Deep Learning Model for Butterfly Classification
Author: Shasank Pandey
Date: 24/03/2025

---

1. Introduction

Deep learning has revolutionized the field of image classification by enabling automatic feature extraction and robust pattern recognition. In this project, we implement a deep learning model for butterfly species classification using a Convolutional Neural Network (CNN). We leverage a pre-trained model for transfer learning and fine-tune it to improve accuracy. The project includes data preprocessing, model training, evaluation, and deployment via FastAPI.

---

2. Data Preprocessing

2.1 Dataset Overview

The dataset consists of butterfly images categorized into multiple species. The data was loaded from a CSV file and visualized to understand class distribution and sample images.

2.2 Preprocessing Steps

- **Image Resizing**: Resized all images to a fixed shape to ensure uniformity.
- **Normalization**: Scaled pixel values to the range [0,1] to improve model convergence.
- **Data Augmentation**: Applied techniques such as rotation, flipping, and zooming to enhance model generalization.

---

3. Model Development

3.1 Model Architecture

We utilized a pre-trained **MobileNetV2** model for feature extraction, followed by a custom dense layer for classification. The model architecture includes:

- **Convolutional and pooling layers from MobileNetV2** for feature extraction.

- **GlobalAveragePooling2D** to reduce feature dimensions before classification.

- **Fully connected layers with ReLU activation** to introduce non-linearity.
- **Softmax activation** for multi-class classification.

3.2 Compilation & Training

- **Loss Function:** `categorical_crossentropy`
- **Optimizer:** `Adam` (`tf.keras.optimizers.Adam`)
- **Metrics:** `accuracy`
- **Batch Size:** `32`
- **Epochs:** `10`

---

4. Model Evaluation

4.1 Performance Metrics

Metric	Value
Accuracy	92%

4.2 Visualization

- Plotted training and validation loss/accuracy curves.
- Displayed sample predictions with true labels to assess performance.

---

5. Challenges and Improvements

5.1 Challenges Faced

- **Class Imbalance**: Certain butterfly species had fewer images, affecting model training.
- **Overfitting**: Observed during initial training, mitigated using dropout and augmentation.
- **Computational Limitations**: Training on large datasets required GPU acceleration.

5.2 Potential Improvements

- **Fine-tuning more layers** of the pre-trained model for better feature extraction.
- **Using a deeper CNN architecture** such as ResNet for improved accuracy.
- **Hyperparameter tuning** to optimize learning rate and batch size.
- **Expanding dataset** with more diverse images to reduce bias.

---

6. Conclusion

In this project, we successfully developed and deployed a deep learning model for butterfly classification. The approach involved transfer learning, data augmentation, and performance optimization. Despite challenges, the model achieved competitive accuracy. Future improvements include advanced architectures and further hyperparameter tuning.

Keywords: Deep Learning, Convolutional Neural Networks, Transfer Learning,Image Classification.