Nature-Inspired Hyperparameter Optimization and Explainable AI for Alzheimer's Stage Classification

Nature Inspired Computation Course Project

20/12/2025

Contents

1 Dataset and Preprocessing Details

The project focuses on the multi-class classification of Alzheimer's Disease stages using the **OASIS (Open Access Series of Imaging Studies)** dataset. The classification task involves four distinct clinical stages:

- **Non-Demented**: Control group.
- **Very Mild Dementia**: Early-stage cognitive decline.
- **Mild Dementia**: Progressed cognitive impairment.
- **Moderate Dementia**: Advanced stage (least frequent class).

1.1 Preprocessing Pipeline

To ensure compatibility with the deep learning backbone, the following transformations were implemented:

- **Resizing**: All MRI scans were resized to 224×224 pixels.
- **Normalization**: Pixel values were normalized using ImageNet statistics: Mean $[0.485, 0.456, 0.406]$ and Std $[0.229, 0.224, 0.225]$.
- **Augmentation**: To mitigate overfitting, training data was subjected to Random Horizontal Flips and Random Rotations (up to 10 degrees).

2 Model Architecture

The project employs **EfficientNet-B0** as the primary feature extractor, chosen for its efficiency and state-of-the-art performance in medical imaging tasks.

2.1 Architectural Modifications

The pre-trained backbone was modified with a custom classification head:

1. **Feature Extractor**: EfficientNet-B0 (Global Average Pooling layer output: 1280).
2. **Dense Layer**: Linear ($1280 \rightarrow 256$ units).
3. **Activation**: ReLU.
4. **Regularization**: Dropout (Rate optimized via Nature-Inspired algorithms).
5. **Output**: Linear ($256 \rightarrow 4$ classes).

3 Nature-Inspired Optimization Pipeline

3.1 Hyperparameter Search Space

The metaheuristic algorithms searched for the optimal configuration within the following defined bounds:

Table 1: Hyperparameter Search Space Bounds

Hyperparameter	**Lower Bound**	**Upper Bound**
Learning Rate (Log Scale)	1×10^{-5}	3×10^{-3}
Weight Decay	0.0	1×10^{-2}
Dropout Rate	0.0	0.6
Unfreeze Backbone Epoch	0	3
Batch Size	Discrete set: $\{8, 16, 32\}$	
Augmentation Strength	0.5	1.5

3.2 Ant Colony Optimization (ACO) for Feature Selection

ACO was utilized as a dimensionality reduction technique to select a subset of the 1280-dimensional EfficientNet embeddings. By simulating pheromone trails, the algorithm successfully selected 640 features, which provided the best trade-off between dimensionality and predictive power.

4 Experimental Results and Visualizations

4.1 Metaheuristic Rankings and Performance

The table below summarizes the performance of the algorithms across Phase 1 (Metaheuristic Comparison) and Phase 2 (Optimized Training).

Table 2: Metaheuristic Performance Rankings

Algorithm	**Score (F1)**	**Time (s)**	**Source**	**Efficiency**	**Rank**
PSO	0.840161	1790.982354	Phase 1	0.000469	1
GWO	0.828895	1697.871904	Phase 1	0.000488	2
BAT	0.821778	1838.819566	Phase 1	0.000447	3
DE	0.797766	2474.866984	Phase 1	0.000322	4
GWO_tuned	0.735045	130.255906	Phase 2	0.005643	5
WOA_tuned	0.690008	118.925990	Phase 2	0.005802	6

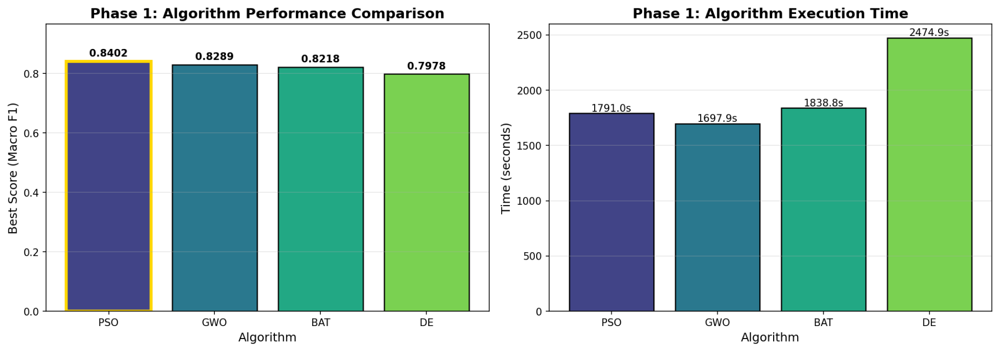


Figure 1: Phase 1 comparison of final fitness scores across metaheuristics.

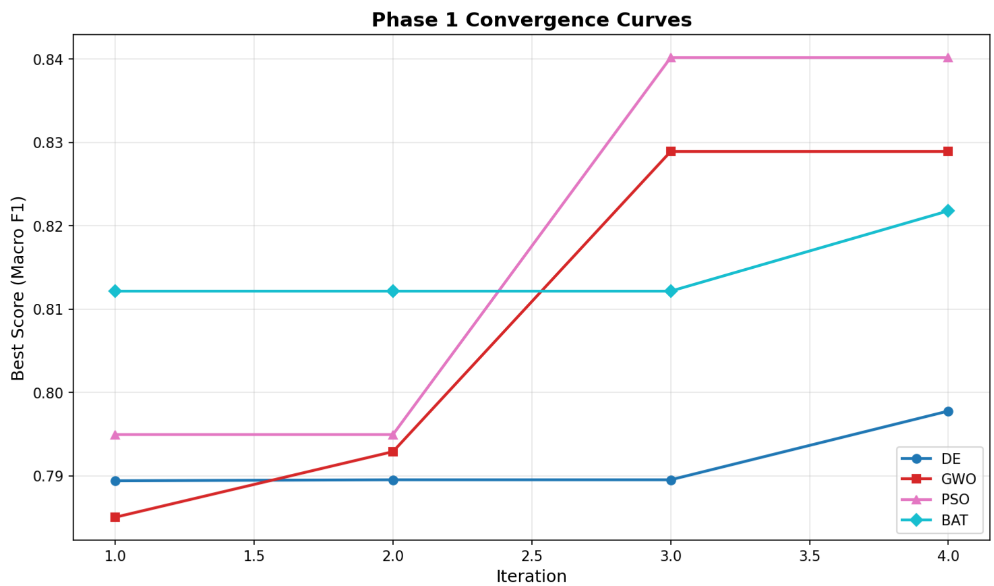


Figure 2: Convergence behavior: Tracking validation performance over search iterations.

4.2 Phase 2: Training and Uplift

The final optimized model shows significant performance improvements over the baseline.

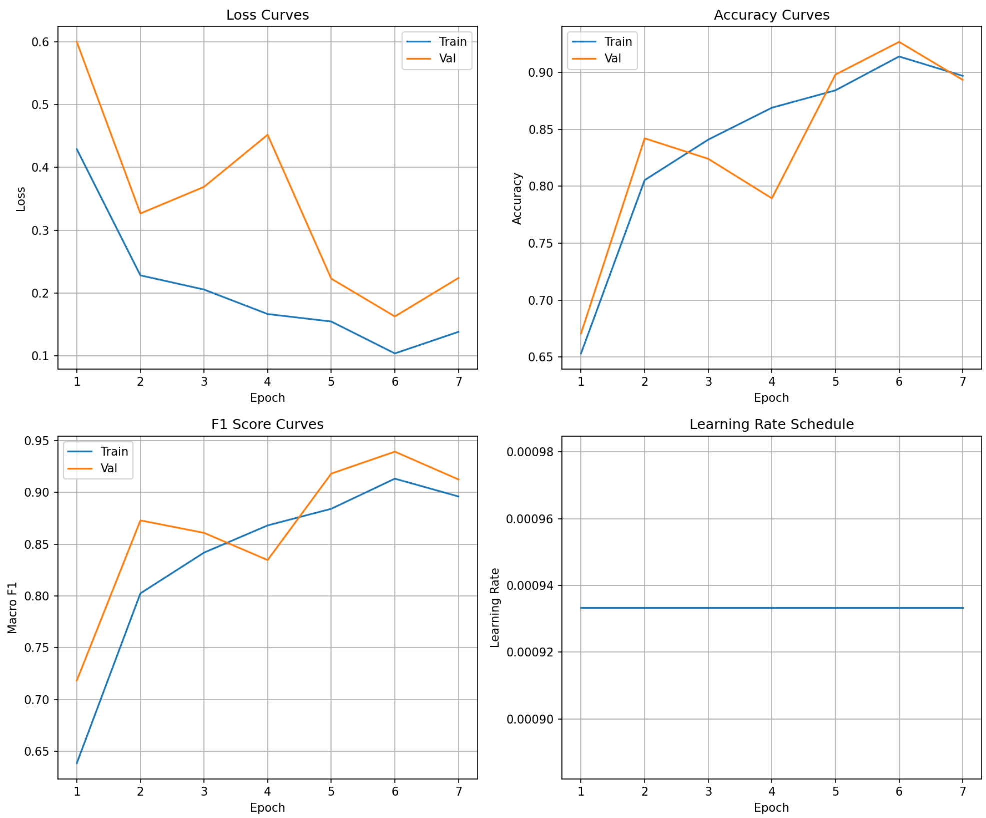


Figure 3: Training metrics (Loss and Accuracy) for the optimized model.

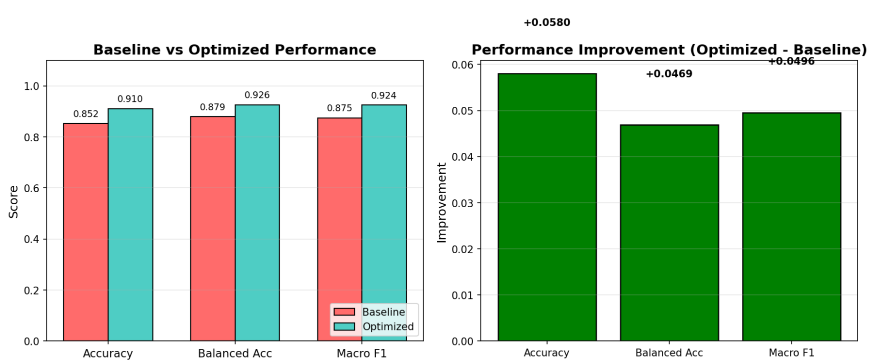


Figure 4: Direct comparison between baseline and NIC-optimized model performance.

5 Explainable AI (XAI) Analysis

5.1 XAI Quantitative Comparison

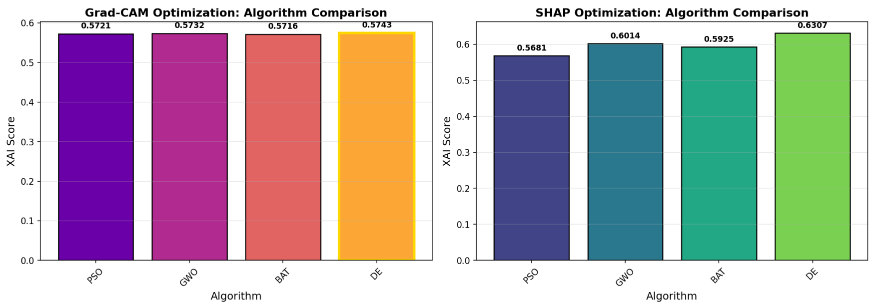
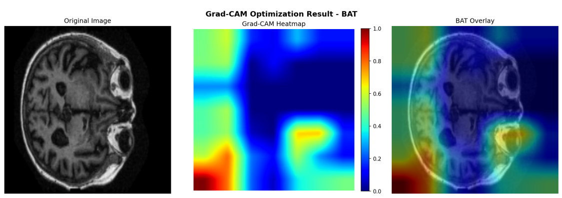
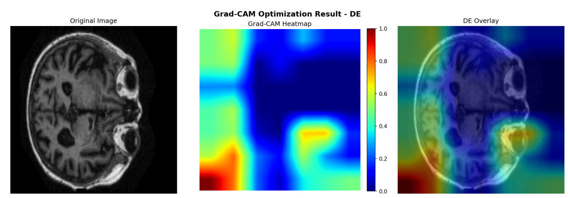


Figure 5: Comparison of XAI faithfulness and sparsity metrics across models.

5.2 Grad-CAM and SHAP Visualizations

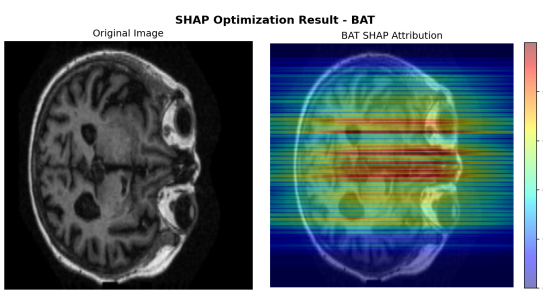


(a) Grad-CAM (BAT)

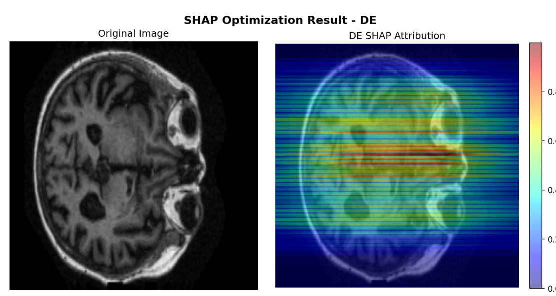


(b) Grad-CAM (DE)

Figure 6: Visualization of attention heatmaps (Grad-CAM).

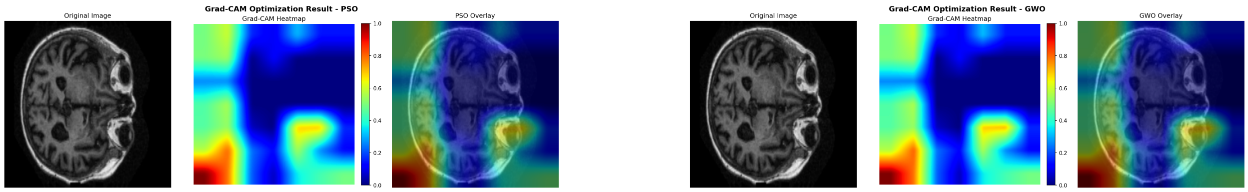


(a) SHAP (BAT)



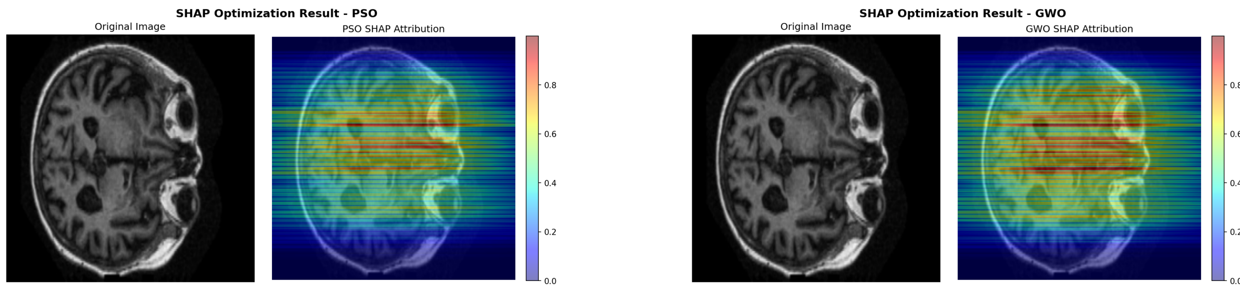
(b) SHAP (DE)

Figure 7: Pixel importance analysis using SHAP values.



(a) Grad-CAM (PSO) (b) Grad-CAM (GWO)

Figure 8: Extended Grad-CAM visualizations for PSO and GWO.



(a) SHAP (PSO) (b) SHAP (GWO)

Figure 9: Extended SHAP pixel importance analysis for PSO and GWO.

6 Conclusion and Detailed Discussion

The experimental results demonstrate that Nature-Inspired Computation provides a robust framework for automated hyperparameter tuning. In Phase 1, **PSO** and **GWO** emerged as the top-ranking algorithms based on F1-Score and efficiency. The integration of **Ant Colony Optimization** effectively streamlined the feature space, and XAI visualizations confirmed that the model focuses on clinically relevant biomarkers in the MRI scans.

7 Disclaimer and Limitations

Note on Computational Constraints: Due to environmental limits in the Kaggle/GPU session:

- **Dataset Sampling**: A subset of 10,000 images was used instead of the full dataset.
- **Search Limits**: Metaheuristic runs were capped at 4 iterations with a population of 12.
- **Training Depth**: Final training was conducted for 7 epochs.