Project Report: Butterfly Species Classifier

Team Details:

Team ID: LTVIP0225TMID42172\ **Team Members:**

- Vanipenta Vidya Reddy (Team Leader)
- Syed Mohammed Siddiq
- S.Supraja

---

1. Introduction

Butterflies are not only beautiful creatures but also indicators of a healthy ecosystem. Identifying their species manually is time-consuming and requires expertise. In this project, we present a deep learningbased solution using the VGG16 convolutional neural network architecture to classify butterfly species from images. The model is deployed via a Flask web application for user-friendly interaction.

---

2. Problem Statement

Manual identification of butterfly species can be error-prone and requires specialized knowledge. This project aims to automate the classification of butterfly species using a deep learning model and provide a simple web-based interface to perform this classification.

---

3. Objectives

- To develop a butterfly species classification model using the VGG16 architecture.
- To train and validate the model using a suitable dataset.
- To deploy the model in a Flask-based web application.
- To provide users with a simple interface for uploading butterfly images and viewing predictions.

---

4. Tools and Technologies

- **Programming Language:** Python
- **Libraries:** TensorFlow, Keras, NumPy, OpenCV, Matplotlib
- **Model:** VGG16
- **Web Framework:** Flask
- **Deployment:** Localhost / GitHub

---

5. Methodology

a. Data Collection

A dataset of butterfly images was collected with labeled species. The images were resized to 224x224 pixels to match the VGG16 input size.

b. Preprocessing

- Resizing images
- Normalization
- Splitting into train and test sets

c. Model Selection

The VGG16 model, pre-trained on ImageNet, was fine-tuned to classify butterfly species. The top layers were modified for our classification task.

d. Training

The model was trained on the dataset using the categorical crossentropy loss and the Adam optimizer. Various metrics such as accuracy and loss were tracked.

e. Evaluation

The model achieved high accuracy on the test set, indicating its effectiveness in classification.

---

6. Project Structure

```
project/
├── static/
│   └── style.css
├── templates/
│   ├── index.html
│   └── result.html
├── vgg16_butterfly_model.h5
├── app.py
├── README.md
└── requirements.txt
```

---

7. Web Application

The application is built using Flask. Users can upload a butterfly image via the UI, and the app displays the predicted species.

Key Features:

- Upload image feature
- Real-time classification
- Display of prediction result

---

8. Challenges Faced

- Large model file size (>100MB) not supported by GitHub directly
- Training time for deep learning model
- Dataset imbalance

---

9. Solutions Implemented

- Used Git Large File Storage (Git LFS) for model file
- Applied data augmentation to improve dataset diversity
- Tuned hyperparameters for better model performance

---

10. Future Enhancements

- Extend support to more butterfly species
- Deploy on cloud for public access
- Add species information alongside predictions

---

11. Conclusion

This project successfully demonstrates the use of deep learning and Flask to classify butterfly species from images. It bridges the gap between AI technology and biodiversity research, providing an accessible solution to a real-world challenge.

---

12. References

- Kaggle Butterfly Dataset
- TensorFlow Documentation
- Keras Applications
- GitHub Docs - Git LFS

---

13. Screenshots (if any)

(Include images of your UI, model training results, prediction output screenshots here in final documentation)

---

14. Appendix

Sample Code Snippet (Model Load)

```
from keras.models import load_model
model = load_model('vgg16_butterfly_model.h5')
```

Sample Code Snippet (Flask route)

```
@app.route('/', methods=['GET', 'POST'])
def index():
    if request.method == 'POST':
        file = request.files['file']
        # Process and predict
    return render_template('index.html')
```

Prepared by:

Vanipenta Vidya Reddy , Syed Mohammed Siddiq , S.Supraja

Team ID: LTVIP0225TMID42172