# Comparative Analysis of MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50 on a Butterfly Image Dataset

Shadman Ahmad Nafee
*Computer Science and Engineering*
*Brac University*
Dhaka, Bangladesh
shadman.ahmad.nafee@g.bracu.ac.bd

Jannatus Sakira Khondaker
*Computer Science and Engineering*
*Brac University*
Dhaka, Bangladesh
jannatus.sakira.khondaker@g.bracu.ac.bd

***Abstract*—One of the major challenges in computer vision and picture categorization is accurately identifying different classes. This study conducts an extensive comparative investigation of four well-known convolutional neural network (CNN) architectures—MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50—applied to a specific image dataset that consists of 75 different categories of butterflies. The dataset, which has been carefully curated with more than 1000 annotated images, including validation samples, is used to assess the performance of the models. Every image is allocated to a distinct butterfly category, highlighting a one-to-one correspondence. The objective of our inquiry is to evaluate the accuracy of classification, speed of training, and resilience of models in the specific context of recognizing butterfly species. Our goal is to use advanced CNNs to provide valuable insights that can improve the effectiveness of picture categorization systems, specifically in the field of natural biodiversity. The results of this study are positioned to provide valuable insights for the creation of sophisticated models that may be used in practical situations, such as monitoring the environment and protecting different species.**

## I. Introduction

The complex and eclectic realm of butterflies has always fascinated both scholars and lovers. The emergence of computer vision and deep learning has the potential to automate the identification process, offering great opportunities for scientific research and conservation initiatives. This study involves a comprehensive investigation of four prominent CNN architectures—MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50—implemented on a specific picture dataset consisting of 75 distinct butterfly categories. The choice of these Convolutional Neural Networks (CNNs) is based on their well-documented achievements in a wide range of image categorization assignments. Through the utilization of these models, our objective is to identify and evaluate the specific advantages and disadvantages they possess in the intricate domain of butterfly species identification. The dataset contains more than 1000 labeled images that have been carefully annotated. This provides a thorough evaluation of the models' skills, with a specific focus on accurately matching images to butterfly categories. This research aims to make a valuable contribution to the wider field of computer vision applications in biodiversity monitoring. The results are anticipated to provide insights into the most effective framework for categorizing butterflies, enabling progress in ecological studies, conservation efforts, and automated species identification systems. During this exploration, our objective is to uncover ideas that connect artificial intelligence with the intricate complexities of the natural world.

## II. Methodology

### A. Dataset Description:

The dataset consists of over 1000 images with labels of butterflies, covering 75 distinct categories. Every image is specifically assigned to only one butterfly category, guaranteeing a direct correspondence for classification tasks. Furthermore, a portion of tagged photos is kept aside specifically for validation.

### B. Model Architectures:

MobileNetV2, renowned for its compact architecture, is selected for its optimal utilization of resources. The architecture is equipped with pre-trained weights from ImageNet, which serves as a robust benchmark for comparison.

DenseNet161 is chosen for its densely connected blocks, which encourage feature reuse and improve learning capabilities. The model is initialized with pre-trained weights in order to utilize the knowledge acquired from a variety of datasets.

EfficientNetV2 is renowned for its ability to optimize across various dimensions, providing a well-balanced compromise between the size of the model and its performance. ImageNet pre-trained weights are used to accelerate convergence.

ResNet50 is a cutting-edge model that revolutionizes the concept of residual learning, offering advanced capabilities for deep neural networks. The utilization of pre-trained weights from ImageNet enhances the model's ability to capture complex characteristics throughout the training process.

### C. Data preprocessing:

Before training the model, the dataset goes through preprocessing stages, which involve scaling the images to the specific dimensions required by each model architecture (e.g., 224x224 for MobileNetV2 and ResNet50). Normalization is utilized to standardize the values of pixels, hence improving the convergence process during training.

### D. Training Process:

The optimization approach used to train all models is stochastic gradient descent (SGD), with a learning rate schedule that dynamically adjusts learning rates during training. The utilization of a categorical cross-entropy loss function is highly suitable for applications involving multi-class classification.

### E. Evaluation Metrics:

The performance of the models is evaluated using a comprehensive set of evaluation metrics. Accuracy refers to the degree of correctness in predictions made across all classes. Precision refers to the proportion of accurately predicted instances with the total number of expected examples for a specific class. The F1 Score is a metric that quantifies the model's performance by taking the harmonic mean of precision and recall. It provides a balanced evaluation of the model's accuracy. Area Under the ROC Curve (AUC-ROC) is a metric that measures the model's capacity to differentiate between classes, which is particularly important for datasets with uneven class distribution. ROC Curve is a visual depiction of the model's accuracy in identifying genuine positives compared to false positives at different threshold levels. These metrics provide a thorough assessment of each model's ability to classify butterfly species, ensuring a detailed grasp of their strengths and drawbacks. The utilized methodology is rigorous and seeks to provide precise insights into the comparative performance of MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50 on the provided butterfly dataset.

## III. Results

The experimental results provide valuable insights into the performance of each model on the butterfly dataset. MobileNetV2 demonstrated outstanding precision, attaining an impressive 99% accuracy in the classification of butterfly species. The weighted precision, recall, and F1 score achieved impressive results, with values of 0.8811, 0.8503, and 0.8523, respectively. MobileNetV2 demonstrated a successful equilibrium between precision and recall for all classes. DenseNet161 and EfficientNetV2 exhibited similar accuracies of 87%, indicating their ability to effectively handle a wide range of butterfly categories. However, ResNet50 has proven to be a leading model with an impressive accuracy of 99.27%. This demonstrates its ability to accurately classify data by capturing detailed elements. Nevertheless, the precision, recall, and F1 score, which were adjusted for weight, were marginally inferior for ResNet50. This suggests that there may be difficulties in accurately distinguishing specific classes. This analysis highlights the subtle advantages and disadvantages of each model. MobileNetV2 stands out for its high accuracy and well-balanced metrics. DenseNet161 and EfficientNetV2 consistently deliver reliable performances. On the other hand, ResNet50 demonstrates exceptional accuracy, although there is potential for improvement in precision and recall.

## IV. Discussion

The results provided from our comparison investigation shed light on the nuanced performance of MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50 in the field of butterfly species categorization. MobileNetV2 demonstrated outstanding accuracy, precision, recall, and F1 score, highlighting its ability to achieve a well-balanced performance across many classes. The impressive performance of DenseNet161 and EfficientNetV2, with an accuracy of 87%, highlights their ability to effectively handle the intricacies involved in butterfly categorization. ResNet50, while excelling in accuracy, demonstrated somewhat poorer weighted precision, recall, and F1 score, suggesting possible issues in fine-grained class differences.

The results have important consequences for practical use, especially in the fields of ecological monitoring and species preservation. MobileNetV2's exceptional precision makes it an excellent option for automated butterfly identification systems, particularly in the field of biodiversity studies. DenseNet161 and EfficientNetV2, though marginally less accurate, are nonetheless dependable alternatives that have wider utility. The remarkable precision of ResNet50, although subtle limitations, suggests its promise for particular applications where detailed categorization may be of lesser importance.

## V. Future Work

Based on the discoveries and suggestions made in this study, there are promising opportunities for improvement and creativity in future research on categorizing butterfly species. A possibility worth exploring is the incorporation of attention processes into current models, which would improve their capacity to concentrate on important aspects that are crucial for precise classification. Furthermore, conducting pretraining specifically on datasets that are closely related to the butterfly domain has the potential to result in models that are more cognizant of the context. It is important to address class imbalances to reduce biases and enhance model performance. This can be achieved by implementing advanced data augmentation techniques or oversampling tactics. Using automated machine learning (AutoML) methods for hyperparameter optimization provides an efficient method for identifying the best model configurations. The inclusion of spatial-temporal factors and interactive learning methods could improve the ecological validity and adaptive capacity of the models, making them more closely aligned with real-world situations. Investing in explainability and interpretability analyses guarantees the trustworthiness and comprehensibility of models, hence promoting transparency in automated butterfly identification systems. The comprehensive strategy for future research intends to develop the discipline by improving accuracy, generalization, and applicability, ultimately contributing to the progress of biodiversity monitoring and species conservation activities.

## VI. Limitations

Although our research on butterfly species classification has provided helpful insights, it is important to acknowledge numerous limitations. The models assessed in this work, specifically MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50, may not include all existing architectures, which could restrict the investigation of better alternatives. Furthermore, although our dataset is carefully selected and includes a wide range of butterfly categories, the number and distribution of examples in each category may lead to biases that can impact the models' ability to apply their knowledge to real-world situations. The study mostly examines static photographs and overlooks the potential advantages of adding the temporal components of butterfly behavior. Moreover, the extent to which these models can be applied to various ecosystems or geographical regions is yet uncertain. Another drawback arises from the lack of a comprehensive analysis regarding the computational resource demands of each model. Furthermore, comprehending the decision-making processes of complicated deep learning models, especially concerning ecological complexity, may be challenging due to the limited extent to which interpretability has been addressed. Recognizing these constraints is crucial for placing the results in a practical context and directing future research toward addressing these obstacles.

## VII. Conclusion

Our research concludes with a thorough assessment of MobileNetV2, DenseNet161, EfficientNetV2, and ResNet50 to classify butterfly species. MobileNetV2 stands out for its high accuracy and balanced metrics, while other models provide dependable alternatives. The results emphasize the significance of selecting models according to particular use cases and dataset features, emphasizing the requirement for a subtle approach in choosing models for tasks connected to biodiversity.