DISORDER DETECTION OF TOMATO PLANT USING IOT AND ENSEMBLE TECHNIQUES

Many countries utilize fertilizers and pesticides to enhance agricultural productivity. Fertilizers provide essential nutrients for plant growth, while pesticides offer protection from destructive pests and pathogens. Traditionally, plant disorders have been identified by visually inspecting crops in the field—a method that can be labor-intensive, time-consuming, and costly. In recent years, numerous approaches have been explored to automate the detection of plant disorders, such as employing IoT networks (particularly tested on grapes), machine learning models with constrained datasets, and spectral imaging techniques.

This research introduces a comprehensive prototype that employs sensor data (for instance, temperature, humidity, and moisture readings) alongside computer vision. The aim is to detect disorders in tomato plants (Solanum Lycopersicum) using a fusion of machine learning, IoT, cloud computing, and image processing methods. By combining multi-model analyses—encompassing both sensor-derived information and visual characteristics—this framework aspires to provide reliable, scalable, and efficient disease detection that can be adapted for real-world agricultural settings.

Index Terms— Image processing, plant disorder, feature extraction.

1. INTRODUCTION

Over the past fifty years, global warming has progressively intensified due to anthropogenic activities, leading to erratic climatic patterns that jeopardize agricultural stability. These unpredictable conditions affect all facets of plant development, including the soil—often rendering it infertile—planting areas, and cropping intensity. In scenarios where the soil loses its fertility, farmers must rely on fertilizers containing fundamental nutrients such as nitrogen, phosphorus, and potassium to sustain crop productivity.

The present study focuses on Solanum Lycopersicum, commonly known as the tomato plant, which belongs to the nightshade family, Solanaceae. A principal challenge with large-scale tomato cultivation is the array of diseases that can influence different parts of the plant; however, scientific evidence indicates leaves are particularly susceptible. Because leaf properties, including color and texture, reveal key insights into a plant's health status, researchers often target leaves when diagnosing plant disorders.

Although large datasets can offer robust modeling capabilities, handling them may demand substantial memory and computational power. Furthermore, excessive or improper training (overfitting) risks poor predictive performance on novel samples. Hence, leveraging effective and efficient feature extraction becomes crucial in machine learning for plant-disease analysis. Plant disease symptoms often overlap in appearance—for example, various fungal infections can present strikingly similar leaf lesions—making it critical to identify subtle differences that accurately classify the disorder. This research aims to address these challenges through an integrated system designed to predict which specific disorder is most likely present on a tomato leaf.

2. RELATED WORK

Considerable work has been conducted on feature identification and classification to diagnose plant health conditions. Stephan Gang Wu et al. used a probabilistic neural network combined with image and data processing to develop a general-purpose automated leaf-recognition system. Harish Velingkar et al. employed image pre-processing followed by K-means clustering for classification and Support Vector Machines (SVM) for feature extraction. Other researchers implemented Convolutional Neural Networks (CNNs), enabling direct learning of raw image features.

Alvaro Fuentes et al. introduced Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully Convolutional Network (R-FCN), and Single Shot Multibox Detector (SSD) architectures for plant disease and pest recognition. Mohammed A. Hussein and Amel H. Abbas explored various feature-extraction strategies—including texture-based descriptors, color moments, and shape-based features—while Sanjay Mirchandani et al. used feed-forward artificial neural networks combined with image processing. Additionally, Jihen Amara et al. developed a LeNet-based CNN for classifying leaf images. These diverse methods highlight how machine learning and deep learning have both evolved to tackle the complexities of plant disease identification.

3. METHODS FOR FEATURE EXTRACTION

A variety of feature extraction approaches focuses primarily on analyzing the visual properties of leaves—such as color, shape, and texture—to gain essential information about a plant's health. The system proposed herein expands on that paradigm by integrating both image data and sensor data, thereby offering deeper insight into environmental conditions influencing plant disorders. Broadly, these features are categorized into:

1. **Sensor-Based Feature Extraction**
2. **Deep Learning–Based Feature Extraction**

3.1. SENSOR-BASED FEATURE EXTRACTION

Environmental variables significantly impact plant health. This study employs three core sensors—moisture, temperature, and humidity—to gather real-time data via IoT. Collected sensor readings are uploaded to the cloud, enabling continuous monitoring and analysis. Once a sufficient volume of sensor data is available, it is used within machine learning algorithms to detect patterns indicative of plant health or illness. The system differentiates between two primary classes: "healthy" and "not healthy." For processing sensor-based inputs, algorithms such as SVM and XGBoost are suitable choices, particularly for structured numerical data, due to their strong performance in classification tasks.

3.2. DEEP LEARNING–BASED FEATURE EXTRACTION

While sensor data provides valuable context, a direct examination of leaf images is key to diagnosing many diseases. However, limited data can impede network performance, making transfer learning appealing. By employing established networks such as VGG19, ResNet, or Inceptionv3, a model can benefit from pretrained weights to boost accuracy on a relatively small dataset. Convolutional Neural Networks (CNNs) excel at extracting pertinent image features—edges, spots, and shapes—that can classify a leaf as either healthy or diseased.

These features, when combined with the sensor-derived parameters (temperature, humidity, and moisture), offer a more holistic profile of a plant's health. In a deep learning framework, each input feature is automatically assigned a weight reflecting its relative importance to the final classification. This end-to-end approach, leveraging both sensor streams and advanced CNN architectures, has the potential to improve both precision and recall in disease detection.

5. PROPOSED SYSTEM

5.1. MATERIALS AND METHODS

To validate the proposed approach, a controlled greenhouse environment was set up with 8–10 tomato plants under careful monitoring. This greenhouse had temperature averages around 29°C and humidity around 69%. The image dataset comprises various sources, including the well-known Plant Village dataset, actual photographs of greenhouse-grown plants, and additional images from the internet. By combining these diverse sets, the system can handle a range of real-world lighting conditions and background complexities.

Figure 1 illustrates the greenhouse setup and the IoT components (moisture, temperature, and humidity sensors) used to collect data.

Fig. 1. Greenhouse and IOT Components

5.2. IMPLEMENTATION METHODOLOGY

5.2.1. Pre-processing

Images are enhanced, filtered for noise, and resized. This step not only lowers computational overhead but also often improves classification performance, as consistently sized and denoised images can reveal disease symptoms more clearly.

5.2.2. Feature Extraction

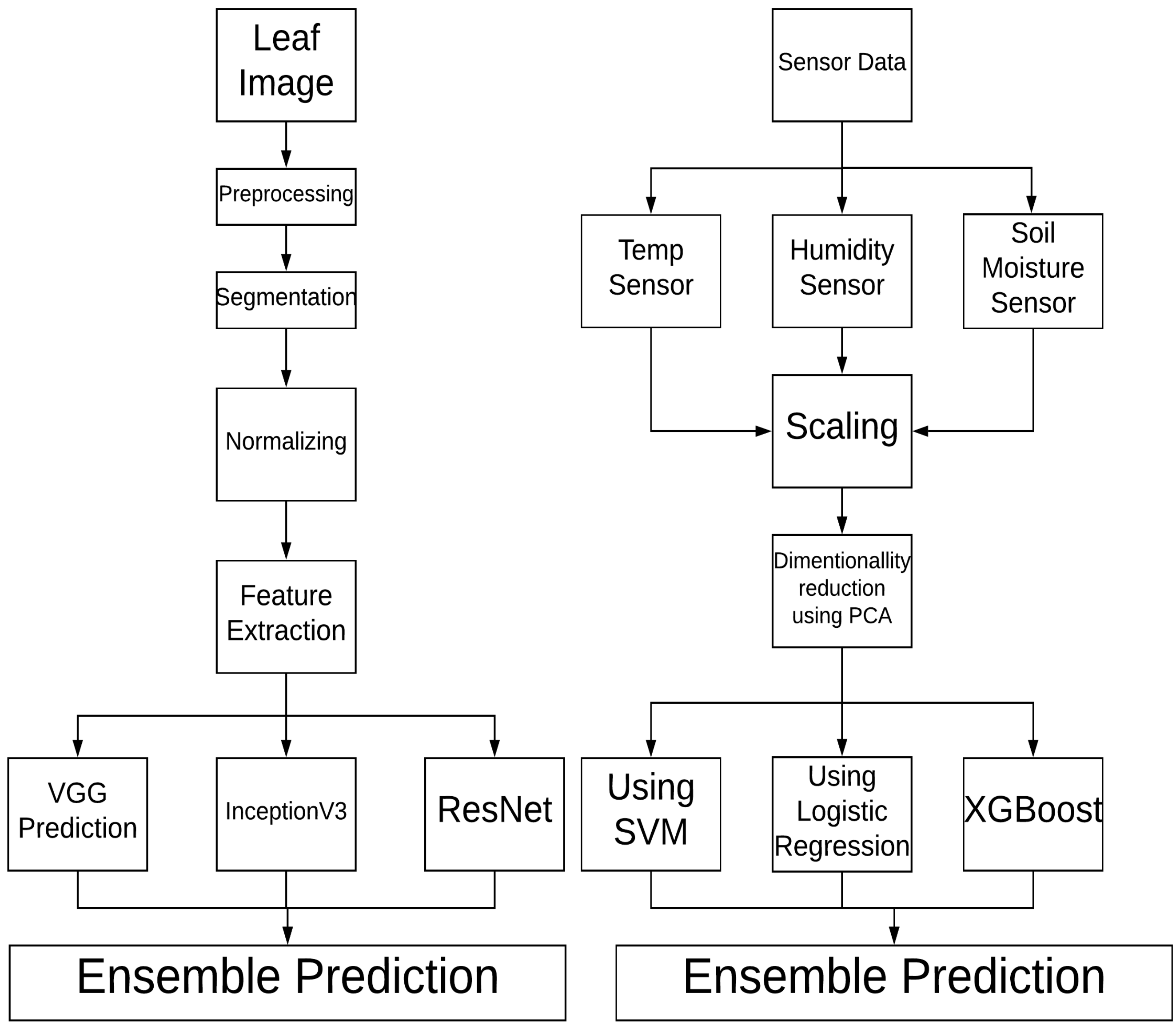
The system extracts two categories of features: (1) sensor-based measurements (temperature, humidity, and moisture), which capture environmental conditions, and (2) visual leaf properties derived from digital images. The inclusion of IoT-based sensor inputs accounts for the broader growth context, thereby strengthening the reliability of predictions.

5.2.3. Ensemble Classification

In the concluding phase, multiple deep neural networks are applied to the same problem. Alongside sensor-based classifiers, CNN-based modules generate predictions on leaf images. By combining these outcomes in a "voting" or "stacked" ensemble, the final diagnostic decision exploits the strengths of each individual model.

This aggregated output often provides higher accuracy and robustness than any single model alone.

Figure 2 presents the system's block diagram, demonstrating how sensor data and leaf images are collectively processed. Each data stream undergoes separate feature extraction, and the outputs are then integrated to achieve the final prediction.



CONCLUSION

Prior research on detecting disorders in tomato plants has largely focused on either image-based techniques or sensor-based monitoring. In contrast, the system outlined in this paper merges leaf images with real-time sensor data—enabling a more complete and accurate diagnosis of plant various types of plant disorders. By providing actionable insights into both environmental conditions and visual disease indicators, the proposed prototype holds promise for scalable, precise, and cost-effective crop protection.