Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 338–347

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Insect Classification Framework based on a Novel Fusion of High-level and Shallow Features

Raye Haarika[a], Tina Babu[a,*], Rekha R Nair[b]

[a] *Department of Computer Science and Engineering, Dayananda Sagar University, School of Engineering, Bengaluru, India*
[b] *Department of Computer Applications, Dayananda Sagar University, School of Engineering, Bengaluru, India*

Abstract

The largest and most diversified group of organisms are insects. Since insects play a crucial role in many ecosystems, it must be precisely identified for efficient management. However, it is difficult and labor-intensive to identify insect species. Due to advancements in deep learning, computer vision, and sensor technologies, there is rising interest in image-based systems for quick, accurate identification. This study investigates a novel hybrid feature set of shallow features from the tiger beetle dataset, such as texture and wavelet features and high-level features from SqueezeNet. The tiger beetle insect is classified into the cicindelini and collyridini classes with a random forest classifier with 97.65 percent accuracy using this hybrid feature set, which considers the texture and structural characteristics of the insect. Thus the technique provides insight into various features and indicates promising future directions for image-based insect identification and species classification relevant to Computer Science, Agriculture, and Ecology research.

Keywords: image classification; insect identification; machine learning; transfer learning; features extraction.

1. Introduction

Nearly 60 percent of the 1.82 million documented species of plants and animals are insects, making them the largest and most diverse class of living things [1]. As pollinators, biologic controls (i.e., carnivores and parasitic organisms) of crop pests, or as food sources for humans and other animals, some insects are loosely referred to as "beneficial." On the other hand, we refer to some insect species as "pests" because of their interactions with people. These insects may contaminate stored grains, harm crops, and decorative plants, or otherwise impair the beauty of our surroundings, or our ability to produce food [2].

* Corresponding author. Tel.: +91-080-4909-2924.
E-mail address: b.tina-cse@dsu.edu.in

10.1016/j.procs.2023.01.016

While other species have been found and described using morphology and morphometrics, others have been found and characterized using molecular methods. However, morphology-based identification has difficulties recognizing cryptic and fewer species, and the application of molecular approaches necessitates a high degree of ability, as well as a significant financial and time commitment [3, 4]. The difficulty of morphological identification is exacerbated by the quantity of visually comparable categories needed to calculate inter-class similarity. Traditional species identification, therefore, requires an in-depth understanding of each species. Additionally, identifying species requires a lot of time. Artificial intelligence-based picture categorization has grown in significance due to the lack of specialized knowledge, the high cost of species identification and the time commitment [5].

A huge portion of research has been conducted on automating the insect detection and classification process to solve these issues, especially in crop/stored-grain insect pests, [6, 7], invasive insect species, and so on. Convolutional neural networks (CNN) for the image categorization are among the newest advances in machine learning models [8, 9, 10]. These models can categorize objects by identifying apparent and invisible traits to the human eye. [11, 12] To know and understand the characteristics that are arranged into a feature vector, an arbitrary length vector that gathers all the useful properties in defining the object under analysis to improve the model's performance, these models, on the other hand, inevitably require a large amount of training data from each class. [13][14]. As a result, to obtain high validation accuracy, ultra-specific classifications such as insect species classification, which exhibit inter-genus morphological commonalities, require a large amount of image data. [15, 16].

Photographs of Tiger beetles require specialized knowledge to classify. These images were compiled from a range of sources and settings. Each image on the beetle has a unique noise level, a different size, and a different zoom level. Therefore, it is challenging to use computational models to categorize these images. The images that need to be classified contain too much noise and fluctuation for conventional computational methods to handle.

1.1. Objectives

The framework for categorizing tiger beetle insects (cicindelidae) into the two species cicindelini and collyridini is presented in this research. This research seeks to synthesize and analyze critical studies on image-based insect categorization systems. The study's main contributions are:

- A large variety of tiger beetle dataset is analyzed.
- The tiger beetle images are pre-processed and segmented from the background for the texture and wavelet based feature extraction
- Makes use of a hybrid feature set consisting of
  - Deep learned features from the SqueezeNet with transfer learning.
  - Texture based features
  - Wavelet based features
- The hybridized features are further fused together for the classification into two classes with a random forest classifier.

2. Methodology

The proposed classification framework consists of two phases of feature extraction to increase the accuracy of the classification, as illustrated in Fig. 1. In phase-I, the images are preprocessed, and feature extraction is carried out from the segmented images. In phase II, the high-level SqueezeNet features are extracted with the help of the transfer learning method. Finally, classification is carried out on the fused features.

2.1. Image Pre-processing

The obtained image data is typically disorganized and derived from various sources. To decrease complexity and improve feature accuracy, the photos are preprocessed. The photos are first transformed from RGB to grayscale, after which they are denoised using a Gaussian and non-local means filter [17] to reduce the unwanted features. Compared to an RGB image, a grayscale image achieves a higher object recognition accuracy. Mexican Hat filters are used to

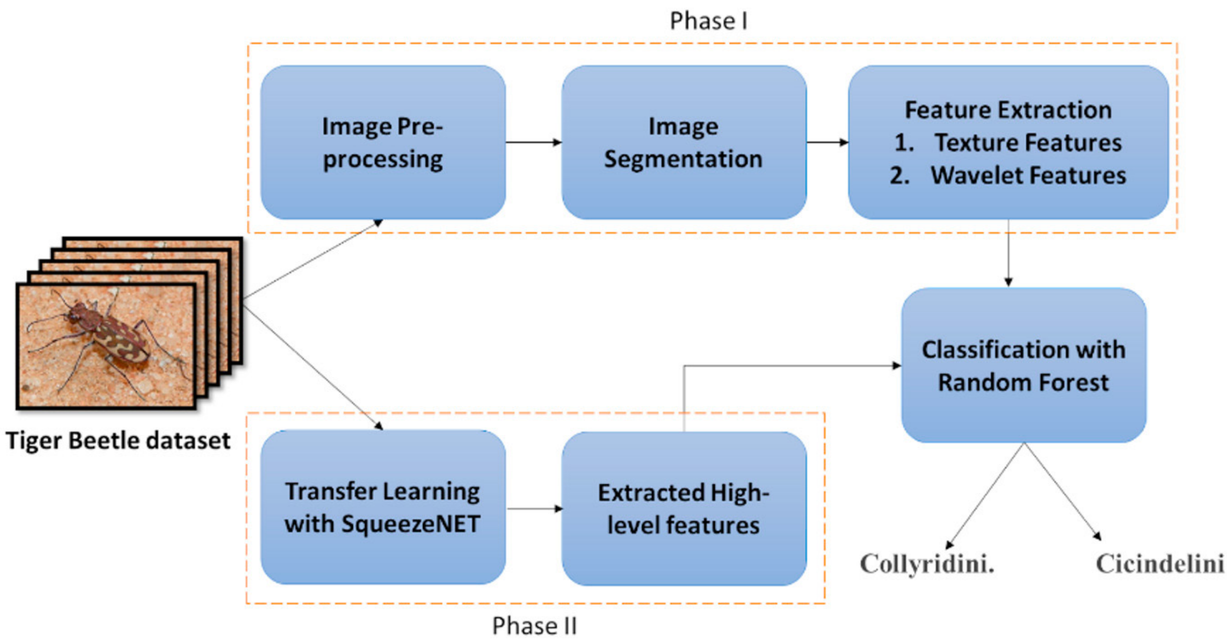


Fig. 1. The proposed classification framework.

sharpen the denoised images, bringing out the edges and minute features of the acquired image [18]. With the aid of the linear contrast adjustment approach, the images which have been sharpened and denoised are treated for enhancement [19]. Fig. 2 displays the tiger beetle photos in their raw and pre-processed forms.

2.2. Image Segmentation

A zero crossing in a derivative expression retrieved from the image helps find out the insect edges from an image. To segment the body of the insect from the preprocessed image, utilized canny edge detection [20] which is a zero-crossing based edge detection method. The segmentation of the tiger beetle image is illustrated in Fig. 3.

2.3. Feature Extraction

Feature extraction helps in extracting the essential features from a large number of features. Feature extraction methods are applied to segmented images to make the classification method accurate and more effortless.

2.3.1. Texture Features

The spatial fluctuation of pixel brightness intensity is one of an image's textural characteristics. This paper used the Gabor filter with the CS-LBP operator to extract the texture characteristics [21]. Local texture features can be extracted using the symmetric and central CS-LBP approach. Since the texture properties of the photographs do not significantly change when the light and posture vary, the technique of using the CS-LBP operators to extract local texture features from the photos is effective. Gabor filter with CS-LBP operator aids in texture extraction in various ways. Fig. 4 illustrates how the Tiger beetle photos were used to extract CS-LBP features.

2.3.2. Wavelet Features

Discrete wavelet transform is a multi-resolution technique to extract high frequency(HF) and low frequency(LF) features from images which help in edge details more accurately[22, 23, 24]. Instead of considering the original segmented images, the wavelet provides the different wavelet decomposition coefficients for extracting features. These

HF and LF features provide minute edge details of the image. The feature extraction results using Wavelets are shown in Fig. 5.

$$Gray\_image = LL + LH + HL + HH \tag{1}$$

where, LL(low low), LH(low high), HL(high low) and HH(high high) decomposed wavelet coefficients of gray image as shown in Figure 5.

2.3.3. High-level SqueezeNet features

Phase II of feature extraction uses SqueezeNet, a pre-trained CNN model that is easy to use and performs well, to extract the features from the image dataset.[25]. Due to the channel projection constraint, this model significantly reduces parameter space and computational cost (squeeze layers). Additionally, the model contains shortcut connections for indemnity mapping that, like residual networks, enable proper training of deep network models. This system consists of "fire modules," which divide the input map into two-channel sets after passing through a bottlenecked channel-projection layer. The first channel set is grown using a 3X3 convolution, while the second is expanded using a channel projection. The resulting convolution map is universally mean into a 512-vector and then given to a fully connected layer with 2048 units. The SqueezeNet model has 68 layers and a 227x227x3 input image size. The model extracts picture features for subsequent processing and comprises 14 convolution processing block elements with various resizing and resampling actions.

2.4. Classification

The binary classification of the picture dataset, which contains two classes of tiger beetles, is the proper activity given that the task falls under the supervised learning category and calls for the detection of dependencies between the goal prediction output and the input features. Texture determines how persistent patterns are in an image. Texture can be selected as a potential feature to extract different properties of each class (which helps in class separation) in the current dataset because combinations of pattern variations can be utilized to classify tiger beetles to the genus level. To classify the insects, the phase-I and phase-II data are combined to create hybrid features, which are then processed using a Random Forest classifier [26]. The Gini index is used for root node identification.

3. Experimental Results

3.1. Image Datasets

The tiger beetle dataset, which is used in [27] was utilized for the experiments and evaluation of the proposed technique. The dataset consists of images from field studies, wildlife and nature photographers using different camera types with other image quality, and tiger beetle books and websites. These pictures were collected from various sources. This dataset includes tiger beetles from the cicindelini and collyridini families. The cicindelini category has 261 images, while the collyridini category has 166.

3.2. Performance Measures

Through the use of benchmark metrics like Accuracy, Sensitivity, Specificity, and F-score, the classification outcomes are assessed. The confusion matrix is used to calculate these metrics. The following equations provide descriptions of these.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Sensitivity/Recall = \frac{TP}{TP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

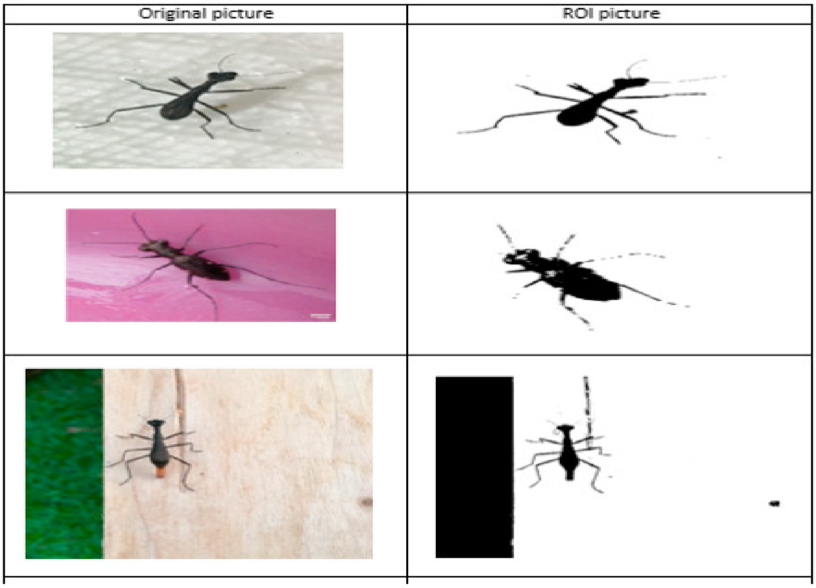


Fig. 2. Pre-processing of the Tiger beetle images.

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

$$F\_Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{6}$$

3.3. Results and Discussion

3.3.1. Experimental Setup

The dataset is divided into training and testing with respective weights of 80% and 20%. The window for LBP feature extraction was specified with 24 and 20. In the forest, ten decision trees were employed to create the RF classifier. Deep artificial neural networks require an extensive training data set to learn efficiently and avoid over-fitting. The dataset was expanded through picture augmentation using various processing techniques or combinations of methods, including random rotation, shifts, shear, and flips, among others, to increase the number of images artificially. To extract the features from the SqueezeNet, the epochs were set to 100.

3.3.2. Results of Image Pre-processing

The pre-processing steps help in detecting the shape of the tiger beetle insects. Fig. 2 shows the raw and pre-processed images of the tiger beetle images. It is clear from the visualization results that Fig. 2 identifies the insect from the original image after the pre-processing step.

3.3.3. Results of Image Segmentation

The segmentation process involves separating the insect's back end with edge details. The segmentation of the tiger beetle image is illustrated in Fig. 3. The visualization results help identify the tiger beetle insect with structure and edge details from its back end.

3.3.4. Results of Feature Extraction

Feature extraction results includes results using texture features and Wavelet features.

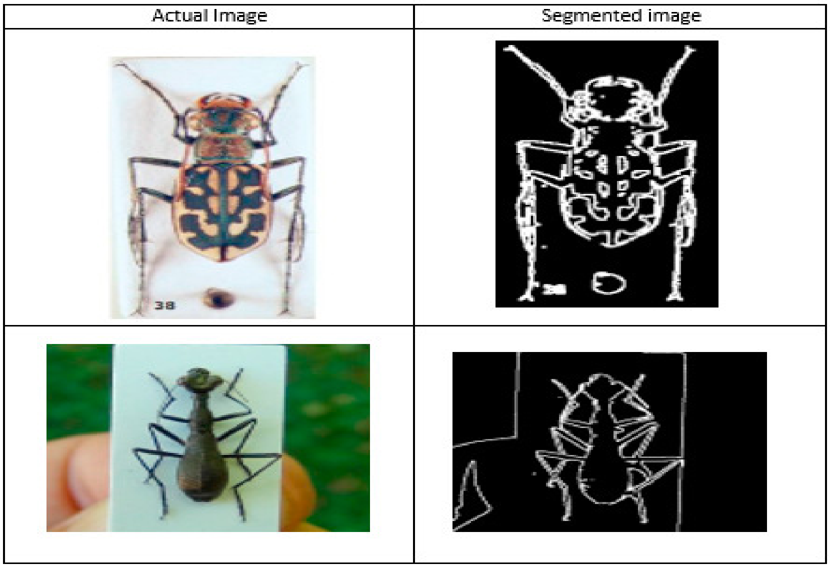


Fig. 3. Image segmentation of the Tiger beetle images.

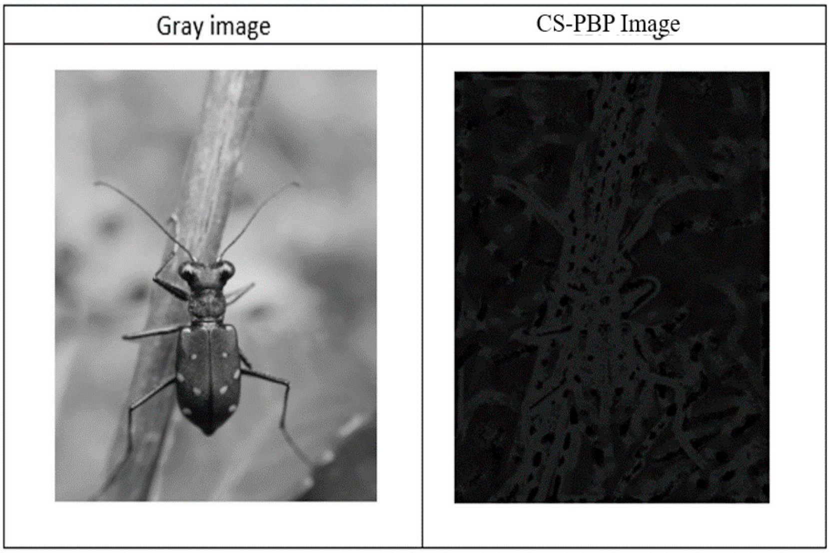


Fig. 4. CS-LBP feature extraction of the Tiger beetle images.

- Results of Texture features
  Gabor filter with CS-LBP operator provides multiple directions of the texture extraction. The illustration of CS-LBP feature extraction of the Tiger beetle images is shown in Fig.4.
- Results of Wavelet features The minute edge details are obtained by the wavelet process as shown in Fig. 5 to support the edge details clearly. The results of feature extraction using Wavelets shown in Fig. 5.

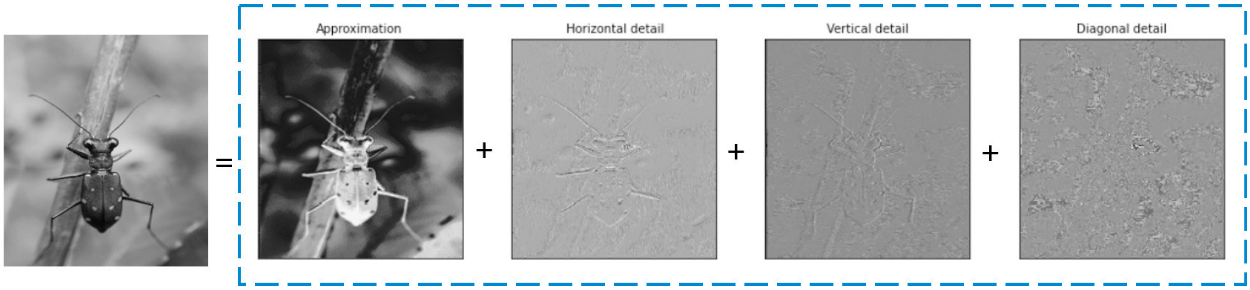


Fig. 5. High and low wavelet features of insect.

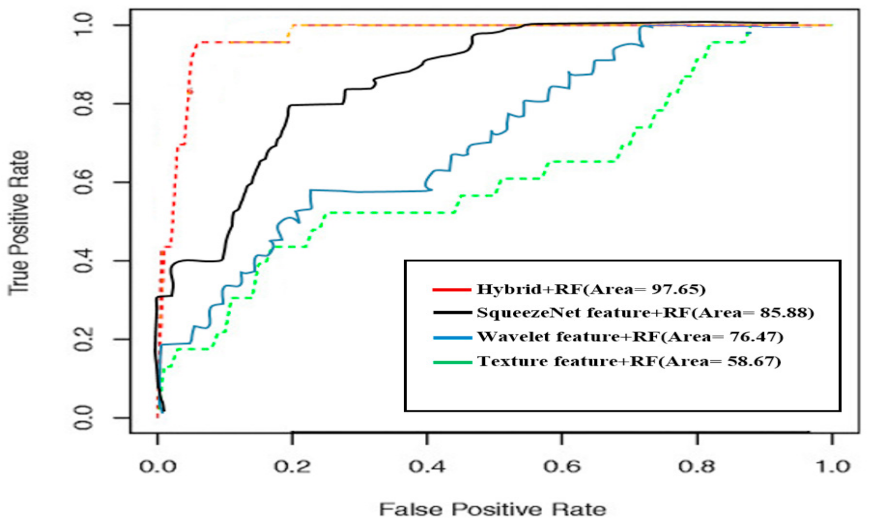


Fig. 6. The ROC curve of the proposed four classification models.

3.3.5. *Results of classification*

Approximately 80% of the total number of photos in each group were placed in the training set, and the remaining 20% were placed in the test set, which was created from the original image collection. The hybrid features extracted from the images were trained using the RF classifier, and performance metrics were assessed.

Table 1. Performance evaluation measures of the proposed model.

Performance Measures	SqueezeNet Features	Texture Features	Wavelet Features	Hybrid Features
Accuracy	85.88	58.67	76.47	**97.65**
Sensitivity	0.88	0.63	0.81	0.98
Specificity	0.82	0.53	0.70	0.97
F_Score	0.88	0.63	0.81	0.98

Table 1 illustrates the suggested model's performance. With the hybrid features, the RF classifier has classified the tiger beetle images with an average accuracy of 97.65%. The individual features accuracy for SqueezeNet, texture, and wavelet features are 85.88%, 58.67%, and 76.47%, respectively. From Table 1, it is observed that SqueezeNet features classify the images with higher accuracy than the other features. Also, when the fused features are taken together, it results in a comparable outcome compared to the individual features. Thus, the proposed framework has shown considerable improvement in identifying the two classes of tiger beetles.

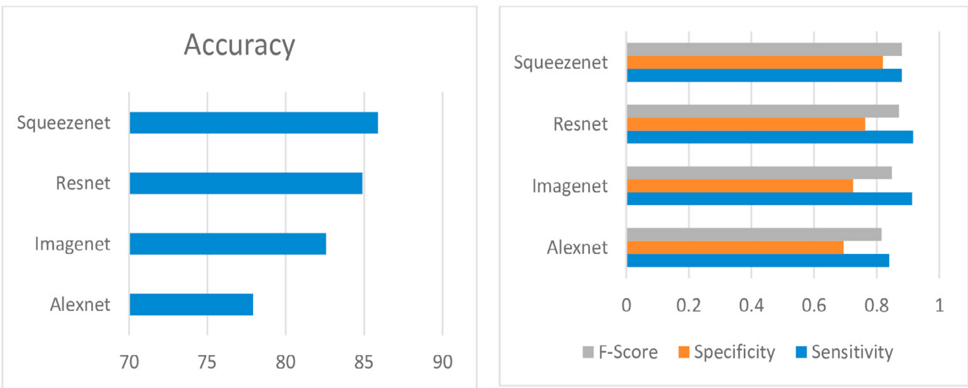
The performance of the proposed model is evaluated using the ROC curve as shown in Fig. 6. ROC curve shows different classifier performances such as texture feature + RF, wavelet feature + RF, SqueezeNet feature + RF, and hybrid feature + RF. It is clear from Fig. 6 that the hybrid feature + RF provides more accuracy than other models.

Table 2. Evaluation of the model with hybrid features on various classifiers.

Performance Measures	SVM Classifier	MLP Classifier	DT Classifier	RF Classifier
Accuracy	96.47	87.06	94.12	**97.65**
Sensitivity	0.98	0.89	0.94	0.98
Specificity	0.94	0.84	0.94	0.97
F_Score	0.97	0.90	0.95	0.98

The comparison of the suggested model with other classifiers (Support Vector Machines (SVM), Multilayer Perceptron (MLP), and Decision Tree (DT)) is illustrated in Table.2. The hybrid features are considered for the classification. By analyzing the accuracy measures for various classifiers, the RF classifier shows the best with an accuracy of 97.65%. The least performance was observed for the MLP with 87.06% as there are many layers to process. SVM has an accuracy of 96.47% which is almost comparable with the RF classifier as it is a binary classification. DT classifier has also given a not bad accuracy of 94.12%. However, by analyzing all the performance measures like accuracy, sensitivity, specificity, and F-score, the RF classifier gives the best performance for the hybrid feature set.

Fig. 7 shows the supremacy of the high-level features extracted from the Sqeezenet compared to the other existing CNN models. The accuracy with SqueezeNet high-level features is 85.88% which is higher than the other CNN models' features. F-score, sensitivity, and specificity of the SqueezeNet also show comparable results with other networks.



(a) (b)

Fig. 7. Comparison of the high-level features of various CNN models with RF classifier a) Accuracy b) F-score, specificity, sensitivity

Table 3. Comparison of the proposed framework with existing model.

Framework	Technique	Classification	Accuracy
Abeywardhana et. al [27]	SqueezeNet + SVM	9 class classification	60.2%
Proposed	SqueezeNet + Texture + RF	2 class classification	97.65%

The proposed framework can be compared with [27], as the same dataset has been utilized and is illustrated in Table 3. In the pre-existing work, high-level features were extracted from SqueezeNet and classified with an SVM classifier for nine classes with an accuracy of 60.2%. At the same time, the proposed model has considered the texture features and the high-level features from SqueezeNet and attained an accuracy of 97.65% for the two-class classification. Thus it can be observed that with the fused features along with the RF classifier, the accuracy is comparable with the existing model.

Thus by analyzing the results obtained, the proposed model with the hybrid features has shown good performance for the tiger beetle two-class classification with an RF classifier. The hybrid features show incremental improvement rather than the individual features alone.

4. Conclusion and Future research directions

An essential step in many agricultural and environmental management systems is insect identification. This study uses image processing and machine learning approaches to divide the dataset of tiger beetles into cicindelini and collyridini classes. 1) The pre-processing and image segmentation methods for the beetle photos were thoroughly examined in this research. 2) A unique hybrid feature set composed of the salient shallow texture and wavelet characteristics from the beetle photos and the transfer learned features or high-level features from the SqueezeNet model. 4) The extracted hybrid feature set and a random forest classifier were used to divide the beetle photos into two categories. The hybrid feature set, which consists of high-level characteristics, texture, and wavelet features, provided the best accuracy (97.65%) for the classification with the RF classifier rather than individual features when evaluating the model with the tiger beetle picture dataset. To demonstrate the model's superiority over the current framework, comparisons were made. Grading the cicindelini and collyridini classes of tiger beetles is one of several intriguing future directions. Additionally, a variety of strategies can be used to increase the accuracy of this kind of dataset.

References

[1] P. W. Price, R. F. Denno, M. D. Eubanks, D. L. Finke, I. Kaplan, Insect ecology: behavior, populations and communities, Cambridge University Press, 2011.
[2] D. C. K. Amarathunga, J. Grundy, H. Parry, A. Dorin, Methods of insect image capture and classification: A systematic literature review, Smart Agricultural Technology 1 (2021) 100023.
[3] A. J. Alvarez, E. A. Hernández-Delgado, G. A. Toranzos, Advantages and disadvantages of traditional and molecular techniques applied to the detection of pathogens in waters, Water Science and Technology 27 (3-4) (1993) 253–256.
[4] S. E. Hamsher, M. M. LeGresley, J. L. Martin, G. W. Saunders, A comparison of morphological and molecular-based surveys to estimate the species richness of chaetoceros and thalassiosira (bacillariophyta), in the bay of fundy, PLoS One 8 (10) (2013) e73521.
[5] N. MacLeod, Automated taxon identification in systematics: theory, approaches and applications, Crc Press, 2007.
[6] K. S. Banga, N. Kotwaliwale, D. Mohapatra, S. K. Giri, Techniques for insect detection in stored food grains: An overview, Food Control 94 (2018) 167–176.
[7] M. C. F. Lima, M. E. D. de Almeida Leandro, C. Valero, L. C. P. Coronel, C. O. G. Bazzo, Automatic detection and monitoring of insect pests—a review, Agriculture 10 (5) (2020) 161.
[8] T. Babu, T. Singh, D. Gupta, S. Hameed, Optimized cancer detection on various magnified histopathological colon images based on dwt features and fcm clustering, Turkish Journal of Electrical Engineering & Computer Sciences 30 (1) (2022) 1–17.
[9] T. Babu, D. Gupta, T. Singh, S. Hameed, M. Zakariah, Y. A. Alotaibi, Robust magnification independent colon biopsy grading system over multiple data sources, CMC-Computers Materials & Continua 69 (1) (2021) 99–128.
[10] T. Babu, T. Singh, D. Gupta, S. Hameed, Colon cancer prediction on histological images using deep learning features and bayesian optimized svm, Journal of Intelligent & Fuzzy Systems (Preprint) (2021) 1–12.
[11] P. Caramazza, A. Boccolini, D. Buschek, M. Hullin, C. F. Higham, R. Henderson, R. Murray-Smith, D. Faccio, Neural network identification of people hidden from view with a single-pixel, single-photon detector, Scientific reports 8 (1) (2018) 1–6.
[12] M. Jangblad, Object detection in infrared images using deep convolutional neural networks (2018).
[13] A. Kamilaris, F. X. Prenafeta-Boldú, A review of the use of convolutional neural networks in agriculture, The Journal of Agricultural Science 156 (3) (2018) 312–322.
[14] J. Wang, L. Perez, et al., The effectiveness of data augmentation in image classification using deep learning, Convolutional Neural Networks Vis. Recognit 11 (2017) 1–8.
[15] O. L. Hansen, J.-C. Svenning, K. Olsen, S. Dupont, B. H. Garner, A. Iosifidis, B. W. Price, T. T. Høye, Species-level image classification with convolutional neural network enables insect identification from habitus images, Ecology and evolution 10 (2) (2020) 737–747.
[16] H.-W. Pang, P. Yang, X. Chen, Y. Wang, C.-L. Liu, Insect recognition under natural scenes using r-fcn with anchor boxes estimation, in: International Conference on Image and Graphics, Springer, 2019, pp. 689–701.

[17] M. Wang, S. Zheng, X. Li, X. Qin, A new image denoising method based on gaussian filter, in: 2014 International Conference on information science, electronics and electrical engineering, Vol. 1, IEEE, 2014, pp. 163–167.

[18] F. Jin, D. Feng, Image registration algorithm using mexican hat function-based operator and grouped feature matching strategy, PloS one 9 (4) (2014) e95576.

[19] T. Kumar, S. Bhushan, S. Jangra, A brief review of image quality enhancement techniques based multi-modal biometric fusion systems, in: International Conference on Advanced Informatics for Computing Research, Springer, 2018, pp. 407–423.

[20] M. A. Ansari, D. Kurchaniya, M. Dixit, A comprehensive analysis of image edge detection techniques, International Journal of Multimedia and Ubiquitous Engineering 12 (11) (2017) 1–12.

[21] B. Yuan, B. Xia, D. Zhang, Polarization image texture feature extraction algorithm based on cs-lbp operator, Procedia computer science 131 (2018) 295–301.

[22] R. R. Nair, T. Singh, Multi-sensor medical image fusion using pyramid-based dwt: a multi-resolution approach, IET Image Processing 13 (9) (2019) 1447–1459.

[23] R. R. Nair, T. Singh, An optimal registration on shearlet domain with novel weighted energy fusion for multi-modal medical images, Optik 225 (2021) 165742.

[24] R. R. Nair, T. Singh, R. Sankar, K. Gunndu, Multi-modal medical image fusion using lmf-gan-a maximum parameter infusion technique, Journal of Intelligent & Fuzzy Systems (Preprint) (2021) 1–12.

[25] A. Ahmed, Pre-trained cnns models for content based image retrieval', International Journalof Advanced Computer Science and Applications 12 (7) (2021).

[26] A. Sarica, A. Cerasa, A. Quattrone, Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: a systematic review, Frontiers in aging neuroscience 9 (2017) 329.

[27] D. Abeywardhana, C. Dangalle, A. Nugaliyadde, Y. Mallawarachchi, An ultra-specific image dataset for automated insect identification, Multimedia Tools and Applications 81 (3) (2022) 3223–3251.