Soil Wetness Classification in Agriculture using Machine Learning Models

Comment [DB1]: Title should be "**Image-Based Soil Wetness Detection in Agriculture Through Machine Learning**" more appropriate for this manuscript

Abstract

Soil wetness is the most important factor for a plant to survive. If the soil is completely dry for a long time, the plants will perish. Many plants will also die if the soil is submerged in the water for a long period of time. Without water, plants will not be able to take nutrients from the soil. Besides, different plants have different soil wetness or moisture requirements. To ensure proper plant growth, soil wetness levels should be monitored and maintained continuously. But most of the time, it is not possible to continuously monitor the water level manually. Therefore, in this work, we have proposed an image-based soil wetness classifier using different machine learning algorithms. We have classified the soil in six different wetness levels and have used five machine learning algorithms for classifying the soil wetness levels as an artificial neural network, convolutional neural network, decision tree, k-nearest neighbor, and support vector machine. We have compared these algorithms and found that the convolutional neural network achieves the highest accuracy which is 97.7%. Our proposed method can be used by the stakeholders to increase crop production by ensuring proper soil water levels for continuous plant growth.

Keywords- ANN (Artificial Neural Network), KNN (K-Nearest Neighbor), DT (Decision Tree), SVM (Support Vector Machine), CNN (Convolutional Neural Network), and CM(Confusion Matrix).

Comment [DB2]: Add "**Soil Wetness and Machine Learning**" key words

1. Introduction

The water content of the soil is one of the most important aspects of agricultural production. It is required for all life on earth including plants and organisms in the soil. Because soil water plays a major role to dissolve nutrients for plant uptake. If there is no water in the soil, plants and organisms cannot take any nutrients and water from it. On the other hand, soil submerged in the soil can prevent many plants and organisms to survive. Therefore, proper water level or wetness level of the soil is very important for plant growth. Although larger trees can draw water and nutrient from deeper parts of the soil, crops and small plants mainly rely on surface soil moisture. But due to environmental conditions and seasonal changes, the soil may dry out which can damage the crops heavily. To prevent soil, dry out, regular irrigation is necessary to maintain the wetness of the soil. On the other hand, excess irrigation can wash out the nutrients of the soil and prevent plant growth. Therefore, balanced soil wetness is required for proper plant growth and this soil wetness requirement also varies with different plants [1]. Traditionally, soil wetness is measured by looking at the soil surface and if the soil looks dry, irrigation is carried out by the farmers or garden owners. This traditional process requires regular monitoring and maintaining of soil wetness. But most of the time, regular monitoring can be hampered and soil may become completely dry. The land owner may also forget to stop the irrigation which may cause the crops to be flooded with water. This over-irrigation can cause huge water loss. To automate this monitoring of soil wetness, soil moisture sensors have been used. But these sensors are placed inside the soil which results in sensor corrosion over time. As a result, sensors may not work or give faulty readings. If the sensors give faulty readings, it is not possible for the farmers to check the actual soil condition remotely from the sensor data alone. He has to visit the land to match the sensor readings with soil conditions. To overcome these sensor limitations, image-based soil wetness monitoring can be used to monitor the ground soil wetness and these images can be collected using different types of cameras (e.g., outdoor CC or IP cameras). Farmers can also check the actual soil surface remotely from the images. Therefore, this image-based system can provide much more reliable soil wetness monitoring. As mentioned in the previous section, an image-based system can provide reliable soil wetness monitoring. Therefore, in this work, we aim to integrate machine learning algorithms to monitor and classify soil wetness levels automatically and efficiently using soil images. This will help to increase crop production, decrease water wastage and retain soil nutrients more efficiently. The proposed machine learning-based soil wetness classification system can be integrated with IoT (internet of things) enabled devices to make smart farming more reliable. Since there are no notable works on soil wetness classification using machine learning algorithms, we have developed this machine learning-based soil wetness classification. We have also compared five machine learning algorithms for this classification. These are artificial neural network (ANN), convolutional neural network (CNN), decision tree (DT), k-nearest neighbor (KNN), and support vector machine (SVM). The main contribution of this work-

- We have collected and captured thousands of soil images on different wetness levels.
- This image dataset will be made public and can be used by other researchers.
- We have applied different machine learning algorithms for automatically classifying the soil image wetness levels.
- We have also used a convolutional neural network that does not require manual feature extraction like other machine learning algorithms.
- We have compared the performance of these algorithms and found that the convolutional neural network performs the best in terms of classification accuracy.

The rest of the paper is organized as follows: In section-2 briefly explains the literature review. Section-3, explains the methodology, and dataset. Section 5 is the analysis of the results. Section 6 represents the future work and conclusion.

2. Literature Review

In this part, we have extensively reviewed the existing works on soil wetness or moisture level classification and related works using image processing techniques. We have briefly explained the contributions and limitations of the existing systems for improving our present work.

In the case of Y. Cal et al. [2] applied a backpropagation neural network (BPNN) for clay soil classification. Here, they have used six types of clay soil images, namely heavy clay, light clay, heavy sub-clay, medium sub-clay, light sub-clay, and sub-sandy soil. These types of soil images were used for training the BPNN. They used only three hundred ninety-six images for training and twenty images for testing the classification of clay soil images. Due to the use of a small number of images, BPNN may produce biased results. On the other hand, Hongli Jiang et al. [3] estimated soil moisture by using the artificial neural network (ANN) model. For soil moisture estimation, they used remotely sensed Infrared Reflective (IR) data. Here, they have compared results between the ANN model and the API (Antecedent Precipitation Index) and found that the ANN model produces a more accurate estimation of soil moisture than API. Their dataset is not publicly available. Besides, images captured from satellites may have different types of noises which may produce errors in the classification results. Moreover, Ahmad et al. [4] also classified soil moisture content using image data from remote sensing. Here, they have used three classifiers, namely support vector machine (SVM), artificial neural network (ANN), and multiple linear regression (MLR) algorithms to classify soil moisture content. These space-borne remote sensed data were collected from the Tropical Rainfall Measuring Mission (TRMM) and Advanced Very High-Resolution Radiometer (AVHRR) and the site was in the Lower Colorado River Basin in North America. From the experiment, they observed that the SVM classifier achieved high accuracy compared to the other two classifiers. The space-borne remote sensed image data are not publicly available. Wireless sensors have been used by X. Gao et al. [5] to classify soil moisture. They also used SVM for the classification. They did not compare their results with any other algorithms and their sample data set was very small. Iftikhar Ali et al. [6] have reviewed machine learning methods like ANN, SVM, and Linear Regression (LR) for biomass and soil moisture classification. Ashwini Rao et al. [7] classified soil and crop using SVM. They considered different soil images, namely red, black, clay, and alluvial for their proposed work. They did not mention the number of soil image samples and their source. SVM is also used by Pethkar et al. [8] for the classification of different soil samples, namely peat, sandy soil, and clay soil from various land sources. They also did not discuss the number of training and testing images used in this study. Furthermore, Abdullahi et al. [9] proposed an image classification model to classify satellite images. These satellite images were captured from various land regions in Nigeria for the proper treatment of different crops. In their study, they used two machine learning algorithms namely SVM, and CNN. Above these two algorithms, the convolutional neural network generated outstanding results from the support vector machine. But they did not discuss how many images were used and how many features are extracted in their study. In addition Chandan et al. [10] classified soil type from the Indian ground surface using SVM. The soil types were clay, clayey peat, clayey sand, humus clay, peat, sandy clay, and silty sand. Here, they have used 24 soil images for training. The result was compared with ANN and KNN. The SVM classifier achieved a high accuracy than ANN and KNN classifiers. The number of images was very small and thus the results are prone to underfitting. Priyanka et al. [11] used Maximum Likelihood classification (MLC), Subpixel classification (SP), and SVM to predict soil images into seven soil class levels, namely clay, clay peat, clay, sand, humus clay, peat, sandy clay, and silty sand. Among these three methods, MLC and SVM exhibit high classification accuracies but the SP method displays poor classification accuracy. They did not mention the number of images used and the source of the images. Hu et al. [12] analyzed the global soil moisture. They used a passive microwave sensor AMSR-E, as well as modeled soil moisture reanalysis from ERA-interim in the ECMWF model which is a global atmospheric reanalysis from 1979, continuously updated in real-time. They used CNN to predict the global soil moisture map and found that the CNN model achieved better performance than the SVM model to predict the global soil moisture. On the other hand, applications of machine learning in agriculture have been reviewed by Liakos et al. in [13]. The authors described the works from different categories including crop management, applications on yield prediction, disease detection, weed detection, crop quality, species recognition, livestock management, water management, and soil management. Riese et al. [14] used the freely available Land Use Cover Area Frame Statistical Survey (LUCAS) soil dataset. This soil dataset includes hyperspectral and soil texture data from measurements all over Europe. This soil dataset was collected in various areas all over Europe between 2009 and 2012. Here, the authors used CNN to classify soil texture.

A Feed-forward neural network was used in [15] for predicting soil moisture. In this work, they used four hidden layers, and each hidden layer contained 26 neurons. Lim et al. [16] discussed the geophysical parameters for classifying the soil. They have employed three machine learning models, namely decision tree, artificial neural network, and Bayesian network to classify the soil characteristics. The Bayesian network showed the highest accuracy among the machine learning models. Moreover, Nuno et al. [17] analyzed the acid sulfate property of soil in a small area of Finland. This acidic property of soil is a major environmental problem in his country. He classified the soil acid property by using random forest, gradient boosting, support vector machine, and convolutional neural network (CNN). In this study, the dataset consisted of 187 non-acidic and 93 acidic soils. He found the highest accuracy using random forest rather than CNN. Here, CNN did not perform well due to the small dataset. Because deep learning algorithms require a lot of images for good results. Finally, decision tree (DT) regression was used in [18] by Pekel to calculate soil moisture considering various parameters including air temperature, time, relative humidity, and soil temperature. The dataset was collected from the

University of Toronto Mississauga campus. This soil data included three different areas namely, pond, field, and forest area. They considered 1000 soil images among which 80% of the images were used for training and 20% for testing. They have only used the DT algorithm for their soil classification. In table 2.1, we have summarized the works which were discussed in this chapter.

Table 1: Summary of the existing works

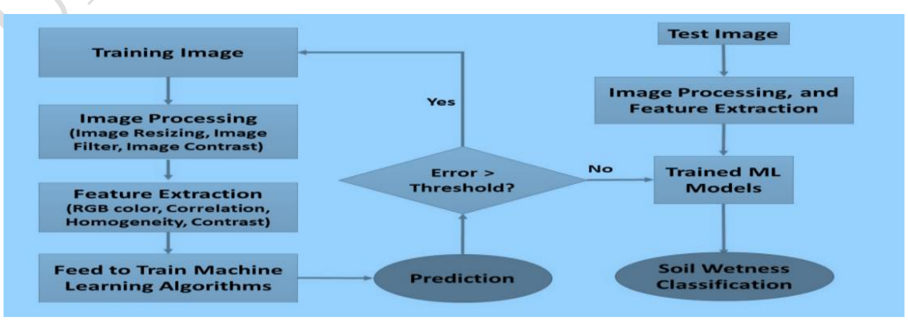
Literature	Year	ML Algorithm	Classification Task
[2]	1995	BPNN	Clay soil classification
[3]	2004	ANN	Soil moisture Estimation
[4]	2010	SVM	Soil moisture Estimation
[5]	2014	SVM	Soil moisture Classification
[6]	2015	ANN, SVM, LR	Soil moisture classification
[7]	2016	SVM	Soil classification and crop Detection
[8]	2016	SVM	Soil layer classification
[9]	2017	CNN	Soil nutrient classification
[10]	2018	SVM	Soil type classification
[11]	2018	SVM	Soil type Classification
[12]	2018	SVM	Soil Classification
[13]	2018	CNN, SVR	Global Soil moisture Classification
[14]	2019	ANN	Different factor classification in agriculture
[15]	2020	CNN	Soil fertility, color, substantially prediction
[16]	2020	CNN	Soil texture classification
[17]	2020	FFNN	Soil and Soil moisture Classification
[18]	2020	SVM	Soil color, PH, texture classification

3. Methodology and Dataset

The algorithms and datasets are explained in this section. To categorize soil moisture, these algorithms are utilized. Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Decision Tree (DT), Support Vector Machine (SVM), and CNN are some of the methods used in this (Convolutional Neural Network). This study focuses on identifying the most qualified pre-trained CNN candidates (Convolutional Neural Network). These algorithms use a dataset made up of soil photographs with varying degrees of moisture.

3.1 Proposed System

Here, a block diagram and an algorithm are used to demonstrate the suggested system. In Figure 3.1, the block diagram is displayed. Because the original photos could include noise and artifacts, the training images of soil are scaled and given a different pre-processing in the first stage. The photographs' contrast is then improved. After extracting various attributes, the machine learning models are trained to categorize the photos. A confusion matrix is used to assess the model's accuracy once training is finished. The suggested system's whole process is broken down into four fundamental parts, which are covered in more detail below: training image, image preprocessing, feature extraction, image classification, and data assessment.



Comment [DB3]: Either keep each word capital or first word capital for all sentences as written in classification task

Comment [DB4]: If possible, add more recent work that were carried out after 2020

Comment [DB5]: Modification is necessary in this sentence

Comment [DB6]: Write the correct figure name instead of 3.1 write 1.

Figure 1: Block Diagram of Proposed System

The algorithm of the proposed system is shown in algorithm 1. This algorithm shows the steps of the proposed model.

Comment [DB7]: Figure

Table2: Proposed System Algorithm

Algorithm : Soil wetness classification	
1	Read training images from the dataset
2	Pre-process images
3	Train the models (ANN, DT, KNN, SVM, and CNN) to classify
4	Read test image
5	Apply the trained model to classify the test image
6	Measure classification_error

7	If (classification_error > threshold)
	Go to step 1.
	Else
	Log accuracy.

3.2 Training Image

Here, a block diagram and an algorithm are used to demonstrate the suggested system. In Figure 1, the block diagram is displayed. Because the original photos could include noise and artifacts, the training images of soil are scaled and given a different pre-processing in the first stage. The photographs' contrast is then improved. After extracting various attributes, the machine learning models are trained to categorize the photos. A confusion matrix is used to assess the model's accuracy once training is finished. The suggested system's whole process is broken down into four fundamental parts, which are covered in more detail below: training pictures, image preprocessing, feature extraction, image classification, and data assessment.

Comment [DB8]: Same paragraph was written in sub-heading **3.1 Proposed System so, correct it**

Figure 2: Different 6 levels (0%, 20%, 40%, 60%, 80%, 100%) of soil wetness

Comment [DB9]: Correct the title

Table 3: Details of Image Dataset

Class Name	% of Wetness Level	Label	No. of samples
Class0	0%	0	1,218
Class2	20%	1	1,091
Class4	40%	2	1,164
Class6	60%	3	1,157
Class8	80%	4	1,282
Class10	100%	5	1,268
		Total Samples	7,180

3.3 Image Pre-processing

Diverse sources provide images with varying noise levels and resolutions. As a result, downsizing the photos and eliminating noise need image pre-processing. The images were resized using the Matlab resize () function to 320x180 pixels, then smoothing, sharpening, and edge enhancement were achieved using the filter() function. Additionally, we have used the rgb2gray () method to remove hue and saturation data, and the gray comatrix() tool to analyze the texture of the photos.

Table 4: Image processing functions

ML Model	Description	Matlab Functions
ANN, SVM, DT, KNN	Resizing images to 320 x 180 pixels	resize()
	Filtering for smoothing, sharpening, and edge enhancement	filter()
	Eliminating the hue and saturation information	rgb2gray()
	Texture analysis of the images.	gray_matrix()
CNN	Automatic resizing and converting of the grayscale image to 224 x 224 RGB image	augmentedImageDatastore()

3.4 Feature Extraction

A quality of an image or object that can be counted or measured is feature extraction. Finding the bare minimum collection of characteristics necessary to distinguish one item from another is the aim of feature extraction. The picture feature is automatically extracted by the CNN model. However, other characteristics from the images must be extracted by other machine learning models. The vertical and horizontal contrast, correlation, energy, homogeneity characteristics, and other features have all been retrieved. Table 5 provides a description of these manual features.

Table 5: Show the manual feature description.

Features Name	Features Description
average red color	The average value of all red pixels on the soil surface.
average green color	The average value of all green pixels on the soil surface.
average blue color	The average value of all blue pixels on the soil surface.
average hue color	The average value of all hue pixels in HSV soil surface.
average saturation color	The average value of all saturation pixels in HSV soil surface.
average values color	The average value of all values pixels in HSV soil surface.
horizontal & vertical contrast	Find the local variations in the gray-level co-occurrence matrix.
horizontal & vertical correlation	Find the joint probability occurrence of the specified pixel pairs.
Horizontal & vertical energy	Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.
Horizontal & vertical homogeneity	Find the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

3.5 Machine Learning Models

In a nutshell, machine learning is the process that enables accurate predictions from computers. Over time, it can reduce errors by learning from its past transgressions [19]. In order to categorize soil moisture in this work, we trained five machine learning algorithms. They are DT, KNN, SVM, CNN, and ANN. The following subsections provide a brief description of these algorithms.

3.5.1 Artificial Neural Network (ANN): The layers of an artificial neural network are input, hidden, and output. Numerous neurons, also known as nodes, make up each layer. Each neuron has a connecting link that connects it to other neurons. The input is mapped to the output using the activation function. Each layer's output passes through an activation function to create the layer's output, which then passes through the following node to create the network's overall output [20]. The mean square error (MSE) between the network outputs and the matching target values is trained into the ANN repeatedly.

3.5.2 Decision Tree (DT): Each node in a decision tree has two or more branches. It's a supervised machine learning approach for categorization issues. The feature or characteristics are represented by an internal node, and the decision rule is represented by a branch. The outcome is represented by each leaf node. The root node is represented as the topmost node in a decision tree [21]. The most prevalent advantage of DT is its ease of understanding and interpretation. The downside of DT is that it is prone to overfitting and instability.

3.5.3 K-Nearest Neighbor(KNN): The simplest ML learning method for classification issues is KNN. A supervised ML algorithm is another name for it. When using KNN, the value of K provides the number of nearest neighbors to take into account when categorizing a new data point. It is crucial to choose the proper value for K since doing so will increase the mode [22]l's accuracy. Faster computation times, a straightforward method, and more accuracy are benefits of KNN. But the quality of the data affects accuracy. Additionally, because all of the training data must be stored, it may be computationally costly.

3.5.4 Support Vector Machine (SVM): A supervised ML approach called a support vector machine may be applied to classification and regression issues. But categorization is the main application. Finding the hyperplane that separates two or more classes of data is the aim of SVM. If a hyperplane separates two classes as closely as possible, it is said to be optimum. This supervised classifier performs quite well. The main benefit of SVM is that it performs effectively in high-dimensional areas and with a distinct margin of separation. SVM has the drawback of not functioning properly when the data set is too big. If the data set is noisy, SVM's performance also suffers.

3.5.5 ResNet-50

A convolutional neural network with 50 layers is known as ResNet-50. It is possible to load a pre-trained version of the neural network that has been trained on more than a million images from the ImageNet database. A trained neural network can classify images into 1000 categories, such as keyboard, mouse, pencil, and many animals. Thus, the neural network has learned rich feature representations for a wide range of images. An image input size of 224 by 224 is used for the neural network.

3.6 CNN Setting

Convolutional Neural Network (CNN) is a machine learning technique that improves on classic Artificial Neural Networks (ANN). A CNN is made up of input layers, hidden layers, and an output layer. In this work, we employed a pre-trained ResNet-50 architecture to classify soil dampness. It is a 50-layer deep convolutional neural network [23]. Figure 3 depicts the overall structure of the CNN model visually.

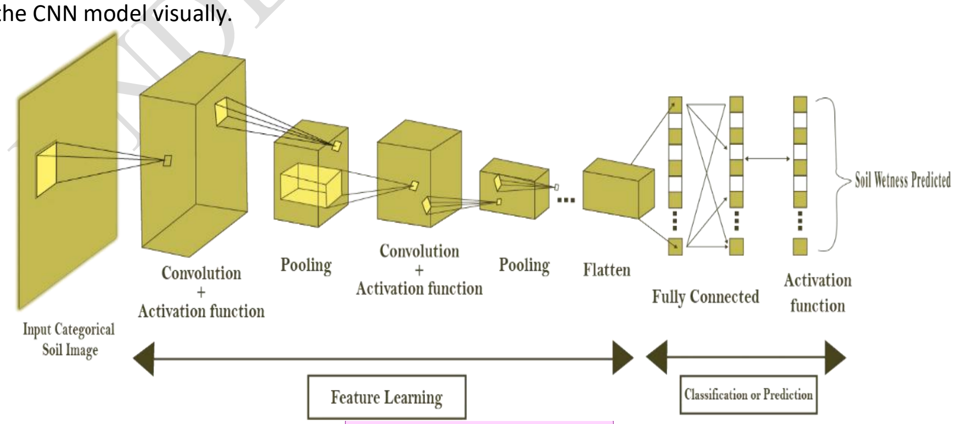


Figure 3: Structure of CNN

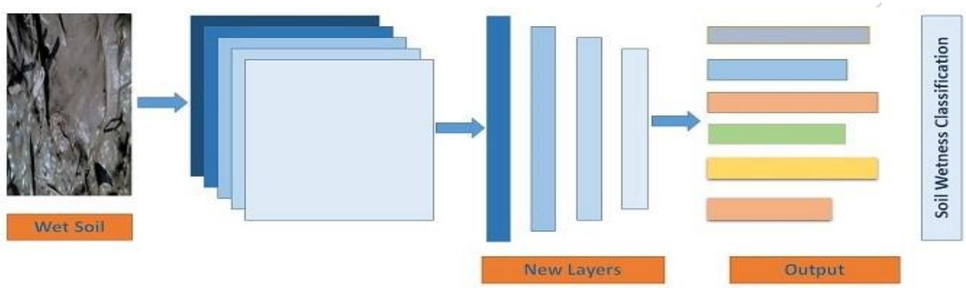
Comment [DB10]: Give space/Keep uniform everywhere from **3.5.1 to 3.5.5**

Comment [DB11]: Mention the source of this Figure

Convolution, pooling, flatten, and fully linked layers of CNN are the most often used layers. Direct feature extraction from the picture is possible. The CNN algorithm uses convolutional layer construction, pooling, and flattening to accomplish automated feature extraction. We have achieved the result in the output layer when the picture feature is entering a collection of dense layers or completely linked layers.

3.7 Soil Wetness Image Classification

After feature extraction has been completed, soil images are used to train machine-learning models based on the extracted features. As soon as the training process is completed, we will test the model using one image from the test dataset, and we will use the confusion matrix to



evaluate

Figure 4: Soil Wetness Classification Phase.

whether or not the model is correct. This figure illustrates the fine-tuning process that took place when the three final layers were replaced by our classification task.

3.8 Evaluate the Classifiers

By contrasting the convolutional neural network model with other machine learning models, we will assess the effectiveness of numerous machine learning classifiers that are employed in our suggested system. Accuracy, precision, recall, and F1-score are the four measures used to evaluate the performance of our suggested method [24]. When categorizing an item into category A or category B,

Accuracy: The ability to identify an object as A or B correctly. Mathematically,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where, TP or True Positive = the number of objects correctly identified as A, FP or False Positive = the number of objects incorrectly identified as A, TN or True Negative= the number of objects correctly identified as B, and FN or False Negative = the number of objects incorrectly identified as B.

Precision: The ratio of correctly predicted positive observations to the total positive observations. It also represents the exactness of the model. Mathematically,

$$precision = \frac{TP}{TP + FP}$$

Recall The ability to determine the A object correctly. It also represents the completeness of the model. Mathematically,

$$Recall = \frac{TP}{TP + FN}$$

F-Score: F-Score represents the harmonic mean precision and recall. Mathematically,

$$FScore = \frac{2 * (recall * precision)}{recall + precision}$$

6. Result Analysis

We reported the findings of five machine learning models for soil categorization in this part. These are the ANN, DT, KNN, SVM, and CNN models. We used 80% of the images for training and 20% for testing in each trial. The images were chosen at random from various soil moisture levels. This accomplishment is now described more below. For cross-validation, we used a 10-fold test, where one fold implies randomly picking 80% images for training and 20% images for testing. Table 6 summarizes the findings.

Table 6:Performance measures (%) for CNN classifier with other machine learning classifiers.

Fold	Accuracy					Precision					Recall					F-Score				
	CNN	SVM	KNN	DT	ANN	CNN	SVM	KNN	DT	ANN	CNN	SVM	KNN	DT	ANN	CNN	SVM	KNN	DT	ANN
1	98.1	79.0	71.6	83.4	70.6	98.1	79.0	71.6	83.4	70.6	98.1	79.0	71.6	83.4	70.6	98.3	79.1	71.6	83.4	70.6
2	97.4	79.1	69.7	83.1	68.6	97.4	79.1	69.7	83.1	68.5	97.4	79.1	69.7	83.1	68.6	97.4	79.1	69.7	83.1	68.6
3	97.1	80.6	71.4	83.0	70.6	97.1	80.7	71.5	83.0	70.6	97.1	80.6	71.4	83.0	70.6	97.1	80.6	71.4	83.0	70.6
4	98.0	79.7	74.0	82.4	69.7	98.1	79.7	74.0	82.4	69.7	98.0	79.6	74.1	82.3	69.7	98.0	79.7	74.0	82.4	69.7
5	98.0	78.9	73.9	85.3	70.9	98.0	79.0	73.9	85.3	70.9	98.0	78.9	73.9	85.3	70.9	98.1	78.9	73.9	85.2	70.9
6	99.0	79.6	72.6	84.9	71.0	99.1	79.6	72.6	84.9	71.1	100	79.6	72.6	84.9	71.1	99.6	79.6	72.6	84.9	71.0
7	98.6	79.6	75.7	84.6	69.4	98.6	79.6	75.7	84.6	69.4	98.6	79.6	75.7	84.6	69.4	98.6	79.6	75.8	84.6	69.4
8	96.7	77.0	71.9	83.1	67.1	96.7	77.0	71.9	83.1	67.0	96.7	77.0	72.0	83.1	67.5	96.7	77.0	71.9	83.0	67.1
9	97.4	79.3	73.1	81.1	69.4	97.4	79.3	73.2	81.1	69.4	97.5	79.4	73.1	81.1	69.4	97.4	79.2	73.1	81.1	69.6
10	97.4	79.6	71.4	83.1	69.9	97.5	79.7	71.8	83.1	69.9	97.6	79.6	71.4	83.0	69.9	97.3	79.6	71.8	83.1	70.9
Avg.	**97.7**	**79.2**	**72.5**	**83.4**	**69.7**	**97.8**	**79.3**	**72.6**	**8.34**	**69.7**	**97.9**	**79.2**	**72.5**	**83.8**	**69.8**	**97.8**	**79.2**	**72.6**	**83.3**	**69.8**

Table 6 shows that, when compared to other machine learning models, the pre-trained CNN model has the greatest average accuracy (97.7%). In contrast, the CNN model performs better than ANN, DT, KNN, and SVM. Additionally, figure 5 shows the confusion matrix for the pre-trained CNN model, which has the top performance rating based on the performance metrics. To assess how well the classification model performed on the testing data, the CNN confusion matrix is created here. The total number of accurate classifications divided by the total number of categories yields the classification accuracy. The greatest number of images are accurately classified using the CNN model. We can observe that diagonally colored boxes show the CNN model's accuracy percentage.

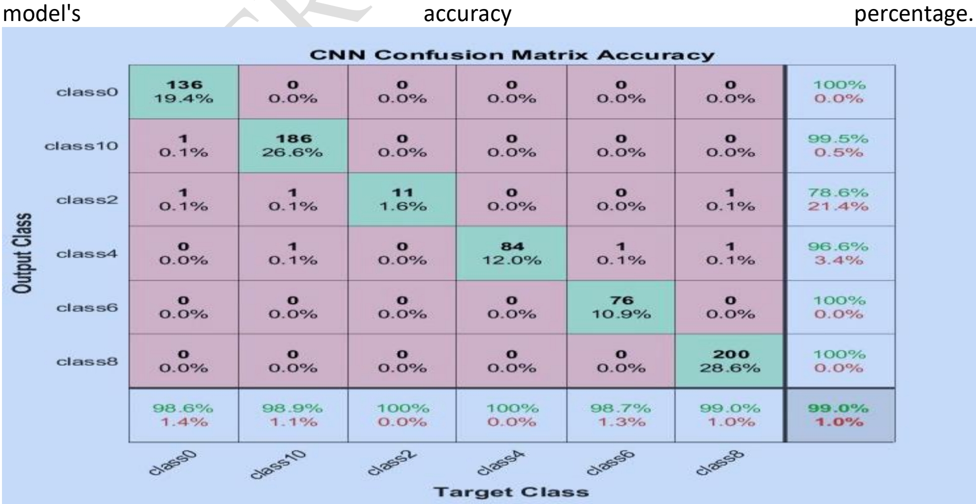


Figure 5: Display the Confusion Matrix of the CNN model

CNN's average accuracy is judged to be 99.0%. We can observe in this classification that (class0, class6, class8), and class2 have the greatest and lowest classification accuracy, accordingly.

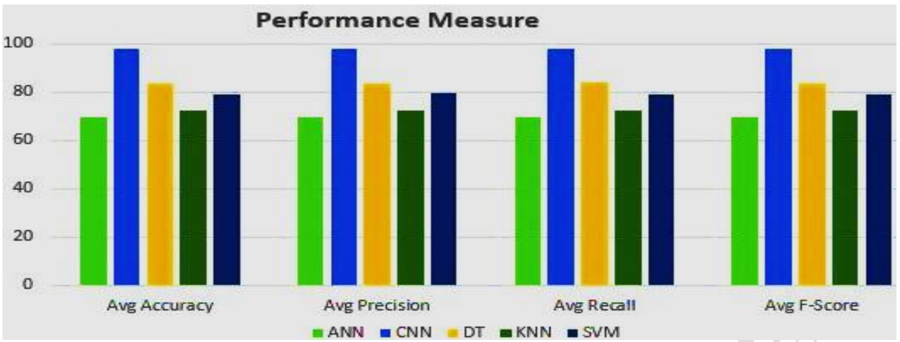


Figure 6: Average Performance Measures for each Classifier

It was then decided to determine the performance measures for each classifier, and a comparison between CNN, SVM, KNN, DT, and ANN was performed. Figure 6 shows the best outcome that was achieved using CNN, SVM, KNN, DT, and ANN.

7. Conclusion and Future Work

Production of crops hugely depends on the water in the soil. At present, machine learning algorithms are used in the agriculture sector for smart farming and agro-industry. In this paper, we have applied different machine learning algorithms for classifying soil wetness levels and found that CNN performs the best if the number of images is sufficiently large. In agriculture, there are many fields where machine learning models can be applied for prediction, classification, and automation. For example, our work can be extended by using images from satellites to measure the overall dryness of soil or flooded areas in different regions. This work can also be extended with the use of GIS for mapping soil wetness in different parts of Bangladesh. Individual crop-based decision support systems can also be developed in the future.

References

Comment [DB12]: References are not in proper format so correct all the references as per this Journal norms/APA style whichever is applicable

[1] D. R. H. Machbah Uddin, "A novel feature based algorithm for soil type classification," Complex & Intelligent Systems, vol. 8, no. 4, pp. 3377-3393, 2022.

[2] Y. Cal, "Soil classification by neural network," Advances in Engineering Software, vol. 22, no. 2, pp. 95-97, 1995.

[3] H. J. a. W. R. Cotton, "Soil moisture estimation using an artificial neural network," Canadian Journal of Remote Sensing, vol. 30, no. 5, pp. 827-839, 2004.

[4] A. K. H. S. Sajjad Ahmad, "Estimating soil moisture using remote sensing data," Advances in Water Resources, vol. 33, no. 1, pp. 69-80, 2010.

[5] T. L. P. L. L. Xiang Gao, "A soil moisture classification model based on SVM used in agricultural WSN," 2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, ITAIC 2014, pp. 432-436, 2014.

[6] F. G. ,. J. S. ,. M. N. ,. C. N. Iftikhar Ali, "Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data," Journal of Remote Sensing , vol. 7, no. 15, pp. 16398-16421, 2015.

[7] J. U. A. G. ,. M. M. B. Ashwini Rao, "Machine Learning in Soil Classification and Crop Detection," International Journal of Scientific Research & Development, vol. 4, no. 01, pp. 792-794, 2016.

[8] S. Sneha Pethkar, "Review on Soil Classification Methods," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 2, no. 6, pp. 8617-8623, 1999.

[9] O. M. Z. Halimatu Sadiyah. Abdullahi, "Advances of image processing in Precision," Journal of Multidisciplinary Engineering Science and Technology, vol. 4, no. 8, pp. 7981-7987, 2017.

[10] R. T. Chandan, "An Intelligent Model for Indian Soil Classification using various Machine Learning Techniques," International Journal of Computational Engineering Research , vol. 8, no. 9, pp. 33-41, 2018.

[11] V. D. Priyanka Dewangan, "Soil Classification Using Image Processing and Modified SVM Classifier," International Journal of Trend in Scientific Research and Development, vol. 2, no. 6, pp. 504-507, 2018.

[12] L. X. B. Y. Zhongzheng Hu, "Soil moisture retrieval using convolutional neural networks: Application to passive microwave remote sensing," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 20, no. 3, pp. 583-586, 2018.

[13] P. B. D. M. P. a. D. B. Konstantinos G. Liakos, "Review of Machine Learning in Agriculture," Sensors (Switzerland), vol. 18, no. 8, pp. 1-29, 2018.

[14] S. K. F. M. Riese, "SOIL TEXTURE CLASSIFICATION with 1D CONVOLUTIONAL NEURAL NETWORKS BASED on HYPERSPECTRAL DATA," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 4, no. 2, pp. 615-621, 2019.

[15] A. D. M. Belete Biazen Bezabeh, "The effects of multiple layers feed-forward neural network transfer function in digital based ethiopian soil classification and moisture prediction," International Journal of Electrical and Computer Engineering, vol. Vol. 10, no. 4, pp. 4073-4079, 2020.

[16] E. T. M. ,. M. R. M. Chee Soon Lim, "Machine Learning Classifiers for Modeling Soil Characteristics by Geophysics Investigations," Applied Science(MDPI), vol. 10, p. 5734;, 2020.

[17] Nuno, "Machine Learning methods for classification of AcidSulfate soils in Virolahti," Arcada University of Applied Science, 2020.

[18] E. Pekel, "Estimation of soil moisture using decision tree regression," tical and Applied Climatology, vol. 139, no. 3-4, pp. 1111-1119, 2020.

[19] Ö. M. Baştanlar Yalin, "Introduction to Machine Learning," Methods in Molecular Biology, vol. 1107, no. 105-128, 2014.

[20] J. J. F. J. F. Mas, "The application of artificial neural networks to the analysis of remotely sensed data," International Journal of Remote Sensing, vol. 29, no. 3, pp. 617-663, 2008.

[21] J. A. K. Anamul Haq Sajib, "A New Algorithm for Classification Based on K Nearest Neighbor and Decision Tree," Dhaka University Journal of Science, vol. 61, no. 2, pp. 179-183, 2013.

[22] B.-S. K. JINHO KIM, "Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines," Applied Mathematics in Electrical and Computer Engineering, pp. 133-138, 2012.

[23] X. L. Z. L. anyang Zhu, "Deep learning for smart agriculture: Concepts, tools, applications, and opportunities," International Journal of Agricultural and Biological Engineering, vol. 11, no. 4, pp. 21-28, 2018.

[24] C. E. G.-T. Valeria Maeda-Gutiérrez, "Comparison of convolutional neural network architectures for classification of tomato plant diseases," Applied Sciences (Switzerland), vol. 10, no. 4, 2020.