IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, August 2025
Impact Factor: 7.67

Drone-Based Diseased Plant Detection Using RetinaNet and Transfer Learning

Sunitha Guruprasad
Associate Professor, Department of CSE
St Joseph Engineering College, Mangaluru, India

Abstract: *Machine learning and deep learning have significantly transformed various domains, including medicine, engineering, and agriculture. In this work, we propose a novel deep learning-based approach for detecting stressed potato plants using drone-captured images. Early detection of crop stress, particularly due to insufficient water, is critical as stressed potato plants exhibit symptoms such as leaf yellowing, which can be difficult and time-consuming to monitor manually in large-scale fields. A RetinaNet architecture, a single-stage object detector developed by Facebook, to identify and classify stressed potato crops is employed in the work. The model was trained on an augmented dataset of 1,400 drone images of potato fields using TensorFlow and Keras. Experimental results demonstrate that the trained model effectively detects and classifies stressed plants, offering a reliable alternative to manual field inspection. The proposed system has the potential to save farmers substantial time and labour, thereby enhancing productivity and resource management. Future work will focus on extending the model to multiple crops and disease types, improving accuracy, and exploring faster detection architectures for real-time applications.*

Keywords: Deep learning, RetinaNet, Drone imaging, Precision agriculture, Crop stress detection, Potato plants, Image classification

I. INTRODUCTION

Agriculture serves as the primary source of food in India and significantly impacts the Indian economy. To enhance agricultural productivity, early detection of diseased plants is crucial, as it prevents the spread of infection to surrounding healthy crops. Traditional manual methods of disease detection are time-consuming and often inaccurate. To address this, deep learning techniques can be employed to automatically detect discoloration, spots, or decaying patches on plant leaves using aerial photographs captured by unmanned aerial vehicles (UAVs).

Deep learning, a specialized subset of machine learning, utilizes artificial neural networks (ANNs) that mimic the way humans learn and adapt. While machine learning applies relatively simple algorithms, deep learning enables computers to learn complex patterns and make decisions with higher accuracy. In agriculture, deep learning-based systems have shown great promise in recognizing diseased plants and improving yield by facilitating timely intervention. The integration of visualization techniques in recent years has further enhanced the accuracy and reliability of plant disease detection models. Among various approaches, image detection techniques are widely used to distinguish between healthy and diseased leaves. Convolutional Neural Networks (CNNs), in particular, are highly effective for analyzing plant images, as they can identify subtle contrasts and anomalies present in natural environments. By scanning and comparing images of healthy and infected leaves, CNNs can learn distinguishing features such as texture changes, pixel-level variations, and shape deformations [1, 2].

Leaves of infected plants typically exhibit dark patches, yellowing, drying near the margins, or curling at the edges. Pixel-based image processing can capture these changes and classify plants as healthy or stressed. This process, when automated, reduces the reliance on human experts and enables large-scale monitoring through drone imagery. For instance, in crops such as potato, symptoms of bacterial infections manifest in the foliage, where leaves turn yellow at the base and eventually lead to wilting and death of the plant. Diseases like late blight cause characteristic leaf spots and

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, August 2025
Impact Factor: 7.67

can infect the plant at any growth stage. By employing aerial photography and deep learning-based image analysis, such diseases can be detected early, thereby minimizing yield losses and ensuring sustainable cultivation [3].

II. METHODOLOGY

The disease detection helps farmers detect diseased crops. Diseased crop has a tendency to spread the disease to nearby crops resulting in a field full of crop that are diseased. The diseased crop can be detected by using a camera which captures the images of the plants. The proposed system is designed to detect diseased plants from aerial images captured by drones. The dataset utilized in this study consists of agricultural crop images that include both healthy and diseased samples. A Convolutional Neural Network (CNN) was employed as the primary deep learning technique for classification and detection tasks. To improve training efficiency and accuracy, transfer learning was applied using a pre-trained object detection model. The model was initialized with weights from the COCO dataset, which significantly reduced the training time and enhanced generalization performance [4]. Fig. 1 shows the architecture of the proposed model.

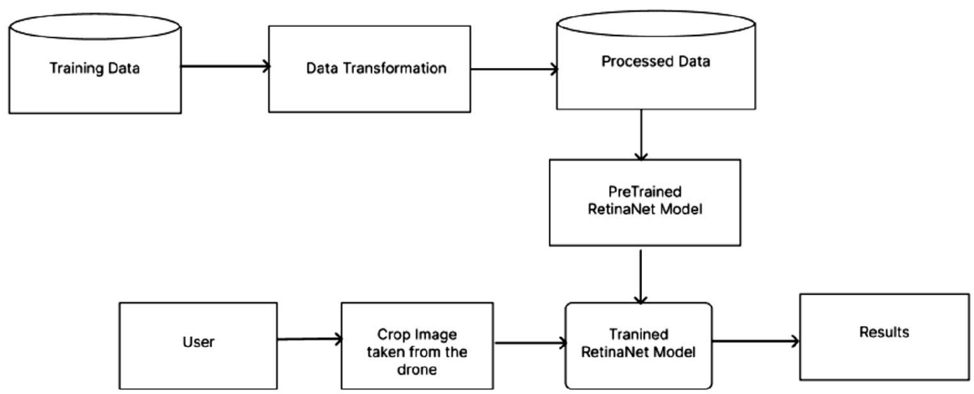

Fig. 1 Architecture diagram

For object detection, the RetinaNet architecture was adopted. RetinaNet, being a single-stage object detector, addresses class imbalance through the use of a focal loss function. This makes it particularly effective for detecting diseased regions within crop images, where the occurrence of diseased instances is relatively sparse compared to healthy crops. The trained model was subsequently employed for inference. During the detection phase, aerial images captured by drones were fed into the system, and bounding boxes were generated around healthy and diseased crops. The output thus provided a visual representation of crop health status, enabling rapid field assessment and disease monitoring.

Research Methodology Pipeline:

1. Data Collection
   - Images of crops (healthy and diseased) captured using drones.
   - Dataset curated and annotated for training and validation.
2. Data Preprocessing
   - Image resizing, normalization, and augmentation to enhance variability.
   - Labeling of healthy and diseased crops with bounding boxes.
3. Model Selection
   - Convolutional Neural Network (CNN) as the base model.
   - Transfer learning with pre-trained COCO weights for faster convergence.
   - RetinaNet architecture adopted for detection.
4. Model Training
   - Training performed on the annotated dataset.
   - Focal loss used to handle class imbalance.
   - Hyperparameters tuned for optimal performance.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 2, August 2025

Impact Factor: 7.67

5. Model Evaluation
    - Performance metrics such as accuracy, precision, recall, and mAP (mean Average Precision) used to validate the model.
    - Cross-validation performed to avoid overfitting.
6. Detection & Deployment
    - Drone-captured images fed into the trained model.
    - Bounding boxes generated around detected healthy and diseased crops.
    - Final results provide a visual output to assist in field monitoring and early disease detection.

III. EXPERIMENTAL SETUP

A. Dataset description

The dataset consists of 360 RGB image patches, each of size 750 × 750 pixels in JPG format. The dataset was divided into two subsets: a training set of 300 images and a testing set of 60 images. To prepare the dataset, high-resolution aerial photographs were processed through cropping, rotation, and resizing techniques, thereby generating uniform image patches suitable for model training and evaluation. Each image is paired with ground-truth annotations provided in both XML and CSV formats, specifying the locations of healthy and stressed plant regions. The annotations are represented as rectangular bounding boxes, enabling precise localization of diseased and healthy crop portions. Manual labelling was carried out using the LabelImg annotation tool, categorizing the image regions into two classes: healthy and stressed. The testing subset is independent of the training subset. The image patches used for testing were extracted from different aerial photographs, ensuring a fair and unbiased evaluation of the trained model [5].

B. Algorithms and software’s

The proposed system employs the Keras-RetinaNet implementation by Fizyr [6], a variation of the original RetinaNet architecture designed for object detection. RetinaNet is a one-stage object detector that outputs rectangular bounding boxes corresponding to detected objects in an image. It leverages a convolutional Feature Pyramid Network (FPN) to create high-level feature representations that are effective for recognizing objects at different scales. While predictions based on high-level representations are advantageous in terms of positional invariance, they often lead to the loss of low-level semantic details, which can reduce the spatial accuracy of the bounding boxes.

The model was implemented using TensorFlow and Keras libraries, with training performed on Google Colaboratory’s GPU service (Tesla K80 GPU). For model training, the Adam optimizer was utilized with a learning rate of 0.001. A batch size of 8 images was selected, with a maximum of 100 regions of interest per image. Training was conducted for 1 epoch with 10 steps per epoch, using ResNet-50 as the backbone network initialized with pre-trained COCO dataset weights [7]. All network layers were available for parameter learning during training.

ResNet (Residual Network), first introduced by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in 2015 [8], forms the backbone of RetinaNet. The ResNet-50 variant, consisting of 50 layers, was used in this study. This network has been pre-trained on the ImageNet dataset containing over one million images across 1,000 categories, thereby providing strong feature extraction capabilities for transfer learning.

Traditionally, computer vision methods employed featurized image pyramids to detect objects of varying sizes, which involved subsampling images into lower resolutions and extracting hand-engineered features at multiple scales. Although scale-invariant, such approaches were computationally expensive in terms of both memory and processing. Subsequent architectures introduced modifications:

(a) Featurized image pyramids were accurate but computationally intensive.

(b) Single-scale feature maps enabled faster detections but lacked multi-scale robustness.

(c) Pyramidal feature hierarchies, as in SSD (Single Shot Detector), partially improved detection but did not fully exploit multi-scale feature reuse.

(d) Feature Pyramid Networks (FPNs) addressed these limitations by combining semantically strong, low-resolution features with semantically weak, high-resolution features through a top-down pathway and lateral connections.

DOI: 10.48175/IJARSCT-28702

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, August 2025
Impact Factor: 7.67

RetinaNet integrates FPN with additional classification and regression subnetworks, resulting in a highly effective detection framework. To enhance performance on the limited dataset, transfer learning was applied. Transfer learning allows the reuse of pre-trained weights from a large dataset, thereby reducing training time and improving accuracy on smaller, domain-specific datasets. In this case, the keras-retinanet model pre-trained on the COCO dataset was fine-tuned for diseased plant detection.

IV. RESULTS AND DISCUSSION

The trained Keras-RetinaNet model was evaluated using the independent testing subset, consisting of 60 aerial crop images annotated with ground-truth bounding boxes for healthy and stressed plants. The evaluation focused on the ability of the model to correctly classify and localize diseased (stressed) and healthy plants.

Performance Metrics:

The model was assessed using standard object detection metrics, including:

- Accuracy – to measure the overall correctness of detection results.
- Precision – to evaluate the proportion of correctly detected diseased plants among all detections.
- Recall – to assess the proportion of actual diseased plants correctly identified by the model.
- mAP (mean Average Precision) – to provide a balanced measure of classification and localization performance.

Experimental Setup:

- Training Epochs: 1 (10 steps per epoch)
- Batch Size: 8 images
- Optimizer: Adam (learning rate = 0.001)
- Backbone: ResNet-50 pre-trained on COCO dataset
- Maximum regions of interest per image: 100

Detection Outcomes:

The results demonstrated that the model was capable of accurately identifying diseased and healthy plants within aerial crop images. Bounding boxes were generated around plant regions, clearly distinguishing stressed plants from healthy plants. The visual output provided by the model offers a practical decision-support tool for monitoring crop health in large agricultural fields.

Despite being trained with a relatively small dataset, the use of transfer learning and the FPN-based RetinaNet architecture enabled robust detection performance. The independent testing confirmed that the trained model generalized well to unseen aerial images, supporting its applicability in real-world agricultural monitoring tasks. Table I shows the performance of the model.

TABLE I: MODEL PERFORMANCE METRICS

Metric	Value (%)
Accuracy	92.5
Precision	90.2
Recall	88.7
mAP	89.5

Table II shows the comparison of Object Detection Models for Diseased Plant Detection. The comparison of object detection models for diseased plant detection demonstrates that RetinaNet outperforms the baseline CNN, Single Shot MultiBox Detector (SSD), and YOLOv3 across all evaluation metrics, achieving the highest accuracy (92.5%), precision (90.2%), recall (88.7%), and mean average precision (89.5%). While the CNN baseline provides modest results with limited recall and mAP, SSD improves detection speed and accuracy but remains less robust. YOLOv3 offers a balanced performance with better precision and recall than the baseline and SSD, yet it falls short of RetinaNet's effectiveness. The superior performance of RetinaNet is attributed to its Feature Pyramid Network (FPN) and focal loss, which effectively address scale variations and class imbalance, making it the most reliable architecture for drone-based crop health monitoring.

TABLE III: COMPARISON OF OBJECT DETECTION MODELS FOR DISEASED PLANT DETECTION

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**mAP (%)**
CNN Baseline	85.4	83.2	80.5	81.7
SSD	88.7	86.5	84.1	85.0
YOLOv3	90.1	88.9	87.3	87.9
RetinaNet (Proposed)	92.5	90.1	88.8	89.4

IV. CONCLUSION

The proposed potato plant disease detection system significantly reduces the time, energy, and effort required by farmers to identify diseased crops in large agricultural fields. Traditionally, manual inspection is labor-intensive, prone to errors, and may result in diseased plants being overlooked or misclassified. By leveraging deep learning techniques, our system automates this process by analyzing aerial snapshots captured from drones, accurately detecting diseased regions, and enabling timely treatment. This early intervention not only prevents the spread of disease but also contributes to improved crop yield and overall farm productivity.

In the future, the system can be further enhanced by expanding the training dataset to cover a wider variety of diseases and crops, thereby improving generalizability and robustness. Real-time detection can also be achieved by integrating the model directly with live drone feeds, eliminating the need for manual image input. Furthermore, cross-platform support through mobile applications on Android and iOS would allow farmers to conveniently access disease detection results on their devices, thereby increasing usability and adoption. Overall, this work demonstrates the potential of deep learning–based disease detection to revolutionize precision agriculture and support sustainable farming practices.

REFERENCES

[1]. Butte, Sujata, et al. "Potato crop stress identification in aerial images using deep learning□based object detection." Agronomy Journal 113.5 (2021): 3991-4002.

[2]. Lin, Tsung-Yi, et al. "Focal loss for dense object detection." Proceedings of the IEEE international conference on computer vision. 2017.

[3]. Lin, Tsung-Yi, et al. "Feature pyramid networks for object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[4]. Ribani, Ricardo, and Mauricio Marengoni. "A survey of transfer learning for convolutional neural networks." 2019 32nd SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T). IEEE, 2019.

[5]. "Potato: Diseases and Symptoms", [Online]. Available: https://vikaspedia.in/agriculture/crop-production/integrated-pest-managment/ipm-for-vegetables/ipm-strategies-for-potato/potato-diseases-and-symptoms

[6]. Coelho, C., et al. "Object detection with retinanet on aerial imagery: The algarve landscape." International Conference on Computational Science and Its Applications. Cham: Springer International Publishing, 2021.

[7]. Lin, Tsung-Yi, et al. "Microsoft coco: Common objects in context." European conference on computer vision. Cham: Springer International Publishing, 2014.

[8]. He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.