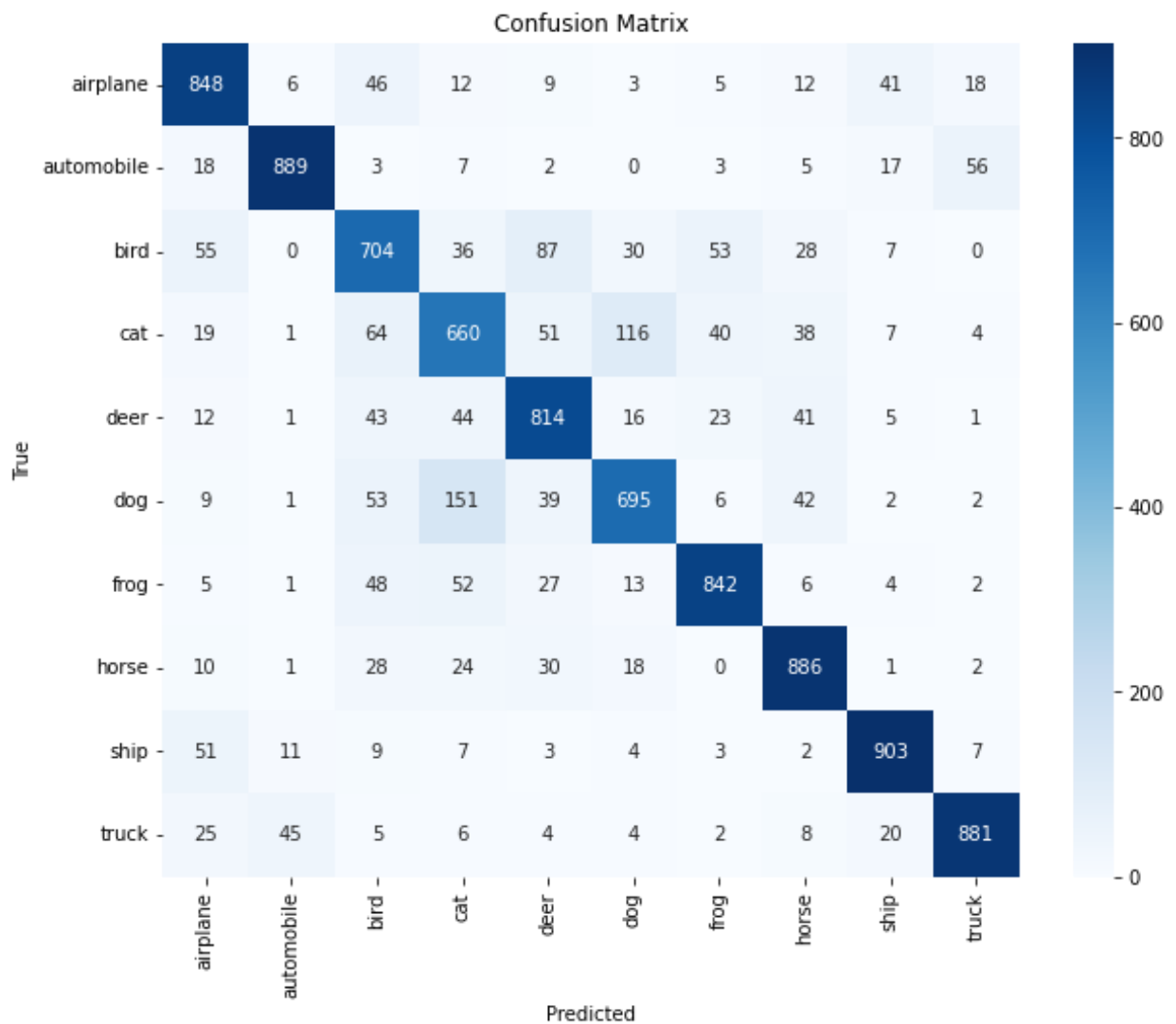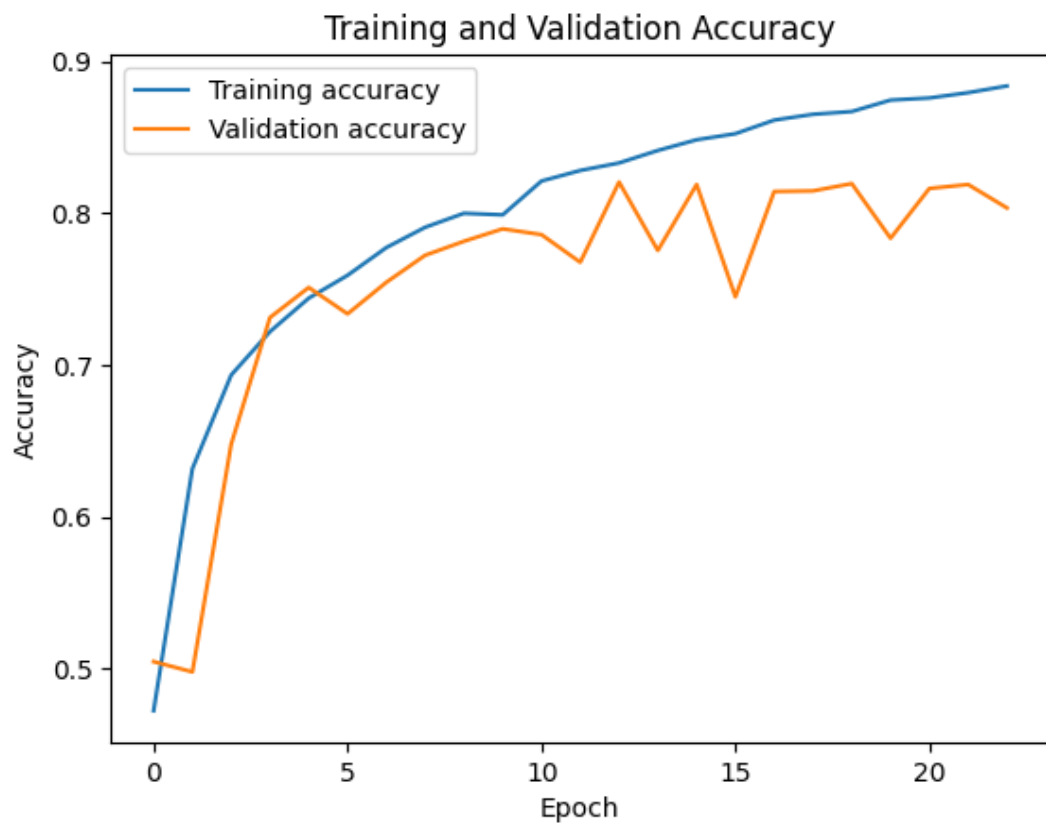
ACO Run 1

Run Information:

Algorithm	Hyperparameters
Ant Colony Optimization (ACO)	colony_size = 20, n_iterations = 20, num_params = 4, lower_bounds = [16, 16, 256, 0.0001], upper_bounds = [128, 128, 1024, 0.01], alpha = 1.5, beta = 4, rho = 0.8, q = 0.3, randomness_factor = 0.5

Results:

Test Accuracy	0.8122	ROC-AUC Score	0.9808

Classification Report:			
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.81	0.85	0.83	1000
1 (automobile)	0.93	0.89	0.91	1000
2 (bird)	0.7	0.7	0.7	1000
3 (cat)	0.66	0.66	0.66	1000
4 (deer)	0.76	0.81	0.79	1000
5 (dog)	0.77	0.69	0.73	1000
6 (frog)	0.86	0.84	0.85	1000
7 (horse)	0.83	0.89	0.86	1000
8 (ship)	0.9	0.9	0.9	1000
9 (truck)	0.91	0.88	0.89	1000
Average	**0.8153**	**0.8122**	**0.8107**	**10000**

Training and Validation Accuracy

Confusion Matrix

ACO Run 2

Run Information:

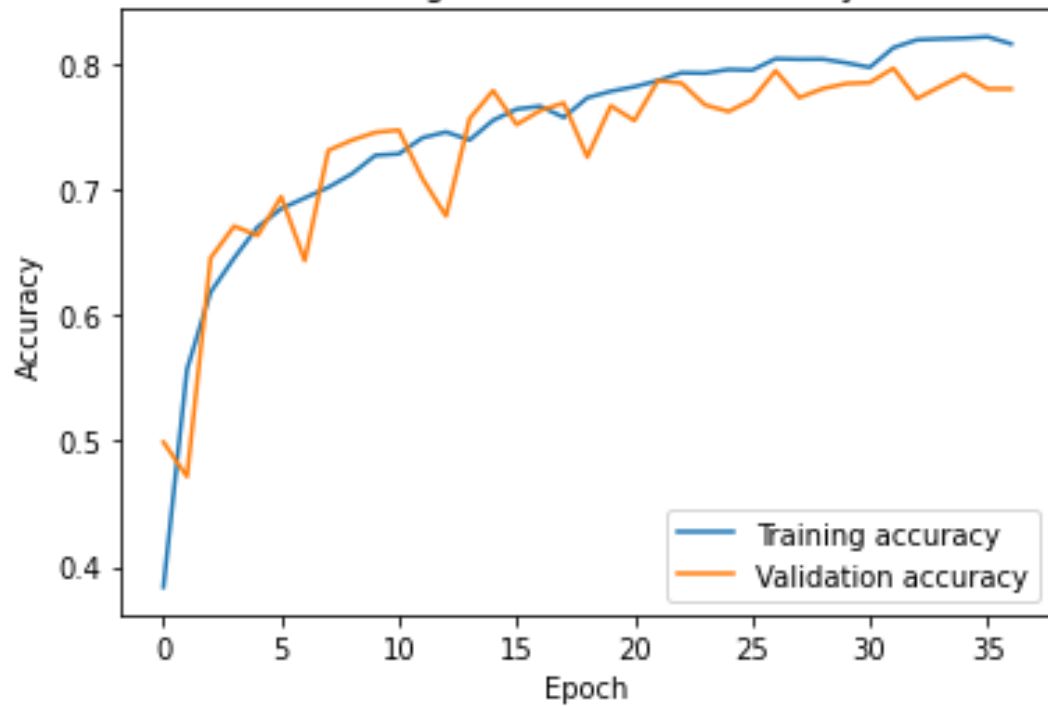| Algorithm	Hyperparameters
Ant Colony Optimization (ACO)	colony_size = 20, n_iterations = 20, num_params = 4, lower_bounds = [16, 16, 256, 0.0001], upper_bounds = [128, 128, 1024, 0.01], alpha = 1, beta = 4.5, rho = 0.8, q = 0.5, randomness_factor = 0.5

Results:

Test Accuracy	**0.7938**	ROC-AUC Score	**0.9763**

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.76	0.85	0.80	1000
1 (automobile)	0.86	0.94	0.90	1000
2 (bird)	0.63	0.77	0.70	1000
3 (cat)	0.66	0.59	0.62	1000
4 (deer)	0.78	0.77	0.77	1000
5 (dog)	0.76	0.66	0.71	1000
6 (frog)	0.82	0.87	0.84	1000
7 (horse)	0.91	0.77	0.83	1000
8 (ship)	0.88	0.88	0.88	1000
9 (truck)	0.93	0.84	0.88	1000
Average/Total	0.8153	0.8122	0.8107	10000

ACO Run 2

Training and Validation Accuracy

Confusion Matrix

GA Run 1

Run Information:

Algorithm	Hyperparameters
Genetic Algorithm (GA)	Population size = 10, Number of generations = 20, Crossover probability = 0.5, Mutation probability = 0.2, Gaussian mutation parameters: mu = 0, sigma = 0.2, indpb = 0.1

Results:

Test Accuracy	**0.8053**	ROC-AUC Score	**0.9799**

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.83	0.81	0.82	1000
1 (automobile)	0.89	0.93	0.91	1000
2 (bird)	0.83	0.62	0.71	1000
3 (cat)	0.61	0.69	0.65	1000
4 (deer)	0.81	0.72	0.76	1000
5 (dog)	0.69	0.77	0.73	1000
6 (frog)	0.82	0.87	0.85	1000
7 (horse)	0.83	0.86	0.84	1000
8 (ship)	0.89	0.90	0.90	1000
9 (truck)	0.88	0.88	0.88	1000
Average/Total	**0.8043**	**0.8044**	**0.8024**	**10000**

Training and Validation Accuracy

Confusion Matrix
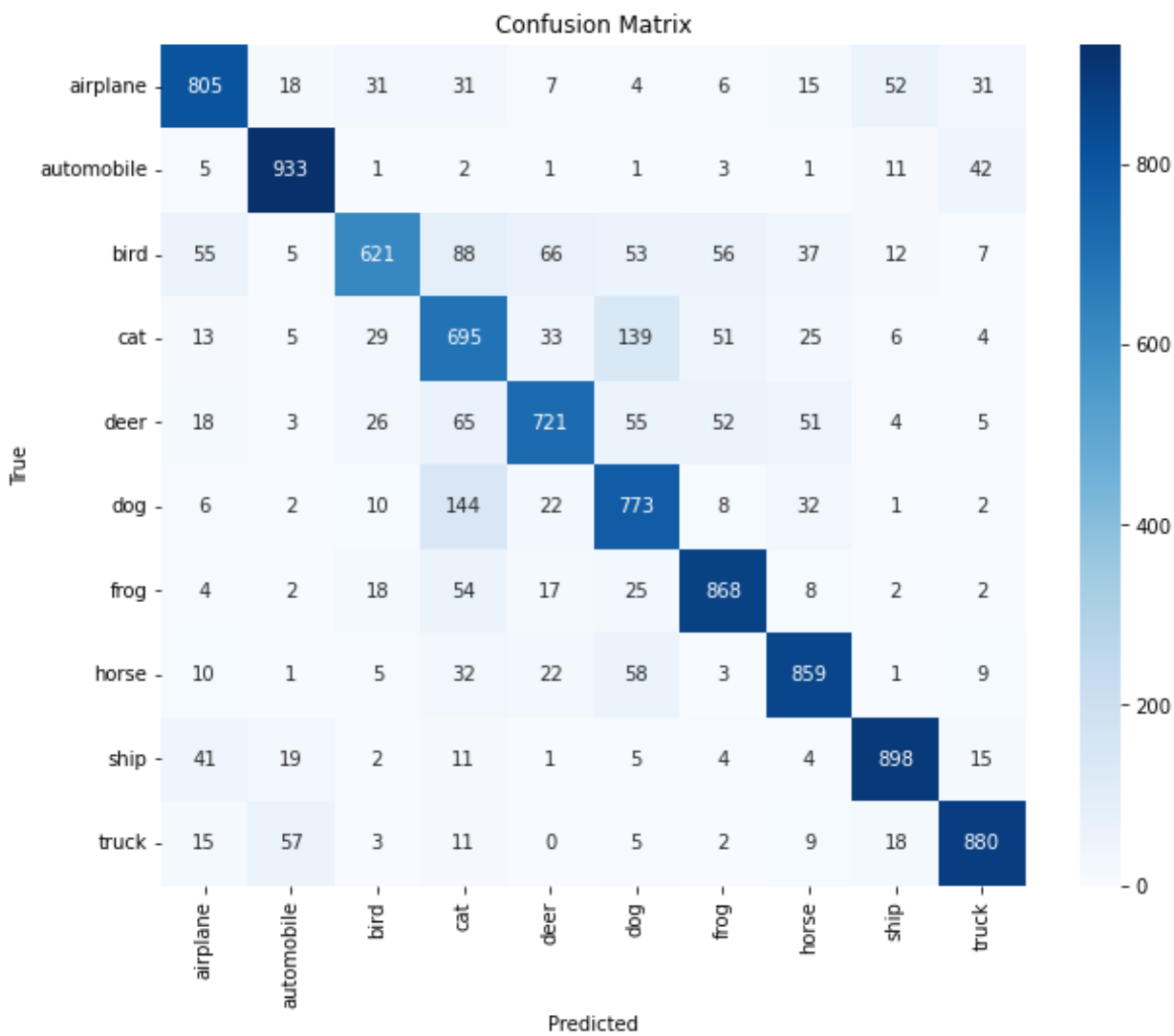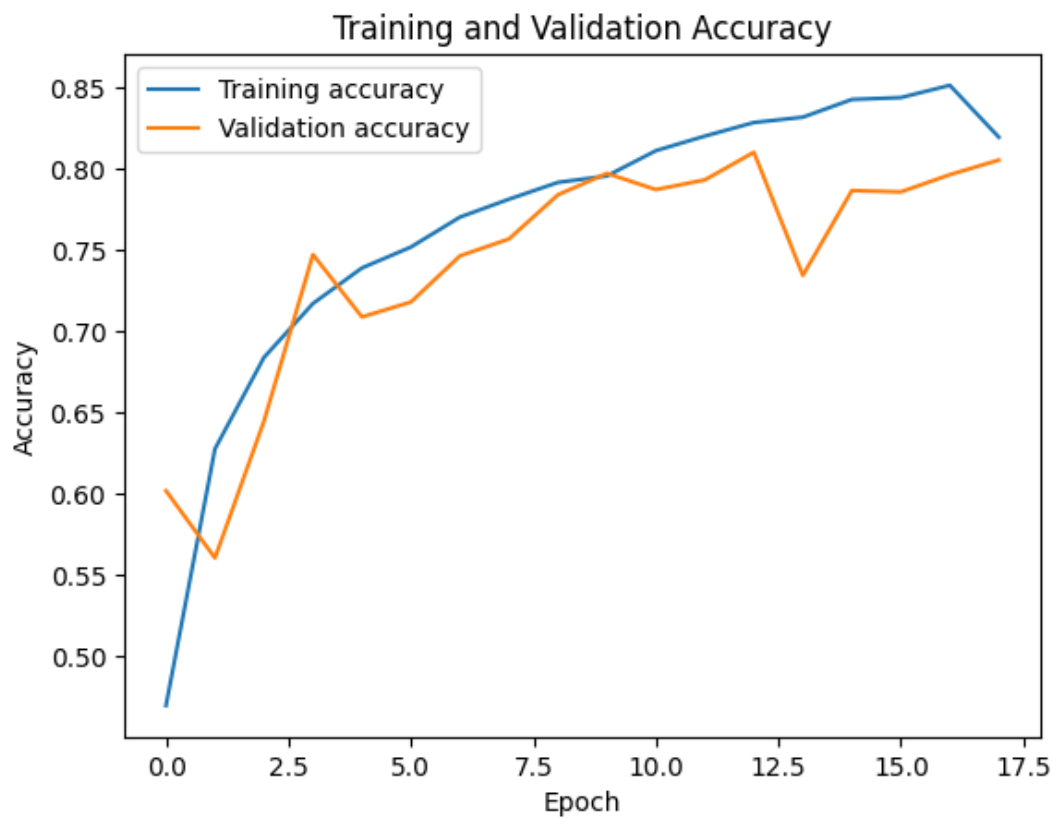
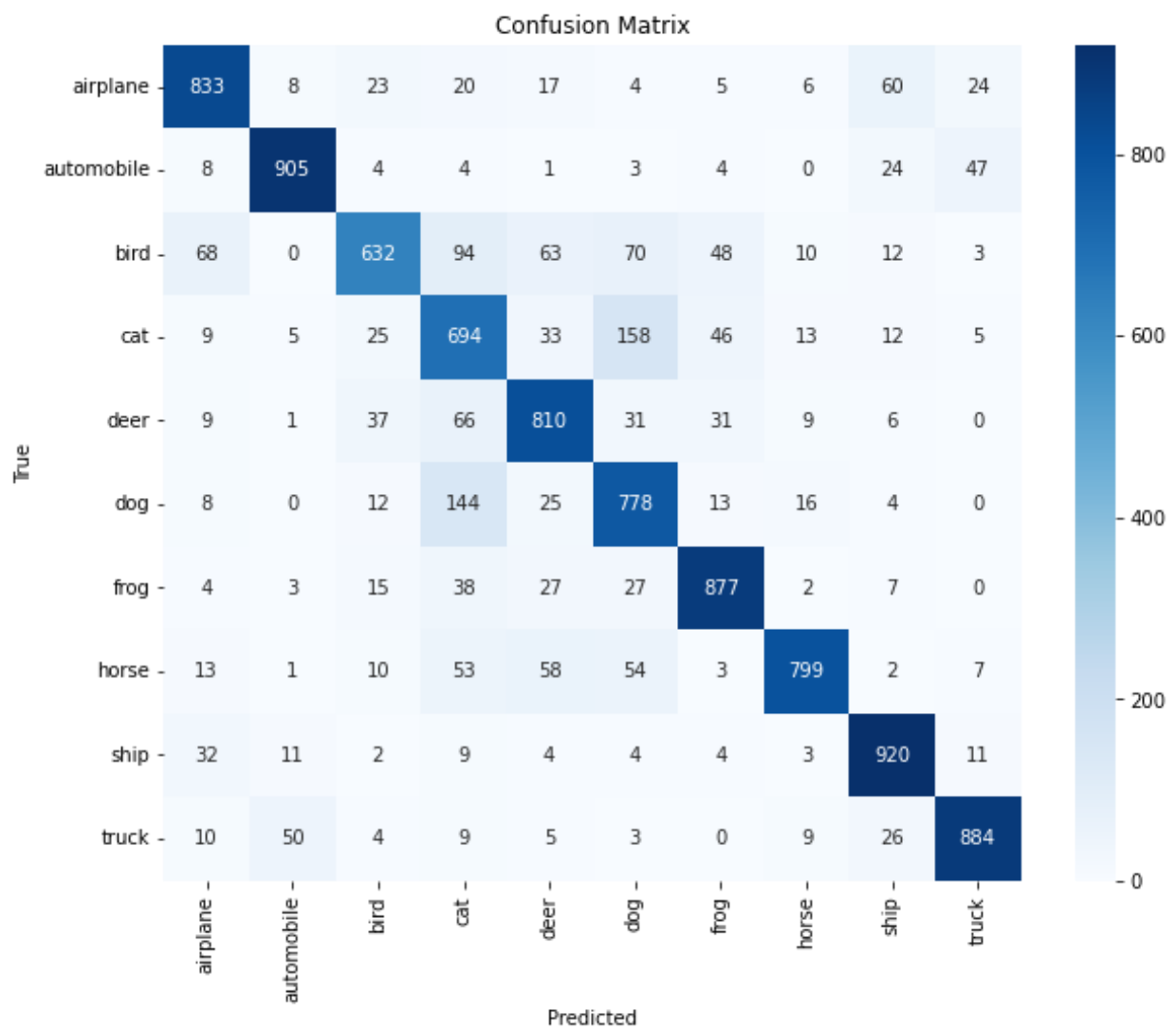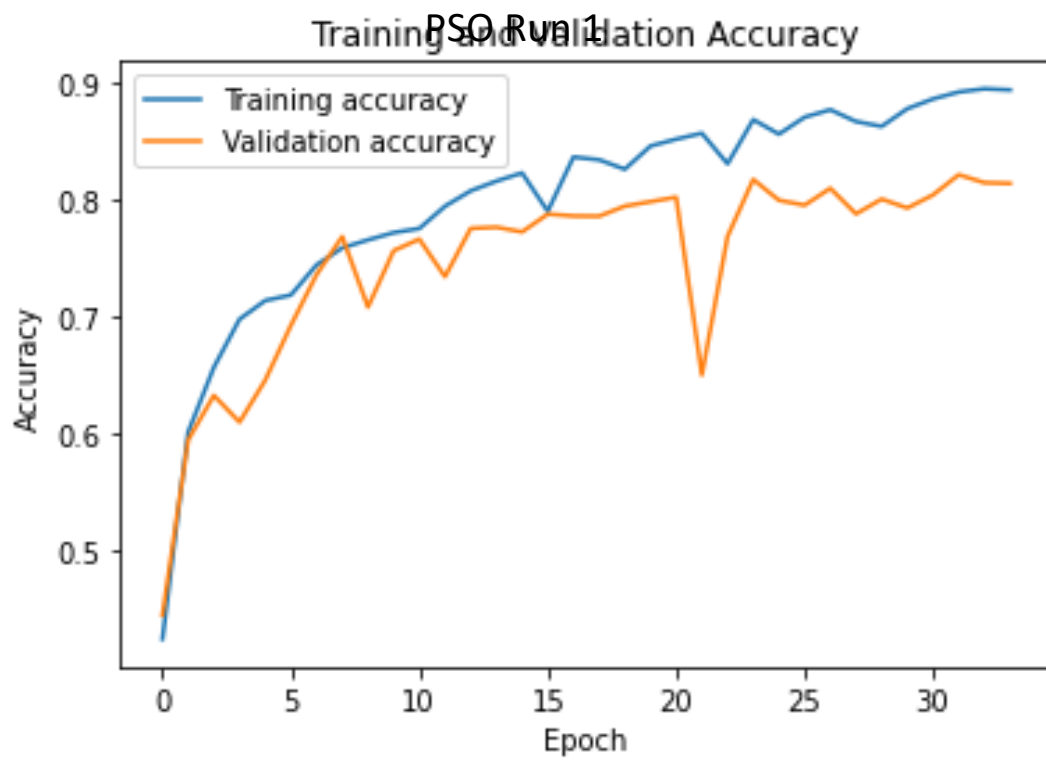# PSO Run 1

Run Information:

Algorithm	Hyperparameters
Particle Swarm Optimization (PSO)	{'c1': 1.5, 'c2': 1.5, 'w': 0.9}, n_particles = 20, optimiser_iters = 20

Results:

Test Accuracy	0.8132	ROC-AUC Score	0.9806

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.84	0.83	0.84	1000
1 (automobile)	0.92	0.91	0.91	1000
2 (bird)	0.83	0.63	0.72	1000
3 (cat)	0.61	0.69	0.65	1000
4 (deer)	0.78	0.81	0.79	1000
5 (dog)	0.69	0.78	0.73	1000
6 (frog)	0.85	0.88	0.86	1000
7 (horse)	0.92	0.80	0.86	1000
8 (ship)	0.86	0.92	0.89	1000
9 (truck)	0.90	0.88	0.89	1000
Average/Total	**0.8153**	**0.8122**	**0.8107**	**10000**

Training and Validation Accuracy

Confusion Matrix
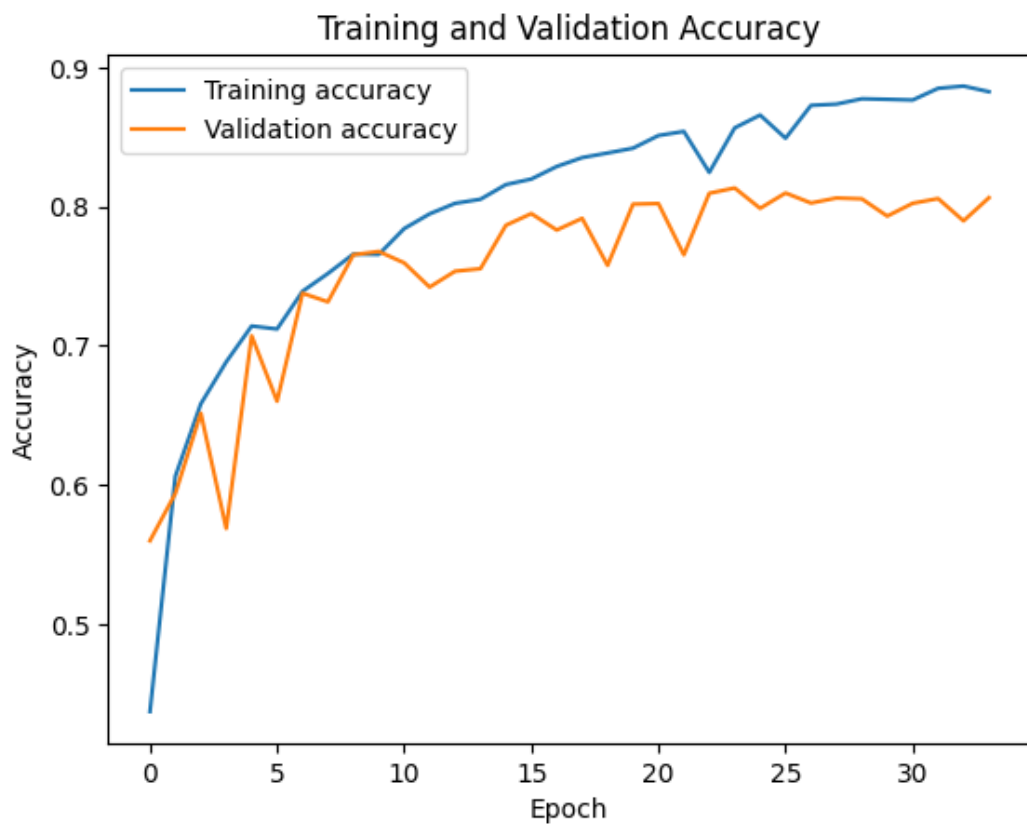
PSO Run 2

Run Information:

Algorithm	Hyperparameters
Particle Swarm Optimization (PSO)	{'c1': 1.5, 'c2': 1.5, 'w': 0.9}, n_particles = 20, optimiser_iters = 20

Results:

Test Accuracy	0.8118	ROC-AUC Score	0.9807

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.84	0.83	0.84	1000
1 (automobile)	0.92	0.91	0.91	1000
2 (bird)	0.83	0.63	0.72	1000
3 (cat)	0.61	0.69	0.65	1000
4 (deer)	0.78	0.81	0.79	1000
5 (dog)	0.69	0.78	0.73	1000
6 (frog)	0.85	0.88	0.86	1000
7 (horse)	0.92	0.80	0.86	1000
8 (ship)	0.86	0.92	0.89	1000
9 (truck)	0.90	0.88	0.89	1000
Average/Total	**0.8153**	**0.8129**	**0.8119**	**10000**

Training and Validation Accuracy

Confusion Matrix
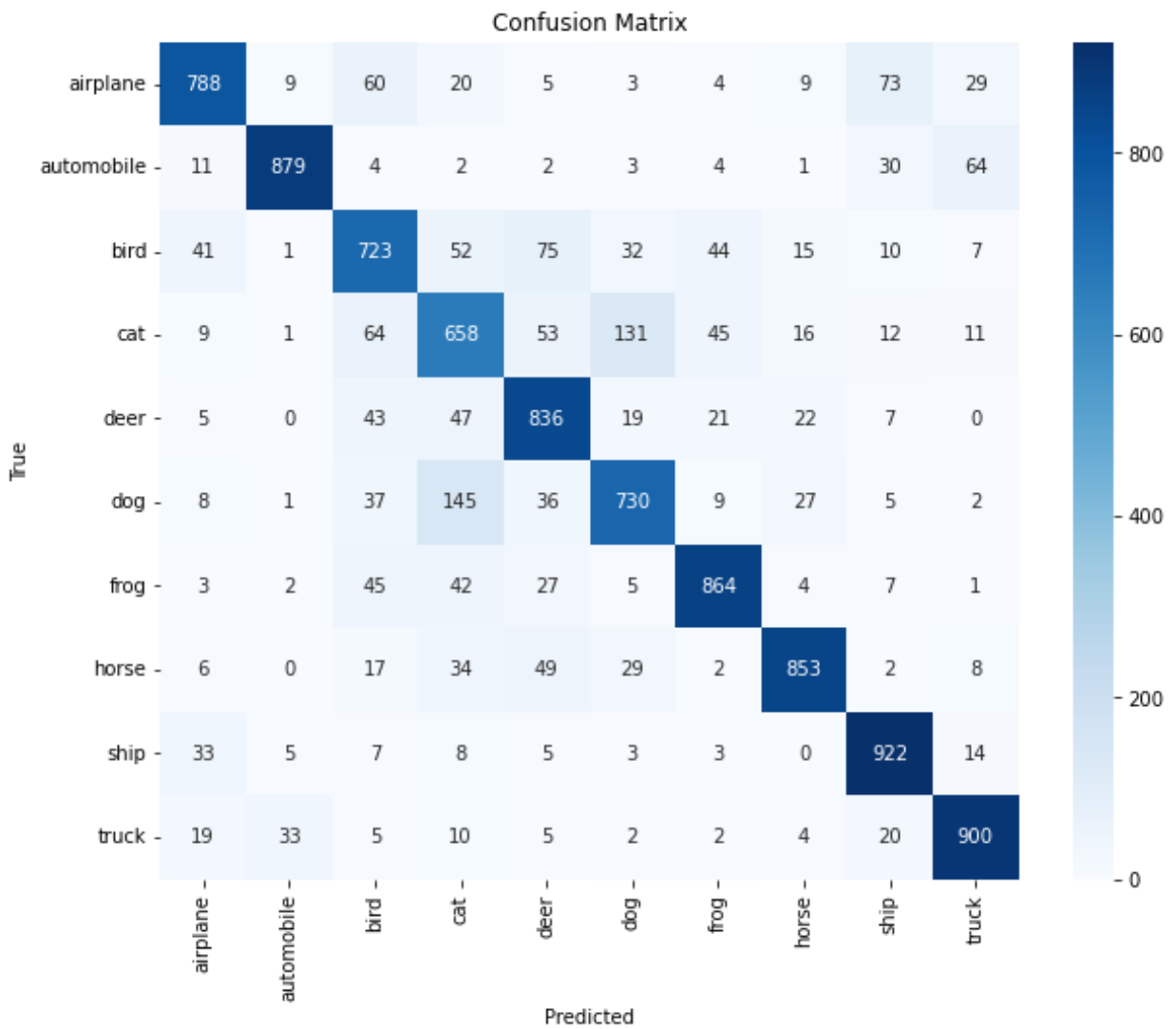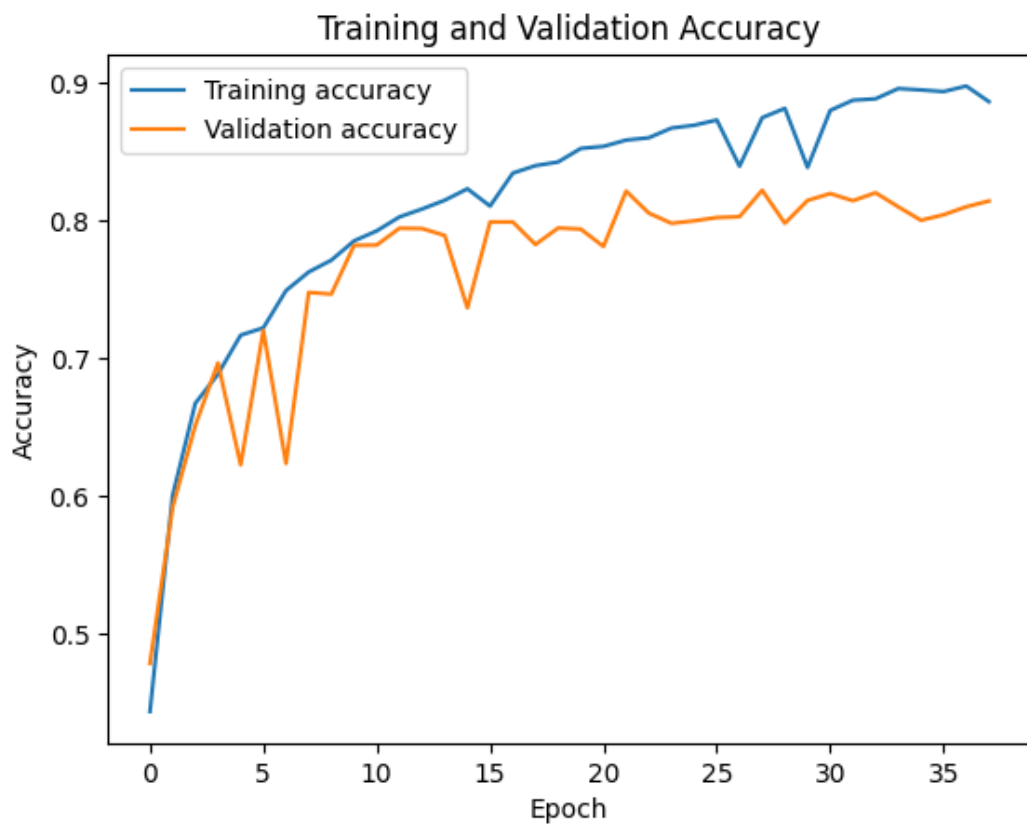
PSO Run 3

Run Information:

Algorithm	Hyperparameters
Particle Swarm Optimization (PSO)	{'c1': 1.0, 'c2': 1.9, 'w': 0.8}, n_particles = 20, optimiser_iters = 20

Results:

Test Accuracy	0.8153	ROC-AUC Score	0.9815

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.85	0.79	0.82	1000
1 (automobile)	0.94	0.88	0.91	1000
2 (bird)	0.72	0.72	0.72	1000
3 (cat)	0.65	0.66	0.65	1000
4 (deer)	0.76	0.84	0.80	1000
5 (dog)	0.76	0.73	0.75	1000
6 (frog)	0.87	0.86	0.86	1000
7 (horse)	0.90	0.85	0.87	1000
8 (ship)	0.85	0.92	0.88	1000
9 (truck)	0.87	0.90	0.88	1000
Average/Total	**0.8203**	**0.8203**	**0.8196**	**10000**

Training and Validation Accuracy

Confusion Matrix
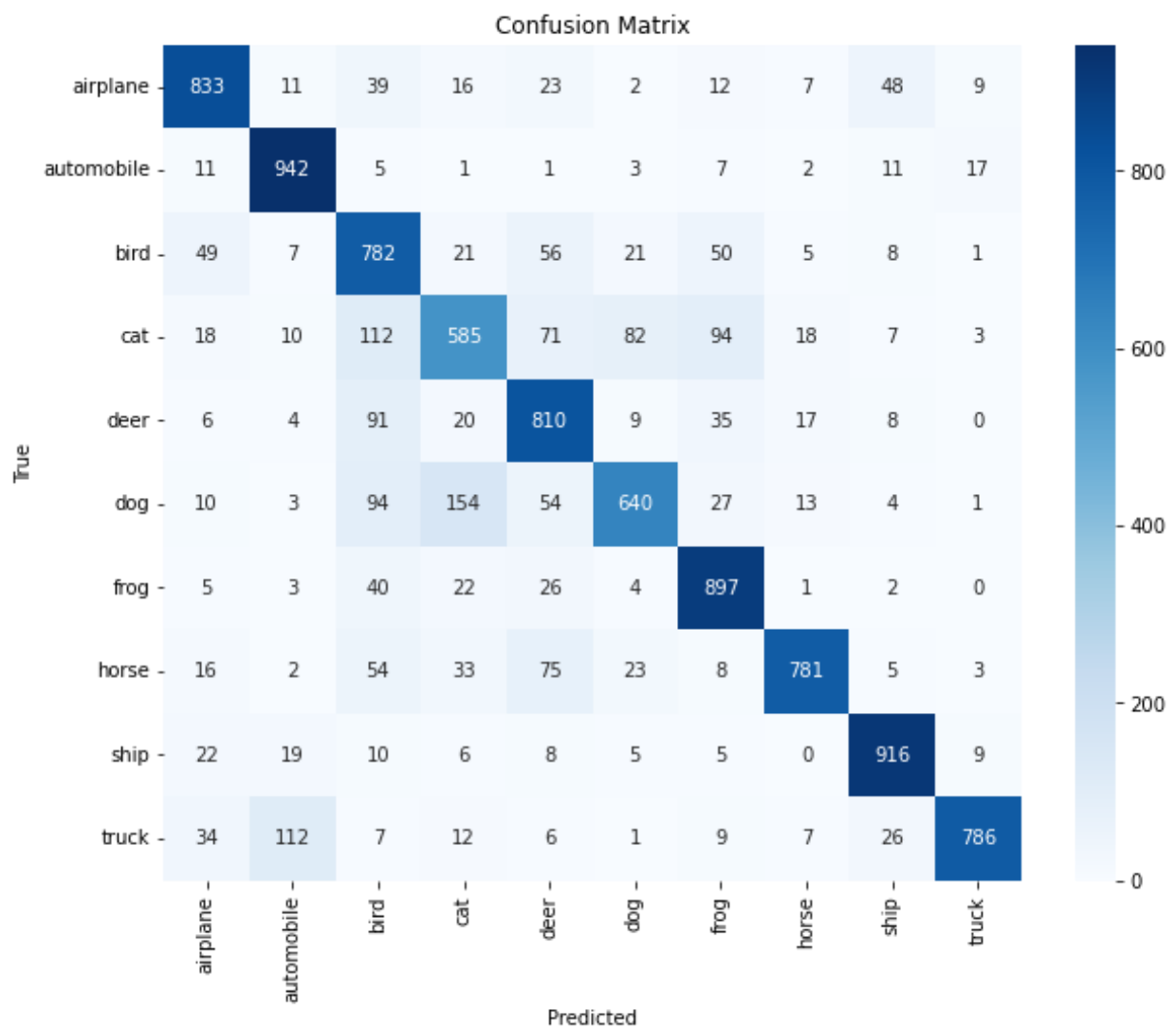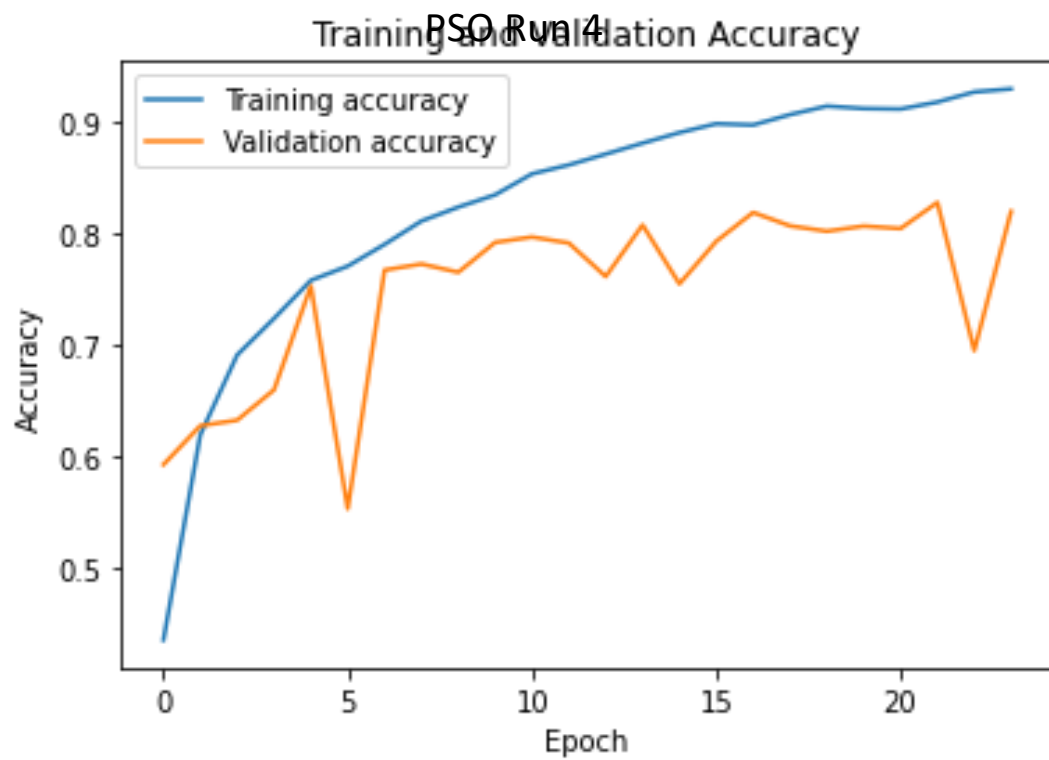
PSO Run 4

Run Information:

Algorithm	Hyperparameters
Particle Swarm Optimization (PSO)	{'c1': 1.2, 'c2': 1.9, 'w': 0.7}, n_particles = 20, optimiser_iters = 20

Results:

Test Accuracy	0.7972	ROC-AUC Score	0.9781

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.83	0.83	0.83	1000
1 (automobile)	0.85	0.94	0.89	1000
2 (bird)	0.63	0.78	0.70	1000
3 (cat)	0.67	0.58	0.63	1000
4 (deer)	0.72	0.81	0.76	1000
5 (dog)	0.81	0.64	0.72	1000
6 (frog)	0.78	0.90	0.84	1000
7 (horse)	0.92	0.78	0.84	1000
8 (ship)	0.89	0.92	0.90	1000
9 (truck)	0.95	0.79	0.86	1000
Average/Total	**0.8216**	**0.8209**	**0.8201**	**10000**

Training and Validation Accuracy

Confusion Matrix

PSO Run 5

Run Information:

Algorithm	Hyperparameters
Particle Swarm Optimization (PSO)	{'c1': 1.7, 'c2': 1.7, 'w': 0.92}, n_particles = 20, optimiser_iters = 20

Results:

Test Accuracy	**0.8147**	ROC-AUC Score	**0.9809**

Classification Report:				
Class	**Precision**	**Recall**	**F1-Score**	**Support**
0 (airplane)	0.84	0.84	0.84	1000
1 (automobile)	0.95	0.85	0.90	1000
2 (bird)	0.80	0.66	0.72	1000
3 (cat)	0.68	0.66	0.67	1000
4 (deer)	0.68	0.88	0.76	1000
5 (dog)	0.77	0.73	0.75	1000
6 (frog)	0.85	0.87	0.86	1000
7 (horse)	0.89	0.83	0.86	1000
8 (ship)	0.90	0.90	0.90	1000
9 (truck)	0.84	0.92	0.88	1000
Average/Total	**0.8306**	**0.8301**	**0.8298**	**10000**

Training and Validation Accuracy

Confusion Matrix
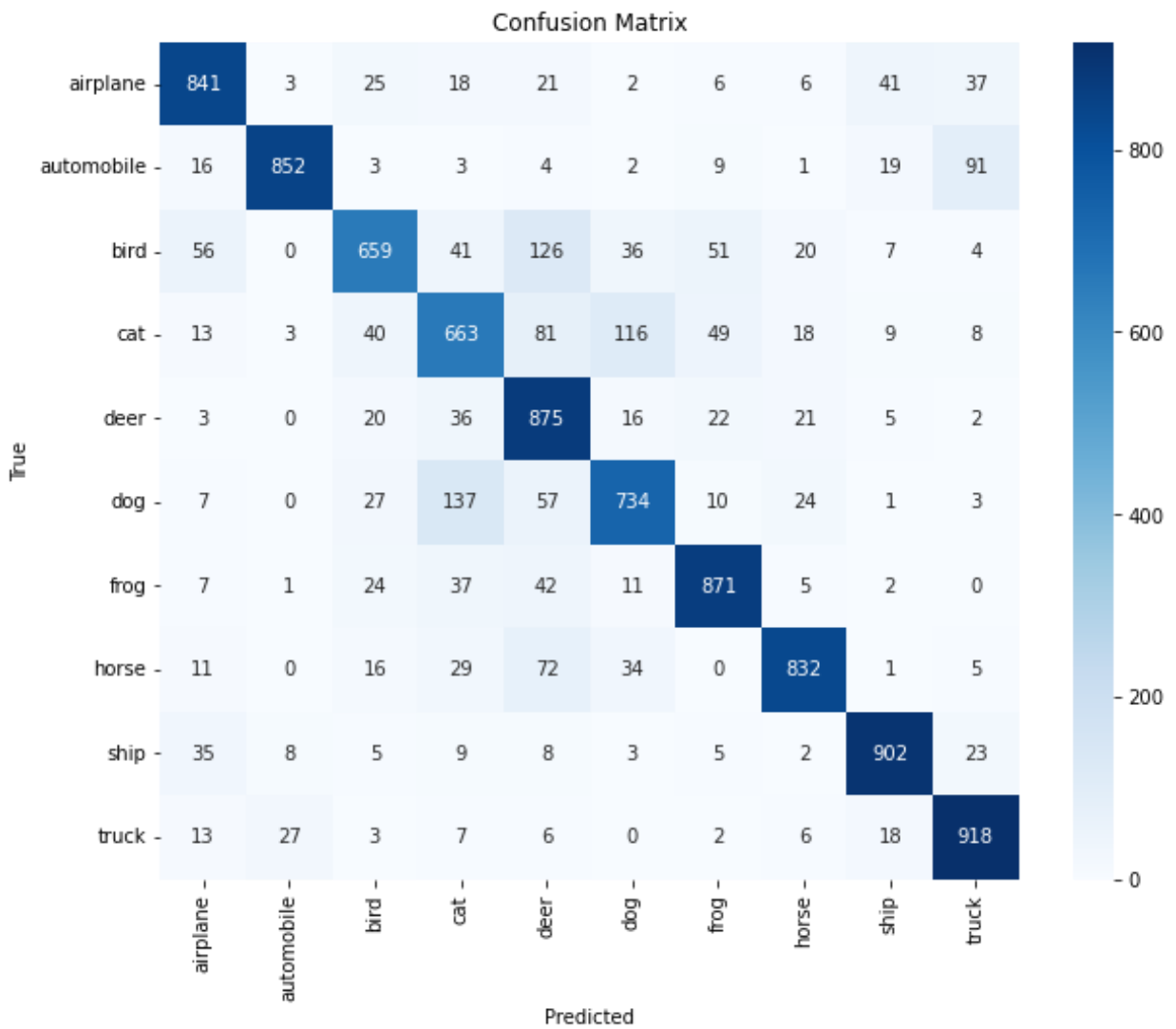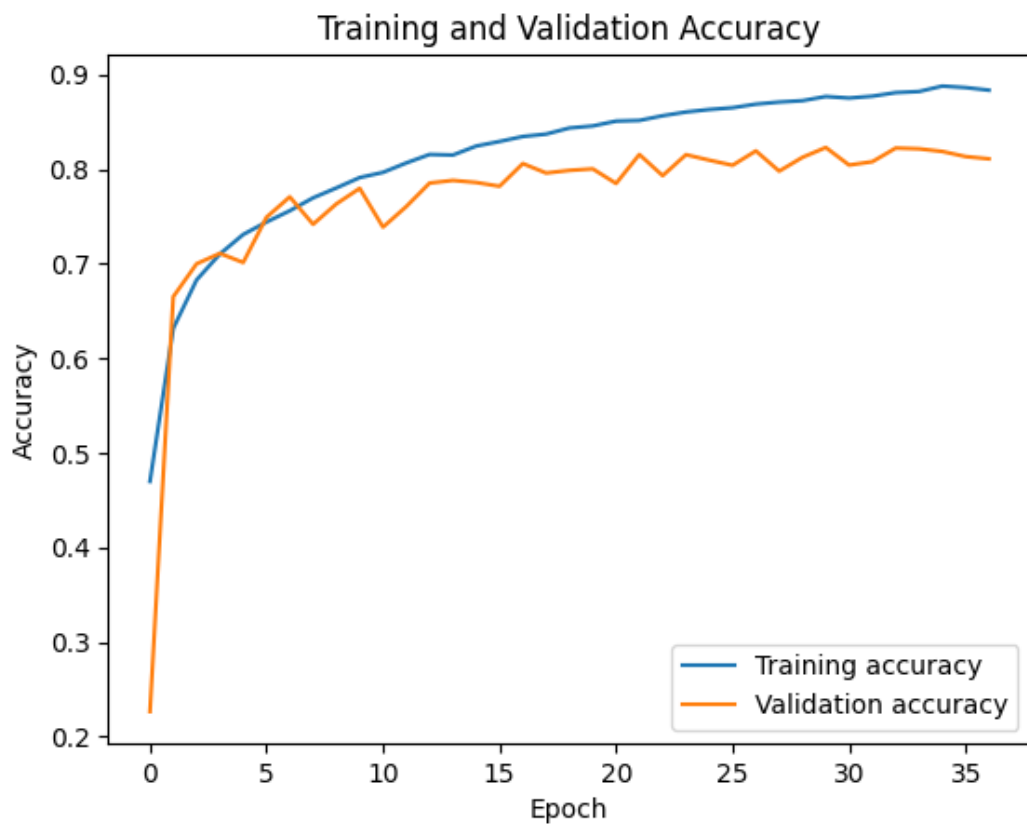
Conclusion:

In conclusion, the evaluation results presented in this report provide valuable insights into the performance of the CNN with different sets of hyperparameters optimized by the EAs. By comparing the test accuracy, ROC-AUC score, and classification report for each run, we can identify the most effective EA configurations and the corresponding hyperparameters that contribute to improved performance in the image classification task.

The epoch vs accuracy plots for training and validation accuracy help visualize the learning progress of the CNN with different hyperparameter sets, enabling further understanding of the training dynamics under various configurations. This in-depth evaluation highlights the potential of EAs to enhance the training process and overall performance of CNNs in image classification tasks. The insights gained from this report can be instrumental in guiding future work, including the exploration of alternative EA techniques, fine-tuning of hyperparameters, and application to other datasets and tasks.