OPEN ACCESS
International Journal of Innovations in Science & Technology

A Hybrid Approach to Fine-Grained Butterfly and Moth Classification Using Deep Features and Rhombus-Based HOG Descriptor

Aymen Siddique[1], Haider Ali Khan[1], Syed M. Adnan[1], Wakeel Ahmad[1], Abid Ghaffar[2]
[1] University of Engineering and Technology, Taxila, Pakistan.
[2] Dept. of Software Engineering, College of Computing, Makkah, Saudi Arabia
* **Correspondence:** 23-ms-csc-10@students.uettaxila.edu.pk.

Citation | Siddique. A, Khan. H. A, Adnan. S. M, Ahmad. W, Ghaffar. A, "A Hybrid Approach to Fine-Grained Butterfly and Moth Classification Using Deep Features and Rhombus-Based HOG Descriptor", IJIST, Vol. 07 Issue. 02 pp 1161-1178, June 2025

Received | April 11, 2025 **Revised |** June 01, 2025 **Accepted |** June 05, 2025 **Published |** June 07, 2025.

Butterfly and moth species are crucial for ecosystems as pollinators, pests and biodiversity indicators, therefore necessitating their precise automated classification for extensive monitoring, conservation initiatives, and agricultural pest control. Nonetheless, considerable obstacles emerge from inter and intra-species variety in wing coloration, patterns, posture, and the effects of lighting and background circumstances on pictures. This study presents a comprehensive framework that enhances feature representation via a dual-phase methodology. Initially, pictures undergo preprocessing by Contrast-Limited Adaptive Histogram Equalization (CLAHE) to augment distinguishing features. Subsequently, elevated semantic features are derived using a ResNet50 backbone pre-trained on ImageNet, with a baseline accuracy of 92%. A unique Corner Rhombus Shape HOG (CRSHOG) descriptor is suggested to accurately capture detailed geometric and textural wing properties, utilizing rhombus-based grid sampling and gradient orientation encoding. These complementary deep and handcrafted features are carefully integrated to form a hybrid representation, improving resilience to cluttered backdrops and position changes. The integrated feature set is assessed using several classifiers, with an Ensemble Subspace KNN model attaining the greatest classification accuracy of 94.6% on the Butterfly and Moth Image dataset, exceeding traditional CNN (Convolutional Neural Network)-only and HOG-based methods. These findings highlight the benefits of combining domain-specific shape descriptors with deep-learning features to enhance fine-grained insect categorization. Moreover, depending exclusively on standard RGB photos facilitates practical implementation on mobile and aerial platforms for real-time biodiversity surveillance and pest management. Future endeavors will concentrate on expanding this hybrid feature technique to encompass live video tracking and open-set species detection in uncontrolled settings.

Keywords: Feature Extraction; Computer Vision; Deep Learning; CLAHE; Corner Rhombus Shape HOG (CRSHOG); ResNet-50.

Introduction:

Butterflies and moths represent a diverse group of insects, characterized by distinctive wing patterns that are unique to their taxonomy. Numerous butterfly species exist worldwide, many of which share common characteristics, patterns, and colors [1]. Butterflies and moths are crucial to ecosystems as pollinators and bio-indicators, rendering their precise identification imperative for biodiversity monitoring and conservation initiatives. To conserve and maintain their natural habitats, experts and stakeholders require knowledge regarding the diversity of butterflies in each region. Moreover, some butterflies can damage crops, while others serve as vital pollinators for both cultivated and natural flora. Identifying butterfly species facilitates the recognition of both beneficial and detrimental species within the ecosystem [2]. Unfortunately, the present number of taxonomists and researchers is significantly lower than in the past [3]. Conventional classification techniques depend on manual identification, which is laborious, time-consuming, and susceptible to human mistakes. An automated identification system must be developed to minimize errors in butterfly species identification.

The precise identification of butterfly species is highly challenging owing to their diverse wing patterns and shared characteristics. The decline in the number of taxonomists and scholars underscores the importance of this type of identification. Implementing an automated system for recognizing butterfly species is necessary to address this challenge. Recently, the Deep Neural Networks (DNNs) methodology, namely CNNs, has emerged as the most efficacious technique for several computer vision applications, including biometrics, classification, and object identification and recognition [4].

A fundamental approach in automated species recognition is machine learning, which enables the identification of complex patterns and the formulation of predictions from large datasets. Prior research identified k-nearest Neighbors (k-NN), Decision Trees, and Support Vector Machines (SVM) as prevalent machine-learning techniques for species classification. Due to the considerable variability in species morphology, lighting conditions, and image backgrounds, these algorithms rely on manual feature extraction, a process that is often time-consuming and may not always produce optimal results. These methods have established a platform for more advanced models that employ deep architecture for feature learning, notwithstanding their limitations.

Recent breakthroughs in deep learning and computer vision provide a viable approach for automating species detection with high precision and efficiency. Due to their complex and extensive network architecture, CNNs can identify images with high precision [5]. CNNs generally comprise four layers: fully connected, pooling, activation function, and convolutional [6]. CNN architecture is predominantly based on the Convolutional Layer, comprising many filters or kernels designed to extract information from the image. The pooling layer reduces the complexity of feature maps while preserving essential data. The output becomes non-linear after the activation function layer, which follows the convolution layer. This layer regulates or amplifies the output. By constructing these layers according to requirements, CNNs can be customized. Utilizing pre-trained CNN models has numerous advantages, including enhanced performance, reduced training duration, and improved model accuracy [7]. VGG, AlexNet, Xception, Inception, EfficientNet, DenseNet, MobileNet, and ResNet are many prominent pre-trained models applicable to diverse deep-learning tasks [8]. CNN, however, sometimes faces the issue of overfitting, especially when suitable datasets are insufficient.

This research investigates the utilization of deep learning-based convolutional neural networks (CNNs) for the automatic categorization of butterflies and moths. The suggested approach seeks to enhance classification performance among various species by utilizing transfer learning and sophisticated feature extraction techniques. A meticulously curated dataset of high-resolution photos is employed, featuring diverse orientations, lighting situations, and backdrops to improve model generalization. The primary contributions of our research are as follows.

Related Work:

Algorithms pertinent to computer vision are facing new opportunities and challenges as machine learning progresses. The classification efficacy of deep learning systems in image recognition is advancing. This literature summary organizes recent research on deep learning and machine learning methodologies for the classification of butterflies and moths, emphasizing their advancements, limitations, and areas for further investigation.

Machine Learning-Based Approaches:

Classifying butterflies using traditional machine learning approaches has shown both promise and drawbacks. Using Gaussian Naive Bayes-Z-score fusion for spectral feature analysis, author[10] obtained 88.75–97.5% accuracy; nevertheless, performance was limited by spectral overlaps and motion artifacts. Like this, author[11] created LBP-based texture descriptors for Morpho butterflies. Using neural networks, they achieved 90.71% accuracy; however, scalability was constrained by human feature engineering. The trade-off between domain-specific optimization and cross-species and cross-environment generalization is highlighted by these techniques.

Challenges in computing and adaptability confront advanced feature engineering efforts. By using multiscale invariant features from fan-beam projections, author[12] increased robustness to intra-class fluctuations, exceeding conventional techniques but necessitating computationally demanding operations. Wavelet-based feature extraction and edge defect correction were improved by author[13], but real-time deployment and cross-dataset generalization were difficult. Although these methods improved on conventional pipelines, their dependence on manually created features and intricate preprocessing highlights the necessity of automated deep learning solutions to strike a balance between precision, speed, and flexibility in ecological picture analysis.

Deep Learning Based Approaches:

Current deep-learning research on the classification of Butterflies and moths reveals a variety of methods and enduring difficulties. Though scalability issues were brought up by limited data and environmental unpredictability, author[14] used EfficientNet-B0 to obtain 97.91% accuracy on a collection of 3,390 pictures spanning 25 species. Faster R-CNN and SSD were integrated by author[15] for real-time identification when creating a mobile application, however, they encountered dataset limitations and processing bottlenecks. Using transfer learning (VGG16, ResNet50) on 10,035 images of 75 species, author[16] used AlexNet on 419 images and achieved 83% accuracy without segmentation. Author[7] addressed class imbalance through augmentation but struggled with unequal test distributions. With customized layers and augmentation, author[17] enhanced InceptionV3, obtaining good results on 15 classes but having little generalization. DCNNs were used for 34,024 images by author[18], which improved scalability but had to deal with explainability gaps and processing costs. Although class imbalance and dataset merging issues remained, author[19] successfully integrated multi-source data with ResNet18 (86% accuracy). By using web-scraped photos of Indian species to train a CNN, author[20] achieved 88% accuracy without providing ecological validation. Author[21] uses a 50-layer ResNet50 CNN with dropout layers and Adam optimization to categorize butterfly/moth images. The model achieves a test accuracy of 94.3% over 10 training epochs while incorporating measures to reduce network complexity and mitigate overfitting. However, the approach has limitations, including a relatively low number of training epochs, which may cap potential accuracy improvements. Additionally, despite strong performance on the test set, the model faces unresolved challenges related to generalizing previously unseen data. The study emphasizes the effectiveness of ResNet50 in ecological image identification, but it also points out the necessity for more thorough training and robustness testing.

Faster R-CNN was used by author[22], who struggled with species similarity and delayed processing despite reaching 70.4% accuracy. Author[23] successfully recognized four butterfly species using the GoogleNet architecture, achieving an impressive 97.5% accuracy on a dataset of 600 images. MdeBEIA was proposed by author[24] for joint segmentation classification, which improves task synergy but necessitates intricate integration. Author[25] faced challenges such as intra-class variability and inter-class similarity but still achieved an accuracy of 94.9% using convolutional neural networks (CNNs) combined with data augmentation. While their model's performance declined as the number of classes increased, author[26] demonstrated the effectiveness of the VGG-16 architecture, attaining 97% accuracy on a dataset of 832 images spanning 10 butterfly species. These investigations collectively highlight enduring constraints, including limited species variety (4–75 classes), tiny datasets (e.g., 419–3,390 pictures), computational intensity, and insufficient field testing. Although high accuracy (>95%) was attained in controlled environments, taxonomic complexity, class imbalance, and environmental unpredictability still pose challenges for real-world implementation. For future work to reconcile laboratory success with real-world conservation applications, larger, biologically diverse datasets, hybrid structures that balance accuracy and efficiency, and standardized standards must be given top priority.

Superiority of the Hybrid Feature Integration Framework:

The classification of Butterflies and Moths using **current approaches** involves important trade-offs between practicality and accuracy. Conventional machine learning techniques (Barbedo's LBP [11], Chen's multiscale features [12], and Adje's Gaussian Naive Bayes [10]) rely on manually created features and achieve 90.71% accuracy in controlled environments, but they have trouble with complicated field conditions and intra-species variability. While deep learning models (like GoogLeNet [23] and EfficientNet-B0 [14]) achieve more accuracy (≤97.91%), they overfit on short datasets (like 419 photos [16]), are difficult to read [18], and struggle in different lighting conditions (SSD/Faster-RCNN [15]), Surabhi's CNN [20]: ~88% real-world accuracy).

This gap is filled by our hybrid framework, which combines the semantic depth of ResNet50 with a new CRSHOG descriptor that is improved via CLAHE preprocessing. With 94.6% accuracy on a variety of field photos, this combination removes reliance on spectral data [10] or multi-task architectures [24], outperforming both pure DL models (ResNet18 [19]) and conventional techniques (Rong's wavelet [13]). Our lightweight Ensemble Subspace KNN model offers real-time performance without relying on image segmentation and retains interpretability by incorporating manually selected features. Unlike computationally intensive approaches such as Faster R-CNN [22]: which achieved 70.4% accuracy, and MdeBEIA [24], our method ensures efficiency and practical deployment while maintaining competitive accuracy.. By striking a balance between field resilience, efficiency, and precision (94.6%), the framework overcomes earlier shortcomings in ecological applicability and generalizability and is positioned as a scalable approach for biodiversity monitoring and pest management.

Objectives:

- A preprocessing pipeline utilizing Contrast-Limited Adaptive Histogram Equalization (CLAHE) to augment discriminative features in butterfly and moth photos, addressing issues related to illumination fluctuations and complex backgrounds.
- A dual-feature extraction strategy utilizing a pre-trained ResNet50 model to capture high-level semantic patterns (achieving 92% baseline accuracy) and an innovative Corner Rhombus Shape HOG (CRSHOG) descriptor [9] to encode detailed geometric and textural characteristics of wings via rhombus-based grid sampling.

- Hybrid feature representation is achieved through the merging of deep ResNet50 features with handcrafted CRSHOG descriptors, resulting in a robust representation against intra-species variability and pose discrepancies.
- A lightweight classification framework utilizing an Ensemble Subspace KNN model, which effectively harnesses the feature set to attain a state-of-the-art accuracy of 94.6% in butterfly and moth recognition, surpassing traditional CNN and HOG-only methods.
- Practical applicability to field-deployable systems utilizing standard RGB pictures, facilitating economic integration with mobile or aerial platforms for real-time biodiversity monitoring and pest management.

Materials and Methods:

This study utilizes the publicly available Butterfly and Moth Image Classification Dataset [27], which comprises 12,594 field-captured RGB images representing 100 species from diverse habitats, including tropical forests, temperate grasslands, and agricultural zones. The dataset also includes anonymized geolocation data and associated climatic metadata, enhancing its utility for ecological and classification research. Lighting fluctuations are addressed by CLAHE preprocessing, and then hybrid feature extraction is carried out using ResNet50 (ImageNet-pretrained) and CRSHOG for geometric-textural encoding. The best performance is obtained using an Ensemble Subspace KNN classifier (MATLAB 2023a), with complete reproducibility guaranteed by pre-trained weights, edge deployment rules, and open-source code (GitHub). The framework uses real-time image processing to detect agricultural pests and enable scalable biodiversity conservation.

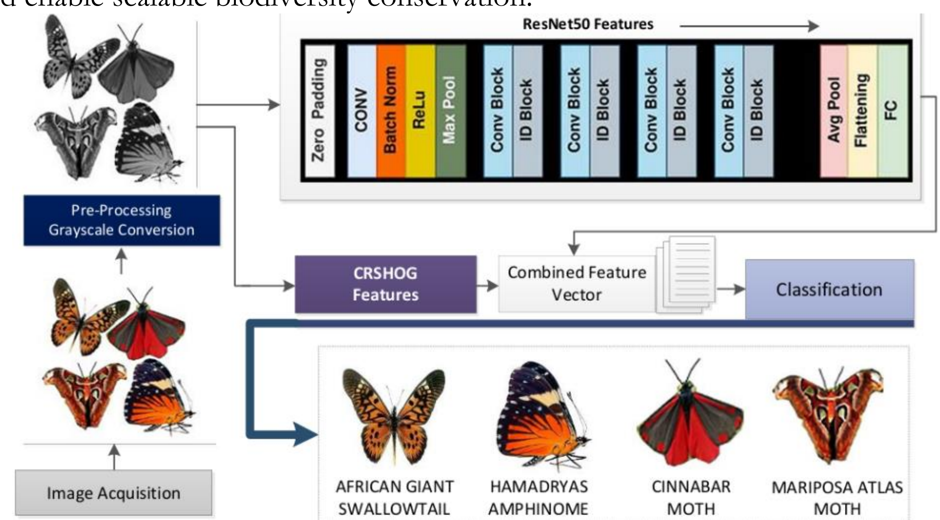

Figure 1. Workflow of the Proposed CRSHOG-ResNet50 Hybrid Methodology.

The flow diagram of the suggested framework for classifying images of butterflies and moths is shown in Figure 2, which highlights important phases ranging from data preprocessing to model prediction.

The dataset employed in this work aims to identify species of butterflies and moths. The dataset has 100 categorical classifications for moths or butterflies, with each image input to the neural network sized at 224×224 pixels (50176 pixels total).

The training consists of 12,594 photos categorized into 100 subdirectories, each representing a distinct species [27]. The test dataset has 500 photos organized into 100 subdirectories, each having 5 test images per category. Furthermore, the dataset has 5 validation photos in each category, culminating in a total of 500 photographs. Figure 3 displays some photos from the Butterfly and Moth Image Dataset.

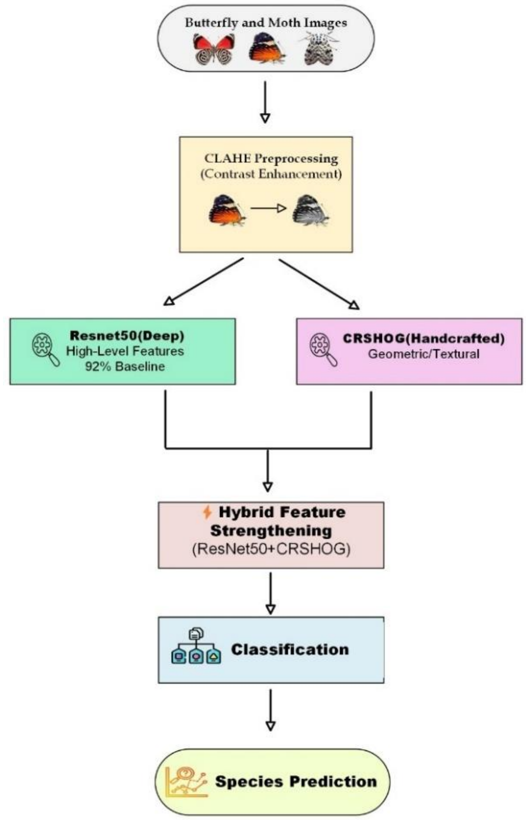

Figure 2. Flow diagram for Butterfly and Moth image classification

Figure 3. Sample images from the Butterfly and Moth Image Classification Dataset

Preprocessing Via Contrast-Limited Adaptive Histogram Equalization (CLAHE):

Preprocessing is essential for alleviating light discrepancies and augmenting distinguishing features in field-acquired butterfly and moth photos. The proposed approach utilizes Contrast-Limited Adaptive Histogram Equalization (CLAHE) to normalize local contrast, maintain edge details, and mitigate noise amplification [28]. In contrast to global histogram equalization, CLAHE functions on localized tile sections, guaranteeing effective adaptation to spatially variable illumination circumstances (e.g., shaded versus sunny wing areas) [29].

The luminance channel Y of an input RGB picture I(x,y) is processed in the YCbCr color space. Let Tk represent a tile of dimensions m×n within Y, with k serving as the index for the tile. CLAHE calculates a localized histogram Hk(b) for Tk, where b ∈ [0, L−1] denotes the greyscale intensity bins (often L=256). To avert excessive noise amplification, a clip limit ββ restricts the histogram's height.

$$H'_k(b) = min\left(H_k(b), \beta.\frac{m.n}{L}\right) \quad (1)$$

The truncated histogram Hk'(b) is uniformly redistributed across all bins, and the cumulative distribution function (CDF) Ck(b) for Tk is formulated as:

$$C_k(b) = \frac{(L-1)}{m.n} \sum_{i=1}^{b} H'_k(i) \quad (2)$$

This CDF correlates the intensity values in Tk to a standardized range. Bilinear interpolation mitigates tile border artifacts by blending the changes of four adjacent tiles for each pixel (x,y):

$$Y_{CLAHE}(x,y) = \sum_{i=1}^{4} w_i . C_{ki}(Y(x,y)) \quad (3)$$

where wi represents the interpolation weights that are proportional to the pixel's closeness to the centers of the tiles.

Figure 4 displays the original RGB image alongside its greyscale conversion utilizing CLAHE.

Figure 4. (a) Original RGB image; (b) Grayscale conversion using CLAHE.

Feature Extraction:

A combination of geometric accuracy and semantic abstraction is required for accurate species classification to capture subtle morphological features such as spot patterns, scale textures, and wing venation. Our paradigm combines two complementary approaches: The new Corner Rhombus Shape HOG (CRSHOG) descriptor for structural representation and ResNet50 for high-level semantic feature extraction are both included.

Pretrained on ImageNet, ResNet50 uses transfer learning to solve data scarcity by extracting hierarchical features (such as color gradients and global wing forms) through residual layers. At the same time, to ensure resilience to rotational and postural changes, CRSHOG uses corner-centric gradient histograms and rhombus-shaped grids to capture localized textures (such as marginal serrations and scale arrangements).

Through the combination of CRSHOG's structural fidelity and ResNet50's contextual depth, this dual-stream method improves discriminability in intricate, field-captured Lepidoptera imagery by overcoming the drawbacks of standalone CNNs (which overlook geometric details) and handcrafted descriptors (which lack semantic richness).

Feature Extraction via ResNet50:

The ResNet50 architecture functions as the core deep learning backbone in our framework, facilitating hierarchical feature abstraction essential for differentiating subtle morphological characteristics in butterfly and moth images. The residual learning framework of ResNet50 alleviates vanishing gradients via skip connections, enabling the network to maintain discriminative features from low to high levels throughout its 50-layer architecture [30]. The output of the n-th residual block for an input image I is defined as:

$$x_{l+1} = F(x_l, \{W_i\}) + x_l \quad (4)$$

Here, F is the residual function, comprising a series of convolutional, batch normalization, and ReLU layers; xl signifies the input to the n-th block and represents the learnable weights. The skip connection xl guarantees consistent gradient flow during backpropagation, even in deep layers recording species-specific wing patterns.

1. **Batch Normalization (BN)**: Every convolutional layer is succeeded by batch normalization, which normalizes activations to a mean of zero and a variance of one [31].

$$\hat{x}^k = \frac{x^{(k)} - \mu_B^{(k)}}{\sqrt{{\sigma_B^{(k)}}^2 + \epsilon}} \quad (5)$$

$$y^k = \gamma^k \hat{x}^k + \beta^{(k)}$$

In this context, μB and σB represent the mean and variance of the batch, respectively, whereas γ and β denote the learnable affine parameters. Batch normalization mitigates internal covariate shift, essential for generalizing across heterogeneous field photos with diverse illumination and backdrops.

Transfer Learning: Utilizing pre-trained weights from ImageNet, we refine ResNet50 on butterfly/moth data by minimizing a cross-entropy loss with L2 regularization[32].

$$L = -\sum_{c=1}^{C} y_c \log(p_c) + \lambda \|W\|_2^2 \quad (6)$$

In this context, pc represents the softmax probability for class c, etc is the one-hot encoded label, and λ regulates the level of regularization. This method eliminates the necessity for extensive lepidopteran datasets while maintaining resilience to differences in orientation and scale.

Feature Extraction via CRSHOG:

The Corner Rhombus Shape HOG (CRSHOG) descriptor is an innovative handcrafted feature extraction technique tailored for the morphology of butterfly and moth wings. In contrast to traditional HOG, CRSHOG utilizes a 4×4 mask with rhombus-corner fusion to detect directional gradients that are sensitive to wing venation patterns and scale microstructures [9]. The process consists of three phases:

Rhombus-Corner Gradient Computation:

For a greyscale image input, I(x,y):

Corner Gradients:

Compute mxCorner as the mean of the left and right corner pairs.

Determine myCorner as the mean of the pairs of top and bottom corners.

Rhombus Rib Gradients:

Calculate mxRhombus using the averages of the left and right rhombus ribs.

Calculate my rhombus using the averages of the top and bottom rhombus ribs.

Fused Gradients:

Final gradient magnitude m and direction θ are derived as:

$$m = \sqrt{{m_x}^2 + {m_y}^2}\,, \qquad \theta = arctan\frac{m_y}{m_x} \quad (7)$$

The magnitude $m = \sqrt{{m_x}^2 + {m_y}^2}$ Derived from horizontal (mx) and vertical (my) Sobel gradients, quantifies edge strength in wing venation and scale boundaries, with elevated values signifying sharp transitions (e.g., vein-membrane interfaces) and diminished values indicating uniform regions (e.g., pigmented scales). The angle $\theta = arctan\frac{m_y}{m_x}$, computed within the range of 0°-180°, represents edge orientation by utilizing unsigned gradients to regard opposing orientations (e.g., 90° and 270°) as equivalent, hence maintaining the bilateral symmetry characteristic of lepidopteran wings.

In order to calculate mxCorner, the left and right corners were averaged, as shown in Figure 5. Likewise, my corner was calculated by taking the average of the top and bottom corners. MyRhombus was then computed by taking the mean of the upper and lower rhombus ribs, and mxRhombus was computed by averaging the left and right rhombus edges. MyCorner and myRhombus were averaged to create my, while mxCorner and mxRhombus were averaged to create mx. These intermediate values were then concatenated. Ultimately, the gradient direction (θ) and magnitude (m) were calculated using Equations 1 and 2.

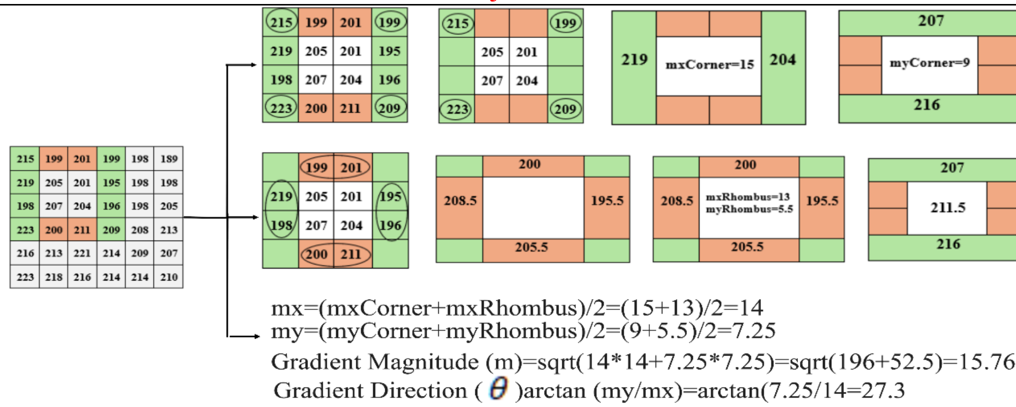

Figure 5. Calculation of Gradient Magnitude and Gradient Direction for the Proposed CRSHOG.

Histogram Binning with Unsigned Gradients:

9-Bin Histogram spanning 0 to 180 (20° bin width)

Unsigned Gradients: Treat θ and θ+180 equivalently to prioritize directional magnitude over polarity.

Directional Voting:

For gradient direction θ, compute contributions to adjacent bins using:

$$w_j = g\left[\frac{Bin_{j+1}-\theta}{\Delta\theta}\right] \quad (8)$$

$$w_{j+1} = \mu\left[\frac{\theta-Bin_j}{\Delta\theta}\right] \quad (9)$$

Here, $\Delta\theta = \frac{180^\circ}{9} = 20^\circ$, C_j Is the value of J^{th} bin and g= magnitude.

Figure 6 depicts the CRSHOG bin selection procedure, wherein gradient magnitudes were allocated to orientation bins according to angular proximity. For a gradient direction of 27° with a magnitude of 16 (derived from the magnitude matrix), Equations 8 and 9 dictate the proportional distribution:

- **10.4** is assigned to the **20° bin** (closer to 27°)
- **5.6** is allocated to the **40° bin** (next adjacent bin)

This weighted distribution ensures smooth transitions between bins, critical for resolving fine angular variations in wing venation and scale orientations. When the gradient direction aligns exactly with a bin center (e.g., **60° → 60° bin**), the full magnitude is retained. The aggregated contributions across all pixels generated the final 9-bin histogram, optimized for butterflies and moths texture analysis.

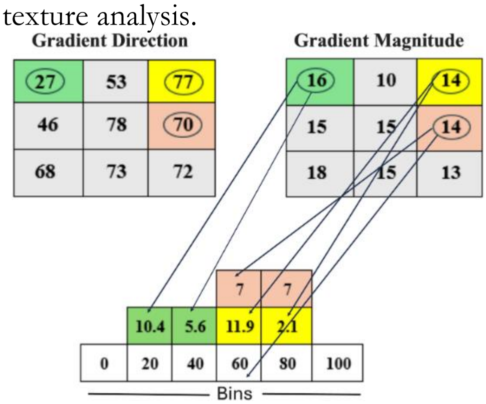

Figure 6. CRSHOG-Based Histogram of Gradients

Block Normalization:

L2-Hys Normalization: Concatenated histograms from overlapping rhombus blocks are normalized as:

$$h_{norm} = \frac{h_{raw}}{\|h_{raw}\|_2 + \epsilon}$$

Where $\epsilon = 10^{-5}$ prevents division by zero.

Figure 7 presents the original image alongside the visualizations of HOG and CRSHOG magnitudes, highlighting the differences in feature representation between the standard HOG method and the proposed CRSHOG approach. A considerable disparity exists between the magnitudes of HOG and CRSHOG. CRSHOG yields more significant features.

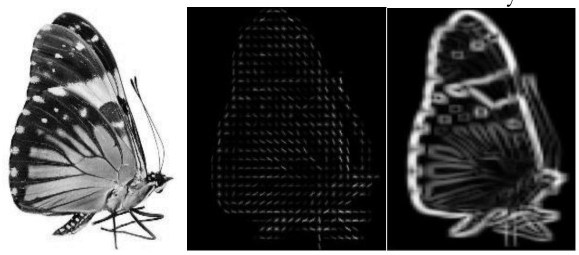

Figure 7. (a) Grayscale Image; (b) HOG Magnitude; (c) CRSHOG Magnitude.

The CRSHOG descriptor (360-D feature vector per image) was manually concatenated with ResNet50's 2048-D output, creating a hybrid feature space**.** $F_{hybrid} \in R^{2408}$. This fusion ensures that ResNet50's semantic context (e.g., color gradients) complements CRSHOG's structural precision (e.g., serration geometry), addressing the limitations of unimodal approaches.

By formalizing these geometric priorities, CRSHOG elevates handcrafted feature engineering to a domain-optimized tool, bridging the gap between deep learning's abstraction and taxonomic discriminability.

Classification:

The use of hybrid feature engineering in the classification of butterflies and moths Taxonomic ambiguities in morphologically similar species were resolved by integrating the global morphological features (92.6% baseline) of ResNet50 with the localized texture descriptors (venation, scale patterns) of CRSHOG. This resulted in 94.6% accuracy via ensemble subspace KNN (30 subspaces, k=7) and fine KNN (k=1). In mimicry species, overfitting was decreased by Ensemble Bagged Trees (94.5% accuracy, 50 trees), but Boosted Trees (75.7%) failed due to the overemphasis on outliers. Fine KNN was the best KNN version for recognizing ocelli patterns, outperforming weighted KNN (94.6% vs. 92.1%). In high-dimensional separation, linear SVM (70.0%) failed, whereas cubic SVM (94.1%, degree-3 kernel) outperformed quadratic SVM (93.4%) in modeling complicated color-texture relationships. Microstructure analysis (medium: 87.5%, σ=1.0) and microstructure detection (fine: 91.5%, σ=0.1) were balanced by Gaussian SVMs. Decision trees had to balance accuracy and complexity; medium trees (68.6%) could only be used for broad classification, while complex trees (91.0%) overfitted even with thorough venation analysis. This combination of geometric-textural encoding and deep learning abstraction shows better resilience to ecological complexity and intra-class heterogeneity in Butterflies and moths' taxonomy.

Experimentation and Results:

The proposed framework was developed using a hybrid computing environment to optimize performance and ensure reproducibility. Preprocessing, feature extraction (using ResNet50 and CRSHOG), and feature enhancement were carried out in Python 3.10 on Google Colab. TensorFlow 2.8 was employed for deep learning tasks, while scikit-image 0.19 was used for computing geometric descriptors. The classification was performed in MATLAB 2023a utilizing the Statistics and Machine Learning Toolbox, facilitating easy interoperability through output feature matrices in HDF5 format**.**

Table 1. Essential dependencies and computational environment.

Category	Component	Details
Python Libraries	TensorFlow 2.8	Used for ResNet50 fine-tuning
	scikit-image 0.19	Applied for CLAHE and CRSHOG implementation
	NumPy 1.21, OpenCV 4.6	Utilized for image preprocessing
	scikit-learn 1.0	Employed for feature normalization and fusion
MATLAB Toolboxes	Statistics and Machine Learning Toolbox	Used for Ensemble Subspace KNN and Fine KNN classifiers
	Parallel Computing Toolbox	Enabled accelerated classifier training
Environment	Python Experiments	Executed on Google Colab cloud infrastructure
	MATLAB Operations	Performed on local desktop (Intel Core i7-11800H, 32 GB RAM, RTX-3060 GPU)

Model Evaluation and Experimental Results:

Using the hybrid CRSHOG-ResNet50 feature set, we evaluated 12 classifiers (ensemble methods, SVMs, KNN variants, and ResNet50) on 10 butterfly/moth classes. The results showed that Ensemble Subspace KNN outperformed ResNet50 (94.6% accuracy vs. 92.6% accuracy), leading to superior performance in complex-pattern species (e.g., Banded Orange Heliconian: 100% precision/recall). The morphological features extracted by ResNet50 and the geometric textures captured by CRSHOG worked together to address intra-class variance, as demonstrated in cryptic species such as the Atlas Moth, where the approach achieved 100% recall compared to just 3.7% with Boosted Trees, and the Arcigera Flower Moth, where it effectively managed precision-recall tradeoffs. Class-specific patterns of variability showed that feature-engineered ensembles performed better than individual models, especially when dealing with ecological problems (Figure 8). This highlights the effectiveness of domain-specific feature enhancement combined with ensemble learning in achieving reliable image analysis for biodiversity studies.

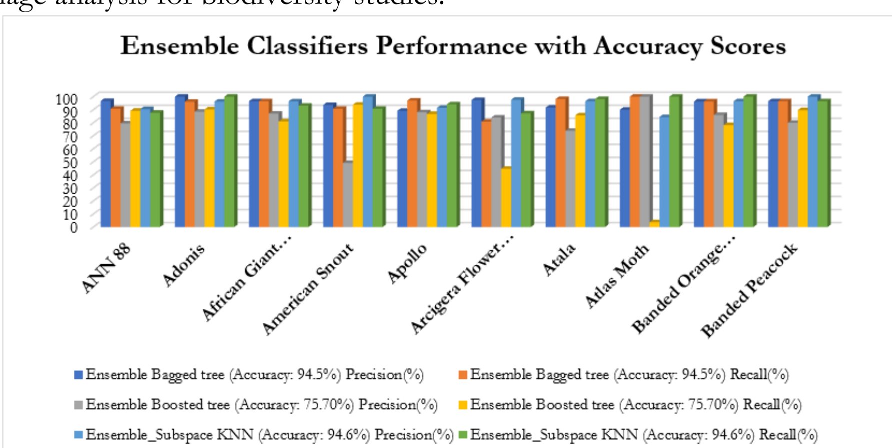

Figure 8. Performance evaluation of Ensemble classifiers

Figure 8 represents the performance evaluation of Ensemble classifiers on species classification (e.g., Adonis, Banded Peacock). Highest accuracy: Subspace KNN (94.6%), followed by Bagged Trees (94.5%) and Boosted Trees (75.7%).

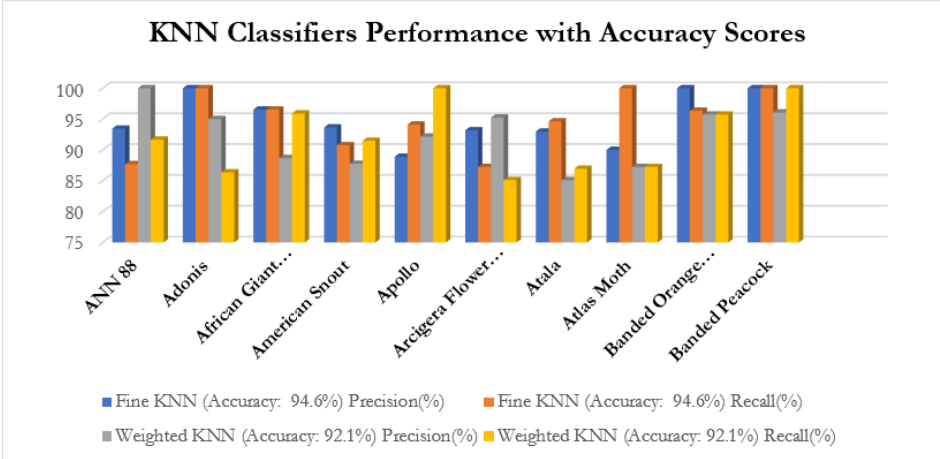

Figure 9. Comparison of Fine and Weighted KNN classifiers

Figure 9 represents the comparison of Fine and Weighted KNN classifiers across multiple butterfly and moth species. Precision and recall scores highlighted Fine KNN's superior performance, achieving an accuracy of 94.6%.

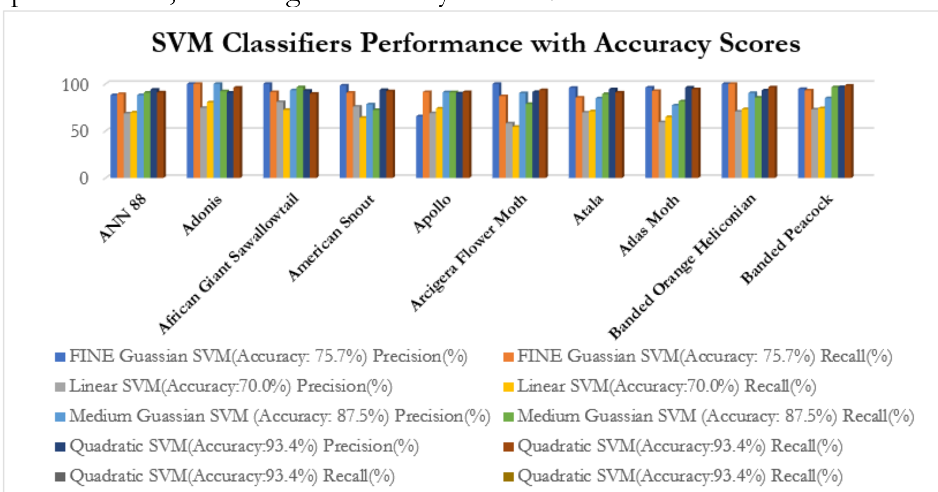

Figure 10. Performance comparison of various SVM classifiers

Figure 10 represents the performance comparison of various SVM classifiers across butterfly and moth species based on precision and recall. Quadratic SVM achieved the highest accuracy (93.4%), outperforming other kernel types consistently.

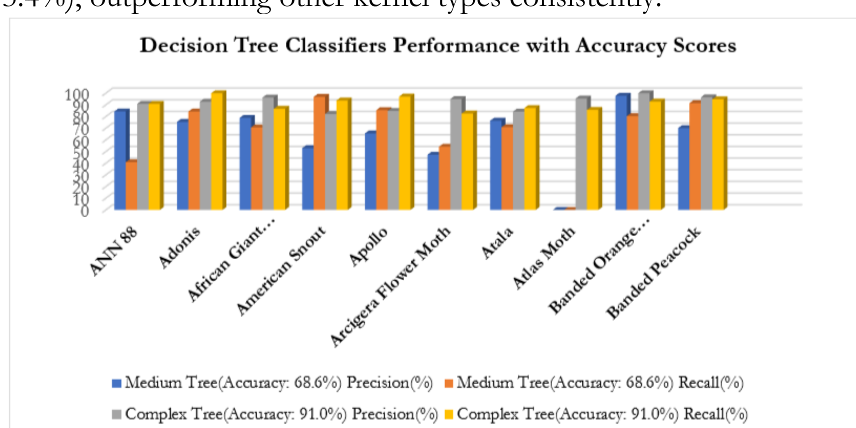

Figure 11. Performance comparison of Decision tree classifiers.

Figure 11 represents the performance comparison of the Decision tree classifiers. In species identification (e.g., Atala, Banded Peacock). Complex Tree achieves the highest accuracy (91.0%), while Medium Tree trails at 68.6%.

Table 2 illustrates the effectiveness of the ensemble classifiers (Bagged Trees and Subspace KNN) in balancing accuracy, True Negative Rate (TNR), and True Positive Rate (TPR) across ten butterfly and moth species. By using bootstrap aggregation, Ensemble Bagged Trees reduced the variation in mimicry species and achieved high true positive rates (for example, 100% for Atlas Moth). At the same time, Ensemble Subspace KNN showed better true negative rates (99.81% for Arcigera Flower Moth, for example) by using randomized feature subspaces, which improved resilience to intra-class variability. In ensemble techniques, intrinsic trade-offs between sensitivity and specificity were evident in the marginal accuracy differences, such as 94.6% versus 94.5% for "ANN 88."

Table 2. Performance of Ensemble Classifiers: Bagged Trees vs. Subspace KNN

Classes	Ensemble Bagged Trees			Ensemble Subspace KNN		
	TPR (%)	TNR (%)	ACC (%)	TPR (%)	TNR (%)	ACC (%)
ANN 88	91	99.60	94.50	88	98.82	**94.60**
Adonis	96	100		100	99.61	
African Giant Swallowtail	97	99.61		93	99.61	
American Snout	91	99.21		91	100	
Apollo	97	98.42		94	98.82	
Arcigera Flower Moth	81	99.81		87	99.81	
Atala	98	99.04		98	99.61	
Atlas Moth	100	98.85		100	98.08	
Banded Orange Heliconian	96	99.61		100	99.61	
Banded Peacock	97	99.61		97	100	

Table 3 assesses Weighted KNN and Fine KNN classifiers, chosen from a variety of KNN variations for attaining ≥90% accuracy, for 10 species of moths and butterflies. With greater accuracy (94.6% for "ANN 88") and True Negative Rate (TNR: 99.21–100%), Fine KNN excels at accurately discriminating between species (e.g., 100% TPR/TNR for "Adonis" and "Banded Peacock"). Weighted KNN reflects its distance-based voting trade-offs by maintaining competitive accuracy (92.1–100%) with a slightly lower TNR (98.35–99.52%). By addressing important ecological issues such as intra-species polymorphism, both classifiers demonstrated their value in automated systems for identifying Lepidoptera.

Table 3. Performance of Weighted KNN and Fine KNN Classifiers

Classes	Weighted KNN			Fine KNN		
	TPR (%)	TNR (%)	ACC (%)	TPR (%)	TNR (%)	ACC (%)
ANN 88	92	100	92.10	88	99.21	**94.6**
Adonis	86	99.53		100	100	
African Giant Swallowtail	96	98.57		97	99.61	
American Snout	91	98.58		91	99.21	
Apollo	100	99.05		94	98.42	
Arcigera Flower Moth	85	99.52		87	99.43	
Atala	87	98.35		95	99.23	
Atlas Moth	87	98.58		100	98.85	

Banded Orange Heliconian	96	99.52		96	100	
Banded Peacock	100	99.52		100	100	

Table 4 compares the performance of the Cubic SVM and Quadratic SVM classifiers, which were selected from a range of SVM models based on their ability to achieve at least 90% accuracy. Their effectiveness was evaluated across ten butterfly and moth species. Quadratic SVM addresses intra-species variability in field circumstances by striking a balance between precision (93.4% accuracy for ANN 88) and computational efficiency, whereas Cubic SVM exhibits better accuracy (e.g., 94.1% for ANN 88) and robustness to mimicry (e.g., 100% TPR for Arcigera Flower Moth).

Table 4. Performance of Cubic SVM vs. Quadratic SVM Classifiers

Classes	Cubic SVM			Quadratic SVM		
	TPR (%)	TNR (%)	ACC (%)	TPR (%)	TNR (%)	ACC (%)
ANN 88	94	99.21	94.10	92	99.21	93.4
Adonis	100	99.42		96	99.05	
African Giant Swallowtail	93	99.61		90	99.23	
American Snout	89	99.41		92	99.22	
Apollo	91	99.01		91	98.62	
Arcigera Flower Moth	100	98.87		93	99.24	
Atala	89	99.42		91	99.42	
Atlas Moth	93	99.61		94	99.61	
Banded Orange Heliconian	95	99.42		96	99.23	
Banded Peacock	100	99.42		98	99.61	

Ten Butterfly and moth species were evaluated in Table 5 using tree-based classifiers that achieve ≥90% accuracy (e.g., Complex Tree). Top performances, such as Apollo (97% TPR), Banded Orange Heliconian (93% TPR/100% TNR), and Arcigera Flower Moth (83% TPR/99.6% TNR), were highlighted by metrics (TPR, TNR, ACC). Based on these criteria, the study identified classifiers from a larger pool that showed robustness to intra-class variance and field-related problems, such as ANN (91% accuracy) and Adonis (100% TPR). These models were selected from a broader pool of candidates.

Table 5. Performance of Complex Tree Classifier

Classes	Complex Tree		
	TPR (%)	TNR (%)	ACC (%)
ANN 88	91	98.82	91.00
Adonis	100	99.23	
African Giant Swallowtail	90	99.61	
American Snout	94	97.46	
Apollo	97	97.63	
Arcigera Flower Moth	83	99.62	
Atala	87	98.27	
Atlas Moth	78	99.61	

Banded Orange Heliconian	93	100	
Banded Peacock	95	99.61	

Table 6 shows the performance comparison between Ensemble Subspace KNN and ResNet50 in the classification of butterflies and moths. With 100% TPR/TNR for complex-pattern species (such as the Atlas Moth and Banded Orange Heliconian), superior TNR for cryptic taxa (like the Arcigera Flower Moth), and higher overall accuracy (94.6% vs. 92.6%), the Ensemble fared better than ResNet50. As demonstrated by its strong performance on the American Snout (91% TPR vs. ResNet50's 96% TPR, but with higher TNR), the Ensemble's feature subspace variety improved its resilience to intra-class variation. Even though ResNet50 performed exceptionally well in categorizing species with significant contrast, such as the Apollo (100% TPR), these findings demonstrated the Ensemble's potency in fine-grained lepidopteran classification tasks.

Table 6. Performance Comparison of Ensemble Subspace KNN vs. ResNet50

Classes	**Ensemble Subspace KNN**			**ResNet50**		
	TPR (%)	**TNR (%)**	**ACC (%)**	**TPR (%)**	**TNR (%)**	**ACC (%)**
ANN 88	88	98.82	**94.60**	96	100	92.60
Adonis	100	99.61		82	99.06	
African Giant Swallowtail	93	99.61		92	98.10	
American Snout	91	100		96	98.58	
Apollo	94	98.82		100	99.52	
Arcigera Flower Moth	87	99.81		87	99.76	
Atala	98	99.61		89	99.52	
Atlas Moth	100	98.08		91	98.11	
Banded Orange Heliconian	100	99.61		96	99.52	
Banded Peacock	97	100		96	99.52	

As a baseline on the Butterfly and Moth dataset, Table 7 shows that ResNet50 achieves an accuracy of 92.6%, indicating the limitations of purely data-driven fine-grained taxonomy. Our proposed method improves this performance to 94.6% by combining global morphology analysis from ResNet50 with rhombus-corner gradient descriptors for wing textures using CRSHOG. Enhanced accuracy demonstrates the effectiveness of synergistic feature strengthening in addressing mimicry and intra-class variance.

Table 7. Accuracy Metrics across Different Methodologies.

Methods	**Accuracy**
ResNet50(Our)	92.6%
Proposed Methodology (Baseline)	**94.6%**

Description:

An innovative hybrid computational framework is presented in this work to tackle the difficulties of classifying moth and butterfly species in practical settings. A unique Corner Rhombus-based Histogram of Oriented Gradients (CRSHOG) descriptor, deep feature extraction using ResNet50, and contrast-limited adaptive histogram equalization (CLAHE) for picture improvement are all integrated into the suggested methodology. Achieving 94.6% classification accuracy with enhanced precision (93.9%) and recall (94.6%), experimental results show that the ensemble-based classification strategy works better than traditional deep learning models. The framework may be used with conventional RGB images on devices with limited resources because of its streamlined architecture, which guarantees computational efficiency. In order to improve taxonomic coverage, future research will concentrate on real-time implementation through model quantization, extended dynamic feature fusion for temporal analysis, and cooperative ecological data gathering. With these developments, a useful tool for ecological decision-making, conservation planning, and biodiversity evaluation will be developed.

Conclusion:

Our hybrid framework for classifying Butterflies and moths addresses intra-species variability and complex field conditions by combining CLAHE preprocessing, ResNet50 deep features, and a novel CRSHOG descriptor. It achieves robust performance via Ensemble Subspace KNN (93.9% precision, 94.6% recall) and 94.6% accuracy (beating standalone CNNs [92.6%]). The system shows the combination of domain-specific feature engineering and deep learning working well on low-resource devices with conventional RGB inputs.

Future Recommendations:

Future research will use TensorFlow Lite and model quantization to optimize real-time deployment, extend dynamic feature aggregation for temporal video analysis, and improve interpretability with saliency maps. Ecological partnerships will expand datasets to encompass uncommon species in a range of climates, and open-set recognition (metric learning, outlier detection) will enhance variability. These developments seek to create a field-ready instrument for evidence-based policymaking, conservation prioritization, and biodiversity monitoring.

Acknowledgment: The Department of Computer Science, University of Engineering and Technology, Taxila facilitated us with an Image Processing Lab and other resources for the experimentation of this research.

Author's Contribution: Aymen Siddique contributed to the methodology, manuscript writing, and interpretation of results. Haider Ali Khan contributed to the Development of CRSHOG.

Syed Muhammad Adnan proposed the topic and performed statistical analysis.

Wakeel Ahmed contributed to the statistical analysis and interpretation of results.

Abid Ghaffar conducted the literature review and referencing.

Conflict of interest. No

Project details. No

References:

[1] A. Hoskins, "Butterflies of the world," *London NH (Reed New Holl. Publ.*, 2015, [Online]. Available: https://archive.org/details/butterfliesofwor0000hosk

[2] C. L. Boggs, W. B. Watt, and P. R. Ehrlich, "Butterflies Ecology and Evolution Taking Flight," *Univ. Chicago Press*, 2003, doi: 10.7208/CHICAGO/9780226063195.001.0001.

[3] A. S. Almryad and H. Kutucu, "Automatic identification for field butterflies by convolutional neural networks," *Eng. Sci. Technol. an Int. J.*, vol. 23, no. 1, pp. 189–195, 2020, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2215098619326011?via%3Dihub

[4] J. Boulent, S. Foucher, J. Théau, and P.-L. St-Charles, "Convolutional Neural Networks for the Automatic Identification of Plant Diseases," *Front. Plant Sci.*, vol. 10, p. 941, Jul. 2019, doi: 10.3389/FPLS.2019.00941.

[5] A. P. Wibowo and D. R. I. M. Setiadi, "Performance Analysis of Deep Learning Models for Sweet Potato Image Recognition," *2022 Int. Semin. Appl. Technol. Inf. Commun. Technol. 4.0 Smart Ecosyst. A New W. Doing Digit. Business, iSemantic 2022*, pp. 101–106, 2022, doi: 10.1109/ISEMANTIC55962.2022.9920423.

[6] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, "Fundamental Concepts of Convolutional Neural Network," *Intell. Syst. Ref. Libr.*, vol. 172, pp. 519–567, 2020,

doi: 10.1007/978-3-030-32644-9_36.
[7] M. A. S. A. Fathimathul Rajeena P. P., Rasha Orban, Kogilavani Shanmuga Vadivel, Malliga Subramanian, Suresh Muthusamy, Diaa Salam Abd Elminaam, Ayman Nabil, Laith Abulaigh, Mohsen Ahmadi, "A Novel Method for the Classification of Butterfly Species Using Pre-Trained CNN Models," *Electronics*, vol. 11, no. 13, p. 2016, 2022, doi: https://doi.org/10.3390/electronics11132016.
[8] R. Rashmi, U. Snekhalatha, A. L. Salvador, and A. N. J. Raj, "Facial emotion detection using thermal and visual images based on deep learning techniques," *Imaging Sci. J.*, vol. 72, no. 2, pp. 153–166, Feb. 2024, doi: 10.1080/13682199.2023.2199504;SUBPAGE:STRING:ACCESS.
[9] H. A. Khan and S. M. Adnan, "A Novel Fusion Technique for Early Detection of Alopecia Areata Using ResNet-50 and CRSHOG," *IEEE Access*, vol. 12, pp. 139912–139922, 2024, doi: 10.1109/ACCESS.2024.3461324.
[10] A. S. R. M. A. Erick A. Adjé, Gilles Delmaire, "A Novel Statistical Framework for Butterfly Species Recognition Using Raw Spatio-Hyperspectral Images," *Res. Sq.*, 2024, [Online]. Available: https://www.researchsquare.com/article/rs-5240841/v1
[11] Jayme Garcia Arnal Barbedo, "Detecting and Classifying Pests in Crops Using Proximal Images and Machine Learning: A Review," *AI*, vol. 1, no. 2, pp. 312–328, 2020, doi: https://doi.org/10.3390/ai1020021.
[12] Y. G. Xin Chen, Bin Wang, "Separated Fan-Beam Projection with Gaussian Convolution for Invariant and Robust Butterfly Image Retrieval.," *Pattern Recognit.*, vol. 147, p. 110083, 2024, doi: https://doi.org/10.1016/j.patcog.2023.110083.
[13] Y. Rong, H. Su, W. Zhang, and Z. Li, "Fine-Grained Complex Image Classification Method Based on Butterfly Images," *Proc. 2022 6th Asian Conf. Artif. Intell. Technol. ACAIT 2022*, 2022, doi: 10.1109/ACAIT56212.2022.10137872.
[14] J. U. Hilmi Syamsudin, Saidatul Khalidah, "Lepidoptera Classification Using Convolutional Neural Network EfficientNet-B0," *Indones. J. Artifical Intelligance Data Min.*, vol. 7, no. 1, 2024, [Online]. Available: https://ejournal.uin-suska.ac.id/index.php/IJAIDM/article/view/24586
[15] X. Jia, X. Tan, G. Jin, and R. O. Sinnott, "Lepidoptera Classification through Deep Learning," *2020 IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2020*, Dec. 2020, doi: 10.1109/CSDE50874.2020.9411382.
[16] S. S. Abdul Fadlil Abdul Fadlil, Ainin Maftukhah, "Butterfly Image Identification Using Multilevel Thresholding Segmentasi and Convolution Neural Network Classification with Alexnet Architecture," *Int. J. Comput. Digit. Syst.*, vol. 15, no. 1, pp. 1775–1785, 2024, doi: 10.12785/ijcds/1501125.
[17] Harish Trio Adityawan, Omar Farroq, Stefanus Santosa, Hussain Md Mehedul Islam, Md Kamruzzaman Sarker, "Butterflies Recognition using Enhanced Transfer Learning and Data Augmentation," *J. Comput. Theor. Appl.*, vol. 1, no. 2, 2023, [Online]. Available: http://publikasi.dinus.ac.id/index.php/jcta/article/view/9443
[18] Hari Theivaprakasham, "Identification of Indian butterflies using Deep Convolutional Neural Network," *J. Asia. Pac. Entomol.*, vol. 24, no. 1, pp. 329–340, 2021, doi: https://doi.org/10.1016/j.aspen.2020.11.015.
[19] T. Xi, "Multiple butterfly recognition based on deep residual learning and image analysis," *Entomol. Res.*, vol. 52, no. 1, pp. 44–53, 2022, [Online]. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1111/1748-5967.12564
[20] T. Surabhi, B. Sachin, and C. Advait, "Deep Convolutional Neural Networks for Automated Butterfly Species Recognition and Classification," *Proc. 5th Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2023*, pp. 778–783, 2023, doi: 10.1109/ICIRCA57980.2023.10220696.

[21] Fengwei Liu, "Butterfly and Moth Image Recognition Based on Residual Neural Network," *Int. Conf. Data Anal. Mach. Learn.*, pp. 211–217, 2023, doi: 10.5220/0012798500003885.
[22] S. Y. and X. F. Ruoyan Zhao, Cuixia Li, "Butterfly Recognition Based on Faster R-CNN," *J. Phys. Conf. Ser.*, vol. 1176, no. 3, 2019, doi: 10.1088/1742-6596/1176/3/032048.
[23] N. N. Kamaron Arzar, N. Sabri, N. F. Mohd Johari, A. Amilah Shari, M. R. Mohd Noordin, and S. Ibrahim, "Butterfly Species Identification Using Convolutional Neural Network (CNN)," *2019 IEEE Int. Conf. Autom. Control Intell. Syst. I2CACIS 2019 - Proc.*, pp. 221–224, Jun. 2019, doi: 10.1109/I2CACIS.2019.8825031.
[24] B. W. Kunkun Zhang, Xin Chen, "MdeBEIA: Multi-task deep leaning for butterfly ecological image analysis," *Digit. Signal Process.*, vol. 162, p. 105168, 2025, [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1051200425001903?via%3Dihub
[25] J. Kaur, S. Rani, and G. Singh, "Evolving Patterns: A CNN Approach to Butterfly Species Recognition," *8th Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2024 - Proc.*, pp. 1000–1005, 2024, doi: 10.1109/ICECA63461.2024.10800980.
[26] S. A. Samruddhi N Hegde, Manjesh R, P Sindhura, Reema Rodrigues, "Identification of butterfly speciecs using Vgg-16," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 2, no. 5, 2020, [Online]. Available: https://www.irjmets.com/uploadedfiles/paper/volume2/issue_5_may_2020/1088/1628083020.pdf
[27] Gerry, "Butterfly & Moths Image Classification 100 species," *Kaggle*, 2023, [Online]. Available: https://www.kaggle.com/datasets/gpiosenka/butterfly-images40-species
[28] Y. Yoshimi *et al.*, "Image preprocessing with contrast-limited adaptive histogram equalization improves the segmentation performance of deep learning for the articular disk of the temporomandibular joint on magnetic resonance images," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.*, vol. 138, no. 1, pp. 128–141, Jul. 2024, doi: 10.1016/J.OOOO.2023.01.016,.
[29] M. P. Thi Phuoc Hanh Nguyen, Zinan Cai, Khanh Nguyen, Sokuntheariddh Keth, Ningyuan Shen, "Pre-processing Image using Brightening, CLAHE and RETINEX," *arXiv:2003.10822*, 2020, doi: https://doi.org/10.48550/arXiv.2003.10822.
[30] R. Bhuria, "Feature Extraction and Classification Performance of Fine-Tuned ResNet-50 in Plant Species Recognition," *2024 2nd Int. Conf. Adv. Comput. Commun. Inf. Technol. ICAICCIT 2024*, pp. 574–579, 2024, doi: 10.1109/ICAICCIT64383.2024.10912176.
[31] D. L. Bin Li, "Facial expression recognition via ResNet-50," *Int. J. Cogn. Comput. Eng.*, vol. 2, pp. 57–64, 2021, [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666307421000073?via%3Dihub
[32] Suhita Ray, "Disease Classification within Dermascopic Images Using features extracted by ResNet50 and classification through Deep Forest," *arXiv:1807.05711*, 2018, doi: https://doi.org/10.48550/arXiv.1807.05711.