2024 Asia Pacific Conference on Innovation in Technology (APCIT)
Mysuru, India. Jul 26-27, 2024

Leveraging Machine Learning and Drone Technology for Effective Insect Pest Management in Agriculture

Mohammed Ali Shaik
School of computer science & Artificial Intelligence
SR University
Warangal, Telangana State, India
niharali@gmail.com

Masrath Parveen
Department of Information Technology,
Vidya Jyothi Institute of Technology,
Hyderabad, Telangana State, India
masrath19@gmail.com

Imran Qureshi
College of Computing and information Sciences,
AlMusanna College of Technology,
Sultanate of Oman
imranqureshi1210@gmail.com

2024 Asia Pacific Conference on Innovation in Technology (APCIT) | 979-8-3503-6153-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/APCIT62007.2024.10673597

***Abstract*- Insect pests pose a serious threat to food security and agricultural profitability, as they cause direct damage to crops and transmit plant diseases. Manual monitoring methods are inadequate to detect and prevent infestations, especially in large-scale farming operations. This article explores the potential of using drone technology and machine learning algorithms to enhance insect pest management in agriculture. Drones can capture high-resolution images of fields, while machine learning can analyze the data and identify signs of insect activity. However, this approach faces several challenges, such as the diversity and complexity of insect species, their behavior, and their impact on crops. Therefore, a multidisciplinary collaboration among researchers, farmers, technologists, and entomologists is essential to develop effective and robust solutions. This article aims to provide an overview of the current state and future prospects of using drone technology and machine learning for insect pest management in agriculture.**

Keywords—Deep learning, machine learning, Insect Pest, drone, decision tree.

I. Introduction

From the inception of agriculture to the present era, humanity has engaged in an unrelenting struggle against the insect kingdom. While certain insects serve as valuable agents of pollination, soil aeration, and natural pest management, a multitude of others have gained notoriety for wreaking havoc upon crops, thereby jeopardizing food security and agricultural profitability. These pernicious insects not only inflict direct losses upon yield, but also function as vectors for plant diseases, thereby compounding the challenges faced by farmers.

Insects have coevolved alongside plants for countless eons, meticulously refining their aptitude for exploiting them as sources of sustenance. The shared evolutionary history between crops and insects has rendered the latter remarkably adept at evading detection, particularly during their nascent stages of infestation. The task of identifying these diminutive invaders amidst the expansive expanse of agricultural fields is akin to searching for a needle in a haystack.

The “Food and Agriculture Organization” has projected that food production must escalate by approximately 70% by the year 2050 to cater to global requirements. Within this framework, every individual grain assumes paramount significance, making it indispensable to minimize losses attributable to insect pests. However, the challenge at hand is immense. Given the minuscule stature of numerous insects and the sheer magnitude of modern agricultural operations, manual monitoring methods are increasingly impractical.

Unmanned aerial vehicles equipped with state-of-the-art imaging systems can meticulously scan vast expanses of fields, identifying telltale signs of insect activity that may elude the human eye. The realm of machine learning algorithms has the remarkable ability to meticulously analyze vast volumes of data, meticulously identifying intricate patterns and formulating insightful predictions regarding potential infestations.

However, it is crucial to acknowledge that this journey has only just commenced. The meticulous and accurate identification of a diverse array of insect species, each with their own distinctive behavioral traits, life cycles, and patterns of destruction, presents itself as an imposing and formidable challenge. To overcome this hurdle, it is imperative to adopt a multidisciplinary approach that encompasses various fields of expertise.

The development of algorithms capable of discerning between beneficial insects and pests, detecting insects at different stages of their life cycle, and adapting to diverse environmental conditions necessitates the collaborative efforts of researchers, farmers, technologists, and entomologists. Only through this collective synergy can we effectively harness the boundless potential of modern technology in our unwavering crusade against the relentless menace of insect pests.

II. Literature Review

The amalgamation of artificial intelligence models and Internet of Things (IoT) strategies for the identification, classification, and enumeration of cotton pests, as well as their beneficial counterparts, was explored in a meticulous evaluation presented in the paper [1]. Within the realm of diverse cotton farming scenarios, the intrinsic capabilities and limitations of these technologies were illuminated [2]. A noteworthy technique entailed attracting insects through agents such as pheromones or light, capturing their visual representations using smartphone cameras, and subsequently subjecting these images to analysis through a Neural Network algorithm guided by the “Single Shot Multi-Task Detector (SSD)” [3]. The authors advocated for future investigations encompassing a broader spectrum of pest types, while considering factors such as variations in mobile phone resolution, inconsistencies in lighting due to crop growth, and the ever-changing weather conditions [4]. Moreover, there was an emphasis on expanding the repertoire of pest species detected through the implementation of AI and IoT, thereby accentuating the necessity for enhancing detection precision [5].

The utilization of wireless sensor networks (WSN) across various sectors, including agriculture, was meticulously examined [6]. WSNs in precision agriculture for the purpose of estimating fertilizer and pesticide requirements, as well as crop protection and the paper effectively categorized sensor node architecture into two types namely flat and hierarchical and further expounded upon WSN security in terms of operational and information security [7]. The significance of ensuring a high level of quality of service in WSNs where the strategy for monitoring pests was proposed involving the utilization of diverse sensors in the field to ascertain insect behavior employed through MATLAB, an image processing tool an artificial neural network-based approach was implemented to detect pests and formulate appropriate control measures. This approach demonstrated an improvement of 3.9% in accuracy compared to existing methods [8].

Reference [9] has conducted a comprehensive analysis of previous studies focusing on image-based recognition systems for mobile devices with a specific emphasis on “Convolutional Neural Network (CNN)” architectures. Drawing insights from these studies innovative real-time insect identification system specifically tailored for mobile devices utilizing the YOLOv5-S model and the system showcased commendable levels of accuracy when tested on two distinct insect datasets [10].

The study has magnificently showcased the immense potential of this model in furnishing farmers with indispensable insect-related data as the valuable information encompasses a wide array of characteristics, distribution patterns, and pest control strategies, thereby significantly contributing to the noble cause of sustainable pest management [11].

The detection and management of pests infesting stored products by examining a plethora of commercially available acoustic devices that exhibit remarkable proficiency in precisely identifying concealed insect infestations, be it within stored commodities, trees, or soil [12]. It is worth highlighting their emphasis on the seamless integration of cutting-edge digital signal processing techniques with statistical tools such as neural networks and machine learning [13]. The harmonious amalgamation of these technologies plays a pivotal role in accurately differentiating specific pests from the general ambient noise and other insects. This groundbreaking capability has paved the way for automated systems that diligently monitor the abundance and distribution of pests within storage environments. The authors also judiciously synthesized insights gleaned from diverse research conducted across various regions spanning Europe, Asia, the United States, and Africa. However, it is significant to acknowledge that the initial focus of this paper was not to introduce novel experimental findings but rather to provide a comprehensive literature review [14].

The existing literature pertaining to the automatic detection and monitoring of the notorious codling moth using ingeniously designed traps equipped with high-resolution cameras [15]. Soto astutely evaluated an assortment of both commercial and prototype automated traps, along with ingenious insect counting techniques specifically tailored for apple orchards, in order to gauge the population of the notorious [16]. This review was conducted with utmost diligence, drawing insights from a multitude of research articles, prior reviews, online materials, and Soto's own personal experiences and the overarching objective was to discern the current state of the art in monitoring the elusive utilizing these revolutionary traps [17].

The efficacy of machine learning in detecting insects on sticky traps within tomato crops was meticulously explored by harnessing images captured from traps nestled within Portuguese tomato fields, the researchers embarked on the arduous task of manually annotating these images using the sophisticated CVAT software. These painstakingly annotated images were then utilized to train a highly sophisticated model, which was seamlessly integrated with Nuclio to yield instantaneous results on the CVAT website. Leveraging the formidable YOLOv5 algorithm, the researchers achieved an awe-inspiring mAP_0.5 score of 94.4%, complemented by an impressive precision of 88% and a commendable recall of 91% [18].

An enlightening exploration of the potential wielded by deep learning's meta-architectures, with a particular focus on the remarkable TensorFlow framework, in the realm of insect identification is performed to implement the pivotal defensive role played by plant secondary metabolites (PSMs) in countering both biotic and abiotic stressors with a special emphasis on the remarkable glucosinolates found within the illustrious Brassicaceae family [19]. The research ingeniously introduced the groundbreaking concept of environmentally acquired chemical camouflage (EACC), wherein plants deftly absorb volatiles emitted by neighboring VOC-emitting plants, potentially rendering themselves impervious to insect detection [20]. To validate this captivating concept, the study subjected broccoli plants to the volatiles emitted by the remarkable Rhododendron tomentosum, meticulously assessed the preference of the illustrious Pieris brassicae between the exposed and control plants, and meticulously studied the consequential impact on the parasitoid wasp Cotesia glomerata [21]. The results proved to be truly enlightening, showcasing that the Rhododendron tomentosum-exposed plants were significantly less appealing for P. brassicae egg-laying, while the orientation of C. glomerata remained unaffected. This intriguing finding strongly suggests the immense potential of EACC as a supplementary tool in the realm of integrated pest management, thereby opening up new avenues for innovative approaches in combating pests [22].

An “ingenious real-time Polymerase Chain Reaction (PCR)” methodology is developed for identification of “Gryllodes sigillatus (GS)” in feed and food and this cutting-edge approach specifically targets the “mitochondrial cytochrome b (CYB)” gene, enabling the detection of even minuscule quantities of GS DNA, as low as 5g/100g, even in complex mixtures [23]. The study unequivocally validates the practicality of this method in confirming the presence or non-presence of GS-labeled material in both feed and food which exhibits exceptional sensitivity, accuracy, and facilitated precise differentiation between insects and GS [24]. While the specificity of the PCR product was confirmed through sequencing, potential challenges such as matrix complexity and DNA integrity should be taken into consideration and this technique consistently identifies the target GS DNA and holds promise for future quantitative assessments with appropriate calibration curves [25-28].

III. Implementation

The precise methodologies we employed are devoted to execution through prognostic approach through fabricate machine learning models, commences once we have concluded the phases of comprehending and preparing the data. The construction of the actual models, both for forecasting and terminology, is what is meant by the term "model implementation." Taking into account the exigencies of the business and the dataset at hand, we have previously determined the appropriate libraries and algorithms for implementation, and we have furnished the specific details pertaining to the actualization of these selections within this section of the textual material. Our application was fashioned with PyCharm, an integrated development environment specifically tailored for Python application development. The chosen integrated development environment streamlines the process of composing Python code, appending new files, executing the model in the terminal, and establishing a virtual environment for the execution of the Python application. Furthermore, the IDE offers an advanced option for incorporating a Python flask web application for this very purpose.

We have selected an "integrated development environment (IDE)" that stipulates a user friendly platform for implementing code in Python. This IDE offers various functionalities, including the ability to add new directories, execute Python scripts in the terminal, and create virtual environments for running Python applications. The resulting application primarily consists of code segments dedicated to forecasting and terminology modeling. It also incorporates a well-designed user interface, enabling users to submit videos through the "left hand panel" and examine the corresponding results on the "right hand panel".

To form the core of the application, we have integrated a powerful ML model. This model serves as the primary component, driving the functionality of the entire application. In addition to the model, we have included a CSS style sheet and JQuery to enhance the visual appeal of the user interface and enable script-based interactivity within the application.

The communication between the application and the model is facilitated through application programming interface (API) methods. This communication channel ensures that user input is seamlessly transmitted to the backend for model execution. The backend then processes the input data and returns the relevant results to the user interface. In our specific scenario, the user input is in the form of video clips, while the returned output consists of significant changes or anomalous behaviors detected in the input data, presented as a GIF image path.

A. *Libraries*

For the execution, we employed libraries like "Keras, Numpy, Pandas, Tensorflow, Scipy, OpenCV, Flask, and Matplotlib". Keras is a specialized Python framework and module for neural network methodologies. It has the ability to execute TensorFlow and other software libraries that rely on Keras as their foundation. It encompasses a greater multitude of commonly utilized neural network features and encompasses a tool for processing visual and textual data. In the context of this thesis, we employed the Keras high-level neural networks API to construct a model utilizing a profound machine learning mechanism.

- Keras: Google Keras serves as a sophisticated deep learning application programming interface (API) for the creation of neural networks. This Python-built framework simplifies the process of constructing neural networks and facilitates the computation of diverse backend neural networks.
- Numpy: is a remarkable Python library, permeating an extensive array of research and engineering domains. Unquestionably, it reigns as the quintessential Python standard for the manipulation of numerical data, serving as an indispensable pillar within the scientific Python and PyData realms.
- Pandas: it's a open-source Python framework utilized for the realm of data science, data analysis, and machine learning, has garnered considerable popularity. This library is built upon Numpy, an additional framework that provides support for multidimensional arrays.
- Tesorflow: is a Python library renowned for its expeditious numerical computation capabilities. Serving as a fundamental framework, TensorFlow empowers developers to craft Deep Learning models either directly or by utilizing supplementary libraries that are built upon it.
- Scipy: presents a remarkable assortment of mathematical methodologies and user-friendly utilities. This invaluable resource expands the potentialities of interactive Python sessions, empowering users with elevated functions and interfaces for data manipulation and visualization.
- OpenCV: it facilitates the implementation of computer vision and image processing tasks. It boasts a multitude of functionalities, including but not limited to object detection, facial recognition, and tracking.
- Flask: it doesn't impose the use of specific tools or libraries, giving developers the freedom to choose what suits their needs best. However, this minimalistic approach comes at the cost of certain conveniences commonly provided by third-party libraries.
- Matplotlib: to craft visualizations that are static, interactive, and dynamic in nature and it simplifies the creation of basic plots while enabling the realization of complex visualizations. With Matplotlib, you can generate visually appealing figures suitable for publication, complete with the ability to zoom, pan, and update them.

B. *Step 1: video of insects*

The current investigation employed a continuous video recording spanning a duration of 30 days, capturing the activity of insects within a specific agricultural region. The rationale behind documenting the insect population in the field is multifaceted. Insects pose a dual threat to agricultural endeavors, contributing to two primary categories of damage.

- Firstly, there is direct harm inflicted upon plants when voracious insects engage in the consumption of leaves or engage in burrowing activities within stems, fruit, or roots. This detrimental impact is observed across a wide array of pest species, encompassing orthopterans, heteropterans, homopterans, lepidopterans, coleopterans, and dipterans, in both their larval and adult stages.
- The second manifestation of damage takes on an indirect nature. In this scenario, the insect in question may not directly inflict harm upon the crop, but rather serves as a vector for the transmission of bacterial, viral, or fungal infections.

Fig. 1. Insect detection in a specific field

Sugar beetroot and potato viral infections, for instance, are transmitted from one plant to another by aphids. It is worth noting that grasshoppers defy the norm observed in most insects, as they not only infest the crops they feed on but also possess a unique ability to survive in a relatively harmless solitary phase for extended periods. During this time, their population can increase substantially. Eventually, they transition into a gregarious phase, giving rise to massive migratory swarms that can travel up to a thousand miles with the assistance of winds or flight.

C. Step 2: preprocessing

The task at hand, in precise terms, pertains to the prognostication of forthcoming frames within videos depicting the realm of agriculture, specifically encompassing the presence of "weeds, insects, and pests". This task, commonly referred to as video prediction, involves the anticipation and projection of future visual content within such videos.

Let $ABP_t \in R_w \times i \times p_{be}$ the t-th frame in the video sequence

$$ABP = (ABP_{t-n}, \ldots, ABP_{t-1}, ABP_t) \quad (1)$$

The objective is to anticipate the forthcoming frames based on the given input ABP, where the variables w, i, and p represent weed, inserts, and pest detections, correspondingly.

$$Yield = (Yield^{t+1}, Yield^{t+2}, \ldots, Yield^{t+m}) \quad (2)$$

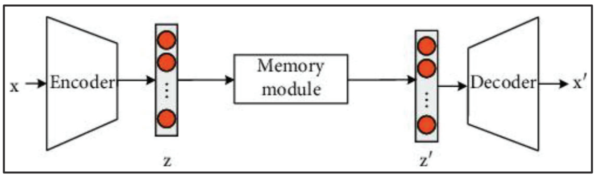

Fig. 2. The flow chart of the proposed data processing

The process of validating the system's understanding of underlying patterns in input data is achieved through learning by prediction. This approach assumes that accurate representations are necessary for effective forecasts, making it a suitable framework for transfer learning. The predictive learning paradigm, along with ML, ARIMA models, and time series forecasting, revolves around predicting future events based on previous inputs. Specifically, in the context of video prediction, the goal is to forecast subsequent frames based on a given series of frames. This involves generating a continuous video using a sequence of image sequences. Unlike unconditioned video creation, video prediction relies on pre-learned representations from a set of input frames.

D. Step 3: Problem identification and detection

The agriculture plays a imperative role in driving financial growth and determining the standard of living. It encompasses both the cultivation of crops and the processing of food, serving as a significant industry in every country. Moreover, it contributes to enhancing the export quality of agricultural and various food products, thereby bolstering the economy. The burgeoning growth of food processing in developing nations can be attributed to the demand from both the domestic market and the opportunity for increased export revenues. However, the agricultural sector faces numerous challenges, such as the need for proper storage facilities, equipment maintenance, and workstations. One of the most pressing issues that plagues agriculture is pest infestation, which leads to a decline in crop quality. Pests, including insects, viruses, and weeds, cause substantial losses in crop yield, resulting in a limited market for the final yield or outcome.

To combat these challenges, there are various "chemical and biological pest control" measures available to ensure these treatments are effective, comprehensive monitoring throughout the entire agricultural site is necessary. Typically, monitoring is conducted reflexively by personnel while carrying out their regular tasks. However, this approach has its limitations, as by the time an incursion is to be detected before considerable damage may have already occurred. Initial detection of pests requires a additional systematic approach over larger farms. Traps are widely used for pest monitoring, capable of effectively sampling insect inhabitants across the absolute area of interest when properly designed.

Considerable efforts have been invested in developing more efficient pest identification and classification technologies. Some systems focus on identifying the damage caused by pests are to be direct detection process that remains the preferred method. Initially the research is explored the use of acoustic analysis to detect and identify insects based on the sounds they produce, but interest in this approach has diminished in recent years. Therefore, proximate images are still widely utilized. While research is underway to investigate the potential of "multispectral, hyperspectral, infrared, and X-ray sensors, traditional RGB

(Red-Green-Blue) sensors" continue to govern due to their affordability.

E. Step 4: Statistical features

- ARIMA Models: ARIMA models, known by the nomenclature ARIMA (p,d,q), encapsulate the key components of autoregressive modeling, differencing, and moving-average modeling. The parameter p signifies the autoregressive model's order, while d indicates the degree of differencing applied to convert a "non-stationary time series into a stationary" one. Lastly, q denotes the order of the moving-average model. By harnessing the power of differencing, ARIMA models have the capacity to forecast forthcoming values by leveraging past data.
- Paired Sample 't' Test: The balancing sample t-test which is referred as the dependent sample t-test as a statistical procedure that requires to assess. The average difference between two sets of data is zero as to test each subject or object is measured twice to a collection of paired observations.
- Multiple Regression: is a statistical technique employed to explore the association between a singular reliant variable and various independent variables where the objective is to analyze and to utilize various known independent variables that tends to predict the individual dependent aspect.

F. Step 5: factoring the prediction space

Our aim is to predict future frames based on an initial image and the positions of objects within the scene. This task is challenging due to the occurrence of numerous entities with diverse dynamics and interactions, as well as the inherent multi modality of the prediction task and to overcome these challenges, we propose a novel approach.

Instead of modeling the overall changes in the scene, we focus on modeling the changes in individual objects. We employ an entity predictor that generates per-entity representations, capturing the expected location and predicted attributes for each object. This factorization allows for efficient prediction of future frames based on these object representations. However, to generate the actual pixel values we need to take an additional step.

To address this, we utilize a "frame decoder" that retains the attributes of each object, respects their expected positions, and resolves conflicts such as occlusions when constructing the final image. Furthermore, we introduce a "global random latent variable" which will capture the inherent ambiguities over videos that consists of factors like weed, insects, and pests.

This latent variable, in turn, determines initially timestamped by latent variables that aid in predicting future frames. The predictor takes the per-entity depiction and the latent variable as input and forecasts the entity representations for the succeeding timestep. Using these predictions along with the initial frame to account for the background, the decoder constructs the anticipated frame. Our model is trained to maximize the likelihood of the training sequences, incorporating terms for both frames and entity locations.

However, directly optimizing the likelihood poses challenges in models with unobserved latent factors. To address this, we exploit a variational lower bound instead. Additionally, we train a latent encoder module using the target video, which forecasts a distribution across the latent variable. This encoder module aids in capturing the underlying uncertainties in the data.

By combining these components and training them effectively, we aim to accurately forecast future frames based on an initial image and the positions of objects within the scene.

It is important to clarify that the future frame or location annotation, as well as the hidden decoder, are exclusively utilized during the training phase. However, during inference, a single image is used as input, along with the predicted or known positions of the existing objects. This enables the generation of multiple plausible future frames. To provide a comprehensive understanding, let us delve deeper into the prediction, decoder, and encoder modules before presenting the complete training objective.

G. Step 6: Machine Learning Methods

RF, SVM, and KNN techniques were applied to analyze films depicting weed, insect, and pest occurrences. The utilization of RF, SVM, and KNN in this investigation stems from several key reasons. SVM, known for its remarkable swiftness, can classify 12-megapixel aerial photographs in a mere 10 seconds, whereas KNN requires approximately 40 to 50 seconds to accomplish the same task. Random Forest, specifically designed for multiclass problems, contrasts with SVM, which is tailored for two-class problems. The multiclass problem of weed, insect, and pest videos was subdivided into various binary classification tasks. Random Forest exhibits remarkable performance when dealing with both numerical and categorical attributes, while SVM excels in handling both linear and nonlinear solutions. In instances where the data possesses a high signal-to-noise ratio (SNR), KNN outperforms linear regression. Lastly, when it comes to robustness and accuracy, Random Forest surpasses decision trees.

H. Step 7: Evaluation metrics

Based on the outcomes of machine learning (ML), moving averages, multiple regression, and paired sample "t" test, it has been observed that the implemented weed control treatments have significantly contributed to the augmentation of crop yield on the specific agricultural land. By extracting the reflectance values from the red, green, and blue bands of an RGB image captured from a small section of the farm and we were able to derive vegetation indicators such as the "normalized red band or normalized green band or the normalized blue band". This normalization process was conducted to mitigate the impact of varying lighting conditions on the color frequencies.

Subsequently, we calculated various indices including the greenness index, "excess green (ExG)", "excess red (ExR)", "excess green and red (ExGExR)", and the "excess green and red (ExGExR)", where Rband represents the reflectance value of a specific band and the ExG index measures the disparity between the noticed light levels in the "green, red, and blue" channels. In order to optimize weed detection results, we incorporated the "r, g, b, GI, and ExGExR" indices into our machine learning methodology.

To retrain our "Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN)" algorithms, we necessitated a labeled image. In this particular case, we employed one of the classic software to classify a small portion of the image into three categories: weeds, crops, and bare land, which served as training examples for the three aforementioned machine learning algorithms. Through the use of Region of Interest (ROI) tools in ENVI, we manually labeled selected regions as “weeds, crops, and bare land” after which we employed a “neural network classifier” to classify the remaining areas of the image. The entire process resulted in a “labeled image” where crops are represented in “green, weeds in red, and barren ground in white”. This labeled images were utilized to train and detect weeds using our RF, SVM, and KNN algorithms.

- YP = Veracious affirmative = The tally of occurrences wherein a weed, insect, or vermin is precisely discerned.
- YN = Veracious negation = When both cultivated land and uninhabited terrain are precisely identified, the tally of occurrences escalates.
- FP = Erroneous affirmative = The tally of records wherein cannabis is erroneously detected.
- FN = Erroneous negation = When the identification of cultivated land and uninhabited terrain is flawed, the tally of occurrences rises.
- P = Complete affirmative = YP + FN.
- N = Complete negation = YN + FP.

We assessed and gauged the presentation metrics for “RF, SVM, and KNN classifiers” by employing the prescribed formulas. These formulas allowed us to determine the accuracy of classification, memory, sensitivity, accuracy, “false positive rate, and kappa coefficient (kc)”.

Accuracy = YP + YN P + N
Recall = Y_P Y_P + F_N
Precision = Y_P Y_P + F_P
Specificity = Y_N Y_N + F_P
FPR = 1 − Specificity
kappa coefficient = max([A−kc 1−kc ∗ kc−A1−A])
where A is accuracy and;

kc = (P(YP+FP))+(N(FN +YN)) $(YP+YN+FP+FN)^2$

IV. Results

The matrices pertaining to these data sets are also displayed within table 5.2, encapsulating a comprehensive array of performance metrics such as “accuracy, recall, specificity, precision, false positive rate (FPR), and kappa coefficient”. It is worth noting that the accuracy of the Random Forest (RF) algorithm on the test dataset was an impressive 0.918, showcasing its superiority over the “K-Nearest Neighbors (KNN) and Support Vector Machine (SVM)” techniques. Furthermore, RF exhibits a commendable recall and specificity value of 0.933, further solidifying its dominance in this domain.

TABLE I. Summary of the Approaches Reviewed

Sl.No.	Performance Metric	RF
1	Accuracy	0.915
2	Recall/Sensitivity	0.931
3	Specificity	0.963
4	Precision	0.923
5	False Positive Rate	0.063
6	Kappa Coefficient	0.943

In the realm of insect detection from video-captured insect photos, the Random Forest (RF) algorithm has proven to be a formidable contender. Its “accuracy, False Positive Rate (FPR), and kappa coefficient” stand at an impressive 0.066 and 0.944, respectively. Not only does RF excel in accuracy, but it also outshines its counterparts across all other evaluation metrics. SVM, another commendable classifier, delivers a laudable performance as well. However, when it comes to accuracy, RF triumphs over SVM.

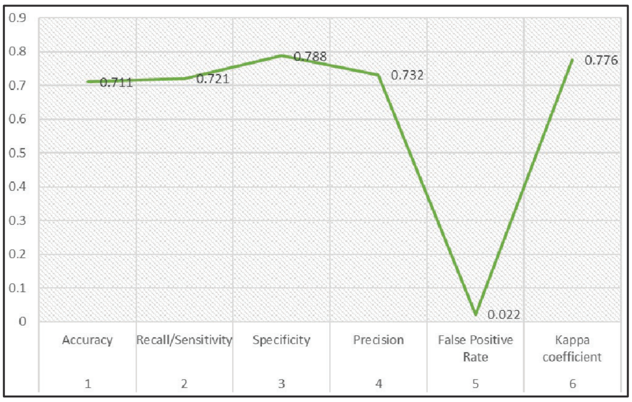

Fig. 3. Comparision of insect detection performance metrics for RF

Regrettably, the K-Nearest Neighbors (KNN) algorithm falls short on our dataset, demonstrating subpar performance. Consequently, both RF and SVM emerge as powerful and efficient classifiers for the task at hand - detecting insects from video-captured insect photos. For a visual representation of RF's insect detection performance metrics, refer to Figure 3.

TABLE II. Comparision of Insect detection performance metrics for the KNN

Sl.No.	Performance Metric	RF
1	Accuracy	0.712
2	Recall/Sensitivity	0.723
3	Specificity	0.765
4	Precision	0.743
5	False Positive Rate	0.023
6	Kappa Coefficient	0.777

These matrices can be observed in table 2, showcasing a plethora of performance metrics such as “accuracy, recall, specificity, precision, false positive rate (FPR), and kappa coefficient”. Notably, the accuracy achieved by K-Nearest Neighbors (KNN) on the test dataset reached a commendable 0.711. However, it is worth mentioning that Random Forest (RF) surpassed the performance of both KNN and Support Vector Machines (SVM) in this regard. KNN demonstrated a recall and specificity of 0.721, highlighting its proficiency in these areas. Conversely, RF boasted an accuracy of 0.022, an FPR of 0.776, and a kappa coefficient of 0.776, underscoring its superiority over the other models. Remarkably, KNN emerged as the frontrunner across all other categories, surpassing its counterparts in terms of performance. SVM, while not to be disregarded, fell slightly short of RF's accuracy prowess.

In stark contrast, the KNN algorithm exhibited subpar performance when applied to our specific dataset. Consequently, it is evident that both the “Random Forest (RF)” and “Support Vector Machine (SVM)” classifiers prove to be highly proficient in accurately detecting insects from photographs extracted from the video footage. The visual representation depicted in Figure 4 succinctly illustrates the comprehensive comparison of performance metrics pertaining to insect detection as achieved by the KNN algorithm.

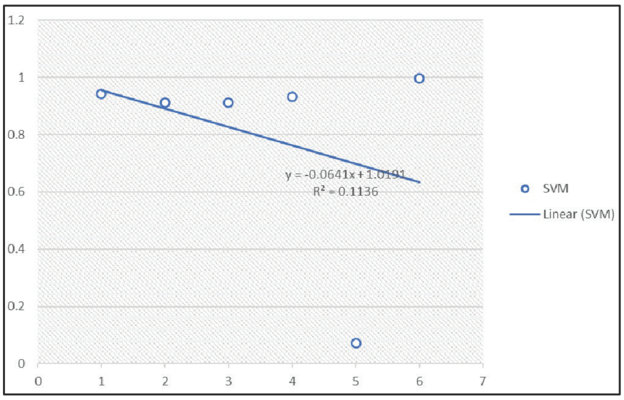

Fig. 4. Comparision of insect detection performance metrics for KNN

TABLE III. SUMMARY OF THE APPROACHES REVIEWED

Sl.No.	Performance Metric	RF
1	Accuracy	0.954
2	Recall/Sensitivity	0.912
3	Specificity	0.912
4	Precision	0.923
5	False Positive Rate	0.072
6	Kappa Coefficient	0.998

The corresponding matrices are also illustrated in table 3, encompassing a plethora of performance metrics such as “accuracy, recall, specificity, precision, false positive rate (FPR), and kappa coefficient”. The test dataset witnessed a commendable accuracy of 0.944 for SVM, with RF surpassing both KNN and SVM in terms of proficiency. Notably, SVM exhibited an impressive recall and specificity of 0.913. As for RF, its accuracy stood at 0.071, with an FPR of 0.996 and a kappa coefficient of the same magnitude. Concurrently, RF emerged triumphant over its counterparts in all other categories, mirroring its superiority to a remarkable extent. SVM, on the other hand, performed admirably, albeit falling slightly short of RF's accuracy. Conversely, KNN failed to deliver satisfactory results on our dataset. Consequently, both RF and SVM serve as highly effective classifiers for the identification of insects in photographs extracted from the video. The visualization in figure 5 effectively conveys the comparison of insect detection performance metrics for SVM.

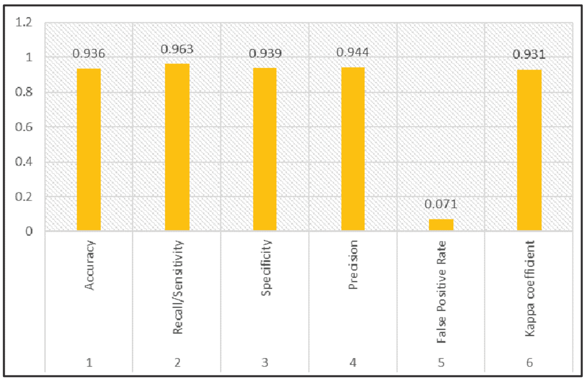

Fig. 5. Comparision of insect detection performance metrics for SVM

V. CONCLUSION

The realm of insect detection from video-captured insect photos, the Random Forest (RF) algorithm has emerged as a formidable contender. It showcases exceptional accuracy, with a False Positive Rate (FPR) and kappa coefficient that are truly impressive, standing at 0.066 and 0.944, respectively. The KNN achieves a commendable accuracy of 0.711 on the test dataset, but it is worth noting that RF surpasses both KNN and SVM in terms of accuracy. KNN demonstrates proficiency in recall and specificity, with values of 0.721. On the other hand, RF exhibits an accuracy of 0.022, an FPR of 0.776, and a kappa coefficient of 0.776, underscoring its superiority. With SVM achieving an impressive accuracy of 0.944 on the test dataset. However, RF surpasses both KNN and SVM in terms of proficiency. SVM showcases an impressive recall and specificity of 0.913. In contrast, RF achieves an accuracy of 0.071, an FPR of 0.996, and a kappa coefficient of the same magnitude. Once again, RF outshines its counterparts, demonstrating its remarkable superiority.

Overall, both RF and SVM emerge as highly effective classifiers for the identification of insects in photographs extracted from the video with the superiority of RF in terms of accuracy.

REFERENCES

[1] W. Swinkels, L. Claesen, F. Xiao and H. Shen, "SVM point-based real-time emotion detection," 2017 IEEE Conference on Dependable and Secure Computing, Taipei, 2017.

[2] K. K. Patro, S. Devipriya, T. Praveen, M. J. Rao, H. V. Kumar and B. Sneha, "Human Facial Emotions Recognition Using Customized Deep Convolutional Neural Network," 2023 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India, 2023, pp. 693-697, doi: 10.1109/AIC57670.2023.10263825.

[3] Mohammed Ali Shaik and Dhanraj Verma, (2022), "Prediction of Heart Disease using Swarm Intelligence based Machine Learning Algorithms", International Conference on Research in Sciences, Engineering & Technology, AIP Conf. Proc. 2418, 020025-1–020025-9; https://doi.org/10.1063/5.0081719, Published by AIP Publishing. 978-0-7354-4368-6, pp. 020025-1 to 020025-9

[4] M. A. Shaik, M. Y. Sree, S. S. Vyshnavi, T. Ganesh, D. Sushmitha and N. Shreya, “Fake News Detection using NLP”, 2023 International Conference on Inovative Data Communication Technologies and Application (ICIDCA), Uttarakhand, India, 2023, pp. 399-405, doi: 10.1109/ICIDCA56705.2023.10100305.

[5] V. Upadhyay and D. Kotak, "A Review on Different Facial Feature Extraction Methods for Face Emotions Recognition System", 2020 Fourth International Conference on Inventive Systems and Control (ICISC), pp. 15-19, 2020.

[6] Mohammad Alsaffar, Abdullah Alshammari, Gharbi Alshammari, Tariq S Almurayziq, Saud Aljaloud, Dhahi Alshammari, Assaye Belay, "Human-Computer Interaction Using Manual Hand Gestures in Real Time", Computational Intelligence and Neuroscience, vol. 2021, Article ID 6972192, 5 pages, 2021. https://doi.org/10.1155/2021/6972192

[7] M. A. Shaik, R. Sreeja, S. Zainab, P. S. Sowmya, T. Akshay and S. Sindhu, “Improving Accuracy of Heart Disease Prediction through Machine Learning Algorithms”, 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), Uttarakhand, India, 2023, pp. 41-46, doi: 10.1109/ICIDCA56705.2023.10100244.

[8] S. Depuru, A. Nandam, S. Sivanantham, K. Amala, V. Akshaya and M. Saktivel, "Convolutional Neural Network based Human Emotion Recognition System: A Deep Learning Approach", 2022 Smart Technologies Communication and Robotics (STCR) Sathyamangalam India, pp. 1-4, 2022.

[9] Patro, Kiran Kumar, Jaya Prakash Allam, Bala Chakravarthy Neelapu, Ryszard Tadeusiewicz, U. Rajendra Acharya, et al., "Application of Kronecker convolutions in deep learning technique

for automated detection of kidney stones with coronal CT images", Information Sciences, vol. 640, pp. 119005, 2023.

[10] Mohammed Ali Shaik, M. Varshith, S. SriVyshnavi, N. Sanjana and R. Sujith, “Laptop Price Prediction using Machine Learning Algorithms”, 2022 International Conference on Emerging Trends in Engineering and Medical Sciences (ICETEMS), Nagpur, India, 2022, pp. 226-231, doi: 10.1109/ICETEMS56252.2022.10093357.

[11] M. A. Shaik, S. k. Koppula, M. Rafiuddin and B. S. Preethi, (2022), "COVID-19 Detector Using Deep Learning”, International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2022, pp. 443-449, doi: 10.1109/ICAAIC53929.2022.9792694.

[12] Kiran Kumar Patro, Allam Jaya Prakash, M. Hammad, Ryszard Tadeusiewicz and Pawel Plawiak, "SCovNet: A Skip Connection-based Feature Union Deep Learning Technique with Statistical Approach Analysis for the Detection of COVID-19", Biocybernetics and Biomedical Engineering, Feb. 2023, [online] Available: https://doi.org/l0.l016/j.bbe.2023.01.005.

[13] Mohammed Ali Shaik, Geetha Manoharan, B Prashanth, NuneAkhil, Anumandla Akash and Thudi Raja Shekhar Reddy, (2022), "Prediction of Crop Yield using Machine Learning", International Conference on Research in Sciences, Engineering & Technology, AIP Conf. Proc. 2418, 020072-1–020072-8; https://doi.org/10.1063/5.0081726, Published by AIP Publishing. 978-0-7354-4368-6, pp. 020072-1 to 020072-8

[14] M. Karnati, A. Seal, D. Bhattacharjee, A. Yazidi and O. Krejcar, "Understanding Deep Learning Techniques for Recognition of Human Emotions Using Facial Expressions: A Comprehensive Survey", IEEE Transactions on Instrumentation and Measurement, vol. 72, no. 5006631, pp. 1-31, 2023.

[15] Mohammed Ali Shaik and Dhanraj Verma, (2020), Deep learning time series to forecast COVID-19 active cases in INDIA: A comparative study, 2020 IOP Conf. Ser.:Mater.Sci.Eng. 981 022041, doi.org/10.1088/1757-899X/981/2/022041

[16] S. Thuseethan, S. Rajasegarar and J. Yearwood, "Deep Continual Learning for Emerging Emotion Recognition", IEEE Transactions on Multimedia, vol. 24, pp. 4367-4380, 2022.

[17] Mohammed Ali Shaik, “A Survey on Text Classification methods through Machine Learning Methods”, International Journal of Control and Automation (IJCA), ISSN:2005-4297,Volume12,Issue-6 (2019), Pp.390-396.

[18] Praveen, P.; Shaik, M.A.; Kumar, T.S.; Choudhury, T. Smart Farming: Securing Farmers Using Block Chain Technology and IOT. In Blockchain Applications in IoT Ecosystem; Springer: Cham, Switzerland, 2021; pp. 225–238.

[19] Mohammed Ali Shaik, “Time Series Forecasting using Vector quantization”, International Journal of Advanced Science and Technology (IJAST), ISSN:2005-4238,Volume-29,Issue-4 (2020), Pp.169-175.

[20] T. K. Kumar and D. S. Tummala, "Artificial Intelligence-Based Real-Time Facial Emotion Monitoring System," 2023 9th International Conference on Computer and Communication Engineering (ICCCE), Kuala Lumpur, Malaysia, 2023, pp. 361-366, doi: 10.1109/ICCCE58854.2023.10246038.

[21] Mohammed Ali Shaik, Dhanraj Verma, (2021), Agent-MB-DivClues: Multi Agent Mean based Divisive Clustering, Ilkogretim Online - Elementary Education, Vol 20(5), pp. 5597-5603, doi:10.17051/ilkonline.2021.05.629

[22] P. Dhope and M. B. Neelagar, "Real-Time Emotion Recognition from Facial Expressions using Artificial Intelligence", 2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), pp. 1-6, 2022.

[23] Mohammed Ali Shaik, Praveen Pappula, T. Sampath Kumar, Battu Chiranjeevi,"Ensemble model based prediction of hypothyroid disease using through ML approaches", AIP Conference Proceedings, 2971, 020038 (2024), https://doi.org/10.1063/5.0196055

[24] Mohammed Ali, P. Praveen, Sampath Kumar, Sallauddin Mohmmad, M. Sruthi, "A survey report on cloud based cryptography and steganography procedures", International Conference on Research in Sciences, Engineering, and Technology, AIP Conf. Proc. 2971, 020040-1–020040-8; https://doi.org/10.1063/5.0196050

[25] Mohammed Ali Shaik, P. Praveen, T. Sampath Kumar, Masrath Parveen, Swetha Mucha, "Machine learning based approach for predicting house price in real estate", International Conference on Research in Sciences, Engineering, and Technology, AIP Conf. Proc. 2971, 020041-1–020041-5; https://doi.org/10.1063/5.0196051

[26] C. Dalvi, M. Rathod, S. Patil, S. Gite and K. Kotecha, "A Survey of AI-Based Facial Emotion Recognition: Features ML & DL Techniques Age-Wise Datasets and Future Directions", IEEE Access, vol. 9, pp. 165806-165840, 2021.

[27] Mohammed Ali Shaik and Dhanraj Verma, (2020), Enhanced ANN training model to smooth and time series forecast, 2020 IOP Conf. Ser.: Mater. Sci. Eng. 981 022038, doi.org/10.1088/1757-899X/981/2/022038

[28] Mohammed Ali Shaik, Dhanraj Verma, P Praveen, K Ranganath and Bonthala Prabhanjan Yadav, (2020), RNN based prediction of spatiotemporal data mining, 2020 IOP Conf. Ser.: Mater. Sci. Eng. 981 022027, doi.org/10.1088/1757-899X/981/2/022027