Insects Sound Identification- Bayesian Classifier

K. Kumaresan[1] M.E., K. Dineshkumar[2] M.E.,
V. Senthilkumar[3] M.E.,
Department of Computer Science and Engineering
K .S .R College of Engineering,
1,2,3 Assistant professor, Tiruchengode, India

K. Sathish[4], T. Siva Sankar Narayanan[5],
S. Smily Deva Kirubha[6], K. Vignesh[7]
Department Of Computer Science and Engineering
K .S .R College of Engineering,
4,5,6,7 UG Students. Tiruchengode, India

***Abstract*—The project is work mainly focusses on developing a complete digitized tool that includes process of detecting and monitoring insect species that threaten biological resources, in both productive and native ecosystems, particularly for pest management and biosecurity. However, the researchers feel the task as a very difficult one to digitize the process of classification and measuring the influence of all possible combinations of environmental, climate and temporal variations on the behavior of insects simultaneously. The data relevant to the flying insects often changes over time, and classification of such data is a central issue. Review on several researchers had shown that the frequency spectrum of flying insects is likely to be within the range 200Hz to 1200 Hz and this has been used for classification of insects / Non insects. In this project work it is proposed to use Bayesian classifier to classify the acoustical sound and recognize the insects and non-insects.**

Keywords— Component; formatting; style; styling; insert (key words)

I. INTRODUCTION

The lifestyle of ecosystem is directly or indirectly associated with insects. This association results in both beneficial and unfavorable response to human life. The entomologists refer the relationship between human and insects as love-hate relationship. The unfavorable response raised by insects is diseases to humans and damages to food crops. At the same time beneficial insects helps in pollination of the food crops. The entomologists have tried to abolish the existence of unwanted species to certain extent with some blanket methods that are costly by creating environmental problems. This necessitated automated tools for detection and monitoring of insect species that threaten biological resources particularly for pest management and biosecurity. Smart sensors are being developed with the aim of protecting the ecosystem just by counting and classifying the species to eradicate the harmful insects deployed on target location.

A lot of novel and relevant applications have emerged with the development of tools and techniques of data mining methods for identification of insects. There exists a sensor that uses a laser and machine learning techniques to classify species of insects. This sensor is being considered as an important step in the development of intelligent traps. Such traps can attract and selectively capture insect species of interest, releasing all other species back into the environment. Then the unwanted insects are tracked and killed by the traps.

For most applications in smart sensors, there is not an assumption that the data is generated by a stationary stochastic process. The important feature that can be considered for the classification of insects is their acoustical sound. Finally, algorithms that classify data streams often assume that once a prediction is made, the actual labels are provided with some delay to assist in updating the classifier. In the case of smart sensors, these labels are rarely available, and the application must adapt to concept drift without assuming that labels from test cases are known.

II. EXISTING SYSTEM

A. Definition of Existing System

Many hardware devices are available to identify insects, their behavior and estimate the population density. Microphones are the sensors used to classify airborne signals, but vibration sensors interface are better useful in soil, grain, or fibrous plant structures. Ultrasonic sensors are used for detecting wood-boring pests as background noise is negligible at > 20 kHz frequencies. And with such new devices and new methods there is great reliability of detection. There exists a new method that considers spectral and temporal pattern features (Mankin, Hagstrum, Smith, Roda, and Kairo, 2015) of insects sound, ignoring the background noise. It observes that active insects are hidden in high-value substrates.

A novel high resolution laser-radar (LIDAR) system (CarstenKirkeby, MarenWellenreuther and MikkelBrydegaard, 2016) quantifies flying insects sound. Totally 22808 insects were recorded, and the relative temporal quantities is matched with the quantities recorded with the light trap within a radius of 5 m. Lidar records small insects with wing size <2.5 mm2 are there near the light trap and larger insects of wing size >2.5 mm2 were most abundant near the LIDAR beam.

An inexpensive optical sensor exist (Yanping Chen, Adena Why, Gustavo Batista, AgenorMafra-Neto, Eamonn Keogh, 2014) for Insect detection, counting and classifying flying insects based on wing-beat frequency from distance. Laser pointers and phototransistors are used in optical sensors. It also includes wing-beat detector that isolates the wing-beat bleeps from the background noise. Mostly the insects have less than 1000Hz of wing-beat frequency and hence this frequency is used as threshold of maximum magnitude of the signal spectrum. This optical sensor identifies Bombus, mosquitos like Aedesaegypti and Culexquinquefasciatus.

B. *Drawbacks of Existing System*

1. Used sensors for recording and recognition
2. Can work for short period of time
3. Uses only limited features
4. Computational expensive

III. P PROPOSED FRAMEWORK

The objective of this work is to give the entomologist an innovative and automated tool to recognize insect or non-insect based on the sound produced by them. To attain the objective, the below mentioned steps are to be followed:

1.Short Time Fourier Transform (STFT) divides a longer time signal into shorter segments of equal length and determines the sinusoidal frequency and phase content of local sections of a signal as it changes over time.

2.Using Discrete Fourier Transform (DFT) on the generated spectrogram by STFT, spectral and temporal features are extracted.

3.Dynamic Time Warping optimally aligns two time series sequences and captures similarities among both the sequences.

4.Then Bayesian classifier assigns class labels as Insect or non-insect after recognition of insects sound.

This work is implemented with the aim of developing a robust classifier to detect and identify insects with high classification accuracy. This work has its scope in the areas like bio-acoustic, birds and marine mammal communication, behavior and conservation learning using audio signal processing and machine learning due to multi-disciplinary_nature.

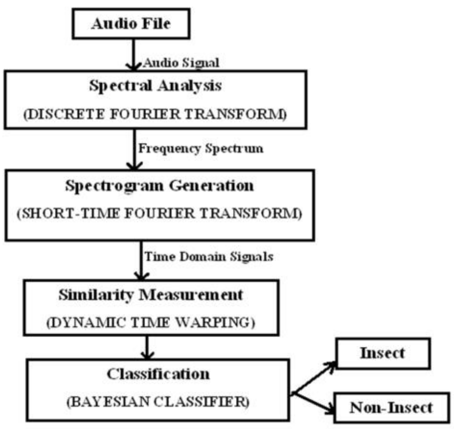

Fig.1.SystemArchitecture

Benefits

- U It is purely digitized method
- It does not depends on any hardware for sensing the acoustical data
- It is inexpensive
- Reduced Computation complexity
- Uses simple method of Bayesian classifier
- Results are more accurate when compared with other classifiers
- It supports the entomologists in identifying the insects
- Preprocessing work removes noise perfectly
- Missing data are also predicted in the preprocessing step
- Uses temporal and spatial features of acoustical sounds
- It accepts any audio file as input
- This software tool can be used in hospital surroundings too.

IV. PERFORMANCE EVALUATION

A. *Feature Extractor*

Discrete Fourier Transformation

The sound snippet of flying insects is recorded. For each insect sound, the frequency spectrum of the sound is computed using the Discrete Fourier Transform (DFT). Truncate the frequency spectrum to include only those data points corresponding to the frequency range: 100 Hz to 2,000 Hz. The DFT transforms signals in time domain to the frequency domain. Discrete Fourier Transformation is used to extract frequency, phase and amplitude of the sinusoidal component of the audio input.

Short Time Fourier Transformation

To remove the noise, Hamming window is applied that smoothens the signal by removing spectral leakages. The time and frequency domain information are too long and lead to complexity in computation and extraction of spectral features and hence Short Time Fourier Transformation is applied. Short Time Fourier Transform (STFT) is then used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. And the data is represented as a spectrogram, the plot includes time and frequency as variables along the orthogonal axes. Using the spectrogram, the spectral and temporal features are extracted and used for classification.

B. *Similarity Estimation*

Dynamic Time Warping

It finds the optimal alignment between two sequences of numerical values, and captures flexible similarities by aligning the coordinates inside both sequences. The distance of the optimal alignment can be recursively computed by:

$$D(A_i,B_j) = \delta(a_i,b_j) + \min\left\{\begin{array}{l} D(A_{i-1},B_{j-1}) \\ D(A_i, B_{j-1}) \\ D(A_{i-1},B_j) \end{array}\right\}$$

C. Training using Bayesian Classifier

Audio Signal Classification has its scope in the areas like signal processing, spectral analysis, psychoacoustics and auditory scene analysis. Classification is the process of extracting relevant features from a sound and then identify a set of classes to which the features matching is most likely to fit. Many classification algorithms are available nowadays. Of which, Bayes classifier performs well with multidimensional acoustic data.Bayes classifier results in highly accurate classification in time-series data Bayes classifier is a binary classifier and in this work it is used to classify the acoustical sound as insect or non-insect.

Use Probabilistic approach (as dataset is large and sensitive) to learn using the spatio-temporal features:

a) Find the prior probability with respect to each feature vector , x1, x2, x3, … x7, p(Ck) where k = 2 (1 – Insect, 2 – Non insect)

b) Find the likelihood p(x|Ck) of each feature vector with respect to each class,

c) Find the evidence, p(x)

$$\frac{1}{\sqrt{2\Pi\sigma^2}} \exp\left(\frac{-(6-\mu)^2}{2\sigma^2}\right)$$

d) Then evaluate the posterior probablity,

p(Ck|x) = (p(Ck) p(x|Ck)) / p(x)

e) The max. posterior determines the class

D. Testing the Bayesian Classifier

Use the same probabilistic approach on the test dataset

- ind the prior probability with respect to each feature vector , x1, x2, x3, … x7, p(Ck) where k = 2 (1 – Insect, 2 – Non insect)
- Find the likelihood p(x|Ck) of each feature vector with respect to each class,
- Find the evidence, p(x)
- Then evaluate the posterior probablity,

  p(Ck|x) = (p(Ck) p(x|Ck)) / p(x)
- Determine the class label of finding which posterior probability matches (close to) with the trained posterior probability.

V. CONCLUSION

This framework has used Dynamic Time Warping (DTW) algorithm for generating the similarity measure between two temporal sequences which may vary in speed. Bayesian classifier produces good performance as it uses the entire training dataset. This digitized tool supports the Entomologists in every aspects of recognizing the insects and it will be very useful in case of preserving the crops and finding Insect while spraying insect pesticide to unwanted place where the Insect isn't available.

REFERENCE

[1] A.Thakur,V. Abrol, P.Sharma and P.Rajan, "Compressed Convex Spectral Embedding For Bird Species Classification".,PosterArea ,DBioacoustics_AND_Medical,Acoustics Conference-Tuesday, April 17, AASP-P2.2

[2] MasatakaFuchida, et al., "Vision-Based Perception and Classification of Mosquitoes Using Support Vector Machine ", Appl. Sci. 2017, Vol 7, 51

[3] V. Abrol, P. Sharma, and A. K. Sao, "Fast exemplar selection algorithm for matrix approximation and representation: A variant oasis algorithm," in Proc. Int. Conf. Acoust. Speech, Sig. Process., 2017, pp. 4436–4440.

[4] S. Mair, A. Boubekki, and U. Brefeld, "Frame-based data factorizations," in Proc. Int. Conf. Mach. Learn., 2017, vol. 70, pp. 2305–2313.

[5] El-Said M. El-Nabawy et al., "The Effect of Organic Fertilizers and Flowering Plants on Sheet-Web and Wolf Spider Populations (Araneae: Lycosidae and Linyphiidae) and Its Importance for Pest Control", Journal of Insect Science (2016) 16(1): 18; 1–8.

[6] Jeffrey Glick, and Katarina Miller, "Insect Classification With Hierarchical Deep Convolutional Neural Networks", Convolutional Neural Networks for Visual Recognition (CS231N) , Stanford University FINAL REPORT, Team ID: 283 13 March 2016.

[7] GlenesConceptaD'mello , Rashid Hussain"Insect Inspection on the basis of their Flight Sound", International Journal of Scientific & Engineering Research, Volume 6, Issue 10, October-2015.

[8] Y. Chen, J. Mairal, and Z. Harchaoui, "Fast and robust archetypal analysis for representation learning," in Proc. Comp. Vision Pattern Recog., 2014, pp. 1478– 1485.

[9] DR. Raman, RR. Gerhardt, JB. Wilkerson, "Detecting insect flight sounds in the field: Implications for acoustical counting of mosquitoes." Transactions of the ASABE, 50(4), 1481, 2007

[10]KS. Repasky, JA. Shaw, R. Scheppele, C. Melton, JL. Carsten, LH. Spangler, "Optical detection of honeybees by use of wing-beat modulation of scattered laser light for locating explosives and land mines." Appl. Opt., 45: 1839–1843, 2006

[11] E. Keogh, M. Pazzani, "Learning augmented Bayesian classifiers: A comparison of distribution-based and classification-based approaches." In Proceedings of the seventh international workshop on artificial intelligence and statistics. pp. 225-230, 1999

[12] A. Moore, "Artificial neural network trained to identify mosquitoes in flight." Journal of the American insect behavior, 4(3), 391-396, 1991 [13] LS. Mian, MS. Mulla, H. Axelrod, JD. Chaney, MS. Dhillon, "Studies on the Bio ecological aspects of adult mosquitoes in the Prado Basin of southern California." Journal of the American Mosquito Control Association, 6(1), 64-71, 1990

[14] P. Frankl and H. Maehara, "The Johnson-Lindenstrauss lemma and the sphericity of some graphs," Journal of Combinatorial Theory, Series B, vol. 44, no. 3, pp. 355–362, 1988.