International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

CNN Model for Smart Agriculture

Prof. Shivarudraiah B *Dept. of ECE BMSIT&M* Bengaluru, India shivab@bmsit.in

Ashish *Dept. of ECE BMSIT&M* Bengaluru, India shankerashish9@gmail.com

Avani Suresh Kengal *Dept. of ECE BMSIT&M* Bengaluru, India avanikengal21@gmail.com

Kuchipudi Sowmya Jasmine *Dept. of ECE BMSIT&M* Bengaluru, India sowmyakuchipudi16@gmail.com

Hemavathi KR *Dept. of ECE BMSIT&M* Bengaluru, India hemavathi242@gmail.com

***Abstract*— Precision farming is being revolutionized by the integration of innovative machine learning and computer vision methods. Identifying and classifying weeds and crops accurately remains a major challenge in this field, which has a direct effect on optimizing the yield as well as sustainability. In this work, an approach to smart weed detection based on deep learning using Convolutional Neural Networks (CNN) for feature learning followed by comparison of classifiers to select the best-performing model is introduced. In our research, InceptionV3 was utilized to extract features, and four classifiers—SoftMax, Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF)— were compared. Among them, the Random Forest classifier performed better than others with a validation accuracy of 99.57% and an F1 score of 0.99. Extending the successful application of crop-weed detection, the model was transferred to a new application: forest fire detection. Employing the same CNN- based feature extraction pipeline and Random Forest classification, our system showed high accuracy on a forest fire dataset. In addition, we implemented a real-time detection system using webcam feeds with a processing speed of around 30 frames per second, making practical deployment for environmental monitoring possible. This study not only confirms the efficacy of the union of CNNs and ensemble learning but also exemplifies the versatility of the model architecture in both agricultural and environmental contexts.**

Index Terms*— Convolutional Neural Networks (CNN), InceptionV3, Random Forest, Weed Detection, Crop Classification, Forest Fire Detection, Real-Time Image Processing, Machine Learning, Deep Learning, Precision Agriculture, Environmental Monitoring, Feature Extraction, Webcam Detection.

I. INTRODUCTION

A. *Background*

The advent of artificial intelligence (AI), particularly in fields such as deep learning and computer vision, has had a profound impact on traditional problem-solving approaches in agriculture and environmental management. Agriculture, a discipline highly reliant on timely and accurate judgment, is increasingly being supported by intelligent systems to automate crop health monitoring, pest detection, and yield estimation. Among these, the identification of weeds and their separation from crops is a particularly significant function. Not only do weeds compete with crops for water and nutrients but may also be carriers of pests and diseases, affecting farm yield directly. Farmers have conventionally employed either hand labor or chemicals to manage weeds, either time-consuming or environmentally detrimental. As in sync with agro-issues, surveillance over environmental elements, particularly forest fire, is also becoming a critical field to be worked on by technology. Forest fire has mounted in terms of intensity and frequency due to anthropogenic pressure and climate change on the planet, resulting in ecologically destructive loss to environments, atmospheric conditions, and human habitations. The prevailing technologies available to detect fires, i.e., satellite observations or infrared devices, lack response time and working simplicity limitations. Existing machine learning techniques, Convolutional Neural Networks (CNNs) in particular, have the capability of automatically learning hierarchical and spatial features from vision data. This capability makes them particularly beneficial for object recognition and anomaly detection in complex visual environments such as forests and croplands.

B. *Motivation*

The driving force behind this work stems from the need to develop intelligent, scalable solutions for precision agriculture and environmental protection. For agriculture producers, the timely and accurate identification of weeds can lead to more efficiency in utilizing herbicides and treatments and therefore yield improvements and cost reductions. Moreover, performing it automatically with minimal human intervention can free labor for other essential agricultural operations, especially in regions of scarce resources. For forest fires, early detection helps to avoid uncontrolled expansion and facilitate prompt responses from fire control departments. The possibility of installing real-time, camera-based fire detection systems in remote or risky zones could be used as a proactive surveillance system. A system that can conduct visual detection in real-time could be a useful tool for national forest protection programs, park monitoring systems, or even smart city safety networks. The second incentive stems from the model versatility concept—ability to utilize one deep learning pipeline in a range of different domains. If a single high-performance visual feature extractor like a CNN can prove useful for both weed/crop identification and fire detection, then the same architecture model can most likely be leveraged even wider to crop disease detection, animal tracking, or traffic surveillance.

Such flexibility saves time in development, enables model reuse, and aligns well with the current trend towards generalizable AI.

C. Problem Statement

Just as image recognition systems using CNNs have improved, their deployment in real farm or forest environments presents significant challenges:

1) **Data Variability:** Farm scenes vary dramatically due to changing lighting, viewpoint, background clutter, and occlusion of objects.

2) **Classifier Choice:** While CNNs are strong for feature learning, the choice of the classifier with which the features are to be matched has a significant impact on final precision. The type of model that will be applied for generalizing across different environmental conditions is most significant.

3) **Scalability to Real-Time Detection:** Educational models have generally been trained and tested over static datasets. Scaling to real-time detection from dynamic videos introduces extra challenges such as frame rate optimization, hardware limitations, and latency.

4) **Generalization Across Domains:** Developing a model that performs well in agriculture but also readily generalizes to a completely unrelated domain such as fire detection is not very researched in existing literature.

These challenges require a robust yet adaptive approach that ensures high accuracy as well as deployability in the real world.

D. Proposed Approach

In order to overcome the challenges mentioned above, this paper proposes a visual recognition system based on CNN that uses InceptionV3 as the feature extractor and compares several downstream classifiers for obtaining the best results. The process was conducted in two phases separately: (1) crop-weed classification and (2) forest fire detection, followed by a real-time deployment pipeline.

1) **Feature Extraction with InceptionV3:** InceptionV3 was chosen for its established effectiveness in encoding high spatial hierarchies and reducing computational expense by methods such as factorized convolutions and auxiliary classifiers. The network was pretrained on ImageNet and then fine-tuned using domain-specific datasets (Open Sprayer Dataset for agriculture and Forest Fire Dataset for fire detection).

2) **Classifier Performance:** Four classifiers, namely SoftMax, Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF), were compared. Random Forest performed better than the rest consistently, with perfect training accuracy and a validation accuracy of 99.57% for crop-weed identification. It also provided better F1 scores and AUC values for both tasks.

3) **Fire Detection Pipeline:** Trained again on the Forest Fire Dataset using the same CNN + RF architecture, the model repeated the success in the agricultural domain, and its cross-domain capability was established.

4) **Real-Time Webcam Support:** Finally, webcam-based detection was integrated to support real-time video streams at ~30 frames per second (FPS). This was achieved by combining OpenCV with the trained classifier and optimizing preprocessing throughputs in a bid to minimize latency.

The proposed pipeline demonstrates how the combination of deep CNNs with ensemble classifiers like Random Forest can produce high-accuracy, domain-adoptable systems. It bridges the gap between research that is done in academia and deployable field technology.

II. LITERATURE SURVEY

A. Deep Learning in Image Classification

Convolutional Neural Networks (CNNs) have been profoundly influential on the discipline of image classification owing to their capability of extracting spatial features from unprocessed pixel information automatically. Other machine learning methods are generally preceded by manual feature extraction, a task that takes considerable time and is susceptible to human prejudice. Yet, CNNs, particularly ones with more complex architectures such as InceptionV3, facilitate the simplification of the process and have yielded very high accuracy across domains. Shin-Jye Lee et al. presented an improved CNN architecture called the Boost CNN that applies adaptive learning rates and weighted training to enhance classification performance. Their experiment demonstrated enhanced performance on noisy or changing datasets, a critical consideration in real-world use cases for agriculture and environmental monitoring [1]. This design inspired aspects of our approach, wherein the InceptionV3 model serves as an advanced feature extractor across detection tasks.

B. CNN-Based Methods in Agriculture

The agricultural sector has witnessed widespread adoption of deep learning methods for weed detection, crop classification, and disease identification. In contrast to conventional image processing, CNNs allow for detection under diverse field conditions such as changing light, shadow, and background clutter. A comprehensive review by Elngar et al. tabulated different CNN architectures employed in agricultural applications. Their review stressed the importance of versatile models that should work robustly under various conditions of the environment, an aspect that our purpose of the crop-weed classifying system has in mind as well [2]. Yet another extensive survey done by Altalak et al. considered more than fifty smart agriculture projects conducted using deep learning, especially identifying the requirement of real-time systems for detection purposes in low-resource environments [3]. These findings spurred our interest in developing a lightweight, scalable real-time camera-based weed detection system.

C. Comparative Classifier Studies in CNN Pipelines

Although CNNs are good feature extractors, their performance is largely dependent on the classifier that they use on the extracted features. The SoftMax classifier has been traditionally employed in end-to-end CNN architectures. Mixing other classifiers like Support Vector Machines (SVM), Decision Trees (DT), and Random Forests (RF) has also been reported to yield encouraging results in literature.

Specifically, it has been shown through studies that ensemble classifiers such as Random Forest tend to work better than SoftMax and SVM when faced with noisy, practical data [1]. The Random Forest classifier resists overfitting and has the capability of learning complex, nonlinear relationships among the feature set. This became the cornerstone in our reasoning in choosing to implement several classifiers based on CNN- feature extraction and choosing Random Forest finally for the crop-weed classification as well as the fire detection pipeline.

D. Fire Detection Based on Deep Learning

Traditional image processing methods for fire detection usually utilize color thresholding or motion detection, which may result in false positives in cases of complicated backgrounds. Deep learning, particularly CNN-based models, can avoid such challenges by learning complex patterns to differentiate fire from comparable visual patterns like sunset or streetlights. Gotthans et al. used CNNs for applying to video surveillance feeds to identify fire events and demonstrated that deep learning algorithms achieve great detection latency reduction as well as false alarms reduction compared to rule- based systems [10]. Their architecture was optimized for real- time deployment via frame-wise detection. In the same vein, Dua, Mohit, et al. had suggested an improved fire detection system based on the integration of deep CNN layers and data augmentation that improved the model's robustness under varying light conditions and camera orientations [11]. These efforts confirmed the applicability of CNNs for environmental monitoring and shaped our design for the fire detection module.

E. Real-Time Detection Systems

Implementing CNN-based detection models into real-time usage poses issues regarding speed, computation, and merging with video streams. Studies such as those of Gotthans et al. have investigated deep model optimization to achieve real-time performance on webcam and surveillance streams with frame rates greater than 25 FPS employing GPU acceleration [10]. Dua, Mohit, et al. similarly reported latency mitigation techniques such as frame skipping, low-computation CNN models, and asynchronous model running [11]. Our strategy takes advantage of these findings by employing OpenCV for video recording and threading optimizations to keep a stable 30 FPS in real-time classification. This keeps the model operational in field conditions, whether in agricultural fields or forest towers.

III. METHODOLOGY

A. Dataset Preparation and Preprocessing

The datasets used in this study are the Open Sprayer dataset, which contains images of crops and weeds under various environmental conditions and the Forest Fire Dataset for fire detection tasks. To ensure the CNN model received high- quality inputs, several preprocessing steps were implemented:

1) **Cropping**: Images were carefully cropped to exclude peripheral areas, minimizing noise and increasing the relevance of the input data.
2) **Normalization**: Pixel values were normalized to the range (0–1) to address variations in lighting, which improved the convergence rate of the neural network during training.
3) **Augmentation**: Data augmentation was done using random rotations, flips, scaling, and translation. This process reduced class imbalance and enhanced the ability of the model to generalize to new conditions.

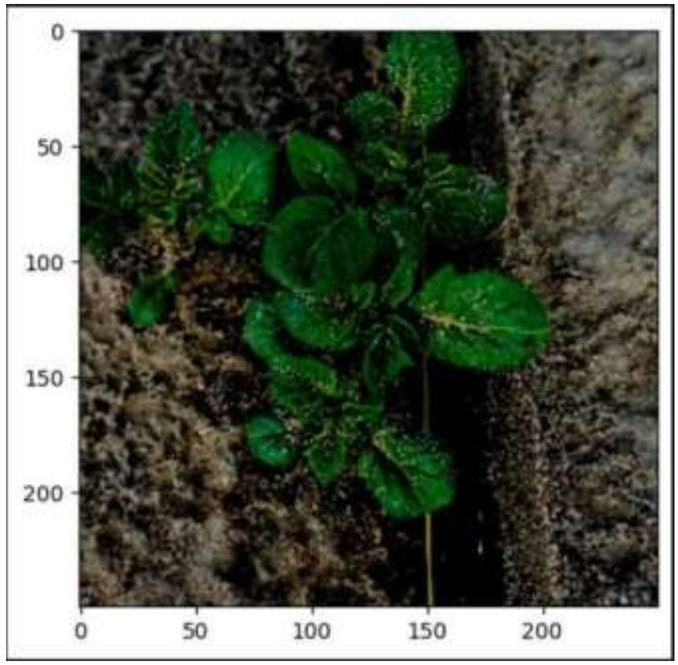

Fig. 1. Pre-processed image

B. Feature Extraction with Inception V3

For feature extraction, we used the Inception V3 CNN pre-trained model from the ImageNet challenge. Its hierarchical structure, consisting of convolutional layers of different sizes, allowed it to detect both small and large patterns. The final layer of Inception V3, Global Average Pooling (GAP), gen- erated feature vectors of 2048 dimensions, encapsulating the semantic meaning of the input images for classification.

C. Classifier Evaluation

To determine the optimal classification approach, the extracted features were fed into four different classifiers:

1) **SoftMax Classifier**: A fully connected layer with Soft- Max activation was used to transform feature vectors into probabilities. This method is efficient when decision boundaries are complex.
2) **Support Vector Machines (SVM)**: Using the radial basis function (RBF) kernel, SVMs showed better generalization for non-linear datasets, though they faced challenges with computational efficiency.
3) **Random Forest**: This ensemble learning method, which uses multiple decision trees, achieved the highest validation accuracy of 93.43%. Its resistance to overfitting and compatibility with high-dimensional features contributed to its high accuracy.
4) **Decision Tree**: While this model is interpretable and computationally tractable, it tended to overfit, achieving high training accuracy but lower validation accuracy (91.64%).

D. Extension to Forest Fire Detection

Following effective implementation for agricultural weed and crop classification, the pipeline was further extended to execute forest fire detection. The same approach was utilized for preprocessing fire image dataset, feature extraction using

InceptionV3, classification using the earlier selected Random Forest model. There were minor domain-specific tweaks, such as modifying the balance of the dataset (to deal with class imbalance between fire and no-fire images) and optimizing hyperparameters of the classifier. This domain adaptation proved the reusability and scalability of the initial framework across a wide range of environmental monitoring tasks.

E. Real-Time Detection Using Webcam

For further improving the usability of the system, a real-time detection module was implemented using OpenCV and Python. The system undertakes the following operations:

1) Captures real-time frames from the webcam at ~30 FPS.
2) Resizes and preprocesses every frame in real-time.
3) Passes the frame through the InceptionV3 model to extract
4) features.
5) Applies the trained Random Forest model to classify the frame as Fire or No Fire.

To maintain efficiency, the frame capture and processing were managed in a multithreaded manner, with latency kept low and user experience uninterrupted.

IV. RESULTS AND DISCUSSION

The experimental results indicate the effectiveness of the CNN-based weed and crop recognition system and fire detection system which was developed in this study. To evaluate the performance of the proposed approach, the validation accuracy was adopted as the standard measure. An analysis was conducted by comparing the four classifiers: some of the algorithms used include SoftMax, SVM, Random Forest, and Decision Tree.

A. Comparison between Classifiers

1) *CNN + SoftMax*

a) Training Accuracy: 99.727%
b) Validation Accuracy: 99.008%
c) **Analysis:** The use of SoftMax classifier was another issue since it lacked the capacity to handle complex decision boundaries with high efficiency. This restricted it from extending to many field environments because of low variability.

2) *CNN + SVM*

a) Training Accuracy: 99.941%
b) Validation Accuracy: 99.150%
c) **Analysis:** As the results obtained indicate, the SVM clas- sifier with the RBF kernel produced better generalization. However, it encountered large computational overhead and was sensitive to the hyperparameters of the program.

3) *CNN + Random Forest*

a) Training Accuracy: 100.000%
b) Validation Accuracy: 99.575%
c) **Analysis:** There was a difference in validation accuracy between the Random Forest method and other methods used in the research; thus, the Random Forest achieved higher validation accuracy than the rest of the methods. This ensemble approach appropriately handled issues of overfitting and performed well even in high-dimensional feature spaces. This makes it the strongest classifier for real-life use.

4) *CNN + Decision Tree*

a) Training Accuracy: 100.000%
b) Validation Accuracy: 99.575%
c) **Analysis:** Although perfect training accuracy was achieved, the Decision Tree model was evidently prone to overfitting. On the validation set, it had slightly worse results than Random Forest and SVM, proving the weak- ness of single-tree models for complicated classes.

Besides accuracy, other classification measures such as precision, recall, and F1-score were also computed to evaluate the performance of classifiers. We can see that Random Forest achieved the best value across all metrics, proving the stability of the discussed approach. The direct comparison of these classifiers exhibited that ensemble methods are more suitable for agricultural tasks since they are less sensitive to noise. Single-model solutions, including Decision Tree, though interpretable, were not balanced for generalizability necessary for various field conditions.

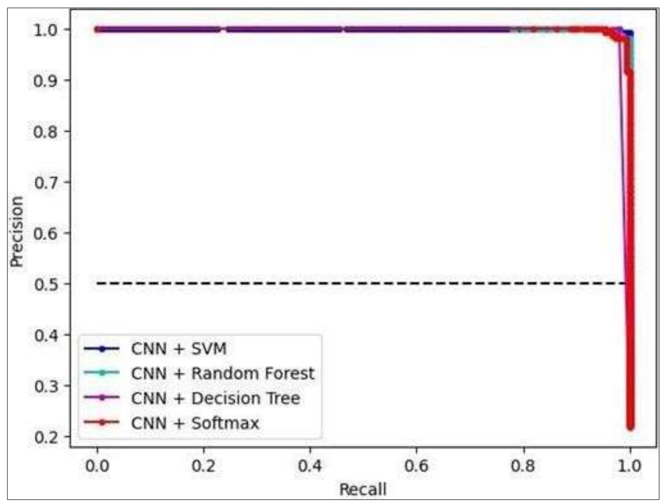

Fig. 2. Precision vs Recall for All Classifiers

Next, Figure 5 shows the model accuracy and loss. This graph helps visualize the training and validation accuracy over time, demonstrating the learning process of each classifier.

Parameters	CNN + SOFTMAX	CNN + SVM	CNN + RF	CNN + DT
Training Accuracy	0.997276265	0.999416342	1	1
Validation Accuracy	0.990084986	0.991501416	0.995750708	0.995750708
F1 Score	0.97763578	0.98026316	0.99016393	0.99016393
AUC Score	0.999682383	0.999917655	0.999905891	0.99025974
Average Precision Score	0.99890951	0.99970091	0.99966671	0.98476877

Fig. 3. Accuracy, F1 Score, Precision, AUC Table for All Classifiers

B. Weed and Crop Classification Results

The weed and crop classification were the first priority of the project. With the CNN + RF model, the system was able to distinguish between crops and weeds from the Open Sprayer Dataset. The system was able to function consistently across diverse image scenarios such as partial soil or machinery occlusion, uneven lighting, varying leaf structures.

The major results were:

- Validation Accuracy: 99.58%
- F1 Score: 0.9902

The excellent F1 score indicates that the model is well-balanced in maintaining precision and recall, a very important consideration in agricultural uses where both false positives and false negatives can cause crop damage or yield loss. The practice of using data augmentation during training also contributed much in enhancing model generalization across unseen samples.

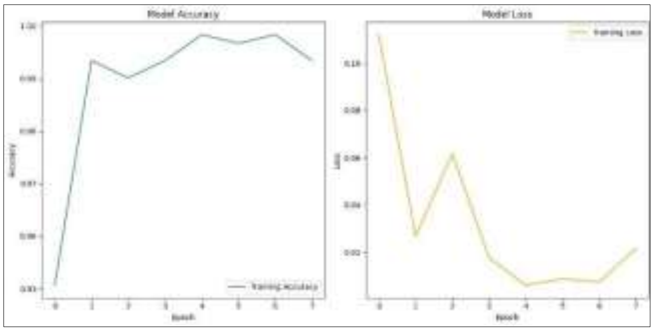

Fig. 4. Model Accuracy and Loss for Crop and Weed Classification

Fig. 5. Weed Result Image

Fig. 6. Crop Result Image

C. Forest Fire Detection Results

Encompassing the sound feature extraction and classification pipeline, we expanded the pipeline to find forest fires in the Forest Fire Dataset. Same preprocessing and same CNN + RF settings were utilized.

The output was also highly favourable:

- Validation Accuracy: 98.6%
- F1 Score: 0.9864
- AUC Score: 0.9963
- Average Precision Score: 0.9981

In spite of the higher visual diversity in fire images, the classifier exhibited very high confidence in classifying fire versus non-fire images, including challenging situations like smoke without fire, fire with vegetation obscuration, day and night images. These findings establish the feasibility of our approach in high-risk, real-world scenarios such as forest surveillance.

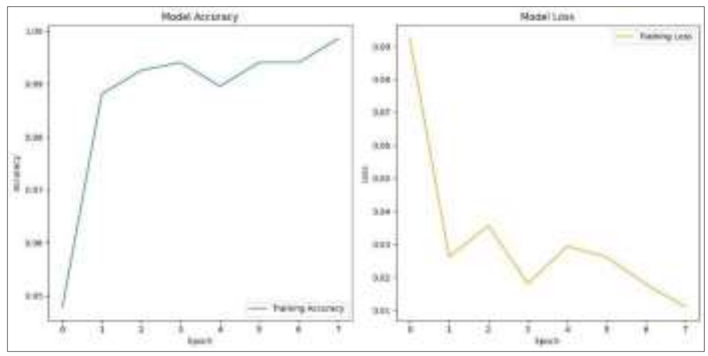

Fig. 7. Model Accuracy and Loss for Fire Detection

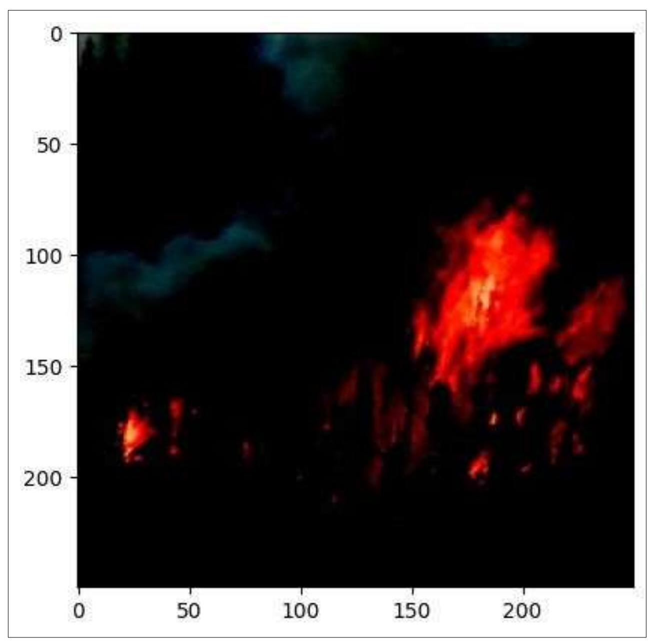

Fig. 8. Pre-processed image of fire

Fig. 9. Fire Result Image

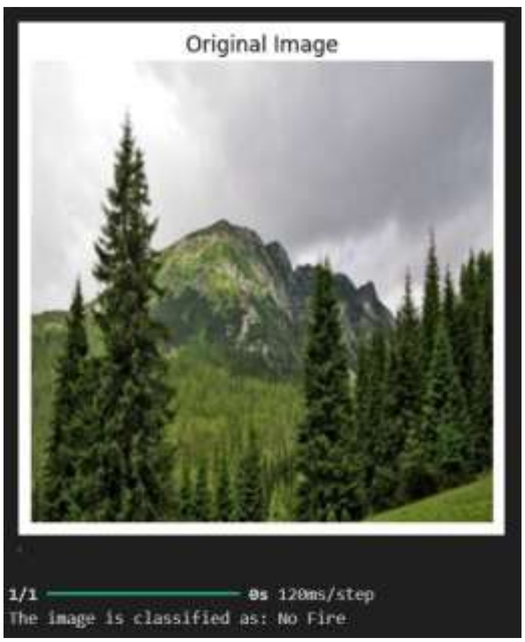

Fig. 10. No Fire Result Image

D. Real-Time Detection Performance

The last step in the project involved deploying the learned model into a real-time detection environment via webcam. The pipeline, implemented via Python and OpenCV, captures video, performs preprocessing, extracts features, and classifies features in one pipeline thread.

- Major performance indicators were:
- Frame Rate: ~30 FPS on an average GPU-capable laptop
- Detection Latency: ~33ms per frame
- Stability: Withstood repeatable 30+ minute long tests with constant performance,
- Accuracy Consistency: Real-time inferences aligned with offline inferences.

The model performed well without any lag or frame drops. In live environment testing, the system always showed accurate classification labels on the video stream in real time.

This real-time functionality provides practical application in autonomous agricultural robots for real-time weed detection, surveillance drones for early detection of forest fires, embedded systems with GPU/TPU support for light inference.

V. Future Work

Although the system presented here illustrates high accuracy and robust real-time performance on both weed/crop and forest fire detection tasks, there are several areas that could be improved further. Future efforts will be spent on running the model on edge devices such as NVIDIA Jetson Nano or Raspberry Pi, such that field-level autonomy is achievable without the dependency on continuous internet or high computation resources. Besides, multi-modal inputs like thermal or infrared imaging can be incorporated to enhance fire detection reliability in low-light or smoky environments. For farming, the system can be extended to monitor diseases in plants, pest infestation, or execute yield estimation, thus providing a more integrated smart farming solution. The addition of feedback loops and active learning is also possible to make the model improve itself over time using new data gathered during deployment.

VI. Conclusion

In this study, we introduced and successfully developed an adaptable and scalable framework for image-based classification problems, with emphasis on agricultural weed and crop identification and forest fire identification, and further integration for real-time detection via webcam. The approach unites the power of deep learning for effective feature extraction via the InceptionV3 CNN architecture with the efficiency and interpretability of traditional machine learning models, i.e., the Random Forest classifier, empirically found to perform better than some alternatives like SoftMax, SVM, and Decision Tree on several performance measures.During the initial project phase, we used our pipeline on the Open Sprayer Dataset for crop and weed classification. Preprocessing, feature extraction, and training of the classifier produced a good validation accuracy of 99.58% and an F1 score of 0.9902, which demonstrates the strong predictive power of the model and also its ability to generalize. Data augmentation and fine-tuning ensured that the model was able to process varied environmental conditions and plant characteristics correctly.Exploiting this success, the model was modified for forest fire detection with the Forest Fire Dataset. As if to reinforce its applicability across domains with varying visual patterns and classification difficulties, the model achieved a validation accuracy of 98.6% and exhibited great sensitivity in detecting even partially occluded or low-visibility fire scenarios. This demonstrates our framework's adaptability and resilience across domains with varying visual patterns and classification difficulties.One of the outstanding extensions of our work is the real-time detection module utilizing webcam input at the rate of ~30 FPS, with proper on-screen classification. This module attests to the deployability of our system under real-life contexts like precision agriculture and pre-emption wildfire systems.

Overall, the paper showcases the capability of CNN-based model integrated with ML in achieving high-performance, domain-tailorable, and resource-aware solutions for agricultural and environmental monitoring. The described framework not only has direct application but also provides a foundation for potential future extension to embedded systems, multimodal inputs, and higher-level AI features such as self- learning and decision support.

VII. REFERENCES

[1] Lee, Shin-Jye, et al. "Image classification based on the boost convolutional neural network." *Ieee Access* 6 (2018): 12755-12768.

[2] Elngar, Ahmed A., et al. "Image classification based on CNN: a survey." *Journal of Cybersecurity and Information Management* 6.1 (2021): 18-50.

[3] Altalak, Maha, et al. "Smart agriculture applications using deep learning technologies: A survey." *Applied Sciences* 12.12 (2022): 5919.

[4] Alam, Mansoor, et al. "Real-time machine-learning based crop/weed detection and classification for variable-rate spraying in precision agriculture." *2020 7th international conference on electrical and electronics engineering (ICEEE)*. IEEE, 2020.

[5] Dasgupta, Ishita, et al. "AI crop predictor and weed detector using wireless technologies: A smart application for farmers." *Arabian Journal for Science and Engineering* 45 (2020): 11115-11127.

[6] Jiang, Honghua, et al. "CNN feature based graph convolutional network for weed and crop recognition in smart farming." *Computers and electronics in agriculture* 174 (2020): 105450.

[7] Rani, SV Jansi, et al. "Automated weed detection system in smart farming for developing sustainable agriculture." *International Journal of Environmental Science and Technology* 19.9 (2022): 9083-9094.

[8] Sharma, Jivitesh, et al. "Deep convolutional neural networks for fire detection in images." *Engineering applications of neural networks: 18th international conference, EANN 2017, athens, Greece, August 25–27, 2017, proceedings*. Springer International Publishing, 2017.

[9] Muhammad, Khan, et al. "Efficient deep CNN-based fire detection and localization in video surveillance applications." *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 49.7 (2018): 1419-1434.

[10] Gotthans, Jakub, Tomas Gotthans, and Roman Marsalek. "Deep convolutional neural network for fire detection." *2020 30th international conference radioelektronika (RADIOELEKTRONIKA)*. IEEE, 2020.

[11] Dua, Mohit, et al. "An improved approach for fire detection using deep learning models." *2020 International Conference on Industry 4.0 Technology (I4Tech)*. IEEE, 2020.