PAPILLECT: AI & Deep Learning-Driven Butterfly Detection & Classification

Anjali S, Assistant Professor, Department of Computer Science
St. Joseph's First Grade College Jayalakshmipuram, Mysore-570012

ABSTRACT

Butterflies play a vital role in maintaining ecological balance and serve as important indicators of biodiversity and environmental health. However, the manual identification and classification of butterfly species can be challenging due to their diverse patterns, similar appearances, variations in wing colors, and environmental conditions. To address these challenges, this research proposes PAPILLECT, an Artificial Intelligence (AI) and Deep Learning-driven system for automated butterfly detection and classification. The proposed system utilizes computer vision techniques and deep learning models to identify butterflies from digital images and classify them into their respective species. Image preprocessing techniques are applied to improve image quality and enhance relevant features, while convolutional neural networks (CNNs) are used to learn distinctive characteristics such as wing patterns, colors, shapes, and textures. The system aims to provide accurate, efficient, and consistent butterfly identification with reduced human intervention. Furthermore, PAPILLECT can support biodiversity monitoring, ecological research, educational activities, and butterfly conservation initiatives by enabling faster analysis of butterfly images. The performance of the proposed approach can be evaluated using metrics such as accuracy, precision, recall, and F1-score. By integrating AI, computer vision, and deep learning, PAPILLECT offers a promising approach to automated butterfly species recognition and contributes to the development of intelligent tools for biodiversity assessment and conservation.

Keywords: PAPILLECT, Artificial Intelligence, Deep Learning, Butterfly Detection, Species Classification, Computer Vision, Convolutional Neural Networks (CNN), Biodiversity Conservation.

1. INTRODUCTION

PAPILLECT: AI & Deep Learning Driven Butterfly Detection and Classification is an intelligent web-based system developed to automatically detect, classify, and analyze butterfly species from images using advanced artificial intelligence and deep learning techniques. The system simplifies butterfly identification, which traditionally requires expert knowledge and manual observation. Users upload images through a user-friendly interface, and a YOLO-based object detection model identifies butterfly regions within the image, supporting multiple butterflies in a single frame. Each detected butterfly is then extracted and processed using a ResNet18-based Convolutional Neural Network (CNN) to predict species along with confidence scores based on visual features such as wing patterns, colors, and textures.

To ensure reliability, the system incorporates a confidence-based filtering mechanism that categorizes predictions into different confidence levels, along with a smart detail control feature that adjusts the amount of biological information displayed. The system is integrated with a structured biological database providing details such as scientific name, habitat, lifespan, and food habits. Implemented as a Flask-based web application, Papillect also includes features like user authentication, butterfly name-based search, and personalized search history storage. Overall, the system provides a complete end-to-end solution combining detection, classification, and intelligent information delivery for applications in biodiversity conservation, ecological research, and education.

The system is integrated with a structured biological database that provides detailed attributes such as scientific name, habitat, lifespan, food habits, classification, and other ecological information. It is implemented using Flask as a web application, offering features such as user authentication, butterfly name-based search functionality, and personalized

search history storage. The system also includes a modern and interactive user interface for image upload and result visualization.

Overall, provides a complete end-to-end solution combining object detection, classification, intelligent decision-making, and user interaction. It is efficient, scalable, and suitable for applications in biodiversity conservation, ecological research, education, and environmental awareness.

LITERATURE SURVEY

1. Title: Insect Detection and Classification Based on an Improved Convolutional Neural Network

Authors: Denan Xia, Peng Chen, Bing Wang, Jun Zhang, Chengjun Xie
Year: 2018

This paper presents an improved CNN-based model for detecting and classifying insects in agricultural environments. The system uses a Region Proposal Network (RPN) to identify object regions and improve detection accuracy. It focuses on extracting deep features from images and classifying insects into multiple categories. The model performs well in complex backgrounds and improves classification performance compared to traditional methods. However, the system is designed for general insect detection and is not specifically focused on butterflies or user-based applications.

2. Title: Identification of Butterfly Species Using Machine Learning and Image Processing Techniques

Authors: Ayad Saad Almryad Year: 2020

This research explores butterfly species identification using both traditional machine learning and deep learning approaches combined with image processing techniques. The study is conducted in two stages: lab-based images and field-based images collected from real environments. Image preprocessing techniques such as noise removal and feature extraction were applied to improve model performance. Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNNs) were used for classification tasks. The system achieved high accuracy in controlled environments and demonstrated good performance with real-world datasets. The research highlights the importance of combining preprocessing techniques with learning models. However, the system mainly focuses on classification and lacks integration with real-time applications or user interfaces.

3. Title: A Novel Method for Classification of Butterfly Species Using CNN

Authors: Halit Çetiner
Year: 2023
This paper proposes a CNN-based method for automatic classification of butterfly species using a dataset of 832 images across 10 different species. The study emphasizes the importance of data augmentation techniques to improve model performance and handle class imbalance. The model was evaluated with and without augmentation to analyze its impact on accuracy. The proposed approach achieved a validation accuracy of approximately 93.41%, showing improved performance with augmented data. The system is designed to be lightweight and efficient compared to transfer learning models. The research highlights the effectiveness of CNN models in species classification tasks. However, it is limited to classification and does not include detection or interactive system features.

4. Title: Automated Classification of Butterfly Species in the Hyderabad Region Using Deep Learning

Authors: Mohammed Uzair Khan, Dr. Tasneem Jahan Year: 2025
This study focuses on classifying butterfly species in the Hyderabad region using deep learning techniques. A dataset was created by collecting images from sources such as Kaggle, iNaturalist, and other online repositories, with around

150–200 images per species. Data preprocessing and augmentation techniques were applied to improve model performance and reduce overfitting. The system uses CNN-based models to classify butterflies based on features like wing patterns and textures. The study highlights the importance of butterfly identification in biodiversity conservation. However, the system mainly focuses on classification and lacks advanced features such as object detection and intelligent system integration.

PROPOSED SYSTEM

PAPILLECT: AI & Deep Learning Driven Butterfly Detection and Classification, is an intelligent and user-friendly web-based application designed to overcome the limitations of existing systems. The system integrates YOLO-based object detection with a ResNet18-based Convolutional Neural Network (CNN) to accurately detect and classify butterfly species from images. Unlike traditional approaches, the system supports multi-object detection, enabling identification of multiple butterflies within a single image under real-world conditions.

The workflow begins with user image upload through a web interface, after which the YOLO model detects butterfly regions and generates bounding boxes. Each detected butterfly is cropped and passed to the classification model, which predicts the top species along with confidence scores. To improve reliability, a confidence-based filtering mechanism is implemented to categorize predictions into different confidence levels. Additionally, a smart detail control feature dynamically adjusts the amount of biological information displayed based on prediction confidence, enhancing user trust and system effectiveness.

The system is integrated with a structured biological database that provides detailed attributes such as scientific name, habitat, lifespan, food habits, classification, and other ecological information. It is implemented using Flask as a web application, offering features such as user authentication, butterfly name-based search functionality, and personalized search history storage. The system also includes a modern and interactive user interface for image upload and result visualization.

Overall, the proposed system provides a complete end-to-end solution combining object detection, classification, intelligent decision-making, and user interaction. It is efficient, scalable, and suitable for applications in biodiversity conservation, ecological research, education, and environmental awareness.

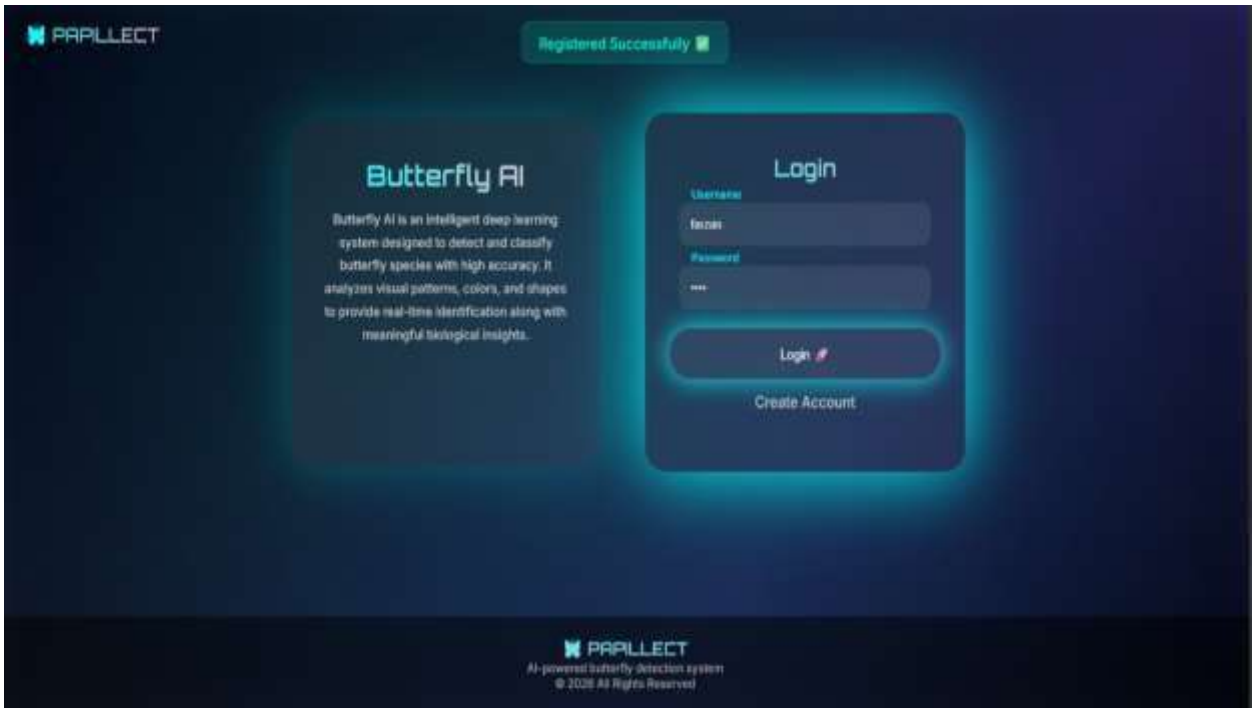
METHODOLOGY

1. Dataset Collection
2. Image Preprocessing
3. Dataset Annotation and Splitting
4. Deep Learning Model Training
5. Butterfly Detection and Classification
6. Model Evaluation
7. Prediction and Result Display

SNAPSHOTS:



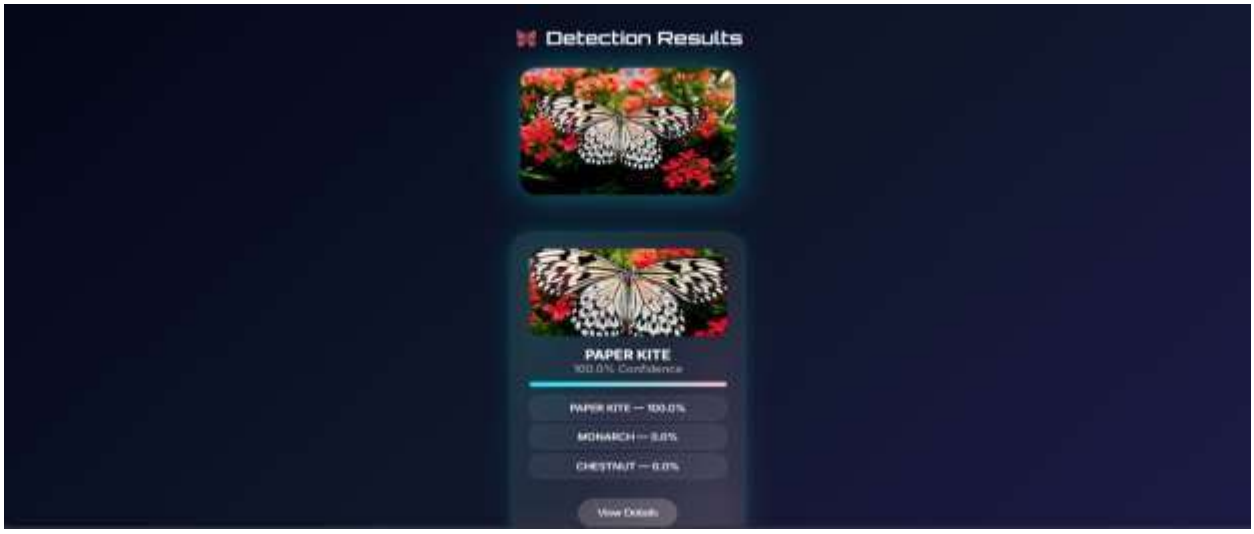
Welcome Page

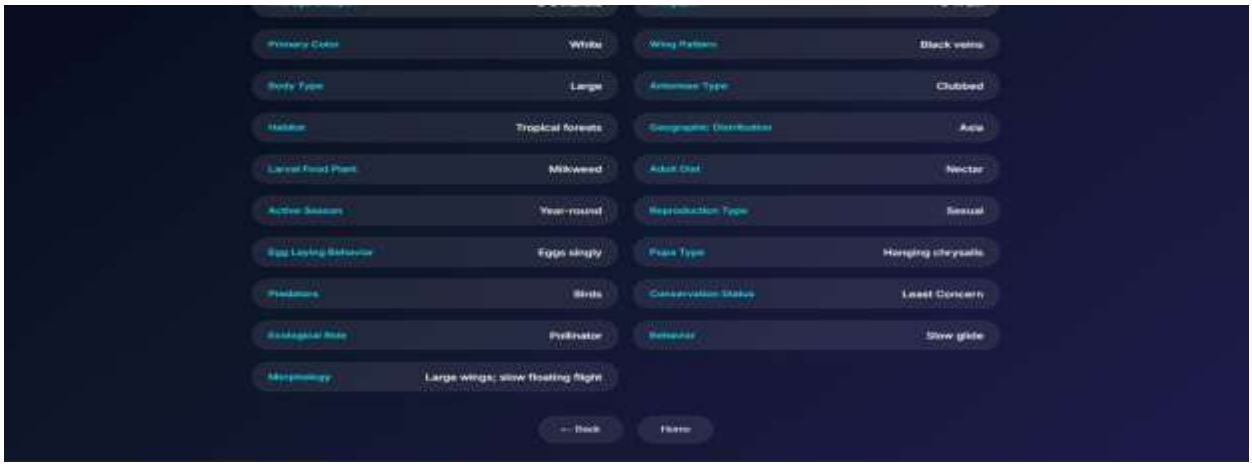


Login Page

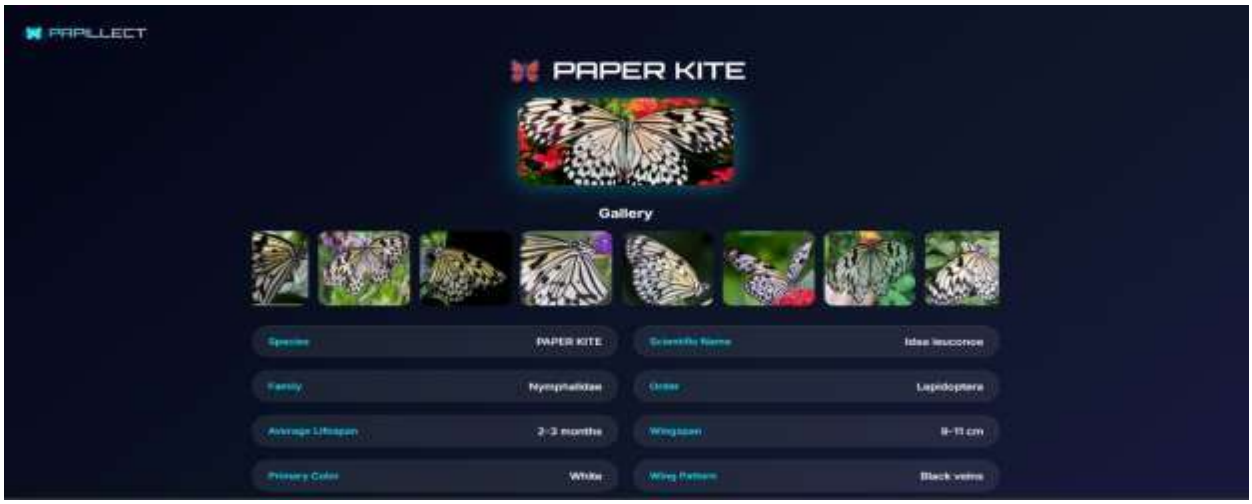


Index/Dashboard Page





Result Page



Details Page

CONCLUSION

The project “PAPILLECT: AI & Deep Learning-Driven Butterfly Detection and Classification System” was developed to automate butterfly detection and species identification using deep learning techniques. It utilizes YOLOv8 for object detection and CNN-based methods for feature extraction and classification, providing species names and confidence scores.

The system reduces manual effort, enables faster identification, and demonstrates the application of artificial intelligence in biodiversity monitoring and environmental conservation. Overall, PAPILLECT offers an efficient and scalable solution for automated butterfly detection and classification.

FUTURE ENHANCEMENT

Although the current system performs effectively, there is significant scope for further enhancements and improvements. One of the primary areas of future work is expanding the dataset to include a larger variety of butterfly species. This will improve the model’s accuracy and allow the system to handle more complex classification tasks.
The system can also be extended to support real-time detection using live video streams, enabling users to identify butterflies directly through a camera without uploading images manually. This will make the application more interactive and practical for real-world usage. Another important enhancement is the inclusion of detailed biological information such as habitat, lifespan, feeding habits, and conservation status for each butterfly species. This will transform the system from a detection tool into a comprehensive informational platform.
Additionally, the system can be developed into a mobile application, allowing users to access the features anytime

and anywhere. Integration with cloud computing services can further improve performance, scalability, and data management capabilities. Future improvements may also include enhancing the model using advanced techniques such as transfer learning, model optimization, and hybrid deep learning approaches to achieve even higher accuracy and faster processing speed.

Overall, the project provides a strong foundation for future developments and can be extended into a more advanced, scalable, and widely usable application in the field of artificial intelligence and environmental studies.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[2] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.

[4] G. Jocher, A. Chaurasia, and Ultralytics Team, "YOLOv8: Ultralytics Object Detection Model," 2023. [Online].