Towards Automatic Butterfly Species Recognition Using a Single Spatio-Hyperspectral Image

Erick Adje[1,2], Gilles Delmaire[1], Arnaud Ahouandjinou[2], Matthieu Puigt[1], and Gilles Roussel[1]
[1]*Univ. Littoral Côte d'Opale, LISIC – UR 4491*, F-62219 Longuenesse, France
[2]*Univ. d'Abomey-Calavi, LRSIA*, Cotonou, Benin
{erick.adje, gilles.delmaire, arnaud.ahouandjinou, matthieu.puigt, gilles.roussel}@univ-littoral.fr

***Abstract*—Butterfly Species identification accounts nowadays for a challenge to evaluate the biodiversity state. Using special compact Hyperspectral Cameras for this task is more attractive. Whereas usual techniques use a sequence of images to compute a datacube, we focus here on a single image resulting in a partial butterfly datacube. With a pre-identification of the features from a butterfly library, we propose a combined probabilistic clustering technique based on a weighted combination of Z-score and Gaussian Naive Bayes probability, which aims to recognize the associated cluster from the particular butterfly species. Results obtained in this context achieve good performance with respect to Gaussian Naive Bayes probability or Z-score-based techniques.**

***Index Terms*—Biodiversity, Species recognition, Hyperspectral imaging, Spatio-spectral camera, Gaussian mixture models**

I. Introduction

Butterflies are pivotal for ecosystems, especially as pollinators aiding plant reproduction and biodiversity. However, while they contribute positively to the environment, some species can also pose significant threats such as pests, which may cause damage to crops and forests. This duality underscores the need for effective butterfly species management, beginning with accurate species identification. Among the methods for their automatic recognition, imaging-based techniques have proven to be highly effective over the years. Traditionally, these systems have relied on RGB images captured by standard cameras [1]. Using traditional machine learning techniques, models are designed with a variety of features extracted from labeled butterfly images. The Gray Level Co-occurrence Matrix (GLCM) is a widely-used method for texture features extraction of butterfly images, which helps to create accurate models using K-Nearest Neighbors (K-NN) [2], Multinomial Logistic Regression (MLR) [3], and Rough Set approach [4]. Texture features using Local Binary Patterns (LBP) have also been explored with a Support Vector Machine (SVM) classifier [5]. Deep learning techniques, through the application of Convolutional Neural Networks (CNN) were studied [6]. Zhu et al. [7] applied CNN on an image database of 10 butterfly species. In contrast, in the case of a large species database—i.e., with more than 80 species—advanced deep learning techniques like YOLO [8] and ResNet [9] have proven to be effective for classification purpose.

The authors thank IITA Benin, for the PhD co-funding.

However, with the advent of Hyperspectral Imaging (HSI), some studies began exploring its use. HSI is a technique that generates a 3D Datacube, where two dimensions are dedicated to spatial dimensions and the third is related to wavelengths. This datacube may be viewed as a collection of spatial maps, each one representing a specific wavelength. HSI enables analysis at more than 100 wavelengths, unlike Multispectral Imaging (MSI) which enables analysis with less than 100 wavelengths. We note that there are very few studies using HSI for butterfly analysis, and those that exist primarily focus on analyzing the physiological aspects of butterfly wings [10], [11]. Moreover, due to the analytical complexity and the abundance of HSI data, butterfly research studies tend to focus on a limited number of species simultaneously. The use of hyperspectral cameras is also limited due to their high cost, but more affordable compact surveillance cameras are available. However, the images they produce require extensive processing before obtaining the hyperspectral cube [12]. Among the alternative solutions, MSI snapshot cameras provide a very limited spectral information and require a demosaicing stage [13]. There are also HSI CASSI cameras which provide a much finer spectral resolution [14], but these are not proposed nowadays in a commercial versions.

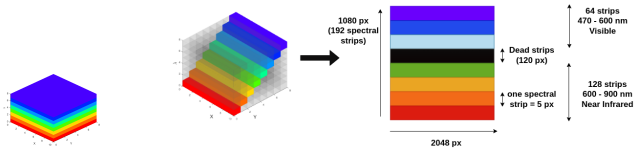
Our study introduces a spatio-spectral camera-based method that precisely identifies species of stationary butterflies captured by a single spatio-spectral image in few spectral layers, focusing on distinguishing among four distinct species. This identification will be done in real-time without the need for datacube reconstruction. For this study, we use the XIMEA MQ022HG-IM-LS150-VISNIR spatio-spectral camera [15] because of its compact size and an excellent spectral resolution. Furthermore, as a hyperspectral camera, it enables identification in more complex environments and can distinguish differences between visually similar species by capturing spectral information in the invisible domain.

The spatio-spectral camera's complexity lies in its need for a complete scan to ensure an object to be captured across all spectral strips. Then this step is usually followed by a datacube reconstruction. This process can be time-consuming and impractical for real-time decision-making in critical tasks, such as early detection of crop-damaging butterflies. Our goal is to set up our stationary camera, which monitors a specific part of a vegetation-covered field. Then, the random movement

of several butterflies across the camera spectral layers enables to gather some partial information. In contrast to conventional methods, we would like to investigate whether we can recognize butterfly species in real-time using efficient methods, without the need for spectral datacube reconstruction. As a preliminary task, we here focus on the ability to identify the species of a stationary butterfly detected on a very limited number of spectral layers with a single raw image captured by the HSI camera. To the best of our knowledge, our work is the first to perform recognition directly on a raw spatio-spectral image. The main contributions of this work are a proposed method for species recognition using a raw single spatio-spectral image and, secondly, a classification model of 4 important butterfly species for biodiversity preservation.

II. Spatio-spectral camera description

A. Characteristics of the Spatio-Spectral Camera



(a) (b)

Fig. 1. Captured HSI content from (a) a conventional and (b) a spatio-spectral camera (structure of our camera).

Unlike conventional HSI data—which consists of a complete 3D cube—we here consider a compact spatio-spectral camera which only acquires a part of the hyperspectral content (see Fig. 1). In particular, our camera is an ultra-compact hyperspectral camera that can capture images for a wide range of wavelengths—i.e., from 470 nm to 900 nm—with a fine spectral resolution. The camera is considered to be spatio-spectral owing to its ability to capture a specific profile of spatial and spectral information for each raw image. Fig. 1-b depicts the structure of a raw image from our camera. The raw image typically has dimensions of 1088×2048 pixels. However, due to the absence of information in the first and last 4 rows of pixels, and due to a dead zone of 120 pixels, the effective informative area is 960×2048 pixels, corresponding to 192 different spectral strips of 5 pixels width. These strips form a staircase pattern in the datacube, with each step representing the wavelength difference between strips

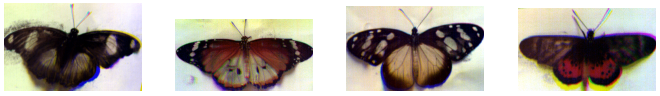
B. Spatio-Spectral Camera Scanning

The normal use of the camera typically involves one of two procedures. Either the target to be captured remains stationary, and the camera executes a slow, linear, and uniform movement above the target using appropriate equipment. Or the camera remains stationary, and the target is moving slowly, linearly, and uniformly below the camera's field of view with a displacement system. In both cases, each spectral strip of the camera sweeps across the object of interest, capturing only a portion of the object's wavelengths at any given time. Obtaining the complete spectrum requires capturing a series of images, which are subsequently stitched together using post-processing methods [12].

III. Data Collection

A. Studied Butterfly Species

Our study focuses on *Hypolimnas Misippus* (HM), *Danaus Chrysippus* (DC), *Amauris Ochlea* (AO), and *Acraea Egina* (AE), four butterfly species belonging to the *Nymphalidae* family and exhibiting a broad distribution across Asia, Africa, and Australia. Fig. 2 provides a visual representation of their distinct appearances. In their butterfly form, these species primarily serve as pollinators and generally do not pose any real danger, subsisting on nectar and select leaves. However, an excessive population of these butterflies could potentially become problematic. These species were captured due to their high predominance in a citrus and legume experimental field at the International Institute of Tropical Agriculture (IITA) [16] station in Benin. In order to construct some spectral datacube for each species to obtain their complete spectrum for their characterization, some complete scans were performed, each one with a different individual specimen for each butterfly species using a conveyor belt system.



(a) species HM (b) species DC (c) species AO (d) species AE

Fig. 2. Butterfly species considered in this work.

B. Hypercube reconstruction

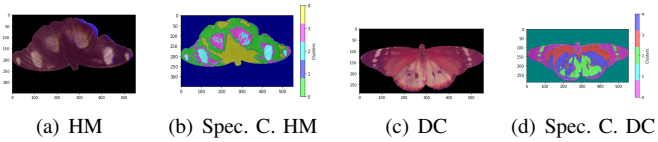
Using images from the complete scans of each species enables us to reconstruct the spectral hypercube for each one. The reconstruction process typically involves three steps, i.e., *spectral strip stitching*—which involves accurately positioning each strip within the datacube based on movement knowledge [12]—a *realignment process*—to correct misalignments across spectral images—and the *reflectance calculation* of each pixel. This process produces 192 reflectance 2D images. These reflectance images are then sequentially concatenated to form the hyperspectral datacube.

IV. Proposed methodology for butterfly species identification

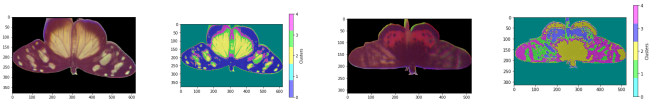
A. Species characteristic features

The first step in getting species characteristic features is to identify regions with similar spectra for each butterfly species, using K-means clustering. The K-means algorithm divides the observations into K groups with the closest mean values, and minimizes the variance within each cluster until it becomes stable. Determining the optimum number of clusters K for each butterfly species is crucial, as it significantly affects the subsequent results. A very small number of clusters could result in overlapping features between species, while a

very large number of clusters could result in unrepresentative features of the species, making them difficult to capture in other butterflies of the same species. Based on visual observation and Elbow curve method investigation, we assume $K = 5$ with one cluster for background, for each of the four investigated butterfly species. Fig. 3 shows the clustering map for each pixel spectrum of the considered species. Let $\underline{s}_i = [r_i^1, r_i^2, ...r_i^{192}]^T$ denote the spectral value of a single pixel p_i where r_i^b denotes the reflectance of a spectral layer $b \in \{1, \ldots, 192\}$ of p_i. We denote $G^{(u,k)}$ as the set of spectra belonging to Cluster k ($k \in \mathcal{K} \triangleq \{1, \ldots, K\}$) for species u. Let $\mathcal{I}^{(u,k)} \triangleq \{i \in \mathcal{N}$ s.t. $\underline{s}_i^u \in$ Cluster $k\}$ be the set of indices of extended pixels indices that belong to Cluster k. Then, $G^{(u,k)}$ is defined as $\forall i \in \mathcal{I}^{(u,k)}$, $G^{(u,k)}(i) \triangleq \underline{s}_{\mathcal{I}^{(u,k)}(i)}$.



(a) HM (b) Spec. C. HM (c) DC (d) Spec. C. DC



(e) AO (f) Spec. C. AO (g) AE (h) Spec. C. AE

Fig. 3. Spectral clusters (Spec. C.) of butterfly species

Having the K spectra clusters for each butterfly species, we assume that for any given species u, the random variable representing the reflectance values within the same spectral layer b in Cluster k follows a Gaussian distribution $\mathcal{N}(\mu_b^{(u,k)}, \sigma_b^{(u,k)})$. We will use the Maximum Likelihood Estimation (MLE) [17] to find $(\mu_b^{(u,k)}, \sigma_b^{(u,k)})$. For this purpose, we assume that all species and different parts of a butterfly represented by our clusters are equiprobable because we used the same amount of data for each species and the area of pattern on their bodies can vary depending on the life stage of the specimen. This Gaussian distribution is identified using the $n^{(u,k)}$ reflectance values $G^{(u,k)}(b,:) = [G^{(u,k)}(b,1) \ldots G^{(u,k)}(b,n^{(u,k)})]$. As one cluster was considered for the background, we didn't take it into consideration resulting into 4 clusters per species. So for each cluster k from species u, we consider the 2×192 matrix $M^{(u,k)}$ defined as

$$M^{(u,k)} = \begin{bmatrix} \mu_1^{(u,k)} & \mu_2^{(u,k)} & \ldots & \mu_{192}^{(u,k)} \\ \sigma_1^{(u,k)} & \sigma_2^{(u,k)} & \ldots & \sigma_{192}^{(u,k)} \end{bmatrix}$$

and its extension $M^u = [M^{(u,1)}, M^{(u,2)}, M^{(u,3)}, M^{(u,4)}]$. Identifying the parameters amounts to search for 1536 features—i.e., 2 parameters · 4 clusters · 192 spectral layers—for each species. This is equivalent to a Gaussian Mixture Model (GMM) with 16 classes per spectral layer. Fig. 4 shows the Gaussian distributions for our four species within spectral layer 154 (located in NIR) with a fixed cluster number $k = 4$. This confirms that our data fits a Gaussian distribution and shows that the distributions for each species are sufficiently distinct, except for occasional overlaps that may challenge butterfly identification. This suggests to propose an approach that is robust to the overlap of multiple Gaussian distributions.

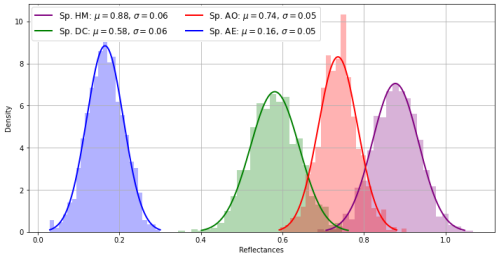
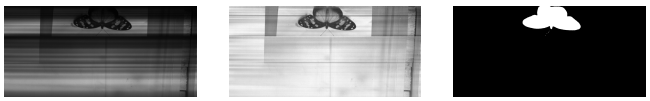


Fig. 4. Gaussian distribution in Spectral Layer 154.

B. Partial reflectance extraction from a single raw image

We define $I(x, y)$ as a raw image from a spatio-spectral camera containing a butterfly specimen. We compute a mask—denoted $Mask(x, y)$—which specifies the location of the butterfly. Fig. 5 describes this process. The mask is obtained by normalizing the image and then subtracting the background using a reference background image. This step is followed by a series of manual adjustments to produce a mask that accurately represents the detected butterfly. The resulting mask corresponds to the amount of information to be used for the species recognition, which may indirectly affect classification performance. After obtaining the mask, it is split into spectral layers, denoted $Mask(x, y, b)$ for Spectral Layer b. We then remove masks from spectral layers with less than 20 detection pixels to prevent false detection. Let b^* be an index of the set $\mathcal{B}^*$ representing the remaining spectral layers of the masks $(Mask(x, y, b^*))$ and $Mask^*(x, y) \triangleq \bigcup_{b^* \in \mathcal{B}^*} Mask(x, y, b^*)$ be the set of remaining masks. For each $Mask(x, y, b^*)$, we calculate the reflectance values $r_{b^*}^i$ of the detected butterfly where $Mask(x, y, b^*) = 255$, using part of spectralon in the same spectral layer. This provides a vector of reflectance values in each spectral layer $b^* \in \mathcal{B}^*$ defined as $\underline{r}_{b^*} = \left[r_{b^*}^1, \quad r_{b^*}^2, \ldots (r^{n^{b^*}})_{b^*}\right]^T$ where n^{b^*} is the number of pixels representing the butterfly in spectral layer b^* .



(a) Raw image (b) Normalized image (c) Butterfly mask

Fig. 5. Butterfly detection

C. Spectral matching

Before carrying out the spectral comparison, it's important to select the features to compare based on the spectral layers b^* identified in the previous step where the butterfly was detected. For this purpose, the corresponding features $(\mu_{b^*}^{(u,k)}, \sigma_{b^*}^{(u,k)})$ obtained for each species during the previous step IV-A were selected. Since we have 4 clusters per butterfly

species, this involves extracting the following feature vectors for each species: $\underline{\mu}^u_{b^*} = \begin{bmatrix}\mu^{(u,1)}_{b^*} & \mu^{(u,2)}_{b^*} & \mu^{(u,3)}_{b^*} & \mu^{(u,4)}_{b^*}\end{bmatrix}$ and $\underline{\sigma}^u_{b^*} = \begin{bmatrix}\sigma^{(u,1)}_{b^*} & \sigma^{(u,2)}_{b^*} & \sigma^{(u,3)}_{b^*} & \sigma^{(u,4)}_{b^*}\end{bmatrix}$, $\forall\, b^* \in \mathcal{B}^*$.

Assuming that the reflectance spectra within the same cluster follow a Gaussian distribution, a classification approach using Gaussian Naive Bayes (GNB) [18] may be explored. This technique uses the Naive Bayes conditional probability formula to estimate the most likely Gaussian distribution to which an observation belongs by finding the maximum probability across all other classes. Let denote $\Theta^{(u,k)}_{b^*} \triangleq [\mu^{(u,k)}_{b^*}, \sigma^{(u,k)}_{b^*}]$, n^{b^*} be the number of detected pixels in spectral layer b^*. In our context, GNB consists in determining, the most likely distribution for a specific $r^i_{b^*} \in \underline{r}_{b^*}$, $\delta^{(u,k)}_{b^*}(i)$ with $\delta^{(u,k)}_{b^*}(i) = \underset{u=1,..,4||k=1,..,4}{\arg\max} p(r^i_{b^*}|\Theta^{(u,k)}_{b^*}),\ \forall\, i \in \{1 \dots n^{b^*}\}$ where $p(r^i_{b^*}|\Theta^{(u,k)}_{b^*}) = \frac{1}{\sqrt{2\pi(\sigma^{(u,k)}_{b^*})^2}} \exp\left(-\frac{r^i_{b^*}-(\mu^{(u,k)}_{b^*})^2}{2(\sigma^{(u,k)}_{b^*})^2}\right)$, is the univariate Gaussian Probability Density Function (PDF) describes the relative probability of $r^i_{b^*}$ occuring within the distribution $\Theta^{(u,k)}_{b^*}$. This means choosing the best distribution between 16 possible ones for each $r^i_{b^*}$. Furthermore, as the distribution is linked to a species, we can attribute this reflectance to the appropriate species. Given that the butterfly detected in an image is represented by multiple reflectances across many spectral layers, to identify a butterfly species in an image, we start with zero scores for the four species, incrementing for each classified reflectance. After, normalizing these scores to ensure their sum equals 1, the species with the highest score is determined to identify as the one present in the image. Let us denote this approach *GNB-scoring*.

Algorithm 1 Proposed scoring algorithm

$gnb_scr \leftarrow \begin{bmatrix}0 & 0 & 0 & 0\end{bmatrix}$ //Scores obtained by GNB-Scoring
$zs_scr \leftarrow \begin{bmatrix}0 & 0 & 0 & 0\end{bmatrix}$ //Scores obtained by ZS-Scoring
for b^* in $\mathcal{B}^*$ **do**
  for $i = 1$ to n^{b^*} **do**
    Find $\delta^{(u,k)}_{b^*}(i)$ and $\gamma^{(u,k)}_{b^*}(i)$
    Find u_δ from $\delta^{(u,k)}_{b^*}(i)$ and u_γ from $\gamma^{(u,k)}_{b^*}(i)$
    $gnb_scr[u_\delta] \leftarrow gnb_scr[u_\delta] + 1$
    $zs_scr[u_\gamma] \leftarrow zs_scr[u_\gamma] + 1$
  end for
end for
$final_score \leftarrow (0.65 * zs_scr) + (0.35 * gnb_scr)$
$normalized_score \leftarrow final_score / \sum_{j=1}^{4} final_score[j]$

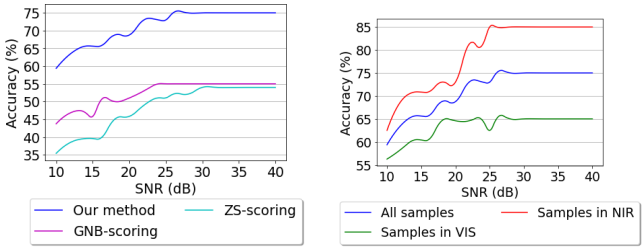
In searching for the best approach to reflectance classification, we also investigated the z-score method [19] widely used for identifying the most probable distribution of an observation. Here, the most likely distribution is the one with the minimal z-score. Following the same logic as GNB, the goal is to identify the most probable distribution $\gamma^{(u,k)}_{b^*}(i) = \underset{u=1,..,4||k=1,..,4}{\arg\min} z(r^i_{b^*}, \Theta^{(u,k)}_{b^*}),\ \forall\, i \in \{1 \dots n^{b^*}\}$ where $z(r^i_{b^*}, \Theta^{(u,k)}_{b^*}) = |r^i_{b^*} - \mu^{(u,k)}_{b^*}| \,/\, \sigma^{(u,k)}_{b^*}$ denotes the absolute z-score measures the distance of an observation $r^i_{b^*}$ from the mean of a distribution $\Theta^{(u,k)}_{b^*}$ in standard deviation units. Following the same idea, as GNB-scoring, we call the z-score approach *ZS-scoring*.

However, these two methods failed to yield satisfactory results. Analysis of successful classifications by both methods revealed that the GNB-scoring was more effective for images with partial butterfly appearances, while the ZS-scoring performed better for images with complete appearances (see Fig. 5). This insight led us to introduce a third scoring method based on a weighted combination of the two previous approaches, as presented in Alg 1. A preliminary study was conducted on a training set of 160 raw spatio-spectral images to find the best weights for the convex combination of the two approaches, where α represents the weight for Z-Scoring and $(1-\alpha)$ the weight of GNB-Scoring with $\alpha \in [0; 1]$. Using a learning rate of $\alpha_{lr} = 0.05$, we evaluated the accuracy of the training set based on the value of α. The best performance was achieved for $\alpha = 0.65$.

V. Result and discussions

To test our model, we randomly selected 96 spatio-spectral raw images that were not used for species characterization, and where the specimens present in the images were not involved in the characterization process. In a sample image, a butterfly is located either in the visible spectrum or in the infrared zone, on average over 25 to 30 successive spectral layers, and the butterfly species is known for the evaluation process. For each image, the scores are estimated as in Alg. 1 according to the four species. Based on our methodology, we also tested the Multinomial Logistic Regression [3] and KNN models [2] (MLR-scoring and KNN-scoring). Table I presents the confusion matrix for all 96 images, revealing that our method achieved an accuracy of 75%, better than GNB-scoring and ZS-scoring methods applied separately. It also performs better than some of the state-of-the-art models used to classify butterfly species in RGB images. The results also indicate significant confusion between the DC and AO species. This confusion is due to the overlap of some Gaussian distributions extracted during the characterization phase of each species. Given the strong correlation between successive and closely related spectral layers, this distribution overlap can occur across several successive spectral layers. From the confusion matrix, we note a slightly higher accuracy on images where the butterfly is detected in the near-infrared spectrum, at 85%, compared to 65% for detection in the visible spectrum.

To evaluate the robustness of our method compare to others, we repeated the previous experiment on our 96 images, but this time adding Gaussian noise to the reflectance vectors $\underline{r}_{b^*}$. This is to simulate the additive noise introduced during acquisition step due to the camera's high temperature or when the illumination is very strong. For this purpose, we evaluate the accuracy of our proposed method compared to others across various signal-to-noise ratio (SNR) values (see Fig. 6). We observe better performance for our method and a good stability of the overall accuracy around 27 dB. Furthermore, better performance is observed for NIR detections compared



(a) our method vs. others (b) our method in all spectral zones

Fig. 6. Accuracy classification given multiple SNR on test images

to those in VIS. This indicates a greater robustness of NIR to noise and demonstrates the relevance of detecting these butterfly species in the near-infrared.

VI. Conclusion

Recognizing butterfly species in an outdoor area with an HSI camera becomes attractive owing to the faculty to provide information in several spectral layers. However, we propose here to use a spatio spectral camera in a non conventional way by using a single image thus providing a partial datacube of the sample butterfly. Using a pre trained probabilistic model for characterization of species, a combined method based on both Z-score and Gaussian Naive Bayes approach is proposed for real time processing. Results on noisy sample images show good accuracy with respect to both Z-score based methods or Gaussian Naive Bayes methods. Perspective work may focus on non binary local decision to keep alive a doubt when two hypotheses are similar and also includes more relevant species.

References

[1] M. Li, G. Zhou, W. Cai, J. Li, M. Li, M. He, Y. Hu, and L. Li, "Multi-scale sparse network with cross-attention mechanism for image-based butterflies fine-grained classification," *Applied Soft Computing*, vol. 117, p. 108419, 2022.

[2] F. Li and Y. Xiong, "Automatic identification of butterfly species based on HoMSC and GLCMoIB," *The Visual Computer*, vol. 34, no. 11, pp. 1525–1533, 2018.

[3] L. Kayci and Y. Kaya, "A vision system for automatic identification of butterfly species using a grey-level co-occurrence matrix and multinomial logistic regression," *Zoology in the Middle East*, vol. 60, no. 1, pp. 57–64, 2014.

[4] K. Yılmaz, K. Lokman, E. Ö. Faruk, and T. Ramazan, "Identification of butterfly species with rough set approach based on textural features," *International Journal of Computer Vision, Machine Learning and Data Mining*, vol. 1, 2015.

[5] D. S. Y. Kartika, D. Herumurti, and A. Yuniarti, "Local binary pattern method and feature shape extraction for detecting butterfly image," *GEOMATE Journal*, vol. 15, no. 50, pp. 127–133, 2018.

[6] A. Fadlil, A. Maftukhah, T. Sutikno *et al.*, "Butterfly image identification using multilevel thresholding segmentation and convolutional neural network classification with AlexNet architecture," *International Journal of Computing and Digital Systems*, vol. 15, no. 1, pp. 1–9, 2024.

[7] L. Zhu and P. Spachos, "Butterfly classification with machine learning methodologies for an android application," in *2019 IEEE Global Conference on Signal and Information Processing*. IEEE, 2019, pp. 1–5.

[8] B. Liang, S. Wu, K. Xu, and J. Hao, "Butterfly detection and classification based on integrated YOLO algorithm," in *Genetic and Evolutionary Computing: Proceedings of the Thirteenth International Conference on Genetic and Evolutionary Computing, November 1–3, 2019, Qingdao, China 13*. Springer, 2020, pp. 500–512.

[9] L. Nie, K. Wang, X. Fan, and Y. Gao, "Fine-grained butterfly recognition with deep residual networks: A new baseline and benchmark," in *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 2017, pp. 1–7.

[10] B. D. Wilts, P. Pirih, and D. G. Stavenga, "Spectral reflectance properties of iridescent pierid butterfly wings," *Journal of Comparative Physiology A*, vol. 197, pp. 693–702, 2011.

[11] J. M. Medina, J. A. Diaz, E. M. Valero, J. L. Nieves, and P. Vukusic, "Detailed experimental characterization of reflectance spectra of Sasakia Charonda butterfly using multispectral optical imaging," *Optical Engineering*, vol. 53, no. 3, p. 033111, 2014.

[12] P. Chatelain, G. Delmaire, A. Alboody, M. Puigt, and G. Roussel, "Semi-automatic spectral image stitching for a compact hybrid linescan hyperspectral camera towards near field remote monitoring of potato crop leaves," *Sensors*, vol. 21, no. 22, p. 7616, 2021.

[13] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel, "Joint unmixing and demosaicing methods for snapshot spectral images," in *ICASSP 2023*, 2023, pp. 1–5.

[14] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2013.

[15] XIMEA, *xiQ, USB 3.0 camera series,Technical Manual*. XIMEA, 2017.

[16] IITA, "International Institute of Tropical Agriculture," https://www.iita.org/, accessed on 07th February, 2024.

[17] S. Girard and J. Saracco, "Supervised and unsupervised classification using mixture models," *EAS Publications Series*, vol. 77, pp. 69–90, 2016.

[18] H. Kamel, D. Abdulah, and J. M. Al-Tuwaijari, "Cancer classification using Gaussian Naive Bayes algorithm," in *2019 International Engineering Conference (IEC)*, 2019, pp. 165–170.

[19] H. Nazeer, N. Naseer, A. Mehboob, M. J. Khan, R. A. Khan, U. S. Khan, and Y. Ayaz, "Enhancing classification performance of fNIRS-BCI by identifying cortically active channels using the Z-score method," *Sensors*, vol. 20, no. 23, 2020.

TABLE I
Confusion Matrix of Test Data

	Result of all samples				**Result of samples in VIS**				**Result of samples in NIR**			
Target Sp. / Predict Sp.	***HM***	***DC***	***AO***	***AE***	***HM***	***DC***	***AO***	***AE***	***HM***	***DC***	***AO***	***AE***
HM	19	0	1	0	8	0	1	0	11	0	0	0
DC	0	15	5	2	0	6	2	2	0	9	3	0
AO	1	9	16	0	1	6	7	0	0	3	9	0
AE	4	0	2	22	3	0	2	10	1	0	0	12
Sum (% good pred)	19/24 (79 %)	15/24 (62 %)	16/24 (67%)	22/24 (92 %)	8/12 (67 %)	6/12 (50 %)	7/12 (67 %)	10/12 (83 %)	11/12 (91 %)	9/12 (75 %)	9/12 (75 %)	12/12 (100 %)
Proposed method	**Accuracy : 72/96 (75%)**				**Accuracy : 31/48 (65%)**				**Accuracy : 41/48 (85%)**			
GNB-scoring [18]	Accuracy : 53/96 (55%)				Accuracy : 26/48 (54%)				Accuracy : 27/48 (56%)			
ZS-scoring [19]	Accuracy : 52/96 (54%)				Accuracy : 22/48 (46%)				Accuracy : 30/48 (62%)			
MLR-scoring [3]	Accuracy : 33/96 (34%)				Accuracy : 16/48 (33%)				Accuracy : 17/48 (35%)			
KNN-scoring [2]	Accuracy : 50/96 (52%)				Accuracy : 25/48 (52%)				Accuracy : 25/48 (52%)			