IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.67
Volume 5, Issue 1, March 2025

Agri-Vision AI Powered Farm Segmentation

Miss. Trupti K. Varma[1], Mr. Gous. D. Bahurupi[2], Mr. Ayan I. Shirgave[3], Mr.Umar S. Shaikh[4], Prof. Mr. A. S. Salavi (Kumbhar)[5]
Students, Computer Science & Engineering Department[1,2,3,4]
Lecturer ,Computer Science & Engineering Department[2]
Yashwantrao Chavan Polytechnic, Ichalkaranji, India
truptivarma1@gmail.com, gousbahurupi5@gmail.com, ayanshirgave08@gmail.com
gousbahurupi5@gmail.com, gousbahurupi5@gmail.com

Abstract*: This project applies deep learning-based semantic segmentation in orchard mapping.The general objective is the accurate detection and localization of tree canopies based on UAV images, under a variety of conditions such as changing seasons, age of trees, and weed cover levels. U-Net architecture is used in training annotated datasets for automatic tree canopy segmentation, which generates masks related to canopy size, shape, and planting gaps, thus helping with precision agriculture. This approach improves decision-making capability in irrigation, pesticide application, and yield estimation. Using advanced computer vision techniques, the project provides a scalable, efficient solution for modern agricultural systems*

Keywords: Sematic segmentation, U-net alogorithm, image processing, deep learning, agricultural technology

I. INTRODUCTION

Precision agriculture is the modern management approach aimed at optimizing field and crop properties by making use of information and communication technology. This method focuses on managing the temporal and spatial variability of these properties through various sensors and technologies, which capture detailed data about the field. RS plays a primary role here in that it describes the noncontact monitoring of crop attributes through an analysis of reflected radiation by the plants. Most often, the RS is deployed for mapping duties in agricultural systems, with data both from the ground and through aerial images.

One of the critical applications of RS in agriculture is the segmentation of orchard canopies to help farmers determine canopy size and plant volume to achieve better decision-making. Semantic segmentation aims at automatically detecting and localizing tree canopies under different conditions such as varied seasons, ages of trees, and levels of weed cover. In accomplishing this, a semantic database of segmentation is developed through test and training datasets generated from UAV images.

By using semantic segmentation, farmers would effectively inspect orchard ecosystems by automatically isolating tree canopies from aerial images. They can derive critical information concerning canopy size and shape, gaps in planting schemes, age of the trees, and consequently, their yields. They could also quantify the amount of water and pesticides applied to the orchard. Ultimately, this tool helps optimize resource use, improves crop management, and enables more informed decisions to improve efficiency and sustainability in farming practices. Semantic segmentation, based on UAV images, can transform the management of orchards by providing farmers with more precise data to enhance their productivity and environmental stewardship.

II. DEEP LEARNING NETWORK ARCHITECTURE

With the fast pace of development in Computer Vision and Deep Learning models and theories, the task of image segmentation and classification becomes simpler compared to the classical methods. This chapter overviews the popular theology neural network architecture in computer vision, such as convolution, antagonist generation, graph, and Transformer networks.

A. Convolutional neural networks (CNNs)

CNN's are one of the types of artificial neural networks that have gained immense usage in a range of computer vision applications. They have attained great success with operations like image recognition, object detection, and image segmentation and have established themselves as one of the key parts of the deep learning methodology. They are more advanced compared to machine learning, CNN also have enhanced speed and accuracy. The backbone of CNNs consists of a convolution layer, pooling layer, and fully connected layer Convolution layers are tasked with feature extraction through convolution operations, while pooling layers strive to decrease the dimensionality of feature maps while enhancing their robustness. Nonlinear layers are used to introduce non-linearity, thus giving neural networks the capability to learn increasingly complex patterns and relationships. The advantages of CNNs are leveraging spatial correlations between neighbouring pixels, parameter sharing, invariance, large data processing, and adaptive feature learning. Some of the prominent CNN architectures are VGGNet, GoogLeNet, and ResNet.

B. Generative adversarial networks (GANs)

Generative Adversarial Networks (GANs) are a type of deep learning model consisting of two adversarial networks: a Generator, which generates realistic data, and a Discriminator, which can tell the difference between real and synthetic. The Generator is trained continuously until its output cannot be differentiated from actual data.

GANs are applied in image creation, deep fakes, data augmentation, and artificial intelligence-based art. They improve image quality, produce artificial datasets, and generate realistic faces. Despite setbacks such as mode collapse and ethical issues, GANs are transforming AI in creativity and data creation.

C. Graph neural networks (GNNs)

Deep learning has been extremely successful in numerous areas. Nonetheless, researchers have been aware of its shortcomings in dealing with and resolving every situation and problem. Specifically, while dealing with graph-structured data in non-Euclidean spaces, there is an inherent difficulty in utilizing structural and semantic information simultaneously. GNNs have come to be a major area of research emphasis in the area of graph data structures because of their ability to simultaneously learn topological data and maintain structural properties. Among many GNN models, Graph Convolutional Networks (GCN) form a unique convolutional neural network structure that is able to directly process graphs while leveraging their in-built structural information. Furthermore, widely used GNN methods include more sophisticated methods like Graph Attention Networks (GAT) and Graph Generative Adversarial Networks (Graphical GAN), further enriching the graph-based learning algorithm landscape.

D. Transformer

In 2017, Google unveiled the revolutionary Transformer model, which took the natural language processing community by storm. In recent years, several pioneering research papers have made successful use of the Transformer technology in cross-domain computer vision tasks, marking a new era in the visual world. Dosovitskiy proposed a model called ViT (Vision Trans
former), which is a pure self-attention-based image classification method. The method of ViT is to split an image into patches of a fixed size, conduct linear transformations and position coding on each patch, and then input the patches into the transform encoder to conduct feature extraction and classification on the whole image. Compared with conventional CNNs,
the ViT model depends entirely on the self-attention mechanism to extract useful information in images, with higher interpretability and transferability.

III. PROPOSED FRAMEWORK

The purpose of this paper is to propose an ensemble model using the bagging technique and UNet network to enhance crop segmentation precision and lower operational complexity. One major issue with aggregation methods is enhanced computational overhead, and our aim is to develop a light-weight semantic segmentation network leveraging aggregation methods with computational overhead in mind. This work addresses a binary classification crop semantic segmentation task in order to endow agricultural robotic operations with perception abilities. We only learn about the semantic segmentation of binary class. Every pixel in the image of the dataset is assigned to only one out of 2 classes/labels, i.e., crop or background. First, we normalize the raw RGB image into an HSV color space. Next, we input the raw and processed images into our ensemble UNet models to get segmentation masks. Next, we add pixel-wise to get the final prediction mask. Lastly, the respective predicted classes are achieved.

A. Data processing

The main task of this part is dataset preparation. The pictures of corn plants in this paper are self-made (refer to the experiment and discussion section for more information), and the real-value pictures are annotated manually. For the purpose of simplifying the calculation and operation, the picture size is uniformly adjusted, and the RGB image is transformed into an HSV image. The geometric transformation techniques primarily consist of flip, rotation, cropping, scaling, translation, etc. For avoiding overfitting, pixel transformation techniques like adding salt and pepper noise, Gaussian blur, etc., can also be applied. Lastly, it is split into training, verification, and test subsets based on a particular ratio.

B. Training and test

The predominant task of this section is the training of the model. High numerical pixel values need more power for computing as raw image pixel values vary between 0 to 255. Hence, we adhere to the image normalization routine to convert pixel values into a range of 0 to 1 by simply dividing every pixel value with the maximum pixel value possible (i.e., 255). We propose a framework that consists of various color spaces to counteract the effect of illumination conditions. RGB and HSV color spaces are capable of producing complementary information, and our framework can capture and combine it to acquire more discriminative information. Our proposed framework is to train UNet-based models (single RGB and single HSV) to produce various predictions. Finally, classical stacked generalization [26] is used to get the ultimate prediction, crop, and background (all others are classified as none crop).

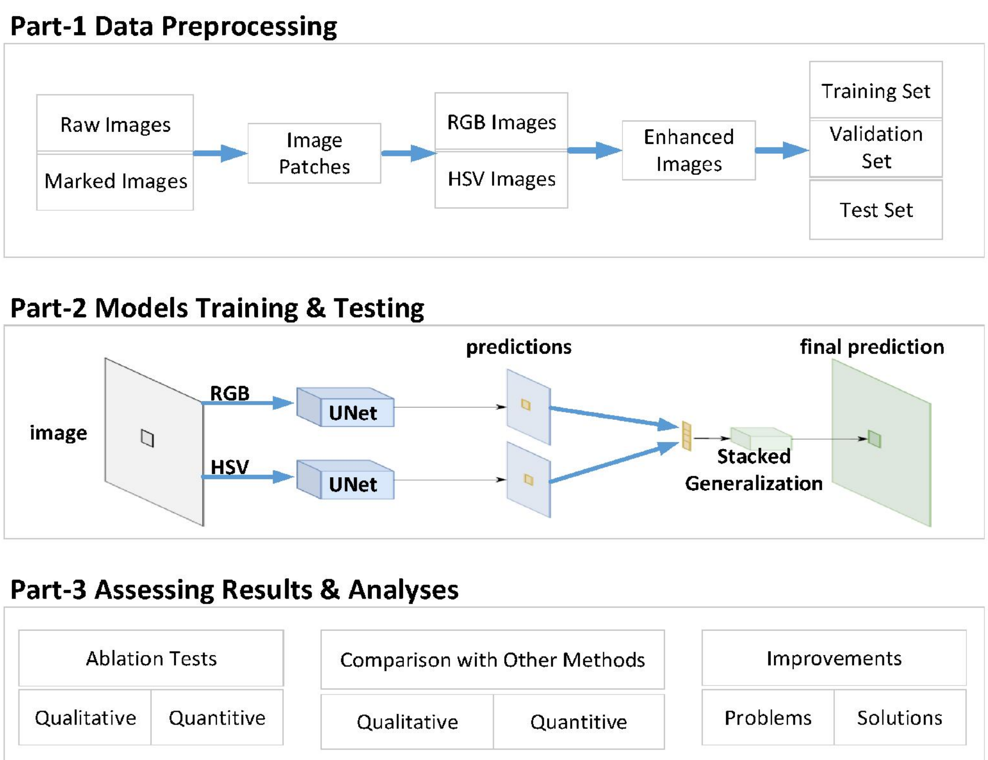

Fig 1: System Architecture for Agrivision AI powered from segmentation

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

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

Impact Factor: 7.67

Volume 5, Issue 1, March 2025

C. Evaluating and analyzing the results

The primary task of this section is framework evaluation. We perform the ablation experiment of the framework in this paper and the experiment of comparative methods and make a horizontal comparison from two qualitative and quantitative angles. The merits and demerits of the framework in this paper are analyzed, and additional improvements are put forward.

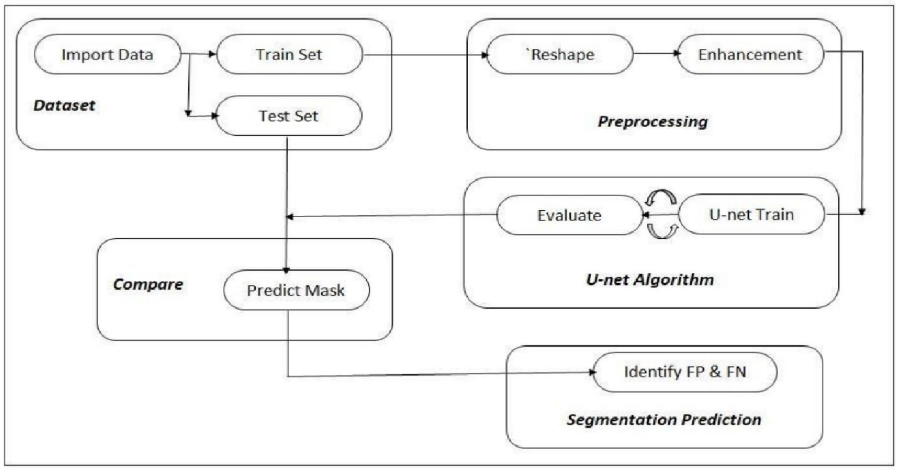

Fig.1.Framework of the system

IV. APPLICATION

The orchard mapping project finds varied uses like maximizing precision farming by offering the health of the trees and availability of resources. It aids crop monitoring, an early detection process for diseases and pests, yield forecasting, and harvest scheduling in order to maintain better productivity. The project provides environmental monitoring, which monitors the effects of the weather and ground, and better farm management is achieved through knowledge-based decisions related to irrigation, fertilization, and pest control. Secondly, it aids in agricultural research: knowledge on crop varieties, growth habits, and farming methods, guiding innovation and sustainability of agriculture.

A. Key Innovation/Features

Our UAV-based image project of our orchard entailed a myriad of features and innovations. When you take UAV image, you get extremely high-resolution imaging that helps generate very detailed orchard maps. The dataset contains training, test, and validation sets that permit accurate model fitting for object detection and classification models, such as tree types, and growth patterns, among others. Advanced algorithms that help in performing accurate mapping are applied to perform the analysis. The use of UAVs provides better coverage, reduced fieldwork time, and enhances accuracy compared to traditional methods, offering a powerful tool for orchard management and agricultural monitoring.

B. Future Scope

Agrivision AI Powered Farm Segmentation addressed the task of accurately segmenting and locating tree canopies in different orchards from imagery collected by a UAS. Tree segmentation is a broad potential application, ranging from mapping orchard terrain to locate tree trunks for self-driving ground vehicle navigation. The method can be utilized as a means to estimate correctly volume of tree canopies in orchards and then identify the age and yield potential of the trees.

V. CONCLUSION

In summary, RS and PA technologies are very promising, particularly in orchard tree canopies by semantic segmentation of UAV imagery. These are able to offer accurate information on spatial variability, canopy size, and crop

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

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

Impact Factor: 7.67

Volume 5, Issue 1, March 2025

health, resulting in resource use optimization, enhanced yield forecasting and early pest and weed detection. Such tools when integrated into farming ensure agriculture that is data-driven, sustainable, and efficient. This will consequently benefit the environment and the farmers.

REFERENCES

[1] Lihong Chang;Hongchao Fan;Ningning Zhu;Zhen Dong. A Two-Stage Approach for Individual Tree Segmentation From TLS Point Clouds Year: 2022 | Volume: 15 | Journal Article | Publisher: IEEE

[2] Seema S. Patil; Yuvraj M. P atil; Suhas B. Patil. Development of Cost-Effectiv Precision Spraying Techniques Using Sensor Technology Year: 2023 | Conferenc Paper | Publisher: IEEE

[3] Yanqi Dong;Zhibin Ma;Fu Xu;Feixiang Chen. Unsupervised Semantic Segmenting TLS Data of Individual Tree Based on Smoothness Constraint Using Open-Source Datasets Year: 2022 | Volume: 60 | Journal Article | Publisher: IEEE

[4] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. October 2015

Wang, A.; Xu, Y.; Wei, X.; Cui, B. Semantic Segmentation of Crop and Weed using an Encoder-Decoder Network and Image Enhancement Method under Uncontrolled Outdoor Illumination. IEEE Access **2020**, 8, 81724–81734.

[5] Kerkech, M.; Hafiane, A.; Canals, R. Deep leaning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. Comput. Electron. Agric. **2018**, 155, 237–243.

[6] Zhu J-Y, Park T, Isola P, Efros AA (2017) unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE international conference on computer vision (ICCV). pp 2242–2251.