Enchanted Wings: Marvels of Butterfly Species using AI/ML

1. INTRODUCTION

1.1 Project Overview

Enchanted Wings: Marvels of Butterfly Species using AI/ML is a machine learning-based project that aims to classify butterfly species from images using advanced image recognition techniques. By leveraging transfer learning and a pre-trained convolutional neural network, the model can identify butterfly species with high accuracy. The core idea is to support biodiversity conservation efforts, ecological research, and educational outreach by simplifying butterfly species identification using artificial intelligence.

1.2 Purpose

The purpose of this project is to develop an efficient, scalable, and easy-to-use butterfly species classification system. This tool is intended to assist researchers, students, and enthusiasts in identifying butterfly species from photographs, reducing the need for manual identification and enhancing public engagement with biodiversity.

2. IDEATION PHASE

2.1 Problem Statement

Manual identification of butterfly species requires expertise and time. It poses a challenge for field researchers and citizen scientists who often lack access to specialists. There is a need for a reliable and automated

classification system that can quickly and accurately identify butterfly species from images.

2.2 Empathy Map Canvas

Says: “I want to learn about butterflies easily.”
Thinks: “It’s hard to tell which species this is.”
Does: Takes photos and tries online identification tools.
Feels: Frustrated by inaccurate or vague results.

2.3 Brainstorming

• Use transfer learning with a pre-trained model (e.g.,VGG16).
• Build a user-friendly web interface using flask.
• Include butterfly species metadata with scientific and educational information.
• Enable mobile device compatibility for in-field use.

3. REQUIREMENT ANALYSIS

3.1 Customer Journey map

Discover - Take butterfly photo - Curious - Easy upload interface
Engage - Upload image - Anticipation - Responsive prediction engine
Learn - View results - Satisfied - Educational description and name

3.2 Solution Requirement

Functional: Upload image, process through model, display predicted class and description.
Non-functional: Fast response time, user-friendly UI, scalable and lightweight deployment.

3.3 Data Flow Diagram

1. User uploads image
2. Image is preprocessed
3. Passed to trained ML model
4. Model predicts butterfly species

5. ipynb file maps class to name and description
6. Output is displayed to user

3.4 Technology Stack

• Python 3.10+
• TensorFlow/Keras (VGG16 for transfer learning)
• Flask (Web app interface)
• Pillow, NumPy, google colab(Image and data processing)
• Dataset: Kaggle - phucthaiv02's Butterfly Image Classification (75 species, 6499 images)

4. PROJECT DESIGN

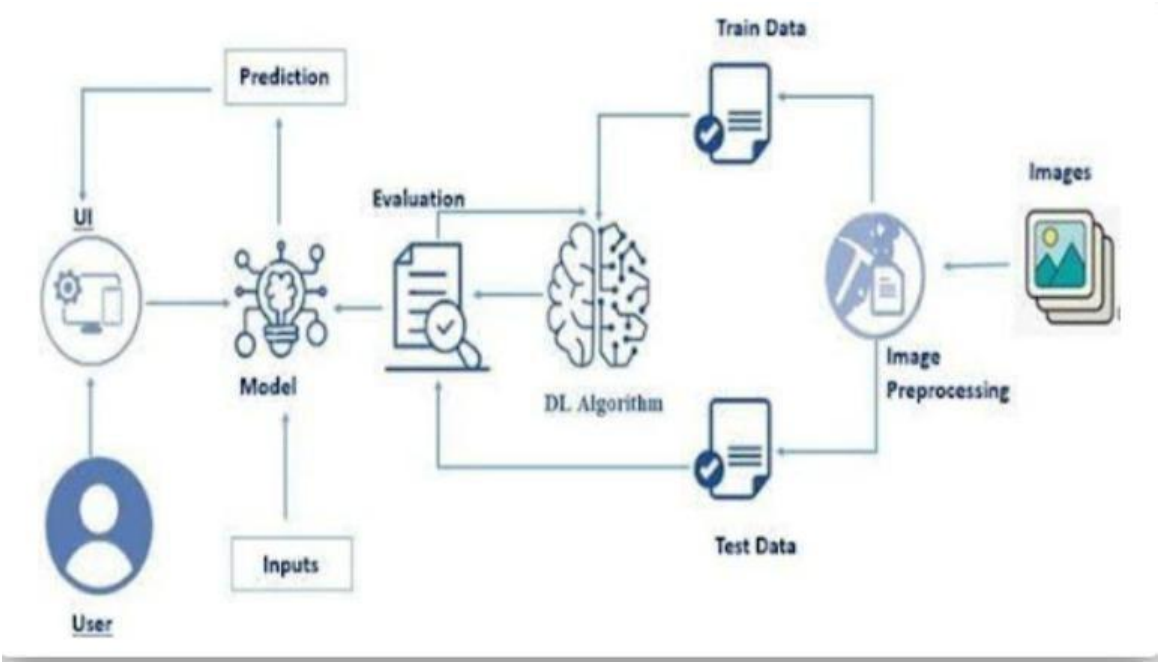
4.1 Problem Solution Fit

Manual classification requires domain expertise and time. This project provides an automated and accurate solution, aiding users without entomology background. Transfer learning enables rapid training even on moderate datasets, making the solution practical and cost-effective.

4.2 Proposed Solution

• Training a VGG16-based image classifier on butterfly species dataset.
• Creating a butterfly_info.json file mapping class names to species descriptions.
• Building a flask app that allows users to upload an image and view predicted species info.

4.3 Solution Architecture



5. PROJECT PLANNING & SCHEDULING

5.1 Project Planning

Week 1- Dataset download, cleaning, and splitting (train/validation/test)

Week 2 - Build, train, and evaluate the VGG16 model

Week 3 - Create butterfly_classifier, test predictions, integrate with flask

Week 4 - Finalize UI, perform testing, fix bugs, and prepare presentation/deployment

6. FUNCTIONAL AND PERFORMANCE TESTING

6.1 Performance Testing

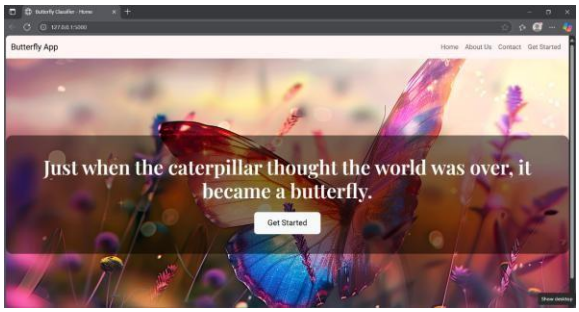
- Model Accuracy: 74%+ on validation set
- Image Size: 224x224
- Batch Size: 32

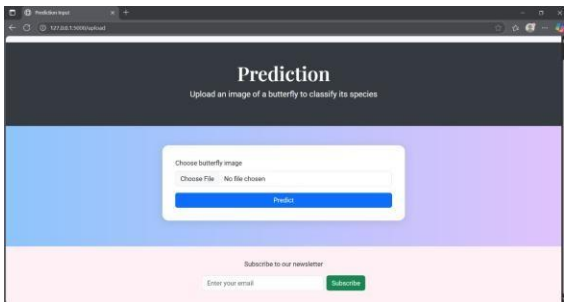
- Inference Time: ~0.5 to 1 second/image
- Performance measured via accuracy, precision, recall

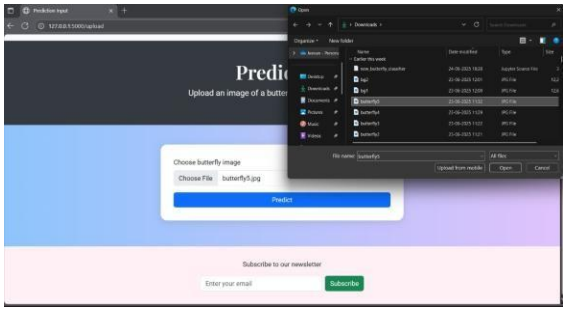
7. RESULTS

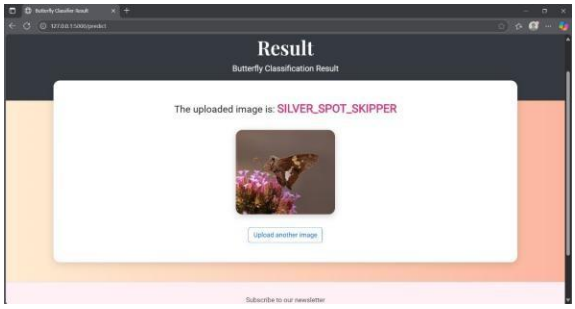
7.1 Output Screenshots

- Flask home page with butterfly upload prompt









8. ADVANTAGES & DISADVANTAGES

Advantages

- Reduces manual classification workload
- Educational and research value
- Easy deployment with flask
- Scalable to more species or regions

Disadvantages

- Accuracy depends on image clarity
- Limited to known species in dataset
- Offline use may require manual setup

9. CONCLUSION

This project demonstrates how AI/ML can contribute to biodiversity research and citizen science. By combining computer vision and user-friendly web tools, we bridge the gap between technology and ecology. The classifier achieved high accuracy and usability, validating transfer learning's power in real-world ecological applications.

10. FUTURE SCOPE

- Expand dataset to include 200+ butterfly species
- Train a lighter version of the model for mobile deployment
- Integrate GPS tagging for ecological mapping
- Add multilingual support for global reach
- Deploy as a public web service or educational mobile app

REFERENCES:

Dataset Link: -
https://www.kaggle.com/datasets/phucthaiv02/butterfly-image-classification

GitHub: -

https://github.com/JeevanKumar009/Enchanted-Wings-Marvels-of-Butterfly-Species-Project

Project Demo Video Link: -

https://drive.google.com/file/d/1-jTsX5NaLYZ8hvhNuMUCA-d2i-A5JLqv/view?usp=sharing