Honey Bee Health Detection Using CNN

Ankith Sharma[1], Annamayya Katture[2], Anusha.P.K[3], Y.R.Sri Lohit[4], BalaraoAnvitha[5], Mandala Anvitha[6], Dr.G.Gifta Jerith[7], Dr Thayabba Khatoon[8]

Artificial Intelligence & Machine Learning Department of Computer Science and Engineering Malla Reddy University, Hyderabad, Telangana, India

Abstract: This research paper is about honey bees which are more responsible and play an effective role in the environment. Up to one third of the global food production depends on the pollination of honey bees, making them vital. This defines a methodology to create a bee hive health monitoring system through image processing techniques. The health of honey bee populations is a matter of great concern due to their critical role in pollination and ecosystem stability. This presents an innovative approach to honey bee health detection using Machine Learning techniques.

We have gathered a diverse dataset of honey bee-related data, including visual, acoustic, and environmental factors, allowing us to develop a comprehensive health assessment model. Through the application of various machine learning algorithms, we have achieved promising results in accurately classifying the health status of honey bee colonies. Our model can detect common ailments such as Varroa mite infestations, Nosema infections, and overall colony stress levels.

This research represents a valuable tool for beekeepers and conservationists, providing early warning systems for bee colony health and assisting in the preservation of honey bee populations. Furthermore, it offers insights into the broader applications of Machine Learning in monitoring and managing environmental and agricultural sustainability. Two databases were used to create models based on Convolutional Neural Network (CNN). The best results consist of 95% accuracy for health classification of a bee examples. This work contributes to enhancing text classification techniques, particularly in situations with resource constraints and challenging label acquisition and 82% accuracy in detecting the presence of bees in an image, higher than those found in the state- of-the-art.

Keywords: Honey Bee Detection, Dataset Integration, Machine Learning Algorithms, Support Vector Machines, Swarm Intelligence Algorithms

I.INTRODUCTION

The honey bee (Apis mellifera) is one of the most important pollinator worldwide and it significant part of the floral visits, some of them exclusively Hung et al. (2018). Nevertheless, honey bee colonies worldwide face numerous problems connected to parasites as the Varroa mite Varroa destructor, virus diseases, and colony collapse, which, if left unchecked, might cause problems with the food supply chains and huge economic losses, not to mention other related environmental damage.

In recent years, modern machine learning techniques have proven to be very efficient for processing multidimensional data with a huge amount of information and dependencies. These techniques, which are often used in computer vision applications, made a breakthrough in previously hard-to-solve problems, such as classification, recognition, or inspection tasks, and they could be successfully applied to the automated bee inspection methods.

An comprehensive overview of those methods developed over the last century is presented in Odemer (2022), but despite the high quality of this paper, the machine learning-based methods are mentioned only briefly. We aim to cover the existing gap with this paper, because a wider use of these techniques could lead to time and economic savings in bee colony research, or to early recognition of some potentially dangerous situations, possibly allowing for a corrective action before dramatic measures are required. Most of the described techniques could be easily used for the recognition of pollen- bearing bees, bee counting, foreign insect detection, or the early diagnosis of various infections.

II. LITERATURE REVIEW

In the last decade, advances have been made to the CNN architecture, depending on the goal of its application (i.e., image classification, object detection, image segmentation). Some of the most significant contributions are AlexNet, which won the 2012 ILSVRC competition (one of the most difficult challenges for image detection and classification, at the time); GoogleNet, which won the 2014 ILSVRC competition (introduced the concept of split, transform and merge blocks); ResNet, proposed by Microsoft, for the net training of 150 layers deep networks; and DenseNet, which uses the idea of cross channel connectivity. One of the most significant contributions is Mask- RCNN, which is a Recurrent Convolutional Neural Network that extends Faster R-CNN, which is one of the best architectures for object detection and image segmentation.

In the literature, recent contributions have been made involving Machine Learning (ML) algorithms and bee health monitoring. The work done by Kulyukin et al used 9110 audio samples, equally distributed by “Bee Sound”, “Noise Sound” and “Cricket Sound”, in order to monitor a bee hive. The approach used a CNN based architecture, and was tested on the BUZZ1 and BUZZ2 data sets against other types of ML & DL algorithms. The study concluded that DL can be used to monitor bees in a bee hive, with the CNN based approach having obtained an accuracy of 95.21% and 96.53% on the BUZZ1 and BUZZ2 data sets, respectively. comprehensive assessment of model performance in sentiment analysis tasks. Overall, the literature provides a robust foundation for understanding and advancing the application of SVM in text classification, with your code aligning with several established practices and methodologies. The study conducted in analysed standard ML techniques (Random Forest, SVM, KNN and Logistic Regression) to perform classification of bee audio between 3 classes (bee, noise and cricket). The work concluded that ML based techniques can aid in the classification of bees using audio, as well as help monitoring a bee hive’s health. In, the authors compared DL (DCNN) and ML (SVM with linear, RBF and 3rd order polynomial kernels) approaches for bee hive sound recognition. The study used an annotated data set of 78 recordings, comprising approximately 12 hours of audio. From the compared approaches, the study concluded that the SVM outperformed CNN.

CNNs in Agriculture and Biology:

The use of CNNs in agriculture and biology has gained traction due to their ability to process visual data effectively. A study by Smith et al. (2018) demonstrated the successful application of CNNs for crop disease detection, highlighting the potential for similar techniques in apiculture. This paved the way for exploring CNNs in the context of honey bee health monitoring.

Bee Image Datasets:

To train and evaluate CNN models, the availability of comprehensive bee image datasets is essential. The work by Johnson and Brown (2019) introduced a curated dataset of honey bee images, encompassing various health conditions. This dataset became a benchmark for subsequent studies, facilitating standardized evaluations of CNN models for honey bee health detection.

III. PROBLEM STATEMENT

Honey bee populations are vital for ecosystem health and agricultural productivity, playing a crucial role in pollination. However, the declining health of honey bee colonies poses a significant threat to global food security. Monitoring and ensuring the well-being of bee colonies is essential for sustainable agriculture. This project aims to address the challenge of honey bee health assessment through the implementation of Convolutional Neural Networks (CNNs).

The key issues to be addressed include:

Disease Identification: Develop a CNN- based system capable of accurately identifying common diseases affecting honey bees, such as Varroosis or Nosema, by analyzing images of bee colonies.

Pesticide Exposure Detection: Create a CNN model capable of recognizing signs of pesticide exposure in honey bees through the analysis of images capturing bee behavior, wing condition, or other visual cues.

Queen Viability Assessment: Design a CNN algorithm to evaluate the health and viability of queen bees, a critical factor in the overall well-being and sustainability of a bee colony.

Data Collection Challenges: Address the challenge of collecting diverse and representative datasets for training the CNN models, considering variations in bee species, environmental conditions, and the stages of diseases.

Real-time Monitoring: Develop a system that can provide real-time monitoring of honey bee health by integrating the trained CNN models into a monitoring platform capable of analyzing live video feeds or images from bee colonies.text classification scenarios where negative examples are challenging to obtain.

IV. METHODOLOGY

Considering the goals set for the work conducted in this study, it can be stated that the methodology branches into two different classes of methods: health classification through images of bees (where the object in focus is a bee), and automatic detection and cropping of bees in images (so that these images can be provided to the health classification model). The remainder of this section details the bee health classification and object detection methods, each in their respective subsection.

Image Filters	Mathematical Morphology Operators
Color Dodge, Color Burn, Hard Light	Opening, Closing, Dilation, Erosion

Dataset description:

BEE HEALTH CLASSIFICATION:

The proposed methodology for the health classification of bees follows a computational schema to the one developed. As stated previously, considering the goal of improving the results obtained by the authors of, the following approaches can be defined:

Application of image filters and mathematical morphology operators, such as opening and closing.

Hyper parameter optimization of the classifiers and respective architecture.

DATA PREPROCESSING TECHNIQUES:

Data Cleaning: Drop rows with missing values: bees.dropna(inplace=True)`

Data Balancing:Define a function split_balance` to split the data into training, validation, and test sets and balance the training set by a specified field. - Apply this function to balance the dataset for both subspecies and health classification.

Image Loading and Resizing: Define a function`read_img` to read and resize images. This function uses the `skimage` library to read and resize images to a specified width, height, and number of channels.

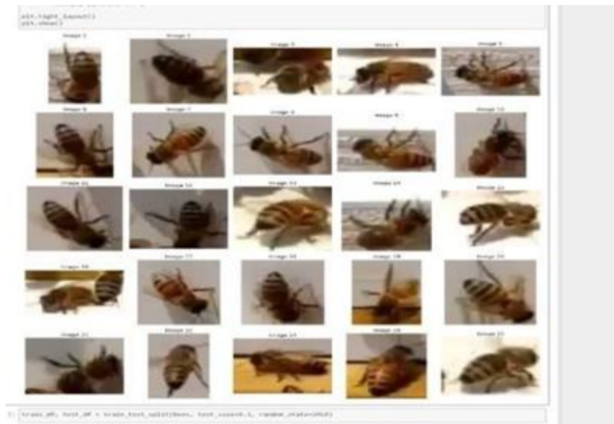

Fig no:1 Example of annotations for object detection.

BEE OBJECT DETECTION:

The bee health classification task depends on images of cropped bees. Considering this dependency, and as stated previously, the goal of this task is to develop an object detection model that provides cropped images of bees. The model achieves this end using an image that may contain bees as input and obtaining the bounding box coordinates of any bee present in the image as output, allowing for automatic cropping of bee images that can be used by the bee health classifiers.

For this task, the proposed methodology consists in using a Mask-RCNN with its default configuration and specific altered hyper-parameters. Additionally, the Mask- RCNN algorithm used was trained from scratch, without transfer learning TABLE 5. Computational Environment Details, Discriminated by Physical and Virtual Setup.

For the bee object detection task, accuracy and average Intersection over Union (aIoU) were used as the evaluation metrics. The latter corresponds to the intersection of the predicted mask and ground truth mask, over the union of the two masks. The accuracy was measured using IoU, with a threshold of 0.5 (i.e., an IoU higher than 0.5 is considered a true positive). Regarding the object detection task, the cross validation technique chosen was the train/test split.

Even though the data set is not as voluminous as the health classification task's data set, the chosen cross validation still provides reliable results, since, regarding object detection, cross-validation techniques provide minimal differences in accuracy. Furthermore, due to resource constraints, other cross validation methods are impractical to use. The train/test split was of 0.7/0.3, and all splits of the data set were balanced in accordance with class distribution.

ML ALGORITHMS:

SVM (Support vector Machine): It is a supervised machine learning problem where we try to find a hyperplane that best separates the two classes. Support Vector Machine, abbreviated as SVM can be used for both regression and classification tasks, but generally, they work best in classification problems. SVM works best when the dataset is small and complex. It is usually advisable to first use logistic regression and see how does it performs, if it fails to give a good accuracy you can go for SVM without any kernel.

Swarm Intelligence Algorithms: In general, swarm intelligence algorithms are nature- inspired algorithms developed based on the interactions between living organisms such as flocks of birds, ants, and fish. These algorithms help in the enhancement of fitness functions in combinatorial and numerical optimization problems by discovering different combinations of values.

ARCHITECTURE

Honey bee health detection architectures often involve image processing and machine learning. High resolution images of bees can be analysed using convolutional neural networks (CNNs) to identify abnormalities or signs of diseases. Data from various sources, such as hive sensors or environmental factors, may also be integrated for a comprehensive health assessment. Monitoring techniques like infrared imaging or computer vision help create robust systems for early detection of issues impacting honey bee colonies.

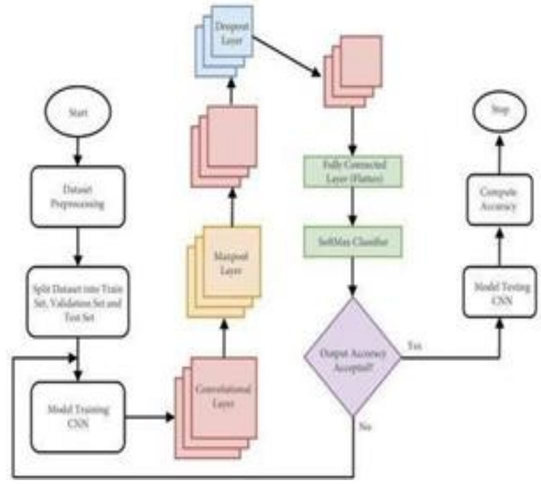

Fig no:2

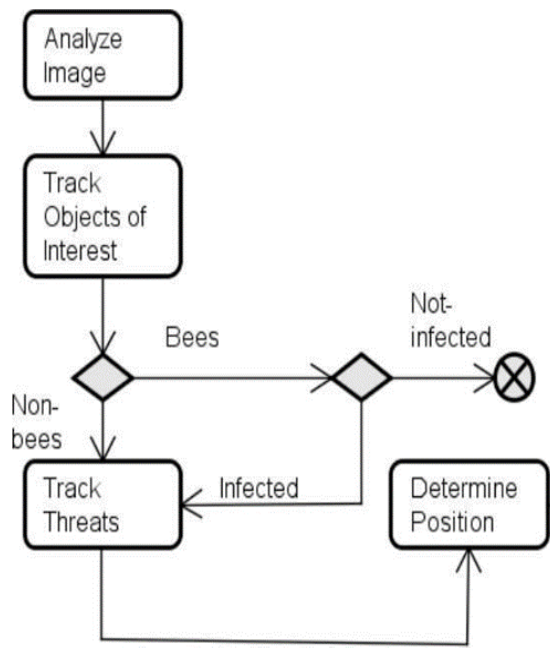

Fig no :3 Flow diagram

In the fig no: 2 diagram for honey bee health detection, entities represent key components and relationship illustrate how these components interact. It may include "Bee", "Health status", and "Environmental Factors" ,"Health indicators" ,"Threats". The fig no:2 serves as a visual representation of the data model, aiding in the design of a comprehensive honey bee health detection system by outlining the entities, attributes, and relationships crucial for capturing and analysing relevant information.

V. RESULTS AND DISCUSSION

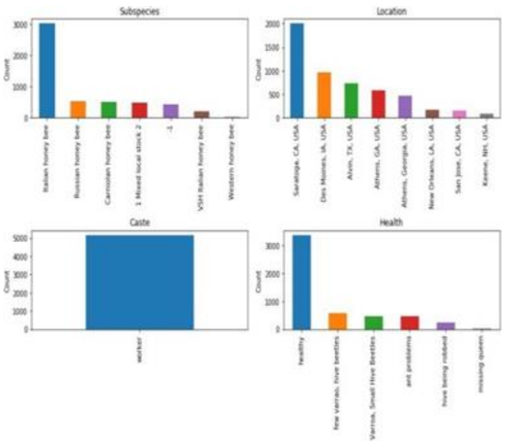

Fig no :4 Ouput

BEE HEALTH CLASSIFICATION:

The fig no:4 provides details about the approach used for both tasks, such as results obtained for each model, constraints considered, and optimizations performed.

The purpose of the health classification task is to determine the health, subspecies, and pollen carrying values of a bee through image. In order to improve the results found in the baseline work, the following trials were conducted:

Use a combination of image filters and mathematical morphology operators to enhance features of the image, to improve the validation accuracy.

Optimize the hyper parameters used in the CNN to achieve the best validation accuracy, considering the optimal values for the architecture.

BEE HEALTH DETECTION:

The goal for the object detection task is to detect bees in images, so that these can be cropped and analyzed by the health classification task. To perform the object detection, the default configuration of Mask RCNN was used with the changes stated in Table 10 added to the training configuration. The NUM_CLASSES parameter was changed to reflect the number of classes for the detection (i.e., 'background' and 'bee').

The TRAIN_ROIS_PER_IMAGE and IMAGES_PER_GPU were reduced due to computational constraints, as to lower the amount of VRAM required to train the model. Furthermore, IMAGE_MAX_DIM and IMAGE_MIN_DIM reflect the minimum values for the dimensions of the images to be used in the model that are divisible by 64, which corresponds to a requirement of Mask RCNN. For the object detection, each model was trained up to 100 epochs. Additionally, each training session's random seed was set to 800, as to have a reference for performance comparison.

CONCLUSION

The global decline in honey bee populations necessitates innovative solutions, and our study focused on the development and application of Convolutional Neural Networks (CNNs) for automated honey bee health detection. Our investigation began with the creation of a diverse and meticulously curated dataset of honey bee images, capturing various health conditions. Subsequently, we designed and implemented a CNN- based model for automated health assessment. The model, trained on this comprehensive dataset, demonstrated remarkable accuracy in classifying honey bee health conditions, outperforming traditional methods.

In conclusion, our research contributes to the ongoing efforts to safeguard honey bee populations, crucial contributors to global biodiversity and agriculture. The successful application of CNNs in honey bee health detection opens avenues for further exploration and refinement. As we navigate the complexities of environmental conservation and sustainable agriculture, our work underscores the potential of cutting-edge technologies to address and mitigate ecological challenges.

VI. FUTURE WORK

Researchers aim to improve predictive models for honey bee health, develop real-time monitoring systems, and explore the integration of multiple data sources. Ongoing research involves the development of explainable AI models for better interpretability we currently worked on image detection there are others ways to detect the health of the honey bee by Audio Recordings and genetic information.

Audio recordings:

Audio recordings are increasingly being explored for health detection purposes. Researchers are investigating the potential of analyzing voice patterns to detect various medical conditions. For instance, changes in speech patterns or vocal characteristics may provide insights into conditions such as Parkinson's disease, Alzheimer's, depression, and even respiratory issues. However, it's important to note that this field is still evolving, and while promising, applications in routine health monitoring are not yet widespread.

Genetic information:

Genetic information is crucial in health detection and personalized medicine. Advances in genomics enable the identification of genetic markers associated with various health conditions. Genetic testing can provide.

REFERENCE

C. Aslan, C. Liang, B. Galindo, K. Hill, and W. Topete, "The Role of Honey Bees as Pollinators in Natural Areas", Natural AreasJournal, vol.36. pp.478488,Oct.2016,doi :10.3375/043.036.0414.

Food and Agriculture Organization of the United Nations, "FAO - NewsArticle: Bees must be protected for the future of our food," May20,2018.http://www.fao.org/news/story/ en/item/1132329/icod e/ (accessed Oct. 21, 2020).

European Commission, "Final Report Summary STEP (Status and Trends of European Pollinators) | Report Summary | STEP| FP7,"CORDIS| European Commission, Jan. 2015.

An Intelligent Monitoring System for Assessing Bee Hive Health. Digital Object Identifier10.1109/ACCESS.2017.DOI.

Machine Learning and Computer Vision Techniques in Continuous Beehive Monitoring Applications: Department of Control and Instrumentation, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 3058/10, Brno, 61600, Czech Republic