Vol. 1 - No. 1
September 2023 E-ISSN: 2584-0657

Design of a Comprehensive Sensor based Soil and Crop Analysis System Implemented using Machine Learning Algorithms

H. Aafreen Sana, P. Pravin Kumar, M. Shabika Fathima, B. Yashwanth Krishnan
Department of Computer Science Engineering
Sri Sairam Engineering College, Chennai, Tamil Nadu, India

I. INTRODUCTION

We live in a world where technology has become an important part of our daily lives, this revolution of technology began when curiosity for technology began, and thus it marked a new niche in human civilization. Soil plays an important role within the Earth's system. While not soil human life would be terribly troublesome. Soil provides plants with foothold for his or her roots and holds the mandatory nutrients for plants to grow. Soil functions are general capabilities of soils that are necessary for numerous agricultural, environmental, nature protection, architecture and concrete applications. Here's an example of how our technology can become an integral part in smart farming and soil monitoring.

A moisture sensor can collect the data about the soil moisture and passes it to the irrigation system, when the irrigation system receives the information it automatically acts upon it by watering the soil. In farmland collect the information regarding the soil moisture and send this, by this the farmer knows how much water is needed to be watered.

Agriculture is the backbone of the Indian economy and around the seventieth of the population depends on this field to run their support. From past agriculture has been a vicinity of the human civilization. It's reworked the approach humans survive. The economy of a specific space was indirectly addicted to agriculture and was a significant thrust behind the economic revolution. Advancements within the field of science and technology junction rectifier to accrued yield. Applying electronic observance systems is one of the technologies for analyzing vital conditions needed for optimum growth of plants. The conditions will be listed as temperature, humidity, carbon oxide, and soil wet and soil hydrogen ion concentration. There are a unit valuable knowledge that may decide the flowers cycle. Economical use of those parameters will increase the output per plant and minimizes crop loss. The quantum of steps taken to watch ne'er ends here, a lot of knowledge assortment successively will increase the accuracy and by effort, no stones right-side-up potency of harvest and output will increase. Agricultural stations have developed novel ways for observance the information, and programs to assist the farmer to generate a lot of output. Desegregation of numerous sensors that area units rugged and capable of generating arduous knowledge in real-time will augment additional analysis. Presently geographical land use patterns, soil parameters area unit determined exploitation satellites, and noninvasive techniques that area unit subtle and generate precise knowledge in real-time.

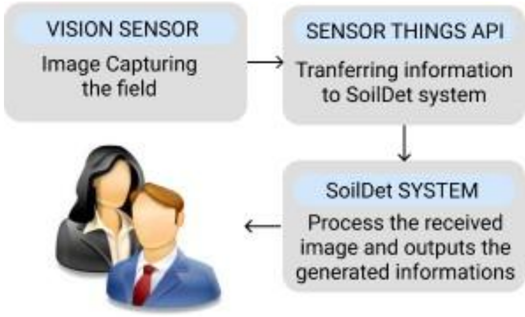

Figure 1: The framework to buil SoilDet System

II. LITERATURE SURVEY

There have been several studies that aim to harness the power of image detection via computer vision to train machine learning models that will be very beneficial to farmers by performing soil analysis. In Agarwal et al. propose a method to analyse soil fertility based on the principle of Colorimetry. The Arduino UNO R3 microcontroller board equipped with a colour sensor is used to analyse an aqueous sample of the soil whose output is calibrated with a database containing information about various soil types. The accuracy of the results obtained is verified using the Naive Bayes Classification Algorithm. Such analysis proves to be highly accurate to analyze the soil fertility and take effective measures to improve the fertility of the soil if required.

In Inazumi et al. discuss the potential of image recognition by artificial intelligence, using a deep learning model of machine learning for the purpose of expanding the cases which employ artificial intelligence. Deep learning was performed with a neural network model. Computer Vision classifies the images as a series of pixels and colour codes to let the machine see and analyse the image input to predict the type of soil.

Radovanović et al. use an existing PlantVillage dataset which comprises images of the leaves of plants taken in a controlled environment to detect fungal, bacterial, viral, mould, and mite diseases affecting the plants. The machine learning analysis was done using Python, scikit-learn, Keras and tensor flow to implement Support Vector Machine, kNN, FCNN, and CNN models. The authors conclude that the deep learning CNN model outperformed other classical machine learning algorithms with high precision and low error rate. Hence while analysing the images of the plants grown in a particular soil deep learning mechanisms can be employed. If it is inferred that the plants grown in a particular soil type are more prone to diseases, our machine learning model can be accordingly trained to ensure that such plants are not recommended to be grown in such soil.

Harnessing the power of Computer Vision and the easy availability of cameras Kumar et al. present an automated system for the categorization of the soil datasets into respective categories using images of the soils. Once the soil has been categorized, crops that can give the maximum yield in the soil are suitably suggested. Similarly, in Han et al. compared and analysed the roles of the visible spectrum and machine vision adopted in soil classification. Subsequently citing the portability of smartphones these days, they propose a new smartphone- based, low-cost, and miniaturized soil colour classification sensor with a low error rate.

With regard to an optimized method of soil classification, Khatti et al. developed a modified textural classification for soil using the concept of two existing soil classifications, namely the international classification system of soil and the textural classification system of soil. The original textural classification systems of soil used triangulation for soil classification based on particle size distribution considering sand, silt and clay as the type of particles. By combining the attributes of two soil classification systems, the proposed modified textural classification method can efficiently be used to train machine learning models with higher precision and less error rate

In Ajdadi et al. develop an algorithm that seeks to evaluate tillage quality in real-time using image processing. Since tillage is an important step in preparing the soil for the growth of crops, the analysis of whether the tilling was performed correctly will have an impact on the end yield of a particular crop in the soil which in turn affects the predictions of a machine learning models which aims to predict crop yield based on the type of soil. To evaluate the quality of tilling photography was performed at three- camera heights and covering nine different sizes of soil to ensure the accuracy of the model with various test cases. The aggregated photos are made accessible to the farmers for real-time soil tracking.

With a wide variety of sensors available in the market at an affordable rate, many IoT devices integrated with sensors and wireless networks have been designed as low-cost smart devices to help farmers and improve crop yield. Sharma et al. have developed a portable microcontroller-based IoT device designed with sensors to detect Electrical Conductivity, pH and colour of the soil. The sensor readings obtained are transferred to a mobile application via Bluetooth for graphical representation and subsequent data upload to a web service which will function as a centralised database for soil comparison. The database can be monitored by farmers and can be used to automate agricultural prediction models using machine learning.

In Ezhilazhahi et al. use the Zigbee technology to transmit data from a soil moisture detection probe to Raspberry Pi 3 to ensure continuous monitoring of soil moisture of the plant. Additionally, they make use of the Exponential Weighted Moving Average (EWMA) event detection algorithm to improve the lifetime of the network. Zigbee network was used for the above device citing its low power requirement and secure nature.

In Reshma et al. propose an IoT system designed with pH, Temperature, Soil nutrient, Humidity and soil moisture sensors connected to a microcontroller designed with WiFi and access to cloud storage. The readings from the sensors are transmitted with timestamps to the cloud database. Subsequently, the Support Vector Machine and Decision Tree Algorithms are used to give suitable crop recommendations based on the comprehensive data stored in the Cloud database. Since a large dataset comprising of vital information regarding the soil is obtained, highly accurate machine learning models can be predicted after the data is cleaned and optimised.

Bacu et al. present the HORUSApp application to make use of satellite imaging in soil analysis. The App intends to integrate the multispectral data coming from the Sentinel -2 satellite images into the soil analysis and classification process which can predict the soil type and its classification. The Sentinel – 2 has been particularly chosen due to its wide range and high-resolution images. Hence a comprehensive image dataset of the soil can be analysed to make accurate predictions.

In Mohapatra et al. propose a machine learning model which can predict soil type and gives correct information to the farmers in audio format for improvised cultivation. It collects different soil parameters such as soil temperature, moisture and nitrogen, phosphorous and potassium values present in the soil by taking the help of different sensors and predict the soil type using Random Forest Classifier, Support Vector Machine and Linear Regression Algorithms. Since the results are presented in the audio format it will be of much use to the farmers who cannot read with ease.

Syed et.al propose a Machine Learning Model using Tensor Flow Object Detection API integrated with IoT to maintain optimum soil moisture for the growth of plants by using a variety of sensors like Humidity, Moisture, Wind and Temperature Sensor. The sensor readings, metrological data obtained using the IFTTT Weather Underground applet service and therealtime crop condition are fed into the machine learning model to suitably perform smart irrigation and suggest suitable crops.

In Malik et al. make use of Naïve Bayes, K Nearest Neighbour and Decision Trees model to predict crop yield by analysis of ambient, soil properties and previous history of crop yield for the given soil sample and geographical area. In their study, the machine learning models were used to predict the yield of tomatoes, potatoes and chillies.

With regard to the machine learning models employed for soil analysis, several studies imply that the Support Vector Machine model produces the most accurate results. In the authors analyse the higher accuracy of the Gaussian Kernel-based Support Vector Machine Algorithm as opposed to the Bagged Tree and weighted K Nearest Neighbour Model to perform soil classification. The soil was classified into classes based on Geographical features and Chemical features like Salinity, Organic Matter, Mineral Composition etc. Based on the soil classification a list of crops suitable for the classified soil is suggested from the database.

Considering the high accuracy of the Support Vector Machine Model Srunitha et al. explain Support Vector Machine based classification of the soil types. Soil classification includes steps like image acquisition, image preprocessing, feature extraction and classification. The texture features of soil images are extracted using the low pass filter, Gabor filter and using colour quantization technique. Mean amplitude, HSV histogram, Standard deviation are taken as the statistical parameters used to analyse the results obtained.

Hence, although there are studies proposing aspects of soil monitoring like analysing soil moisture content, ambient temperature checking, predicting irrigation patterns, clustering of plants that are highly suitable to be grown in a particular soil etc. as individual components, there is no comprehensive system that performs all the above functions. Therefore, it is highly essential to ensure the culmination of the above predictions to enable efficient farming using machine learning and IoT along with an intelligent pricing system to set marginal prices based on local and international demand for the product. Such a system would be highly beneficial for the farmers by ensuring better yields and fair crop procurement rates.

 E-ISSN: 2584-0657

III. Sensors

We live in a world where we are connected to multiple devices which in turn contains various kinds of sensors. In offices, house, schools, cars etc. we could find these sensors which make our life more convenient. We have smart homes where we could even access temperature, lightning, opening garage doors etc. The whole system works by transferring the input received from the user to the system and which in turn processes and takes a quick decision as output. In general, a sensor is a device that measures physical input from its environment and converts it into data that can be interpreted by either a human or a machine. Most of the sensors are electronics. A sensor could be either hardware or software type. Now let us see some of the sensors used in our system in Figure 2,

TYPES OF SENSORS	TYPE	DESCRIPTION
Pressure	Hardware	Measures the ambient air pressure
Temperature	Hardware	Measures the ambient temperature
Vision	Hardware	Captures the entire area with clear magnification
Water	Hardware	Measures the amount of water level
GPS	Software	Provides location, velocity and time synchroniztion
Relative Humidity	Hardware	Measures the relative ambient humidity

Figure 2: Types of Sensors

IV. Methodology and Implementation

In this research, we have planned to achieve the following objectives.

1. To detect moisture content, temperature checks, predicting irrigation needs, and collecting data about rainfall.
2. To detect the pH base on captured soil image and find the soil type
3. To cluster a variety of plants that are suitable for the given soil type.
4. To classify and suggest what sort of suitable crop to be planted based on their soil properties.
5. To set marginal prices based on local and international demand for the product (Raw materials).

Soil monitoring is one of the most excellent features used in IoT farming. By using the concept of IoT and machine learning we are able to achieve the above objectives. Here the soil monitoring sensor checks the moisture content, temperature checks, predicting irrigation needs, and collecting data about rainfall, by placing sensors in the field. Next, to detect the pH of the soil, vision sensors are used to capture the field image and transfer it to the SoilDet system, “SoilDet” is an IoT-based system that identifies the pH value of the soil using machine learning i.e., identifying whether the given soil is acidic, alkaline, or neutral and classifying what sort of plants/crops suitable for planting in the given soil. SoilDet is categorized into two distinct categories. The first and foremost category is the identification of the soil type and producing information based on the soil's nature. The classification of plants and crops based on soil's nature using soil classifier.Using the concept of supervised learning we are able to set the marginal price of the crop based on local and international demand.

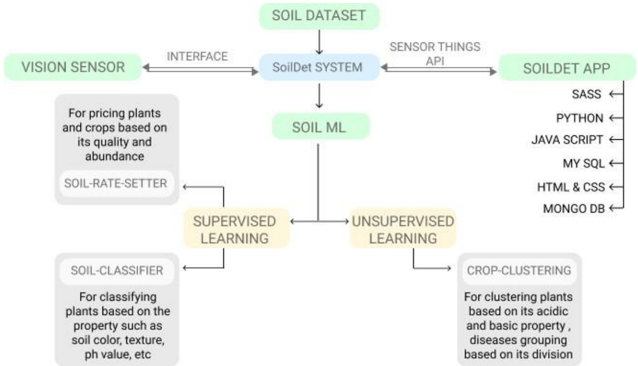

Figure 3: Soil Monitoring System Backend Flow Design

1. Detecting moisture content, temperature checks, predicting irrigation needs, and collecting data about rainfall

Knowing the information about the soil is not a piece of cake, as technology evolves, many things work on the internet. Instead of working primitively in a farm field. The concept of smart farming has come into focus, the use of sensors has been highly popularized. In soil monitoring, the use of sensors, such as water sensors to check the moisture content present in the soil, and temperature sensors to check the temperature of the soil. Helps to reduce the manpower involved in the process. Making the whole process less strenuous and producing excellent results

2. detecting the pH base on captured soil image and find soil type

With the implementation of vision sensors in the farm field. We are able to get a clear 2D image of the soil, with the help of the interface connecting the sensor and SoilDet system, the image of the soil is transferred to SoilDet system. Here SoilDet system works have been incorporated with machine learning, a large amount of soil dataset has been fed to the system. The dataset contains images of different soil, which are labeled with information about the pH of the soil, soil type, soil texture, soil nature, soil color, plants that are suitable to be grown on the given soil, the rarity of the soil, and also it gives information about diseases that are plants prone to, etc. Based on the information given by the soil dataset. The SoilDet system is able to identify the pH of the system. On the identification of pH value, the system classifies soil into three types, if the pH is less than 7 the given soil is classified as acidic. If the pH is more than 7 then the given soil is basic. If the pH is equal to 7 then the given soil is neutral.

3. Clustering a variety of plants that are suitable for the given soil type

Planting the same crops repeatedly will lead to a decrease in the efficiency of the crop and which in turn will also reduce the crop yield. In order to avoid this, we use the concept of crop rotation to increase the crop yield and to also produce good quality crops. By using the method of soil clustering, we are clustering a large variety of crops based on the soil property, plant diseases, plant nature etc. using the concept of clustering a model of unsupervised learning we are able to cluster a large variety of crops, using the clustered information we receive, we can plant.

4. Classification and suggesting what sort of suitable crop to be planted based on their soil properties

During the classification process, the soil classifier implements the process of classification, which is a model of supervised learning which is a type of machine learning model. Here a group of crops is classified based on soil properties, ex: alluvial soil is more suitable for the growth of sugarcane, rice, cotton, etc. with the help of classification model, the system filters through a large amount of data feed to the system, and produces a list containing crops suitable for planting based on the soil properties.

5. Setting marginal prices based on local and international demand for the product

Knowing the rate of the crop is a very important factor for farmers who sell their crops to the middle man, who often scams/ tricks the farmer into selling their crops for a very low price leading to the farmer suffering a very heavy loss. In order for the farmers/ businessmen not to get scammed. We use Crop-Rate- Setter, which works on the process of regression which is a model of supervised learning. Here the estimated marginal rate of the crop is set, based on the rarity of both the soil and crop, demand in the market for the crop, and local availability of the crop. Ex: Darjeeling tea leaves are sold at a very high price because of their unique nature, rarity, and taste. 1 kg of first flush Darjeeling tea is sold at Rs 8000. With the data fed to the system, the soil-rate-setter sets the rate of the crops.

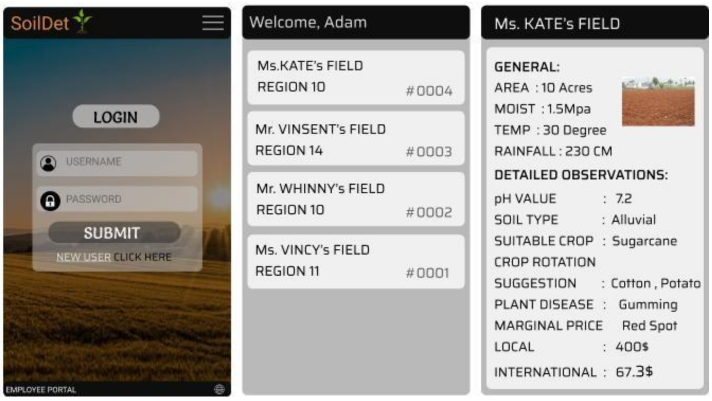

Figure 4: SoilDet Employee Portal UI Architecture

V. CONCLUSION

With the world battling a crippling pandemic like the Coronavirus, the inadequate access to soil testing laboratories has generated a booming requirement for the remote, accessible, and real-time analysis of soil and crops. With the minimal use of sensors we have proposed a low- cost IoT farming companion that comprehensively collects data about rainfall and temperature, checks soil moisture content, and predicts irrigation needs as well as the detection of pH using the high-resolution images captured from a Vision sensor instead of using additional sensors. By aggregating the data mentioned above we develop a robust dataset that can be used to predict the type of soil and the plants which can be grown in it. Such a prediction is of utmost importance since the crop yield highly depends on the compatibility of the crop with the soil. One of the important features of our model is the determination of crop prices based on demand which is of great benefit to farmers as well as the consumers. With regard to the future scope of the study, we intend to measure the exact amount of fertilizers and pesticides required by the soil depending on the nutrient analysis to combat the excessive usage of chemicals. With the right mechanical implementations, our software model can be used to automate the tasks of irrigation and spraying of pesticides. Since our model intuitively provides all the above functionalities via the SOILDET app with a simple user interface it will be easily accessible and beneficial to farmers. We believe that our system will be the catalyst for the prosperity of farmers and in turn the welfare of entire nations.

VI. References

[1] Agarwal, Surili, et al. "Application of Colorimetry to Determine Soil Fertility through Naive Bayes Classification Algorithm." *International Conference on Computing, Communication and Networking Technologies*, 2018.

[2] Inazumi, Shinya, et al. "Artificial Intelligence System for Supporting Soil Classification." *Results in Engineering*, vol. 8, 2020.

[3] Radovanović, Draško, and Slobodan Đukanović. "Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms." *International Conference on Information Technology (IT)*, 2020.

[4] Kumar, Sandeep, et al. "Automated Soil Prediction using bag-of-features and Chaotic Spider Monkey Optimization Algorithm." *Evolutionary Intelligence*, vol. 14, 2018.

[5] Han, Pengcheng, et al. "A smartphone-based Soil Color Sensor: for Soil Type Classificatio." *Computers and Electronics in Agriculture*, vol. 123, 2016, pp. 232-241.

[6] Khatti, Jitendra, et al. "Modified Textural Soil Classification." *Proceedings - Geotechnical Characterization and Modelling*, 2020.

[7] Ajdadi, Fatemeh Rahimi, et al. "Application of Machine Vision for Classification of Soil Aggregate Size." *Soil and Tillage Research*, vol. 162, 2016, pp. 8-17.

[8] Sharma, Prachi, and D.V. Padole. "Design and Implementation Soil Analyser using IoT." *International Conference on Innovations in Information, Embedded and Communication Systems*, 2017.

[9] Ezhilazhahi, A.M., and P.T.V. Bhuvaneswari. "IoT Enabled Plant Soil Moisture Monitoring Using Wireless Sensor Networks." *International Conference on Sensing, Signal Processing and Security*, 2017.

[10] Reshma, R., et al. "IoT based Classification Techniques for Soil Content Analysis and Crop Yield Prediction." *International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, 2020.

[11] Bacu, V., et al. "Building Soil Classification Maps using HorusApp and Sentinel-2 Products." *International Conference on Intelligent Computer Communication and*

Processing, 2019.

[12] Mohapatra, Suman, et al. “Soil Analysis and its Type Prediction with Speech Enabled Output using IoT and AWS.” *India Council International Conference*, 2020.

[13] Syed, Fahad Kamraan, et al. “Low-cost IoT+ML Design for Smart Farming with Multiple Applications.” *International Conference on Computing, Communication and Networking Technologies*, 2019.

[14] Malik, Pranay, et al. “Comparative Analysis of Soil Properties to Predict Fertility and Crop Yield using Machine Learning Algorithms.” 2021.

[15] Rahman, Sk Al Zaminur, et al. “Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series.” *International Conference of Computer and Information Technology*, 2018.

[16] Srunitha, K., and S. Padmavathi. “Performance of SVM Classifier for Image based Soil Classification.” *International Conference on Signal Processing, Communication, Power and Embedded System*, 2016.