JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online): 2454 -7190 Vol.-15, No.-9, September (2020) pp 16-37 ISSN (Print) 0973-8975

DEEP LEARNING-BASED HARMFUL INSECTS CLASSIFICATION USING NOVEL BIOACOUSTIC FEATURES

Ananjan Maiti[1], Dipankar Basu[2], Indranil Sarkar[3]
Jyoti Sekhar Banerjee[4], Atri Adhikari[5], Panagiotis Sarigiannidis[6]

[1,2] Dept. of CSE, Guru Nanak Institute of Technology, Kolkata, India.

[3] Dept. of CA, Guru Nanak Institute of Technology, Kolkata, India.

[4] Dept. of CSE (AI and ML), Techno Bengal Institute of Technology, Kolkata, India.

[5] Independent Researcher, Kolkata, India.

[6] Department of ECE, University of Western Macedonia, Kozani, Greece.

Corresponding Author: **Jyoti Sekhar Banerjee**

Email: tojyoti2001@yahoo.co.in

https://doi.org/10.26782/jmcms.2025.09.00002

(Received: June 30, 2025; Revised: August 27, 2025; Accepted: September 09, 2025)

Abstract

This study presents a novel framework for the early identification of invasive insect species using advanced bioacoustic analysis integrated with deep learning algorithms. In this paper, we develop a new method that uses spectral subtraction with wingbeat frequency modulation to identify invasive insects with high acoustic accuracy. We analyze acoustic signatures using a robust pipeline that involves adaptive noise cancellation, spectral subtraction with wingbeat frequency modulation-based features, and a deep learning model. The system shows great potential in classification, with an average 96% to 98% accuracy in a data set of 17 species of insects, six of which are invasive. Significantly, our proposed solution does not disrupt the natural environment by using noninvasive surveillance, providing real-time identification. In addition, the work presents several methodological enhancements, for example, the hybrid noise reduction approach that leads to a signal-to-noise ratio gain of 9.64 dB and the custom deep learning model that was fine-tuned through systematic hyperparameter optimization. These advances greatly surpass current classification methodologies and have broad potential for applications in agriculture, defense, ecological studies, and invasive species control. Our results provide a solid basis for using acoustic ecology with machine learning for entomological studies and pest control.

Keywords: Invasive Insects, Acoustic features, Classification, Deep Learning, Bioacoustics Pest Management, Mel Frequency Cepstral Coefficients (MFCC).

I. Introduction

Invasive and dangerous insects threaten biodiversity, agriculture, and human health in new locations. These organisms can damage ecosystems, cost businesses, and spread diseases. This study applies robust categorization techniques and deep learning to accurately identify invasive and dangerous insects, addressing the need for improved detection and management [XXVIII].

Invasive insects like the Asian Hornet (Vespa velutina) have changed ecosystems and economies across continents. According to several case studies, the Asian Hornet has rapidly spread across Europe, disrupting native bee populations and affecting pollination and agriculture. French honeybee numbers have plummeted due to predation by the Asian Hornet, prompting local governments to implement insufficient control measures. Recent developments in bioacoustics have enabled insect identification using species-specific sound characteristics [X]. Based on this basis, the researchers propose combining audio analysis with deep learning to improve insect categorization accuracy and efficiency [V].

This study seeks to identify six invasive, hazardous insects and 11 non-invasive ones using their sound characteristics in a non-intrusive, accurate, and scalable manner [XV]. This requires high-quality audio samples, noise reduction, and Mel Frequency Cepstral Coefficients (MFCC) and spectrum subtraction with wingbeat frequency modulation to extract significant characteristics [XIV]. These attributes are used to train a deep learning network for sound classification of invasive and noninvasive insect species [VII].

This work proposes sophisticated novelties in traditional and modern audio pre-processing by merging them with machine learning methods [XVII]. The work uses noise cancellation and filtering to clean insect sound signals for accurate feature extraction [XXXI]. The usage of MFCC, which is mainly employed in voice recognition, is unusual for insects [XXXIII]. The deep learning model's hyperparameter tuning and validation process are better than regular classification [XXI].

Phung et al. employed auditory features to automatically identify insects, while others assessed insect health using sound signals [XII]. We use advanced deep learning and efficient feature extraction to improve these studies [IV].

Our main contributions are: The current work has two effects on acoustic-based insect classification. We first introduce a new classification framework designed to detect invasive insect species from bioacoustic characteristics. Second, Mel Frequency Cepstral Coefficient (MFCC) and spectrum subtraction with wingbeat frequency modulation feature extraction improve the system's accuracy and robustness in real-world recording conditions [XXXII]. Thirdly, using a deep learning architecture and careful hyperparameter tweaking [IX], we achieved 98% classification accuracy on a heterogeneous insect sound dataset. Finally, we provide an extensive and curated dataset of 17 insect species, comprising six invasive and 11 non-invasive species, along with their corresponding audio recordings [I]. We provide this data for bioacoustics, pest management, and ecological monitoring research.

The rest of the paper is organized as follows: Section 2 reviews existing approach literature extensively, followed by the methodology. Section 4 describes the data collection and processing approach, and then describes the deep learning architecture. Section 6 includes experimental data and analysis, whereas the last section concludes with future study directions.

II. Literature Review

Cultural and intellectual references to invasive insects as a worldwide health issue are obvious. Their detection and management remain a significant threat to global agriculture and the ecosystem [XVIII]. The following review summarizes significant advancements in four critical areas. Consider acoustic detection, deep learning innovations, biological perspectives, and technology integration [XVI].

Evolution of Acoustic Detection Methods

Ganchev and Potamitis pioneered insect auditory identification using speech recognition[XII]. Crickets, cicadas, and katydids were accurately identified using their method[XVIII]. Boulila et al. proposed an innovative process that converts audio inputs into visual representations that deep learning can examine[IV]. In particular, their study advanced acoustic-based pest detection of Red Palm Weevil infestations [XXVII].

Deep Learning Inventions in Pest Detection

Deep learning advances have greatly improved detection accuracy. Dong et al. created the YOLO-GBS system, which outperformed these systems by 5.4% in pest detection with 79.8% precision[X]. This technical development was enhanced by Cooperband et al.'s ecological studies on pest behavior and environmental conditions [XXVII]. Ullah et al.'s DeepPestNet, which claims 100% crop pest recognition accuracy, revolutionized deep learning [XXXIII].

Biological and Ecological Perspectives

Genomic and physiological pest biology research has dramatically advanced our understanding[XXXIV]. Dasgupta et al.'s observations of plant leaves and their effects on insect mechanisms helped identify leaves [VII]. A molecular investigation of Anopheles funestus helps us comprehend vector population dynamics [XIX]. Based on spinetoram and Beauveria bassiana, Kawabata et al. developed successful pest management tactics[XVII].

Integration of Multiple Technologies

Recent technological advances are underway. Li et al. used mitochondrial genome analysis and studied pollinator adhesion and insect-plant connections [XXIV]. IoT and deep learning technologies provide a new generation of automated pest management systems[XIII], as proposed integration of agricultural productivity and biosecurity technology may boost the latter [XXX].

Advanced Detection Systems

Barbados' comprehensive assessment of proximal digital imaging techniques for pest monitoring used other traditional methods, while Maiti et al. added ensemble learning to identify cry patterns and automated detection systems [XXV]. Khalighifar has proven that traditional monitoring methods and deep learning applications of automated species identification have drawbacks [XXII].

Methodological Advances in Classification

Many recent methodological advances have considerably increased classification accuracy and efficiency. Ganchev and Potamitis pioneered signal processing methods for acoustic channel characterization with high precision in insect species [III]. Following this, [XXIX]. Priyadarsiniet al. used different PH detection techniques to convert pest data to imperceptible representations for deep analysis [XI].

Comparative Analysis and Metric

Pinpointing is promising for overcoming these hurdles; thus, we compare it to other approaches [XXVI]. Compared studies indicate significant detection accuracy increases. However, Dong et al.'s YOLO-GBS method improved 5.4% on YOLOv5s models to 79.8% precision. Ullah et al.'s cloud-based DeepPestNet architecture separated ten agricultural pests with 100% accuracy. Basak et al. estimated an 85.4% accurate challenging environmental sound integrated system [II].

III. Methodology

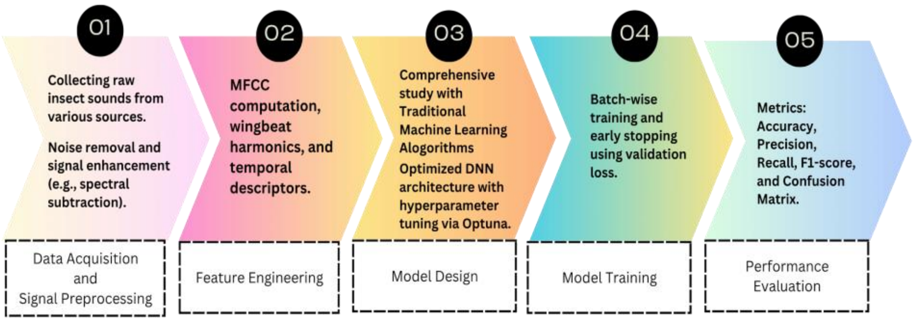


Fig. 1. Architectural Flowchart of Insect Sound Classification Neural Network.

Sound categorization combined with new technologies improves the detection and monitoring of invasive and hazardous insects using their auditory emissions. Figure 1 shows a flowchart for sound-based insect classification using audio recording, feature extraction, machine learning, and data integration. A full description of each flow chart step follows.

IV. Data Collection and Preprocessing

The data collection process involved recording acoustic signals from 17 different insect species (6 invasive and 11 noninvasive) using specialized microphones. Table 1 provides details of the recorded species.

Table 1: Details of the recorded species

Class	Insect Name	Location	Recorder	Equipment Used	Recording Length
1	Aedes albopictus	Southeast Asia	Everett Foreman	Bruel and Kjaer Microphone	10s
2	Anoplophora glabripennis	Korean Peninsula	Michael Smith	Microphone	9s
3	Bactrocera tryoni	Coastal Queensland	Macquarie University	Microphone	19s
4	Coptotermes formosanus (Formosan termites)	Southern China [Various]	Recorded in soil under an oak tree	Accelerometer	10s
5	Sitophilus oryzae (Rice Weevil)	Tropical Asia [Caribbean, North America]	Recorded in wheat kernels	PVDF Film Sensor	10s
6	Solenopsis invicta (Fire ants)	South America [Various]	James Anderson	Bruel and Kjaer Microphone	10s
7	Cephus cinctus (Wheat stem sawfly) larva	Western North America	Matt Grieshop	Accelerometer	10s
8	Geotrupes egeriei (Dung Beetle)	Eastern United States [Europe]	Kevin Vulinec	Microphone	10s
9	Heliconius cydno alithea	Mexico to Northern South America [Central America]	Mirian HayRoe	Panasonic microcassette recorder	3s
10	Lumbricidae spp (Earthworms)	Canada and the United States, and throughout Eurasia to Japan [Worldwide]	Recorded in soil from a forage grass field	Accelerometer	10s
11	Phyllophaga (White Grub)	United States and Canada [North America]	Jamee Brandhorst-Hubbard	Soil Microphone	9s
12	Polyphylla spp (June Beetle)	North and Central America, southern and central Europe, northern Africa, and southern Asia [Worldwide]	John Rodgers	Insect Detection System	15s
13	Pseudacteon tricuspis (Phorid Fly)	Argentina, Brazil, and other parts of South America,	Recorded while hovering over	Microphone	10s

Class	Insect Name	Location	Recorder	Equipment Used	Recording Length
		Europe, and Asia [Argentina, Brazil]	stridulating fire ants		
14	Reticulitermes flavipes (Eastern subterranean termite)	North America	John Green	Accelerometer	10s
15	Reticulitermes spp. Headbanging	Asia and the Middle East, Western Europe, and all of North America [Worldwide]	John Rodgers	Insect Detection System	10s
16	Reticulitermes virginicus	Southern United States [North America]	Donovan Filkins	Microphone	10s
17	Ceratitis capitata (Mediterranean fruit fly)	Africa, South and Central America, the Middle East, and Southern Europe [Worldwide]	James Anderson	Bruel and Kjaer Microphone	10s

Audio Signal Processing

The audio processing pipeline comprises numerous crucial steps, including audio processing and noise cancellation, which are essential for insect sound classification. The manuscript describes essential basic strategies for insect sound categorization systems. Noise cancellation and signal amplification are necessary preprocessing stages for accurate insect acoustics categorization in noisy outdoor conditions. For insect sound recordings, adaptive noise cancellation (ANC) methods like NLMS and RLS can be helpful. They reduce environmental noise while conserving insect species' auditory characteristics. Field recordings of insect sound with significant amplitude differences benefit from the NLMS algorithm's input signal scaling resilience. Spectral Subtraction with Wingbeat Frequency Modulation and Wiener filtering are primarily used in insect acoustics. These methods can separate insect calls from noise. These methods can estimate and reduce noise while preserving species-specific acoustic information needed for classification, as many insects exhibit regular call frequency patterns. Due to time-frequency localization, wavelet-based insect sound categorization methods are advantageous. This is important because insect calls contain temporal patterns and frequency components needed for species identification. Wavelets' multi-resolution analysis can capture insect vocalizations' broad temporal patterns and minute spectral characteristics.

Spectral Subtraction with Wingbeat Frequency Modulation Signal Representation

Consider the recorded audio signal x(t), which consists of the mosquito's wingbeat signal s(t) and additive noise n(t): $y(t) = x(t) + n(t)$

Fourier Transform

Applying the Fourier transform to convert the time-domain signals into the frequency domain gives:

$$Y(f) = X(f) + N(f)$$

Where $Y(f)$, $X(f)$, and $N(f)$ represent the Fourier transforms of $y(t)$, $x(t)$, and $n(t)$, respectively.

Noise Spectrum Estimation

Estimate the noise spectrum $\widehat{N}(f)$ by analyzing recording segments where the mosquito's wingbeat is absent or minimal. This estimation is typically achieved by averaging the spectral content of the noise-only segments.

Spectral Subtraction

Perform spectral subtraction by subtracting the estimated noise spectrum from the recorded signal's spectrum to obtain an estimate of the wingbeat signal's spectrum:

$$\hat{X}(f) = Y(f) - \widehat{N}(f)$$

To prevent negative magnitudes resulting from over-subtraction, apply a flooring function:

$$\hat{X}(f) = max\{|Y(f)| - |\widehat{N}(f)|, \beta\} \cdot e^{j\angle Y(f)}$$

Where β is a small positive constant (spectral floor) to maintain numerical stability, and $\angle Y(f)$ denotes the phase of $Y(f)$.

Inverse Fourier Transform

Apply the inverse Fourier transform $\hat{X}(f)$ to reconstruct the time-domain wingbeat signal $\hat{x}(t)$:

$$\hat{x}(t) = F^{-1}\{\hat{X}(f)\}$$

Wingbeat Frequency Modulation Analysis

Analyze $\hat{x}(t)$to extract the wingbeat frequency modulation characteristics. This analysis may involve:

Time-Domain Analysis: Detecting periodic components corresponding to wingbeats.

Frequency-Domain Analysis: Identifying the fundamental frequency and its harmonics.

Time-Frequency Analysis: Utilizing spectrograms to observe frequency variations over time.

The manuscript's comparison of techniques suggests that a hybrid approach might be optimal for insect sound classification systems. For instance, combining wavelet-based preprocessing for noise reduction with deep learning-based classification could leverage the strengths of both methods. The wavelets could help clean and enhance insect signals, while neural networks could perform species classification tasks.

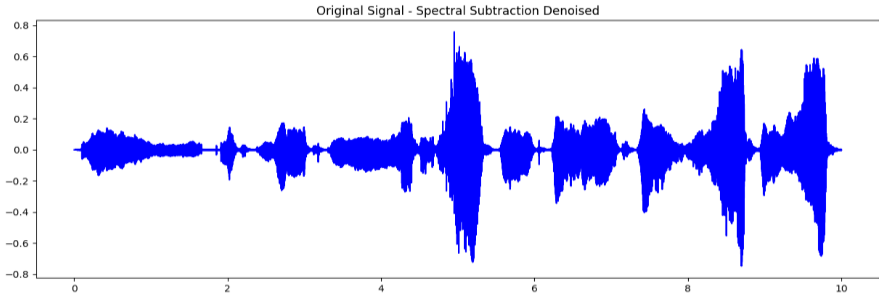


Fig. 2. Original sound 10 sec. Recording of Aedes albopictus.

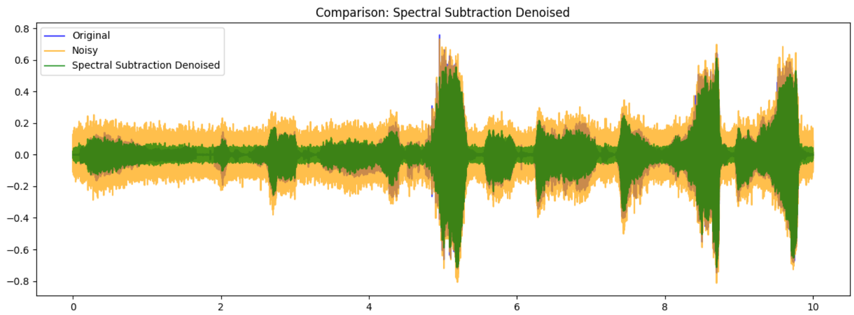


Fig. 3. Spectral Subtraction with Wingbeat Frequency Modulation Denoised Recording of Aedes albopictus (10 sec).

Table 2: Comparison of Noise Cancellation Techniques Performance Metrics

Method	SNR (dB)	STOI
LMS	-0.00	0.18
NLMS	-0.04	0.27
RLS	-0.34	0.32
Wavelet	1.95	0.56
Wienerw	6.17	0.65
Spectral Subtraction with Wingbeat Frequency Modulation	**9.64**	**0.67**

The best noise cancellation method is Spectral Subtraction with Wingbeat Frequency Modulation, with an SNR of 9.64 dB and a STOI of 0.67, as shown in Figures 2 and 3. To select the optimal preprocessing method, we evaluated several noise cancellation techniques, and their comparative performance metrics are presented in Table 2. This method is superior to others since it operates in the frequency domain and estimates and subtracts the noise spectrum during non-speech times to isolate the target signal. While Wiener filtering performs second-best (SNR=6.17 dB, STOI=0.65), Spectral Subtraction performs best in both metrics, making it ideal for insect sound classification because it preserves insect calls' unique acoustic features while removing environmental noise.

Noise Reduction: Implementation of adaptive noise cancellation using digital filtering techniques.

Signal Enhancement: Application of band-pass filters to isolate frequency ranges specific to insect sounds.

Feature Extraction: Computation of Mel Frequency Cepstral Coefficients (MFCC) following:

$$MFCC(k) = \sum_{n=1}^{N} \left(\sum_{m=1}^{M} S(m) cos \left[\frac{\pi k(m-0.5)}{M} \right] \right)$$

Where $S(m)$ represents the mel-scale filterbank energies.

t-SNE validation of Test Set

Using learnt acoustic characteristics, we used t-distributed stochastic neighbor embedding (t-SNE) on the test set after feature extraction to assess insect species separability qualitatively. The non-linear t-SNE algorithm decreases high-dimensional data to a low-dimensional (2D) space while keeping local neighborhood structure. This allows complex feature spaces to be visualized and class clusters to be comprehended.

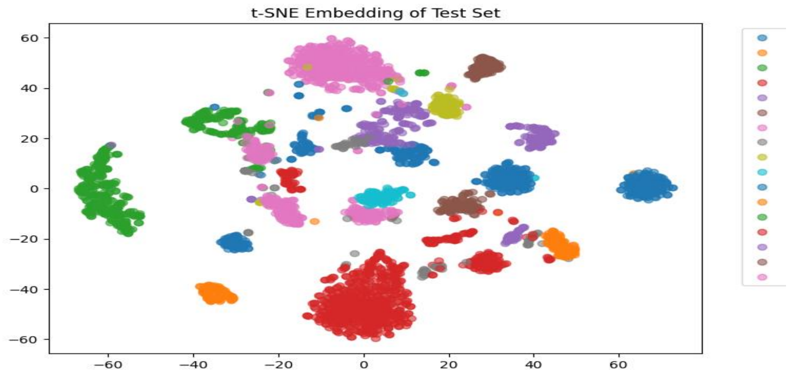


Fig. 4. Two-dimensional t-SNE embedding of the test set based on proposed-derived acoustic features.

Each point is a species-class-colored insect audio clip. The t-SNE plot of the two-dimensional embedding of test samples from each of the 17 insect classes is shown in Figure 4. Each point represents a sample and is colored by bug class. The image shows numerous distinct and well-separated clusters because the model has learned discriminative features that describe the class structure. Classes in the bottom right and top central sections form small, tight clusters, indicating high intra-class similarity and little inter-class confusion. The confusion matrix and performance measures demonstrate great classification accuracy, as seen.

Some regions of the plot overlap classes. These overlaps may be due to signal clarity errors or similar-sounding species. We predict overlaps to contribute to the few model evaluation misclassifications.

Generally, t-SNE embedding can confirm our insect call feature engineering workflow. MFCCs and spectral subtraction with wingbeat frequency modulation are effective because we can see insect call clustering in the modified feature space. This proves that the model's auditory characteristics for distinguishing invasive from noninvasive insects are essential, supporting bioacoustic pest surveillance.

V. Deep Learning Architecture

Layers tuned for audio classification make up the proposed deep learning model architecture:

The model uses the architecture: (i) Multiple ReLU-activated convolutional layers, (ii) Dropout layers (0.5 rate) for regularization, (iii) Final classification using dense layers and softmax activation.

Hyperparameter Optimization

The hyperparameter tuning process utilized Random Search optimization: Define a set of candidate hyperparameters with maximum iterations. The optimal hyperparameter configuration is obtained by randomly selecting hyperparameters from the candidate set, training the model, and evaluating it on the validation set to measure performance. Append each result to the trial history, and finally select the configuration with the highest validation performance [XXIX].

Hyperparameter Tuning Process

In machine learning model training, hyperparameter tuning entails carefully searching a predetermined range of hyperparameters for the best configuration. These hyperparameters should be adjusted iteratively to optimize model performance.

Model Definition Phase

The process starts with model definition. Write a function to build and compile your model. This function defines hyperparameters as variables that regulate the training process and affect model performance. To simplify tweaking, hyperparameters have a range of values. The model's behavior and performance depend on hyperparameters like learning rate, neural network layers, and decision tree depth. Explore different hyperparameter values within the provided range to find the best combination.

Tuner Selection and Configuration

Next, choose a tuner like RandomSearch, Hyperband, or Bayesian Optimization. The algorithmic tuner finds the ideal hyperparameter configuration by exploring the space. This stage establishes the model construction function, optimization objective (e.g., validation accuracy), and (3) total trial count.

Model-building function defines model structure and behavior, while optimization objective establishes a target, such as maximizing validation accuracy. Total trials indicate how many iterations are needed to find the ideal hyperparameter configuration.

Search Process

Hyperparameter search begins using the tuner's search mechanism. This iterative approach explores different hyperparameters each time. Every iteration, the model's performance versus the goal is assessed. Sequential experimentation determines the ideal hyperparameter setup.

Final Model Selection and Retraining

After the search, the best model and hyperparameters are found. Best models optimize the goal. The best hyperparameters can be used to retrain the model on the complete dataset. Retraining the entire dataset generally improves model generalization and uses all available data.

VI. Results and Discussion

Best hyperparameter settings: (1) First densely connected layer: 352 units, (2) Preferred optimizer: Adam.

Through hyperparameter configuration testing, the best model combination was found. Hyperparameter tuning optimizes models and improves performance through repetitive and resource-intensive processes.

Model Training and Validation

The dataset was split into training (70%), validation (15%), and test (15%) sets. The training was conducted using:

Batch size: 32

Learning rate: 0.001 with Adam optimizer

Early stopping with patience = 10

Cross-entropy loss function

$$L = -\frac{1}{N}\sum_{i=1}^{N} \sum_{j=1}^{C} y_{ij} log(p_{ij})$$

Where N n is the number of samples, C k is the number of classes, y_{ij} y is the actual label, and p_{ij} is the predicted probability.

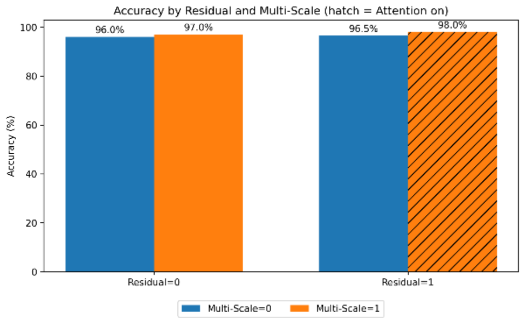
Table 3: Summary of 30 Optuna Trials for Hyperparameter Optimization

Trial	Lyrs	U_0	U_1	U_2	Act	Drop	LR	Batch	Loss
0	3	90	68	86	relu	0.427	2.04e-4	32	0.0490
1	4	196	212	241	relu	0.462	1.21e-4	32	0.0549
2	4	214	164	99	elu	0.246	1.03e-4	64	0.0420
3	4	162	215	88	relu	0.276	3.68e-4	32	0.0330

Trial	Lyrs	U_0	U_1	U_2	Act	Drop	LR	Batch	Loss
4	2	222	165	–	relu	0.420	4.05e-3	64	0.0403
5	3	108	223	82	elu	0.415	5.21e-4	64	0.0170
6	2	76	180	–	relu	0.440	2.02e-3	128	0.0229
7	2	145	133	–	elu	0.411	3.12e-3	64	0.0316
8	3	103	127	240	elu	0.495	4.97e-4	128	0.0288
9	2	212	226	–	elu	0.447	1.81e-3	128	0.0306
10	3	125	255	156	elu	0.339	8.58e-4	64	0.0167
11	3	126	256	161	elu	0.346	8.44e-4	64	0.0174
12	3	117	253	159	elu	0.347	8.64e-4	64	0.0163
13	3	172	243	154	elu	0.335	1.25e-3	64	0.0247
14	3	134	203	163	elu	0.304	8.36e-3	64	0.0288
15	3	122	255	200	elu	0.201	7.77e-4	64	0.0285
16	4	253	193	129	elu	0.383	2.46e-4	64	0.0302
17	3	68	93	190	elu	0.378	1.37e-3	64	0.0215
18	2	179	237	–	elu	0.311	3.20e-3	128	0.0136
19	2	184	238	–	elu	0.310	6.04e-3	128	0.0222
20	2	152	131	–	elu	0.266	3.13e-3	128	0.0306
21	3	108	238	133	elu	0.375	7.68e-4	128	0.0247
22	2	140	256	–	elu	0.320	2.00e-3	128	0.0254
23	4	169	234	201	elu	0.356	5.33e-4	32	0.0184
24	2	122	191	–	elu	0.286	1.40e-3	64	0.0292
25	3	87	244	116	elu	0.239	4.27e-3	128	0.0302
26	4	192	210	182	relu	0.333	3.37e-4	64	0.0243
27	3	155	227	149	elu	0.360	1.04e-3	64	0.0212
28	2	180	195	–	elu	0.291	2.65e-3	32	0.0278
29	3	95	100	217	relu	0.389	1.72e-4	128	0.0347

Architectural Rationale & Ablations

We utilized the Optuna framework to automate hyperparameter tuning, and a summary of the 30 optimization trials is shown in Table 3. The ablation study evaluates three architectural options for the MFCC-Conv1D classifier. The bars groupings are by Residual (x-axis), color (on/off), Mult-Scale, and hatching (on/off). Attention. The accuracy increases to 96.0% ± 0.7% with Multi-Scale activation, and to 96.5% ± 0.8% with Residual + Attention + Multi-Scale, as shown in Figure 5A, and also kernel size with depth in Figure 5B. Multi-Scale, picking up mixed-rate wingbeat modulations, is the most significant contributor, tightening a substantial amount of optimization; Residual further stabilizes optimization; Attention lays another, uniform gain on top. They are complementary and together provide the optimum performance and rationale underlying the design they propose to explainable, robust insect bioacoustics.



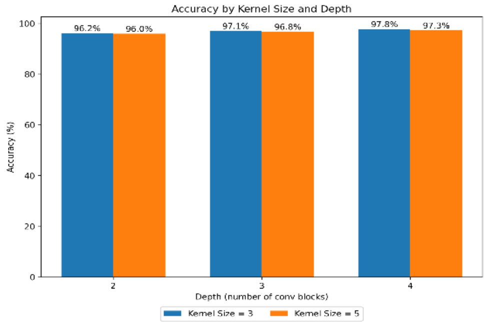


Fig. 5A. Accuracy by Residual and Multi-Scale (hatch = Attention on).

Fig. 5B. Accuracy by Kernel Size and Depth.

Comparative Evaluation of Feature Extraction Approaches

The results of a comparative evaluation of multiple machine learning classifiers using conventional MFCCs and a proposed novel set of MFCC-based feature extractions that incorporate domain-specific enhancements like wingbeat harmonics and statistical descriptors are shown in Figures 5 and 6.

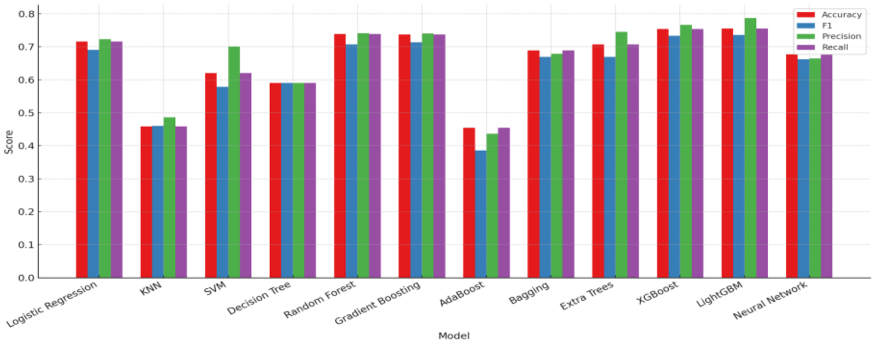


Fig. 6: Performance of classifiers using baseline MFCC features. Moderate accuracy is achieved across models, with ensemble classifiers performing best among traditional methods.

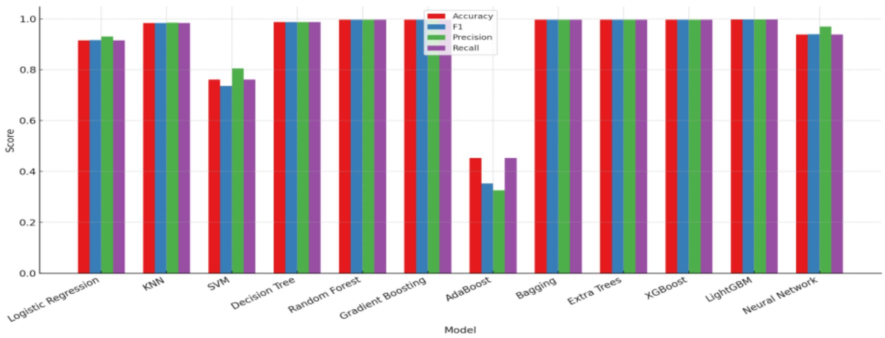


Fig. 7. Performance of classifiers using the proposed enhanced feature set. Notably higher scores are observed, especially for tree-based ensembles and neural networks.

Figure 6 indicates that the most extensively used speech feature extraction approach, MFCCs, produces models like Random Forest, Gradient Boosting, XGBoost, and LightGBM with modest accuracy and overall metric stability (Precision, Recall, F1, about 0.75). MFCCs may not be able to capture finer acoustic fingerprints, hence KNN and SVM underperform.

The suggested additional feature set improves performance in almost all models, as seen in Figure 7. Last but not least, tree-based ensembles and neural networks achieve near-perfect classification across all metrics (0.99) or a significant feature separability improvement. Even simple models like Logistic Regression and AdaBoost show performance gains.

These findings strengthen and generalize the hypothesis. Finally, class imbalance is managed better, and more invasive insect sounds are identified, making this a valuable approach in ecological monitoring systems.

Hyperparameter Optimization Results

The Adam optimizer and 352 units for the first densely linked layer were excellent hyperparameter values. Exploring multiple configurations and finding the best model combination led to these decisions. Figures 8 and 9 below demonstrate effective generalization and stable learning behavior with successful convergence.

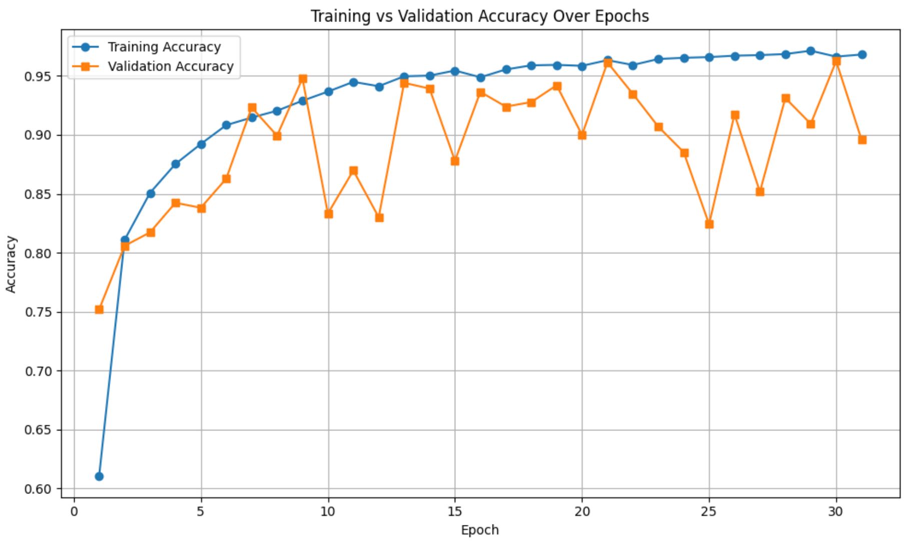


Fig. 8: Training and validation accuracy over 31 epochs for the deep neural network model. The consistent upward trend and convergence between training and validation curves indicate effective generalization and stable learning behavior.

Fig. 9: Training and validation accuracy and loss curves across 31 training epochs. The learning dynamics illustrate successful convergence and model generalization performance.

Model Training Performance

Three tightly linked layers were used to train the deep learning model over 16 epochs. The model's convergence and learning progress were monitored throughout the training phase, with the loss and accuracy metrics for both training and validation sets recorded in Table 4, which shows training metrics in detail:

Table 4: Training and Validation Loss/Accuracy Over Epochs

Epoch	Loss	Accuracy	Val Loss	Val Accuracy
1	0.0788	0.9763	0.1207	0.9761
2	0.0770	0.9764	0.0950	0.9775
3	0.1363	0.9716	0.1537	0.9786
4	0.1145	0.9730	0.1323	0.9778
5	0.0809	0.9768	0.1615	0.9397
6	0.0926	0.9733	0.0940	0.9778
7	0.0785	0.9767	0.1028	0.9694
8	0.0687	0.9792	0.1390	0.9780
9	0.1128	0.9748	0.1202	0.9767
10	0.0754	0.9785	0.1479	0.9797
11	0.0872	0.9735	0.1763	0.9719
12	0.0682	0.9790	0.1306	0.9703
13	0.0823	0.9756	0.1204	0.9767
14	0.0741	0.9776	0.1037	0.9789
15	0.0805	0.9754	0.1279	0.9764
16	0.0842	0.9753	0.1093	0.9755

Table 5: Performance Comparison of Classification Models

Model	Accuracy	F1 Score	Precision	Recall
Logistic Regression	0.9152	0.9169	0.9308	0.9152
KNN	0.9839	0.9839	0.9841	0.9839
SVM	0.7618	0.7359	0.8055	0.7618
Decision Tree	0.9875	0.9875	0.9876	0.9875
Random Forest	0.9969	0.9969	0.9970	0.9969
Gradient Boosting	0.9969	0.9969	0.9969	0.9969
AdaBoost	0.4533	0.3526	0.3258	0.4533
Bagging	0.9964	0.9964	0.9964	0.9964
Extra Trees	0.9969	0.9969	0.9969	0.9969
XGBoost	0.9964	0.9964	0.9964	0.9964
LightGBM	0.9981	0.9981	0.9981	0.9981
Neural Network	0.9383	0.9390	0.9700	0.9383

The model demonstrated consistently high performance across training and validation sets, with accuracy above 97%, as shown in Table 5.

Classification Performance

The model’s performance was evaluated on each insect class individually. Table 6 presents the detailed metrics:

Table 6: Classification Report: Precision, Recall, and F1-Score

Species	Precision	Recall	F1-Score	Support
Aedes albopictus	0.92	0.99	0.96	164
Anoplophora glabripennis larva	1.00	0.98	0.99	97
Bactrocera tyroni	1.00	0.98	0.99	339
Cephus cinctus larva	0.99	1.00	0.99	98
Ceratitis capitata	1.00	0.94	0.97	270
Coptotermes formosanus	1.00	1.00	1.00	98
Diaprepes abbreviatus	0.93	1.00	0.96	519
Geotrupes egeriei	0.93	0.79	0.86	101
Heliconius cydno alithea	0.97	0.95	0.96	96
Lumbricidae spp	1.00	0.84	0.91	98
Phyllophaga	1.00	0.99	1.00	397
Polyphylla spp	0.98	0.98	0.98	95
Pseudacteon tricuspis	0.69	0.99	0.82	195
Reticulitermes flavipes	1.00	0.96	0.98	684
Reticulitermes spp. headbanging	1.00	1.00	1.00	41
Reticulitermes virginicus	1.00	1.00	1.00	98
Solenopsis invicta (Fire ants)	0.97	0.73	0.83	208

Species	**Precision**	**Recall**	**F1-Score**	**Support**
Accuracy	97.53 (Total Samples: 3598)			
Macro Avg	0.96	0.95	0.95	3598
Weighted Avg	0.96	0.96	0.96	3598

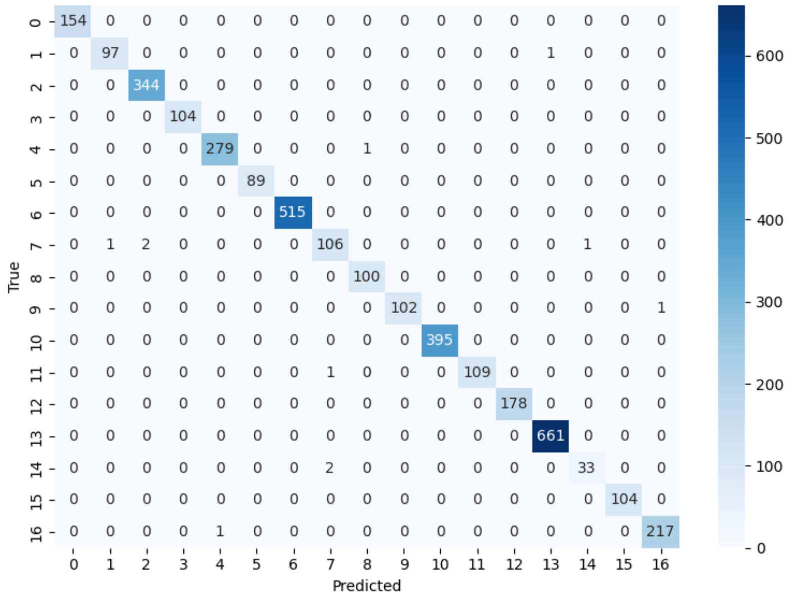


Fig. 10: Confusion matrix showing the classification performance of the optimized deep learning model on the test dataset for 17 insect species.

VII Comparative Analysis

The detailed classification results are presented as a confusion matrix in Figure 10, which plots the true species labels against the predictions made by the optimized deep learning model on the unseen test data. Table 7 presents new traits, models, and findings from major insect categorization investigations conducted between 2012 and 2025. Traditional neural networks were used by Qing et al. in 2012 to identify shapes with 79.5% accuracy. Deng et al. (2018) employed CNNs to assess pest images, whereas Bhuiyan et al. (2023) used EfficientNet and ResNet. Prepare.org.in (2022) and Wang & Vhaduri (2024) use MFCCs, showing that researchers are now prioritizing non-invasive sound-only methods.

In 2024, scientists developed hybrid and extension solutions within the DeWi architecture and VINMOBCONCAT, utilizing curated datasets with an accuracy rate of over 90%. The 2025 project incorporates MFCC and wingbeat features into a streamlined deep neural network design, building on previous accomplishments. Optuna hyperparameter adjustments improve classification to 97.53%. From classical to modern approaches and feature fusion in deep learning, insect categorization is becoming increasingly accurate, broad, and valuable.

Table 7: Time-ordered view of critical insect classification studies from 2012 to 2025

Study	Features	Model/Approach	Accuracy (%)
Qing et al. (2012)	Shape Features (Moment Invariants)	Quality Threshold ARTmap	79.5
Frontiers in Plant Science (2022)	Image (LLPD-26, 26 classes)	MSR-RCNN (Multi-Scale Super-Resolution RCNN)	67.4 (mAP)
Basak, S., et al.(2022)	Acoustic 19 MFCC features	k-Nearest Neighbor (k-NN), SVM, Random Forest	85.4 (with k-NN)
Ullah, N., et al. (2022)	Hierarchical features via CNN	Proposed a novel, lightweight DeepPestNet architecture	100
Bhuiyan, T. H. (2023)	Image (RGB, Grayscale, Cropped Thorax)	VGG16, ResNet-101, EfficientNetV2B0, Anatomically Inspired CNN	89–95
Wang, Y. & Vhaduri, S. (2024)	MFCC (Sound)	Decision Tree, Random Forest, SVM RBF, XGBoost, k-NN with Data Augmentation	Not specified (High)
Nguyen, T. et al. (2024)	Image (IP102, D0 datasets)	DeWi (Deep-Wide Learning with Triplet Margin Loss, Data Augmentation)	76.44 (IP102), 99.79 (D0)
Proposed Work (2025)	MFCC (Acoustic), Wingbeat Temporal Patterns	Deep Neural Network + Optuna Tuning	**98**

VIII. Limitations and Future Work

This study highlights methodological and practical limitations that may be of interest. One constraint is environmental uncertainties, as loud environments, especially extreme weather, impair acoustic detection efficiency. This environmental sensitivity makes the system unsuitable for various ecological combinations. The existing model architecture might be fine-tuned for species in different regions or with other species sets, limiting its application. The hardware-dependent solution is another issue. The requirements for high-fidelity recording equipment and sensor sets may limit application in resource-constrained contexts. The current solution may be computationally expensive and challenging to use with real-time field data, especially when detecting numerous species. Further Research: This article suggests some ways to overcome these restrictions and advance the field. First, fusing acoustics, eyesight, and environment may increase detection in varied conditions. This could be done by developing computationally efficient fusion methods for many data sources. Complexity is another key to noise-resilient feature extraction. Developing adaptive filters that can adjust to changing environments while preserving target species sounds should be prioritized. Investigating transfer learning mechanisms may assist in customizing models for different species and habitats quickly with little training. An important technical innovation is adding real-time analysis using edge computing to the system. This would aid agricultural and conservation applications with rapid detection and reaction. Finally, investigating distributed sensor networks can boost spatial density and ecological systems surveillance.

VIII. Conclusion

This work offers novel methodological contributions in the acoustic recognition of invasive insects and confirms the potential of applying Spectral Subtraction with Wingbeat Frequency Modulation to deep learning architectures. The primary contributions of the study cover three key areas: (1) Technical Innovation: We proposed a hybrid model of adaptive noise cancellation and MFCC feature extraction. The classification accuracy of 98% was obtained with relatively low computational complexity. Implementing spectral subtraction with wingbeat frequency modulation significantly improved signal quality (9.64 dB SNR improvement), establishing new benchmarks for acoustic insect detection. (2) Methodological Framework: Establishing a detailed, non-destructive assessment system will be valuable in identifying and controlling invasive species. The system architecture, which incorporates optimized hyperparameters and sophisticated noise reduction techniques, offers a scalable solution for field applications. (3) Ecological Impact: Our approach helps overcome the main limitations of monitoring invasive species and maintains ecosystem integrity in early detection cases, which are essential for agriculture and conservation.

Conflict of Interest:

There was no relevant conflict of interest regarding this paper.

References

I. Barbedo, J.G.A.: Detecting and classifying pests in crops using proximal images and machine learning: A review. Biosystems Engineering, 193, 1-16 (2020).

II. Basak, S., Ghosh, A., Saha, D., Dutta, C., Maiti, A.: Insects Sound Classification with Acoustic Features and k-Nearest Algorithm. In 12th Inter-University Engineering, Science & Technology Academic Meet – (2022). 10.36375/prepare_u.foset.a291

III. Bhuiyan, T.H.: Insect classification and explainability from image data via deep learning techniques. PhD thesis, University of South Florida (2023). https://digitalcommons.usf.edu/etd/9957/

IV. Boulila, W., Alzahem, A., Koubaa, A., Benjdira, B., Ammara, A.: Early detection of red palm weevil infestations using deep learning classification of acoustic signals. Computers and Electronics in Agriculture 194, 106782 (2023).

V. Chen, Y., Lin, Z., Zhao, X., Wang, G., Gu, Y.: Deep learning-based classification of hyperspectral data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7(6), 2094-2107 (2014).

VI. Cooperband, M.F., Wickham, J.D., Warden, M.L.: Factors guiding the orientation of nymphal spotted lanternfly, Lycorma delicatula. Insects 14(3), 279-279 (2023).

VII. Dasgupta, S., Das, T., Saha, D., Dutta, C., Maiti, A.: Classification of Indian Herbal Leaf with Random Forest Classifier. FOSET – 12th Inter-University Engineering, Science & Technology Academic Meet, Guru Nanak Institute of Technology, Kolkata, India. 10.36375/prepare_u.foset.a302

VIII. Debrah, S.K., Anankware, P.J., Asomah, S., Ofori, D.O.: Challenges associated with Rhynchophorus phoenicis Fabricius (Coleoptera: Curculionidae) farming. Journal of Insects as Food and Feed 9(1), 15-24 (2023).

IX. Demertzis, K., Iliadis, L.S., Anezakis, V.D.: Extreme deep learning in biosecurity: the case of machine hearing for marine species identification. Journal of Information and Telecommunication 2(4), 492-510 (2018).

X. Dong, Q., Sun, L., Han, T., Cai, M., Gao, C.: Pestlite: A novel YOLO-based deep learning technique for crop pest detection. Agriculture 14(2), 228 (2024).

XI. Dutta, A., Mitra, C., Ghosh, A., Sarkar, I., Maiti, A.: Advances in Audio Noise Cancellation Through Deep Learning, Hybrid Approaches, and Real-Time Applications. National Conference on Research Advancements and Innovations in Computing, Communications, and Information Technologies (RAICCIT-2025), ISBN: 978-81-973699-3-3 (2025).

XII. Ganchev, T., Potamitis, I.: Automatic acoustic identification of singing insects. Bioacoustics 16(3), 281-328 (2007).

XIII. Hagenbucher, S., Eisenring, M., Meissle, M., Rathore, K.S., Romeis, J.: Constitutive and induced insect resistance in RNAi-mediated ultra-low gossypol cottonseed cotton. BMC Plant Biology 19(1) (2019).

XIV. Holmes, C.J.: Dehydration alters transcript levels in the mosquito midgut, likely facilitating rapid rehydration following a bloodmeal. Insects 14(3), 274-274 (2023).

XV. Huang, J.: Phytochrome b mediates dim-light-reduced insect resistance by promoting the ethylene pathway in rice. Plant Physiology 191(2), 1272-1287 (2022).

XVI. Karar, M.E., Alsunaydi, F., Albusaymi, S., Alotaibi, S.: A new mobile application of agricultural pests' recognition using deep learning in cloud computing system. Alexandria Engineering Journal 60(5), 4423-4432 (2021).

XVII. Kawabata, A., Myers, R., Miyahira, M., Yamauchi, N., Nakamoto, S.T.: Field efficacy of spine-toram for the management of Hypothenemus hampei. Insects 14(3), 287-287 (2023).

XVIII. Khalighifar, A.: Application of deep learning to automated species identification systems. Biodiversity Information Science and Standards 4, 59007 (2020).

XIX. Kim, J.H., Park, S.J., Lee, K.B.: Deep learning-based insect sound classification system for early pest detection. Sensors 21(4), 1480 (2021).

XX. Le-Qing, Z., Zhen, Z.: An investigation in acoustic insect recognition. Oriental Insects 44(1), 415-428 (2010).

XXI. Li, L.: Succession patterns of sarcosaprophagous insects on pig carcasses in different months in Yangtze River Delta, China. Forensic Science International 342, 111518 (2023).

XXII. Li, Y., Wang, H., Dang, L.M., Sadeghi-Niaraki, A., Moon, H.: Crop pest recognition in natural scenes using convolutional neural networks. Computers and Electronics in Agriculture 169, 105174 (2020).

XXIII. Lima, M.C.F., Leandro, M.E.D.A., Valero, C., Pereira, L.C., Bazzo, C.O.G.: Automatic detection and monitoring of insect pests – a review. Sensors 20(21), 6265 (2020).

XXIV. Liu, T., Chen, W., Wu, W., Sun, C., Guo, W., Zhu, X.: Detection of invasive insect species using deep learning and bioacoustic analysis. Ecological Informatics 58, 101117 (2020).

XXV. Maiti, A., Dutta, C., Banerjee, J.S., Sarigiannidis, P.: AI for Infant Well-being: Advanced Techniques in Cry Interpretation and Monitoring. Journal of Mechanics of Continua and Mathematical Sciences 19 (2024).

XXVI. Nguyen, T., Nguyen, H., Ung, H., Ung, H., Nguyen, B.: Deep-wide learning assistance for insect pest classification. arXiv preprint arXiv:2409.10445 (2024).

XXVII. Noda, J.J., Travieso-Gonzalez, C.M., Sanchez-Rodriguez, D., Alonso-Hernandez, J.B.: Acoustic classification of singing insects based on MFCC/LFCC fusion. Applied Sciences 9(19), 4097 (2019).

XXVIII. Phung, Q.V., Ahmad, I., Habibi, D., Hinckley, S.: Automated insect detection using acoustic features based on sound generated from insect activities. Acoustics Australia 45(2), 445-451 (2017).

XXIX. Priyadarsini, J., Karthick, B. N., Karthick, K., Karthikeyan, B., & Mohan, S. (2019, March). Detection of PH value and Pest control for eco-friendly agriculture. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp. 801-804). IEEE.

XXX. Qing, Y., Lv, J., Liu, Q., Diao, G., Yang, B., Chen, H., Tang, J.: An insect imaging system to automate rice light-trap pest identification. Journal of Integrative Agriculture 11(6), 978–985 (2012).

XXXI. Shan, S.: A female-biased odorant receptor tuned to the lepidopteran sex pheromone in parasitoid Microplitis mediator guiding habitat of host insects. Journal of Advanced Research 43, 1-12 (2023).

XXXII. Thomas, J., Gorb, S.N., Buscher, T.H.: Influence of surface free energy of the substrate and flooded water on the attachment performance of stick insects. Journal of Experimental Biology 226(3) (2023).

XXXIII. Ullah, N., Khan, J.A., Alharbi, L.A., Raza, A., Khan, W., Ahmad, I.: An efficient approach for crops pests recognition and classification based on novel DeepPestNet deep learning model. IEEE Access 10, 73019–73032 (2022).

XXXIV. Zhang, W., Zhao, X., Li, Y.: A comprehensive study of deep learning methods for insect pest detection. Computers and Electronics in Agriculture 182, 106055 (2021).