Soil pH Determination Using Mobile Phone Captured Image

Gangasagar HL[1], Jovin Dsouza[2], Bhagyashree B Yargal[3], Arun Kumar SV[4], Anuradha Badage[5]

[1,2,3,4] *UG Student, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore, Karnataka, India*

[5] *Assistant Professor, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore, Karnataka, India*

---***---

Abstract - *pH is used for the determining the concentration of acidity or alkalinity in the soil along with the presence of the essential nutrients that the soil bares within themselves. It starts at 0 and goes all the way up to the value 14, where 0 states that the soil is extremely acidic, 14 stating the soil is extremely alkaline and 7 representing the soil is neutral, which is an ideal pH values for the growth of plants and crops. Many horticultural crops grow in soils having a pH of 6 which is slightly acidic and 7.5 which is considered to be slightly alkaline, but there are some which can grow in extreme conditions as well. Initially, multiple papers were referred to know how the pH can be determined using digital image processing, but later found to be just using RGB colours, and their relation to find the pH of the soil using images.*

Key Words: Soil pH, Image Processing, Digital Image, Deep Learning (ANN), Machine Learning

1.INTRODUCTION

Agriculture is a traditional occupation that has been in practice since time immemorial. The concept of agriculture is not only gaining prominence in the Indian farming community, but also among the city dwellers who are vastly venturing into this field, thus transforming the way humans survive. The economy of a particular area was indirectly dependent on agriculture and was a major thrust behind the industrial revolution.

Soil, like air and water, is a primitive natural resource supporting a variety of ecosystem goods and services to the benefit of the mankind. Soil is the loose surface material comprising a mixture of rock debris and organic materials, providing the structural support to plants used in agriculture further being their source of water and nutrients. Processes such as leaching, weathering and microbial activity combine to make a whole range of different soil types with each type having particular strengths and weaknesses for agricultural production.

Soil was recognized anciently, but the importance of conservation and enhancement of ecosystem services has been realized only in the recent past. Soils can be naturally alkaline or acidic, and inheriting the correct pH is crucial for growth of healthy plants and can be measured by testing their pH value. Awareness of the long-term effects of different soil management practices and taking extensive care of soil quality assume great significance for ensuring agricultural sustainability which is inevitable to feed the burgeoning population.

2. LITERATURE REVIEW

In the paper [1] titled "**Determination of Soil pH by using Digital Image Processing Technique-A Review**", the authors captured a digital image, resized the image and removed all the unwanted noises. The image is then portioned into multiple parts for extraction of meaningful region of interest. All the necessary features were extracted from the captured image and classification was applied where each pattern is classified into one of the distinct classes making use of KNN, PCA, ANN, SVM algorithm based on the requirement.

In the paper [2] titled "**Determination of Soil Ph and Nutrient Using Image Processing**", states that the user must first upload an image and a histogram from various values ranges is plotted. Using this histogram, the digital values (RGB) is found and the pH index is calculated followed by which the colour composition is found using Bayer's Filter. This pH index value is then compared with the training and testing dataset images after which the result is predicted.

In the paper [3] titled "**Analysis of Agricultural soil pH using Digital Image Processing**", the authors captured the soil images collected from various parts and carried out image processing, from which the main features were extracted. PCA being a feature-based classification technique serves best for this type of prediction. The RGB values obtained was then multiplied by a constant weighted average value and the mean value obtained was concluded to be the pH of that soil. The captured image was they compared with the images collected in the database and the category to which the soil belongs is given as the result.

In the paper [4] titled "**Determine the Ph. of Soil by Using Neural Network Based on Soil's Colour**", the sample of soil were taken from many lands and its RGB composition was calculated. The authors then made use of the data that were already collected by another study which consisted of the pre-calculated RGB values and the corresponding pH value for that soil. The two RGB values of the sample and study were compared to find the value of pH.

In the paper [5] titled "**Determination of Soil Nutrients and pH level using Image Processing and Artificial Neural Network**", the authors mixed the soil samples with the chemicals given with Soil Test Kits (STK) and Rapid Soil Testing (RST). This sample was put in test tubes and placed in a controlled light module box for capturing of images. The images were then cropped to the desired region of interest, applied with thresholding and masking, and given as input into backpropagation neural network to enhance the accuracy of the results and the performance was measured in terms of mean square error.

In the paper [6] titled "**Evaluation of pH Value by Using Image Processing**", the authors carried out neutralization of HCl using NaOH and phenolphthalein indicator. ImageJ program was used to identify the colour change occurring due to titration in the solution. The average colour values of the digital image were determined for each channel of the RGB colour space followed by which the percentage of red colour was computed and correlated with the existing value of pH. A linear relationship was obtained when plotting pH-red value percent graph using Excel software. Thus, the authors arrived at a conclusion that the plotted graph can be considered as an alternative to predict the pH values.

In the paper [7] titled "**Testing of Agriculture Soil by Digital Image Processing**", soil samples were captured from ten different regions of Nagpur and its soil pH values was stored in a database after calculation using a pH meter. The images of the soil were then clicked, uploaded into a customized MATLAB software and its soil pH index was calculated using the formula: [AVG G/ AVG B/ AVG R]. Since there were only few samples, the authors added +0.01 and subtracted -0.01 to the appropriate results. The newly obtained value was compared with the collected values in the database and the value within a nearby range was returned as the pH of the soil.

In the paper [8] titled "**Determination of Soil pH using Digital Image Processing**", the authors provided the image file of soil sample as input and then RGB value of each image of soil sample was calculated by using digital image processing techniques and then compared with the data stored in database, thus giving soil pH factor and the according to which the soil pH was given as output along with the constituents of soil.

In the paper [9] titled "**Soil pH parameter Estimation Using Image Processing and Regression Analysis**", the soil samples were collected from soil fields and three images of the sample were taken and pre-processing of the colour images was done to remove the noise. The various colour features like RGB, Lab, HSV and GLCM were extracted and correlation between the calculated features and the pH values were found out and thus the best regression analysis with low MSE and high R^2 value was computed.

In the paper [10] titled "**Predicting Soil Texture Using Image Analysis**", the soil samples were dried at 45° C for 48 hours then grinded to less than 2mm using a soil grinder and collected in labelled bags. Images of the soil were captured and converted into HSV and grayscale formats respectively followed by image masking which unfolded the R,G,B,H,S,V colours separately which helped in plotting a colour histogram and the values were recorded in a data matrix. The percentage of silt was enumerated by subtracting the composition of clay with that of sand which was then feed into MIA model which resulted the textural composition of the soil using which the pH was determined.

3. RELATED WORKS

3.1 pH Recording

Computing the pH requires two 15mL Borosil test tubes, black box, pH meter, 12V LED strips, 10K potentiometer and IRF540N Mosfet.

We gathered soils from various parts of Karnataka in small plastic bags and labelled them with the area they were procured from. Once the soils were collected, the next main step was to find the pH of the soils. 2 spoons (approx.) of the soil sample were taken in each of the test tubes and ¾ th of the test tube was filled with water. The test was shaken well until the soil gets dissolved in water. The pH Meter was then immersed in both the test tubes until the tip of the pH Meter was fully touching the bottom. After the computation of each soil pH, the test tubes were thoroughly washed and cleaned to prevent the miscalculation of the pH values.

Fig -1: Soil pH Meter

3.2 Black Box

A completely sealed black box with the optionality to vary the light intensity was built. The main purpose of this black box was to capture clear soil images with varying light intensity. Each of the soil samples were taken in a small container.3 different images of each soil were captured by varying the intensity of light each time.4 mobile phones camera's having different mega-pixels were used to capture these images.

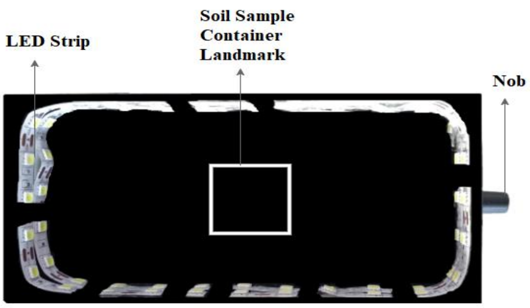


Fig -2: Black Box

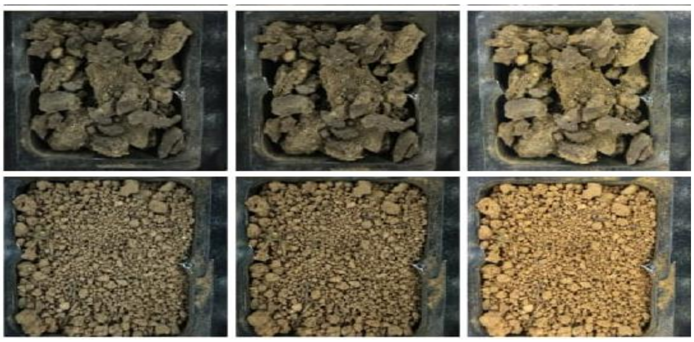
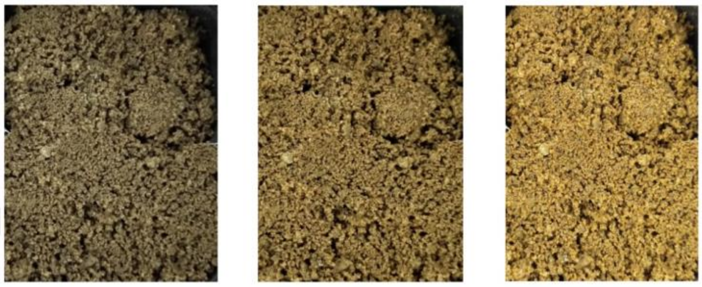


Fig -3: Image Captured using Black Box



Low Intensity Medium Intensity High Intensity

Fig -4: Cropped Images

3.3 Data Processing

The image names were used to extract the pH image, which was measured using pH meter, for the particular soil. The captured image were read through OpenCV python, and converted to many channels such as (RGB,HSV,HLS,LUV,YaCrCb,LAB). These channels were separated and mean of each and every channel was extracted and were held in DataFrame and saved to .csv format.

There are three different ways to extract information from the image.

i) Raw: the captured images were directly included without any processing.

ii) Processed:

Step 1: Spatial Adaptive histogram Equalization

This technique is used to improve contrast in images. It differs from ordinary histogram equalization in the respect that the adaptive method computes several histograms, each corresponding to a distinct section of the image, and uses them to redistribute the lightness values of the image. It is therefore suitable for improving the local contrast.

Step 2: Inpaint
It is used for image restoration, so that any image artifacts can be removed.

Step 3: Wavelet Transform
The image being captured in the form of signal and the presence of noise is a common scenario. To remove this unwanted noise, we make use of wavelet transform. There are 2 types of transforms namely, Bayes and Visushrink, here Visushrink has been implemented.

iii) Combination
The data obtained from raw and pre-processed technique to produce a single large data file.

3.4 Choosing the Model

In order to compute the pH values from the soil, we can use two methods:

i) Deep Learning: On each of the extracted dataset, more than 1200 ANN models were built and trials were conducted to adopt the model with the highest R2 score. After multiple experiments, the highest r2 score obtained for combined was 0.2, which was not considered to be a good one.

ii) Machine Learning: Many regression techniques such as Linear regression, Isotonic Regression, SVR, Gradient Boosting Regressor, KN Regressor, Decision Tree Regressor and Random Forest Regressor.

4. EXPIREMENTS

The dataset which was prepared from the previous steps was included to make experiments, based on which the model for final deployment is used. Screen shots of this can be seen below.

1. No Scalar's in action
This shows that the data was not scaled before it was fitted into the model.

```
No Scaler is in action

[INFO...] RAW Data
(500, 9) (500,) (56, 9) (56,)
1 .  Done.... with r2 score -0.16987800608859294
2 .  Done.... with r2 score -0.12322464517048526
3 .  Done.... with r2 score -0.20235624305884237
4 .  Done.... with r2 score -0.2887383662580525
5 .  Done.... with r2 score -0.22874471575856536
6 .  Done.... with r2 score -0.5227808101120366
7 .  Done.... with r2 score -2.0152255099109455
The model with count  2   with r2score:  -0.12322464517048526

[INFO...] Combined Data
(1000, 9) (1000,) (112, 9) (112,)
1 .  Done.... with r2 score 0.32654715406483474
2 .  Done.... with r2 score 0.29400812572995454
3 .  Done.... with r2 score 0.34004039399831036
4 .  Done.... with r2 score 0.399774048831726
5 .  Done.... with r2 score 0.41132921227555002
6 .  Done.... with r2 score 0.02183832036090072
7 .  Done.... with r2 score -0.2973815538167705
The model with count  5   with r2score:  0.41132921227555002

[INFO...] Brightness Removed Data
(500, 9) (500,) (56, 9) (56,)
1 .  Done.... with r2 score 0.26975082833101609
2 .  Done.... with r2 score 0.2524670751145041
3 .  Done.... with r2 score 0.005818834899547154
4 .  Done.... with r2 score 0.30900257022214965
5 .  Done.... with r2 score 0.32815771545576415
6 .  Done.... with r2 score 0.28830654327001075
7 .  Done.... with r2 score 1.0798830979055034
The model with count  5   with r2score:  0.32815771545576415
```

Fig -6: Results without Scalar's

2. Standard Scalar
Data was scaled applying standard scaler and then fitted into the model.

```
StandardScaler's in action

[INFO...] RAW Data
(500, 9) (500, 1) (56, 9) (56, 1)
1 .  Done.... with r2 score -0.16987800608860315
2 .  Done.... with r2 score -0.12300086476448446
3 .  Done.... with r2 score -0.18142240760725992
4 .  Done.... with r2 score -0.2895299354083454
5 .  Done.... with r2 score -0.32981350186727876
6 .  Done.... with r2 score -0.31530020109164036
7 .  Done.... with r2 score -1.8244757253662742
The model with count  2   with r2score:  -0.12300086476448446

[INFO...] Combined Data
(1000, 9) (1000, 1) (112, 9) (112, 1)
1 .  Done.... with r2 score 0.42654715406482486
2 .  Done.... with r2 score 0.38548625844709257
3 .  Done.... with r2 score 0.5301170251245881
4 .  Done.... with r2 score 0.49337022360858103
5 .  Done.... with r2 score 0.51064615173981271
6 .  Done.... with r2 score 0.34652140280259254
7 .  Done.... with r2 score -0.1667132168210355
The model with count  3   with r2score:  0.5301170251245881

[INFO...] Brightness Removed Data
(500, 9) (500, 1) (56, 9) (56, 1)
1 .  Done.... with r2 score 0.36975082833102653
2 .  Done.... with r2 score 0.32117947172936885
3 .  Done.... with r2 score 0.32637343708059381
4 .  Done.... with r2 score 0.42026326450236702
5 .  Done.... with r2 score 0.46534945269794898
6 .  Done.... with r2 score 0.44895275207014103
7 .  Done.... with r2 score -0.578145802890079
The model with count  5   with r2score:  0.46534945269794898
```

Fig -7: Results applying Standard Scalar's

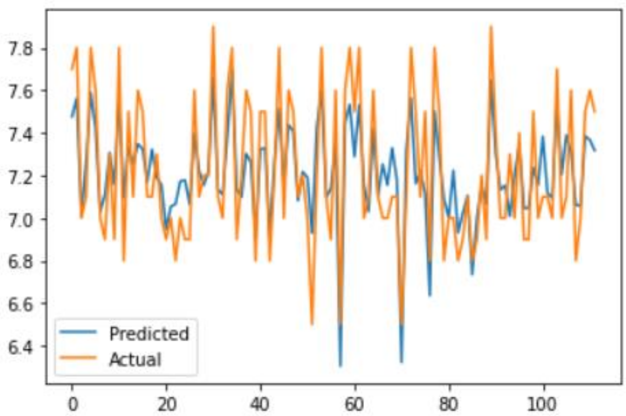
3. MinMax Scalar

Data was scaled using MinMax Scaler.

```
MinMaxScaler's in action

[INFO...] RAW Data
(500, 9) (500, 1) (56, 9) (56, 1)
1 .  Done.... with r2 score 0.36987800608860182
2 .  Done.... with r2 score 0.39730607785929823
3 .  Done.... with r2 score 0.39545603733999162
4 .  Done.... with r2 score 0.2852275362606118
5 .  Done.... with r2 score 0.1198683711577148
6 .  Done.... with r2 score 0.1082217753519092
7 .  Done.... with r2 score -1.4612467681700666
The model with count  2   with r2score:  0.39730607785929823

[INFO...] Combined Data
(1000, 9) (1000, 1) (112, 9) (112, 1)
1 .  Done.... with r2 score 0.52654715406482286
2 .  Done.... with r2 score 0.48668651143774075
3 .  Done.... with r2 score 0.63611169722321101
4 .  Done.... with r2 score 0.6002495267961527
5 .  Done.... with r2 score 0.63884627085344434
6 .  Done.... with r2 score 0.43241396778455452
7 .  Done.... with r2 score -0.101981804366131
The model with count  5   with r2score:  0.63884627085344434

[INFO...] Brightness Removed Data
(500, 9) (500, 1) (56, 9) (56, 1)
1 .  Done.... with r2 score 0.46975082833099933
2 .  Done.... with r2 score 0.425476336800236643
3 .  Done.... with r2 score 0.426034621230911936
4 .  Done.... with r2 score 0.52011014672997933
5 .  Done.... with r2 score 0.59380329253486273
6 .  Done.... with r2 score 0.58377658710829659
7 .  Done.... with r2 score -0.66908589056665
The model with count  5   with r2score:   0.59380329253486273
```

Fig -8: Results applying MinMax Scalar's

5. RESULTS



Graph -1: Results of Actual vs Predicted pH Values

The primary goal was to identify the best model which can be adopted to help in computing the soil ph. The cleaned data was tested on various models to identify which models perform well for the given requirements. On multiple trials over existing models, the best regression model for the prediction of the pH values was found out to be minimax random forest regressor having a r2_score of 0.638. Using this model, we predicted the pH values from the images in virtue of the existing actual values and then plotted a comparison graph thus obtaining an r2_score of 0.686.

6. CONCLUSIONS

Soil taking hundreds of years to form measures the relative amount of acidity or alkalinity supply the essential water, oxygen that is required by plants to grow and flourish and managing the soil pH maximizes the effectiveness of the nutrients in the soil. The contents of the paper suggest a simple but effective method to predict the pH of the soil using digital image processing using which anyone can find the pH for the particular soil of their choice and take measures to improve the pH to yield better growth as per their needs.

REFERENCES

[1] Bhawna J. Chilke, Neha B. Koawale, Divya M. Chandran, Determination of Soil pH by using Digital Image Processing Technique-A Review, International Conference on Recent Trends in Engineering Science and Technology (ICRTEST 2017), Volume: 5 Issue:1 (Special Issue 21-22 January 2017).

[2] Sudha.R, Aarti.S, Anitha.S, Nanthini.K, Determination Of Soil Ph And Nutrient Using Image Processing, International Journal of Computer Trends and Technology (IJCTT) – Special Issue April – 2017.

[3] Gurubasava, Mahantesh S.D, Analysis of Agricultural soil pH using Digital Image Processing, International Journal of Research in Advent Technology, Vol.6, No.8, August 2018.

[4] Makera M Aziz, Dena Rafaa Ahmed, Banar Fareed Ibrahim, Determine the Ph. of Soil by Using Neural Network Based on Soil's Colour, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 6, Issue 11, November 2016.

[5] John Carlo Puno, Edwin Sybingco, Elmer Dadios, Ira Valenzuela, Joel Cuello, Determination of Soil Nutrients and pH level using Image Processing and Artificial Neural Network, IEEE 2017.

[6] A. Beycioglu, B.Comak, D.Akcaabat, Evaluation of pH Value by Using Image Processing, 3rd International Conference on Computational and Experimental Science and Engineering Volume 132, 2016.

[7] Umesh Kamble, Pravin Shingne, Roshan Kankrayane, Shreyas Somkuwar, Prof.Sandip Kamble, Testing of Agriculture Soil by Digital Image Processing, International Journal for Scientific Research & Development, Vol. 5, Issue 01, 2017.

[8] Maneesha G Nair, Neethu Prathapan, Sethulakshmi R, Syamamol T,Determination of Soil pH using Digital Image Processing, International Journal of Information Systems and Computer Sciences, Volume 8, No.2, March - April 2019.

[9] Anup Vibhute, M S Koli, Soil pH parameter Estimation Using Image Processing and Regression Analysis, Journal of Xi'an University of Architecture & Technology, Volume XII, Issue VIII, 2020.

[10] Balakrishna K, Mahesh Rao, Anupama K P, Chaitra B, Pooja L, Automatic testing of Soil Moisture, pH using Arduino and Selection of Specific Crop, International Journal of Computer Sciences and Engineering, Vol.-6, Issue-6, June 2018.

BIOGRAPHIES

ANURADHA BADAGE
Assistant Professor, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore

GANGASAGAR HL
UG Student, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore

JOVIN DSOUZA
UG Student, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore

BHAGYASHREE B YARGAL
UG Student, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore

ARUN KUMAR SV
UG Student, Department of Computer Science Engineering, Sapthagiri College of Engineering, Bangalore