Journal of Mines, Metals and Fuels, *71(12C) : 407-417; 2023. DOI: 10.18311/jmmf/2023/47319*

Print ISSN : 0022-2755

Journal of Mines, Metals and Fuels

Contents available at: *www.informaticsjournals.com/index.php/jmmf*

Application of Machine Learning and Deep Learning Models in the Supervision of Apiculture

C. G. Raghavendra, H. H. Adithya*, H. J. Lekhashree, R. Anaghaa and H. Guha

Department of Electronics and Communication, Ramaiah Institute of Technology, Bengaluru – 560054, Karnataka, India; hadithyaraj820@gmail.com

Abstract

Honey bees are well known for crop pollination and honey production which in turn plays a vital role in global agriculture. They face various threats from parasites, ants, hive beetles, pesticides, pollutants, toxins, and also hive robberies which could finally lead to the collapse of colonies. These major problems have led to the worldwide decline of honey bees. Thus, determining colony strength has gained importance in sustainable beekeeping. Machine Learning (ML) techniques provide various solutions for these problems by accurately monitoring and detecting the hive status. This paper reviews such techniques which help in monitoring and detection of other features. This helps the beekeepers to drive control over effective honey extraction and increase productivity.

Keywords: *Artificial Intelligence, Beehive, Deep Learning, Internet of Things, Machine Learning*

1.0 Introduction

Honey bees are popular pollinators and are also the fundamental producers of honey in the ecosystem. Approximately 73% of crops in the world are pollinated by bees. Honey bees face various threats throughout their life cycles such as parasites, hive beetles, ants, pests, diseases, hive robberies, and many others which may lead to a collapse of bee colonies. It is reported in that, pollinator diversity is declining due to urban expansion, deforestation, and monoculture crops[1]. Varroa destructor attaches to the honey bee and sucks the fat body tissue, which in turn leads to the collapse of the entire bee colony. Small hive beetles also cause damage to the beehive. Hive robbery is a major problem to be concerned about, where robber bees tend to grab food from other hives. Using pesticides on crops is leading to the death of honey bees, which to a great extent can lead to a pollination crisis. Due to these problems, beekeepers need to inspect the status of the hive frequently and take measures to overcome them. Using traditional bee monitoring techniques such as visual observation, beekeepers can detect the problems of the hive, but it is a highly tedious and time-consuming process. These limitations can be resolved with the help of technological developments. Monitoring beehives throughout the year is a complex and laborious process, instead, this can be controlled by using the Internet of Things (IoT) which uses sensors to measure temperature, humidity, colony mass, image pattern, audio, and video analysis. The introduction of these new technologies in beekeeping will help in reducing the production costs, enhance production and facilitate the breeding process for the bee colony. ML provides various applications in modernizing the field of Apiculture. Machine Learning, a part of Artificial Intelligence (AI) is the study of computer algorithms by using the data and imitates the way humans learn and gradually improves the accuracy by itself. It is transforming all sectors including transportation, food, education, healthcare services, agriculture, and entertainment in a better way. It helps in increasing

**Author for correspondence*

the efficiency of production, reducing wastage of food and biofuel, quality assessment, crop disease, and weed detection. Bee monitoring developed can be more efficiently and robustly achieved by implementing ML, computer vision, signal processing, and information technology[2]. These technologies help beekeepers to take better decisions in several situations and reduce human error, thereby improving the quality of bee products. The ML techniques help to predict the causes of bee decline and it is possible to develop conservational strategies to preserve the biodiversity and prevent the decline of the pollinator. The count of comb cells containing brood as well as food reserves helps to study and monitor the colony dynamics, colony strength, and health status. It also helps the beekeeper to know the nutritional status of the queen and honey yield. This paper reviews various approaches to such applications. This includes the study of morphological features, monitoring the infestation level, air pollution detection, swarming phenomena, hive robbery, bee pollen foraging behavior, chemical poisoning of honey, honey harvest prediction, queen body mass, hive sound recognition, and significantly more[3].

2.0 Monitoring of Honey Bee Colony

Hive monitoring can be successfully carried out using different ML techniques to monitor various features of the beehive. Researchers have previously built many ML and DL models for monitoring. Monitoring of the hive is carried out for various purposes like detection and classification of comb cells, and analysis of data from sensors to predict the health status of the colony. Some of these monitoring techniques are mentioned below.

2.1 ML and DL Model for Monitoring the Hive

The deep learning-based classification models are proposed to identify different hive conditions and abnormalities[4]. Deep learning (DL) is a subset of ML which is based on Artificial Neural Networks (ANN) which has many hidden layers between input and output layers. Deep convolution Neural Networks (DNN) are used for image recognition. Convolution layers are responsible for extracting features from the images given by the input layer. Then DNN is trained with an enormous set of data. Transfer learning uses pre-trained DNN, this is done in order to reduce the time consumption while training the model. Besides transfer learning, the other three Support Vector Machine (SVM) models are demonstrated. The classification results were analyzed for three different datasets, firstly the Varroa dataset in which the highest accuracy was found with Transfer Learning, and second the bee image dataset in which the highest accuracy was observed with the SVM model. Third, the pollen dataset, even with this the Transfer learning achieved the highest accuracy. Apart from the classifying process, monitoring the study of honey bee colony circadian rhythm also helps in the development of a monitoring system. A monitoring system was developed for studying bees' behavior during the night[5]. For this, a monitoring device with temperature, humidity, and audio sensors was considered. The complete system was designed with a server, client, and embedded module. Then the classifier training was done by using Mel Frequency Cepstral Coefficient. The process was repeated for several expected bee nights' start and end hours. The SVM model was considered as the classifier for this purpose. The time between 2 a.m. and 3 a.m. was considered as bee midnight as the maximum accuracy was observed at this time.

2.2 DL Model for Detecting and Classifying Comb Cells

Comb cells give intuition to beekeepers regarding the colony's health status, nutritional status, quality of the queen, honey production, and so on. A free software named DeepBee© was developed which can be used by bee researchers and apiculturists to know the information of colony strength. It evaluates the comb cell images with high speed and accuracy. Here, OpenCV v.4.0 library is used for the implementation of Circle Hough Transform (CHT) which is a basic feature extraction technique. Data pre-processing was performed for the removal of noise and to enhance cell edges. Scale-invariant detection methods with two main stages were used to find the mean distance between the cells and the average of their cell size. Later, semantic image segmentation was applied for the image detection of comb cells and for removing false cells which fall outside the comb region. Then cell classification was performed using Convolutional Neural Network (CNN). Suitable window size was selected and 13 different CNN architectures were selected for training. Then data augmentation was used to enlarge the dataset. After training different architectures of CNN, it was noticed that the MobileNet CNN model gave them better results than other models and InceptionResNetV2 showed the best performance in accuracy and recall.

2.3 Automatic Bee Identification System (ABIS)

ABIS, implemented using the image processing technique is a software tool created for the identification and monitoring of bees. It is reported that agrochemical disease causes the decline of the bee population[6]. So, it is necessary to protect the bee species by monitoring them for maintaining valuable ecological resources. Live bees captured in the wild are cooled and kept in an icebox without causing them long-term harm. The images of these samples are used in the recognition system to study the wings by their venation for the generation of feature vectors. This system provides fully automated image processing and recognition engines to classify bees using their wing data by applying SVM or Kernel Discriminant Analysis (KDA) to perform the final phase, which is recognition. The study shows that the venation of the bee's wing can be used to perform the recognition and monitoring of honey bees. The final model can successfully recognize the different members of bees with an approximate recognition rate of 95%.

2.4 Forecasting Health Status of Honey Bee Colony through ML using Sensors and Apiary Inspection

The use of in-hive sensors helps the beekeeper to reduce frequent physical hive monitoring. The study helps to monitor and inspect the hive with the data collected from both external and internal sensors along with the colony health status prediction[7]. The algorithms of Neural Networks (NN), k-Nearest Neighbors (kNN), and Random Forest (RF) predicted the health status of the colony. Initially, in-hive sensors were used to obtain the temperature of the hive and weather conditions of the external environment. The inspections were performed on the basis of the health status of the Healthy Colony Checklist (HCC) which includes features of the hive, for example, presence of all stages of brood, sufficient adult bees, presence of the young productive queen, proper nutrition, and enough space. Later, the data was preprocessed and classification models were developed by ML techniques using a labeled dataset. It was concluded that the RF algorithm showed better performance compared to kNN and NN for the classification of the health status of the bee colony. If the internal temperature of the hive is very less, it leads to the death of brood, and also, the dehydration of nectar doesn't happen to produce honey in time. In high temperatures, an increase in the internal temperature of the hive leads to wax melting, brood death, and quick honey dehydration. This ML approach can be used for monitoring apiaries and predicting the health status of the honey bees by customizing and integrating it into a computer system. A comparison of the above-stated methodologies is mentioned in Table 1.

3.0 Various Application of DL and ML Models in Apiculture

ML and DL techniques are not only used for monitoring honey bees, but also for multiple other applications

Table 1. Comparison of various ML models for Apiculture monitoring

S. No.	Authors	Algorithm	Application	Maximum accuracy
1	S. Kaplan Berkaya, E. Sora Gunal and S. Gunal[4]	SVM Transfer Learning with pre-trained DNN	Classification of honey bee status	0.9907
2	T. Cejrowski, J. Szymanski and D. Logofatu[5]	SVM	Study of honey bee colony circadian rhythm	0.8114
3	T. Alves[1]	CNN	Detection and Classification of comb cells	0.943
4	Arbuckle, Tom & Schröder, Stefan & Steinhage, Volker & Wittmann, Dieter[6]	SVM KDA	Bee species Identification	0.95
5	A. Rafael Braga, D. G. Gomes, R. Rogers, E. E. Hassler, B. M. Freitas and J. A. Cazier[7]	kNN Random Forest Neural Network	Forecasting the health status of honey bee colony	0.98

such as the study of bee classification, pollen foraging behavior, identification of swarming process, infestation level of Varroa destructor, hive robbery, chemical toxicity of honey bees, and beehive sound recognition. These applications are discussed in the following section and briefly listed in Table 2.

3.1 Honey Bee Classification

Deep Recurrent Neural Networks (DRNN) are considered dynamic models in DL. They make use of large datasets for training purposes. DL involves various optimization steps, for example, proper network architecture selection and choosing the hyper-parameters for them, and selection of proper training algorithm for the model. Additional difficulties are observed in training deep neural networks like non-convex nature, minimization of the nonlinear loss function with more than one local optima and saddle points. During these kinds of problems, some commonly used training algorithms like Stochastic Gradient Descent may get trapped at the local optima and the convergence to the global minima is not guaranteed. The DL models cannot be implemented to perform better by only improving the training algorithm, it also requires the proper selection of initial design parameters. A Ternary bees' algorithm BA-3+ for sentiment classification using a deep recurrent neural network with optimum weights is discussed[8]. The BA-3+ algorithm was the improved version of the BA-3 algorithm. The BA-3+ algorithm uses the worst solution, in-between solution, and global test solution as three individual solutions. Each time during the time of training only three individual bees was selected as three initial solutions. Then the search strategies like stochastic gradient descent stabilized by Singular Value Decomposition were applied. The optimization procedure was continued until the stopping criteria were encountered with a new set of bees. This model was implemented with Python programming language using various tools and libraries. Even though many studies are concentrating on training artificial neural networks, the Vanishing and Exploding Gradients (VEG) problem arise due to the insufficiency of hidden layers. The author has also mentioned the work towards handling the problem of VEG. The BA-3+ algorithm performed better when compared to the accuracy of the traditional Stochastic Gradient Descent algorithm. The reports for time consumption of Differential Evolution, BA-3+, and Particle Swarming Optimization (PSO) are also provided for further analysis. By the report, it was observed that BA3+ was more stable and also worked faster compared to PSO. It was also spotted that the model worked better when trained with huge datasets improving the efficiency of the model. The experimental results indicate that BA-3+ can be used to overcome the disadvantages of the Stochastic Gradient Descent algorithm and to handle the problem of VEG.

3.2 Identification of Morphological Features

Identifying and collecting the body features of bees help us in understanding the population dynamics and in monitoring the health status of the beehive. These procedures can be carried out by the Morphological key-based features concentrating mainly on the wing data. But they are complex and need a lot of knowledge about various bee species. The different ML-based visual features with traditional Morphological features were compared and it was found that the ML models with unsupervised learning were error-prone in various other instances. This problem was faced as the ML models were made to focus on the overall features. So, the process of morphological study is carried out by using different DL models. The images for training the model were collected from the public domain. Then the deep networks were trained with a mini-batch Stochastic Gradient Descent Optimizer. Later the gradients of the best performing model were utilized to obtain a Gradient Weighted Class Activation Mapping (Grad-CAM) for the visual indication of the discriminative region to help the models to classify bee species. After classification, the highest accuracy was found to be 94.27%. Further, the classification process can be done by using smaller data set of bee images with outlined morphological features of each bee to get robust results.

3.3 Study of Pollen Foraging Behaviour of Honey Bees using ML

The bee colony consists of a queen bee, drones, and workers. The worker honey bees are quite smaller than other bees and they do almost all the activities like collecting water, nectar, and pollen. For a beekeeper, it is very important to observe the colonies foraging behavior. Beehive strength and pollination efficiency are evaluated by the detection of pollen-bearing bees. Honey bees' pollen foraging behavior is affected by multiple factors for example pollen availability, brood storage, and weather conditions. The pollen traps can predict the pollen production of a beehive by manually monitoring the weight of collected pollen. An efficient method to analyze the environmental condition and foraging behavior of pollen is studied[9]. The honey bees with and without pollen were detected, classified, and tracked. Data processing was done in four steps: image acquisition, object detection and recognition, object tracking, and counting. The detection, recognition, and counting of honey bees were carried out using DL. The multiple pollen and non-pollen-bearing bees were tracked by using Kalman filter and Hungarian algorithm. For further analysis of the results of the classifier, Gradient-weighted Class Activation mapping (Grad-CAM) was used. The Grad-CAM showed that the pollen sacks in the hind legs were used as a necessary component for image recognition. The performance of the mentioned algorithm was assessed using Mean Absolute Percent Error (MAPE) for counting purposes. It was observed that there was also a significant decrease in pollen foraging activities due to gentle breeze or heavy rain conditions. These results also help the beekeepers to know the condition under which the extra pollen has to be fed.

3.4 Tracking all Members of Honey Bee using Learned Models

Manual tracking of bees is highly difficult without marking them individually. Previously, many researchers have worked on manually tracking bees for different activities like analysis of food exchange interaction, analysis of colonies' proximity network, and foraging behavior. With the development of technology like computer vision software it has become easy to automatically identify and track animals. Tracking the bee within the colony is a tedious process due to its dense population as well as similar target appearance. This system which robustly detects the identity of the bee just by connecting matching IDs helps in tracking using cameras of high-resolution and low noise. Approximately 2000 marked honey bees are automatically tracked with recording hardware using markers[10]. It includes an iterative tracking approach where two steps are involved: consecutive detections using tracklets and connecting tracklets over longer gaps. ML models are used in both steps i.e., the tracking step uses a Hungarian algorithm to assign likely matches between detection and manually labeled data using SVM. A colony consisting of 1953 bees was marked during the 2 days' session. The young bees which were bred in the incubation chamber were also marked and a total of 2,775 bees were marked. This approach presents detailed reflections of the individual activity of bees in the colony. Incorrect ID decoding is reduced from 13% to 2%. The entire lifespan of many hundreds of honey bees from their brood cell to their death is covered. All the activities of a self-sustaining colony can be observed, it may be from the egg-laying queen, colony defense, brood-rearing workers, and food collection. The tracking framework provides an approach to identifying the correct IDs of approximately all honey bees and this system provides an accurate movement path of the bee.

3.5 Determining Parameters that Impact the Queen Body Mass using an ML Technique

Queen's body mass is considered one of the best indicators of the queen's quality. The parameters like "breeder" and "ovary mass" were the key factors in the body mass prediction of a queen bee. The body mass of a queen bee varied throughout its lifecycle, for example, body mass increases after mating and decreases after hatching. It is also influenced by different practices in queen rearing. In fertilized queen's body, approximately 40% of body mass is covered by ovaria. The availability of food sources greatly influences the composition of royal jelly as well as the body mass of the queen. ML is used to obtain the relationship of body mass with other physiological, anatomical, and rearing parameters that determine the queen's quality[11]. To configure the complex relationship between queen quality parameters, the Gradient Boosting Machine algorithm (GBM) was used with an open-source software H2O. Three years of data were collected and three different models were built, which enabled the utilization of all available data for each year. Queen's quality is also determined by several queen characteristics such as the queen's success in mating, genetic quality, and developmental conditions. This helps the beekeeper to predict the performance of the queen before purchasing it since they can predict brood production.

3.6 Bee Hive Sound Recognition using ML Approach

The potential of ML in beehive sound recognition systems where both SVM and CNN have experimented[12]. Computational sound scene analyses research is developed to automatically analyze sounds in the natural environment. The computational bioacoustics scene analysis will greatly help beekeepers to do automatic beehive monitoring by identifying natural cycles occurring in the beehive, bee swarming as well as queen missing scenarios. The identification of the sound of the beehive involves many steps. The first step is to build a system that is capable of recognizing and differentiating bee and non-bee sounds. Non-bee sounds are the sounds of the surrounding environment it may be urban sounds, rain, or animal sound. The main aim of this model is to differentiate between sounds produced by bees and external non-related sounds and automatically detect beehive sounds. Data of pure beehive sounds and external sounds have been considered. The dataset is based on the selected set of recording from two projects, one The Open-Source Beehive (OSBH) Project and the other NU-Hive project. The audio recordings were processed at a sample rate of 22050Hz. The model implemented random splitting between train and test set and also a "hive-independent" splitting scheme in which training samples belong to only certain hives. The result showed that the SVM implementation achieved better results when compared with the CNN approach.

3.7 Monitor Infestation Level of Varroa Destructor

The Varroa destructor mite is recognized as the main cause of the mortality of honey bee colonies due to virus transmission. A novel method to automatically count bees and mites to reduce the damage caused to the colony by estimating the infestation level is described[13]. By recording the video sequence of beehives, the Varroa destructor can be monitored using a portable computer vision system. The model consisted of a multispectral illumination and camera along with a video monitoring system. A deep learning-based analysis, a computer vision algorithm named Infestation Level Estimator was adopted to track the number of honey bees and find the spots of Varroa mites. The bees were passed through a narrow passage over an illuminated window and were recorded from underneath the window. As it was difficult to differentiate bees and Varroa mites with white illumination, the light of different wavelengths was used. A classifier was trained and tested using a Linear Discriminant Analysis with a three-dimensional feature vector or SVM. The images with pixel areas containing mites and bees were used to train the classifier. The classifier's output was divided into white and black images where the Varroa destructor was represented by the white region. It was noted that most of the mites were found in the abdomen region of the bees. During the final test and evaluation Shift Invariant Feature Transform was used for feature extraction. While counting bees the Infestation Level Estimator exhibited an F1 score of 0.97 and in the detection of infected bees, an F1 score of 0.91 was observed.

3.8 ML Regression Model for Predicting Honey Harvest

Using data from weather stations and remotely sensed data of *Corymbia calophylla* (marri) trees from the Southwest, it is observed if any ML methods can be developed for predicting honey harvest using weather and satellite data. This helps the beekeeper to adaptively manage the apiaries and it also improves the efficiency along with industrial safety of apiary management by which they can estimate areas of higher and lower honey production.ML regression models are applied to the dataset and identify the capability of both satellite-driven vegetation as well as moisture or weather data. This approach is implemented and a model to predict honey harvest for marri honey is developed. The study also shows that there is a non-linear relationship between temperature and flowering intensity[14]. It explains the factors that influence the production of marri honey. Classification and regression models are commonly used ML methods for predictive models. Random forest is popularly used for remote sensing and ecological applications. The output of testing showed that the Random Forest functions worked well for classification. Gradient Boosted functions via regression models worked well for predicting honey harvest weights. The classification approach was done based on 'good year', 'moderate year', and 'bad year' ratings. The meaning of 'good year' is that the prediction for this is accurate and robust along with a 0% error. Whereas 'poor year' cannot assist apiaries because their bees may starve due to lack of nectar. So, apiary management can prepare for a 'good year' in advance by which they can increase the honey production. Rainfall before the flowering period is also a key factor that influences honey harvest. By this regression tree analysis, different factors influencing honey harvest of marri trees and time frame factors can be predicted a year before honey flow starts, which helps beekeepers to manage the honey collection. This regression model predicted honey yield per hive with a Mean Average Error

(MAE) of 10.58kg with 92% accuracy. Gradient Boosted Regression (GBR) algorithm could predict honey yield per hive with an MAE of 11.72kg. It was observed that cooler weather and lower rainfall in summer are good for marri honey harvest.

3.9 Detection of Air Pollutants

The problem of air pollution is becoming a serious issue due to rapid industrialization and urbanization. Many studies are carried out to monitor atmospheric pollutants. Although, there are many ways to obtain accurate data they come out to be more expensive and bound by special resolution. Bees are sensitive to different chemicals present in their surroundings and emit typical sounds. These sounds can be analyzed and used for monitoring atmospheric pollutants. *Apis mellifera* was chosen and four compounds Acetone, Trichloro methane, Glutaric dialdehyde, and Ethyl Ether were used for the detection[15]. It was noticed that the beehive sounds were different in the morning and afternoon for the same chemical compound exposure. A sound acquisition system was developed using some of the IoT devices to collect sound data. Then the preprocessing of data was carried out with Signal Processing Library in python. Later Short Time – Fourier Transform was used to find the time-frequency features of the beehive's sound. To classify the collected data, different ML techniques were implemented. The Mel Frequency Cepstral Coefficients with 39 dimensions were considered as the features for the classification models. Since large dimensional features are more tedious to calculate acoustic signals, the dimensions were reduced using Principal Component Analysis. To effectively classify the processed data three algorithms were implemented. Firstly, Support Vector Machine with three types of kernels, here Bayesian optimization algorithm was used to optimize the kernel function parameters. The second classifier used was kNN, here the training data points are simply stored. The third classifier used was Random Forest. This is an integrated classifier developed on a Classification and Regression Trees (CART) decision tree. Here the higher accuracy is achieved by taking the average of the final results obtained by combining different weak classifiers. Eventually, after the classification of the data, it was observed that Random Forest had an accuracy of 83.6% and kNN had an accuracy of 83.8%. The SVM classifier reached the highest accuracy of 93.7% with the RBF kernel. From the results, it was observed that beehive sound analysis can provide information of chemicals in the surroundings of the beehive. This work can also be extended to several other compounds.

3.10 Determining Chemical Poisoning of Honey Bee using GACNN

Chemical pesticides are leading to the decline of insect pollinators. Deep Graph Attention Convolutional Neural Network (GACNN) is used to determine the chemical poisoning of the honey bees[16]. The model is trained with a dataset of 720 pesticides. Beekeepers and scientists have predicted that pesticides are weakening bee-colony. Computational prediction of bee poisoning will help scientists to detect harmful chemicals. The potential toxicity of chemicals is predicted by quantitative structure – toxicity/activity relationship (QSTR/QSAR) models. Deep GACNN successfully predicts the features of chemical structure as well as the poisoning of bees. It performs best on neonicotinoid compounds since it can differentiate poisonous and non-poisonous compounds. The data of chemicals with bee toxicity was collected from public databases. They took 900 pesticides in the final dataset. The toxicity data were divided into low, moderate, and high toxicity. The Convolutional Neural Network (CNN) was implemented for image classification. Three models were compared i.e., Radial Basis Function (RBF) kernel SVM model, DNN classification model, and Logistic Regression model. Undirected Graph(UG) and CNN are combined with attention graph mechanisms to build a model of deep GACNN for beekeeping. Using the test set, GACNN showed an accuracy of 83.72% and an accuracy of 69.84% for the recognition of toxic chemicals and 89.09% for recognizing nontoxic chemicals. The toxicity mechanism was divided into signal pathway influence, behavioral influence, and enzymes influence to form a map of the toxicity mechanism. Therefore, GACNN is a powerful model to predict the capability of chemical poisoning in the honey bee colony.

3.11 Chemical Acute Contact Toxicity of Honey Bee Prediction using in silico Techniques via ML Methods

In silico tools, like the Quantitative Structure-Activity / Toxicity Relationship (QSAR/QSTR) model is used to predict the biological and physicochemical properties of the chemical. About 6 ML methods were combined to develop 54 classification models[17]. This experiment was carried out on the basis of honey bee acute contact toxicity data with 676 structurally different kinds of pesticides. This study shows that the combination of the SVM algorithm and Cyclin-Dependent Kinase (CDK) was observed to be the best among other proposed models. This analysis gives nine structural alerts if any honey bee acute contact toxicity is noticed. Many ML approaches have already been implemented, for example, Probabilistic Neural Network (PNN), k-Nearest Neighbour (kNN), QSAR model, and many more. The data was collected from three public databases. The high-quality dataset was selected by filtering unwanted data, removing the mixtures, inorganic and organometallic compounds. The data was then classified as highly toxic, moderately toxic, and non-toxic. Molecular fingerprints were used to predict chemical toxicity and activity. Six ML models used in the classification process are listed as follows: k-Nearest Neighbour (kNN), Decision Tree (DT), Naïve Bayes (NB), Random Forest (RF), Artificial Neural Network (ANN), and Support Vector Machine (SVM). The 9 important substructures like phosphoric acid, pyrethroids, allyl chloride, nitramide causes neurotoxicity, altering the nervous system and death of bees.

3.12 Identification of Honey Bee Swarming Process

During the swarming process, they undergo their reproductive cycle. The old queen bee with a large number of workers and drones leaves the nest to form a new colony. In the springtime, many changes can be observed in the hive, which is demonstrating the occurrence of the swarming process. Even then the prediction of the actual date and time of swarming cannot be done accurately. The communication pattern of honey bees in particular with vibrations and acoustic noise is mentioned. The majority of studies concentrate on the vibration data from the microphone placed in the hive. The vibrations were sensed from two separate hives, approximately 3 meters away, each comprising of two *Apis mellifera* colonies using two accelerometers mounted on the walls of the hive[18]. The vibrations mainly from the acoustic noise of honey bees and also the body motion of individual bees was sensed. The signal thus obtained is too noisy. To overcome this issue, the extraction of the instantaneous feature was carried out. The Principal Component Analysis (PCA) method is used for the extraction of features and was mainly used for dimensionality reduction. The spectrogram of two hives was analyzed, four to five bands were observed clearly and 2000Hz was the strongest. Principle Component scores of relatively higher orders were extracted for the purpose of analysis. The excellent reproduction of the original spectrum was obtained by the linear combination of the first ten scores and Eigen spectra.

3.13 Detection of Bee Hive Robbery

Hive robbery is a commonly observed problem in which a hive gets robbed by its neighboring hive for the accumulation of stored resources like nectar and honey. The scarcity of these resources drives a stronger hive to ambush its neighboring hive which cannot defend itself. The main reasons for the weakening of a hive are the absence of a queen bee, a weaker queen bee, a diseased bee colony, and so on. Such occurrences of robbery are devastating for both the bee colony and the beekeeper. Thus, the beekeeper tries different methods to avoid such circumstances. The analysis of audio and video recordings was used for the prediction of the beehive robbery[19]. This method involves a Raspberry Pi-based system called Beemon. This system was placed in front of the hive which records the videos of the bees that enter and leave the hive. Along with this system, other sensors such as microphones, humidity, and temperature sensors were also used. The collected data was used to estimate the bee traffic using an application named Beevee. This application uses change detection, and object tracking, alongside a simple neural network. From the graph of traffic, it was observed that the bee population was decreasing gradually. A sharp spike was detected on the day when the hive was robbed massively. Thus such a monitoring system can be implemented to detect the occurrence of robbery. Further improvements in this system through the application of ML can be carried out to predict such occurrences in advance. This helps the beekeeper to take preventive measures by introducing a stronger queen bee which can strengthen the hive.

3.14 Determining the Sudden Drop of Temperature in Bee Colonies

The temperature drops in bee colonies were predicted using an ML model considering the long short-term memory (LSTM) algorithm which had the input parameters like mean fanning, internal and external temperature, internal humidity, mean noise, and mass

Table 2. Various algorithms and applications of ML in Apiculture

S. No.	Authors	Algorithms	Applications
1	S. Zeybek, D. Pham, E. Koç and A. Seçer[8]	DRNN Singular Value Decomposition (SVD)	Honey bee classification
2	T. Ngo, D. Rustia, E. Yang and T. Lin[9]	CNN	Analysis of Honey bee pollen foraging behavior
3	F. Boenisch, B. Rosemann, B. Wild, D. Dormagen, F. Wario and T. Landgraf[10]	SVM	Tracking all members of Honey bees
4	J. Prešern and M. Smodiš Škerl[11]	GBM	Determines the parameters that impact queen body mass
5	Nolasco, Ines and Benetos, Emmanouil[12]	CNN SVM	Beehive sound recognition
6	K. Bjerge, C. Frigaard, P. Mikkelsen, T. Nielsen, M. Misbih and P. Kryger[13]	SVM	Determines the infestation level of Varroa destructor
7	T. Campbell, K. Dixon, K. Dods, P. Fearns and R. Handcock[14]	Random Forest	Prediction of Honey harvest
8	Y. Zhao, G. Deng, L. Zhang, N. Di, X. Jiang and Z. Li[15]	SVM kNN RF Decision Tree	Detection of Air pollutants
9	F. Wang *et al.*[16]	GACNN SVM DNN Logistic Regression	Determines chemical poisoning of honey bees
10	X. Xu *et al.*[17]	SVM CDK	Prediction of Chemical acute toxicity of Honey bee
11	M. Bencsik, J. Bencsik, M. Baxter, A. Lucian, J. Romieu and M. Millet[18]	PCA	Identification of swarming process in Honey bee
12	R. Tashakkori, G. B. Buchanan and L. M. Craig[19]	-	Detection of Hive Robbery
13	A. Braga, B. Freitas, D. Gomes, A. Bezerra and J. Cazier[20]	LSTM	Determine sudden drop of temperature in bee colonies

that results in forecasting the internal temperature of the honey bee colony. If the unwanted phenomenon like queen loss which is dangerous to colony health could be automatically detected. The internal temperature of the hive is controlled by a particular mechanism developed by honey bees. The heat of the colony is removed by ventilation or fanning of wings and the colony is kept warm by pooling and metabolic heat. Here, homeostasis loss is predicted by ML. This helps the beekeeper to predict the internal temperature during the temperature loss. The dataset was collected from five individual beehives during the autumn season and the Arnia system uses three sensors i.e., temperature, humidity, and sound sensors to describe the internal condition of hive[20]. The *Apis mellifera* species from healthy colonies were considered for this study. For data pre-processing, the data was split into 50% of training, 20% of validation, and 30% of testing. Hyperparameter optimization was initially done with 50 epochs and since this was an underfitting model they took 150 epochs where the output was considerably better. The "error function" was minimized using Adam optimizer. LSTM model was built by adding the layers of NN. Three different architectural models were designed. The temperature and state of a beehive are predicted by LSTM algorithm. The declining pattern in the high-temperature range indicates a loss of control over the temperature of the hive. Many health problems in the colony can occur due to this drop in internal temperature. Bees spend most of the time collecting food and increasing colony size to manage themselves in the winter season. Later, they can start the swarming process. ML model-based algorithm was used to detect the decrease of temperature in a honey bee colony. After implementing these model beekeepers can be instantaneously alerted.

4.0 Conclusion

Apiculture, the practice of beekeeping, and the world of metals might appear disparate, but they intersect in some unexpected ways. One connection arises from the materials used in beekeeping equipment, such as hives and frames, which often involve various types of metal, like galvanized steel or aluminum. These materials provide durability and structural integrity to beekeeping infrastructure, ensuring the well-being of honey bee colonies. Furthermore, the environmental impact of metal extraction and processing can affect the health of honey bee populations, as pollutants from mining and metalworking industries can potentially contaminate the environment and disrupt ecosystems in which bees forage for nectar and pollen. The responsible management of metal production and disposal plays a part in preserving the natural environments where bees thrive. Thus, while apiculture and the metals industry may appear unrelated, they are connected through the materials used in beekeeping and the environmental considerations that affect bee health. Also, the conventional methods of beekeeping are practically laborious and have limited scope for colony inspection. Thus, the need for an effective hive monitoring technique is increasing significantly. This can be achieved by application of technologies like IoT, image processing and ML in the field of Apiculture. Various ML and DL models were implemented to analyse the processed numerical and visual data extracted through the sensors. The models were successful in forecasting the colony health, honey harvest and in identifying several threats encountered by them. Besides these applications, they also help in monitoring atmospheric pollutants. Therefore, ML models help beekeepers to avoid losses of the colony and maintain the hive effectively.

5.0 References

1. Alves T, Pinto M, Ventura P, Neves C, Biron D, Junior A, *et al*. Automatic detection and classification of honey bee comb cells using deep learning. Comput Electron Agric. 2020; 170:105-244. https://doi.org/10.1016/j.compag.2020.105244
2. Moussa M, Guedri W, Douik A. A novel metaheuristic algorithm for edge detection based on artificial bee colony technique. Traitement du Signal. 2020; 37(3):405-12. https://doi.org/10.18280/ts.370307
3. Kuraparthi S, Kollati M, Kora P. Ro-bust optimized discrete wavelet transform-singular value decomposition based video watermarking. Traitement du Signal. 2019; 36(6): 565-73. https://doi.org/10.18280/ts.360612
4. Berkaya SK, Gunal ES, Gunal S. Deep learning-based classification models for beehive monitoring. Ecol Inform. 2021; 64:101-353. https://doi.org/10.1016/j.ecoinf.2021.101353
5. Cejrowski T, Szymański J, Logofătu D. Buzz-based recognition of the honeybee colony circadian rhythm. Comput Electron Agric. 2020; 175:105-586. https://doi.org/10.1016/j.compag.2020.105586
6. Tom A, Schröder S, Steinhage V, Wittmann D. Biodiversity informatics in action: Identification and monitoring of bee species using ABIS. International Symposium Informatics for Environmental Protection; 2001. p. 425-30.
7. Braga AR, Gomes DG, Rogers R, Hassler EE, Freitas BM, Cazier JA. A method for mining combined data from in-hive sensors, weather and apiary inspections to forecast the health status of honey bee colonies. Comput Electron Agric. 2019; 169:105-61. https://doi.org/10.1016/j.compag.2019.105161
8. Zeybek S, Pham D, Koç E, Seçer A. An improved bees algorithm for training deep recurrent networks for sentiment classification. Symmetry. 2021; 13(8):13-47. https://doi.org/10.3390/sym13081347
9. Ngo T, Rustia D, Yang E, Lin T. Auto-mated monitoring and analyses of honey bee pollen foraging behavior using a deep learning-based imaging system. Comput Electron Agric. 2021; 187:106-239. https://doi.org/10.1016/j.compag.2021.106239
10. Boenisch F, Rosemann B, Wild B, Dormagen D, Wario F, Landgraf T. Tracking all members of a honey bee colony over their life-time using learned models of correspondence. Front Robot AI. 2018; 5: 1-10. https://doi.org/10.3389/frobt.2018.00035 PMid:33500921 PMCid:PMC7805663
11. Presern J, Skerl MS. Parameters influencing queen body mass and their importance as determined by ML in honey bees (*Apis mellifera carnica*). Apidologie. 2019; 50(5):745-57. https://doi.org/10.1007/s13592-019-00683-y
12. Nolasco I, Benetos E. To bee or not to bee: Investigating ML approaches for beehive sound recognition. Detection and Classification of Acoustic Scenes and Events Workshop 2018; 2018. p. 1-6.
13. Bjerge K, Frigaard C, Mikkelsen P, Nielsen T, Misbih M, Kryger P. A computer vision system to monitor the infestation level of Varroa destructor in a honeybee colony. Comput Electron Agric. 2019; 164:1-10. https://doi.org/10.1016/j.compag.2019.104898

14. Campbell T, Dixon K, Dods K, Fearns P, Handcock R. ML regression model for predicting honey harvests. Agriculture. 2020; 10(4):1-17. https://doi.org/10.3390/agriculture10040118
15. Zhao Y, Deng G, Zhang L, Di N, Jiang X, Li Z. Based investigate of beehive sound to detect air pollutants by ML. Ecol Inform. 2021; 61:1-9. https://doi.org/10.1016/j.ecoinf.2021.101246
16. Wang F. Graph attention convolutional neural network model for chemical poisoning of honey bees' prediction. Sci Bull. 2020; 65(14):1184-91. https://doi.org/10.1016/j.scib.2020.04.006 PMid:36659148
17. Xu X, Zhao P, Wang Z, Zhang X, Wu Z, Li W, *et al*. *In silico* prediction of chemical acute contact toxicity on honey bees via ML methods. Toxicol In Vitro. 2021; 72:1-7. https://doi.org/10.1016/j.tiv.2021.105089 PMid:33444712
18. Bencsik M, Bencsik J, Baxter M, Lucian A, Romieu J, Millet M. Identification of the honey bee swarming process by analysing the time course of hive vibrations. Comput Electron Agric. 2011; 76(1):44-50. https://doi.org/10.1016/j.compag.2011.01.004
19. Tashakkori R, Buchanan GB, Craig LM. Analyses of audio and video recordings for detecting a honey bee hive robbery. 2020 SoutheastCon, 2020 Mar 28-29, USA: Raleigh, NC; 2020. p. 1-6. https://doi.org/10.1109/SoutheastCon44009.2020.9249684
20. Braga A, Freitas B, Gomes D, Bezerra A, Cazier J. Forecasting sudden drops of temperature in pre-overwintering honeybee colonies. Biosyst Eng. 2021; 209:315-21. https://doi.org/10.1016/j.biosystemseng.2021.07.009