Android-based Soil Series Classifier Using Convolutional Neural Network

Joe G. Lagarteja

Abstract: Classification of soil based on its series is one of the challenging tasks among farmers. Precise classification and identification of soil would lead to an appropriate use of farm land and appropriate crops to grow. Manual classification is tedious and time-consuming process to be done by farmers. This paper introduces a new way of advancing the classification process of soil series using Android phone and Artificial Neural Networks (ANN). The Back Propagation was utilized in this study and it was composed of 3 layers, the Input, Hidden, and Output layer. The network layer contains the image of the soil taken from an Android phone served as an input for the ANN. The middle layer has 3 neurons and the output layer has 16 neurons which has the task of giving the probable series of the soil. For the purpose of training and testing the network model, about 1,000 image test results were collected from Department of Agriculture, Cagayan Valley, Philippines. The network model obtained an accuracy rate of 91% which may be considered as an alternative tool to classify soil according to its series in Cagayan Valley, Philippines.

Index Terms: Android App, Artificial Intelligence, Classification, Convolutional Neural Network

1 INTRODUCTION

Classification is one of the important functions in data mining. Classification techniques have been widely applied to many areas, in agriculture, chemistry, medicine, social studies, and so on [1][2]. Manual method for classification of soil series has been designed to help farmers identify and characterize the soils in the farm [3]. The behavior of different soils when grown to different crops is determined including the inherent limitation of the soils to its different uses. It can help identify appropriate use of the land especially in selecting the most suitable crops to grow [4]. Unfortunately, manual classification is tedious and very time-consuming process to be carried out and confronted with many constraints. Now, Android System in the electronics market is becoming more and more popular. This inspired people to use Android system since it is an open source and plenty of mobile applications are created. This even pondered many individuals on how to carry out the application of a machine learning algorithm on mobile phones for quality advancement. Convolutional Neural Network is one of the machine learning algorithms that is widely used today [8]. CNN have significantly created an impact on image and many pattern recognitions [10]. CNN added some extra advantages which allow it to automatically learn the features from the training data which makes it effective for image recognition and exhibited many practical applications in agriculture, health, and security [12]. In this study, an Android-based image recognition system using Convolutional Neural Networks for classification of series is introduced and its accuracy is evaluated using a Confusion Matrix.

2 MATERIALS AND METHODS

2.1 Data Representation

Figure 1 shows the different steps on how the soil data are converted to train, test, and validate the accuracy of the network model.

• *Joe G. Lagarteja is from Isabela State University, Echague Isabela*

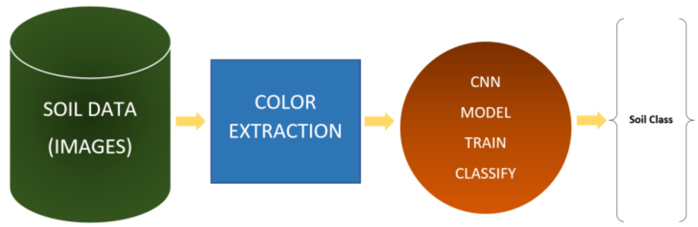

***Fig. 1**. Data Representation.*

The initial step is the color extraction of soil image (dataset) from the database. In order to lessen the effect of noise, filter is applied using Gaussian. The last step is the training of the neural network using back propagation.

2.2 Image Acquisition

The image acquisition system consisted of an Android smartphone Huawei P30 which is powered by a 1.8GHz octa-core HiSilicon Kirin 980 processor that features 2 cores clocked at 2.6GHz, 2 cores clocked at 1.92GHz and 4 cores clocked at 1.8GHz. It comes with 6GB of RAM. The Huawei P30 runs Android Pie and is powered by a 3650mAh battery. Its camera is a Tri-lens camera: 24 MP (Wide Angle Lens, f/1.8 aperture) + 8 MP (Ultra Wide Angle Lens) + 2 MP (Bokeh Lens), supports autofocus.

2.3 Dataset

Data are vital in every study. The dataset used in the study comprises of 1,000 image of soil test results acquired from the Department of Agriculture, Cagayan Valley, Philippines. The dataset was divided into two using random split, 80% for training set data and 20% for validation set data

***Fig. 2**. Soil Samples Taken from the Department of Agriculture, Cagayan Valley, Philippines.*

Figure 2 shows the soil samples collected from the Department of Agriculture, Cagayan Valley, Philippines which are classified into 16 series (shown in Table 1).

TABLE 1
SOIL SERIES

Soil Series	*Number of Samples Collected*
Annam	*79*
Alaminos	*60*
Bago	*55*
Bantog	*69*
Bigaa	*94*
Cuayan	*78*
Faraon	*55*
Guimbalaon	*84*
Ilagan	*49*
Quinga	*58*
Rugao	*61*
San Juan	*46*
San Manuel	*43*
Sibul	*28*
Sta. Rita	*83*
Tagulod	*68*

Table 1 shows the different soil series and the number of samples collected from the Department of Agriculture, Cagayan, Valley Philippines.

2.4 User Design
In this stage, the researcher systematically designed a mobile app layout that would meet the needs of the mobile application. The researcher focused on the simplicity of use of the mobile application.

2.5 THE CONVOLUTIONAL NEURAL NETWORK

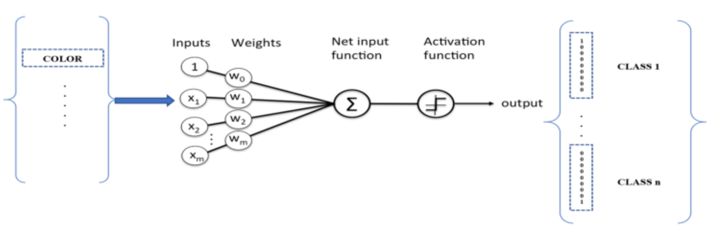

***Fig. 3**. The Convolutional Neural Network*

Artificial Neural Network (ANN) determines the associations between known set of observations and includes two essential methods, training and classification [23]. ANN comprises three interconnected layers namely, the Input Layer which consists of several nodes, Hidden Layer which includes various artificial neurons each representing an activation function, and Output Layer which contains artificial neurons responsible for transmitting a predicted output value [18]. During the training process, the ANN receives the values which correspond to known colors of soil images through the Input layer. The input variables are transmitted to the Hidden layer which computes and extract essential information to predict output values. The weights are defined in each connection through an activation function. The weights are adjusted using a learning function so that the predicted output value matches the known target output for each training point. On the classification process, the network uses the unknown point data for the whole study area and classifies these points using the calibrated weights. Therefore, the network predicts output of the series of the soil. The CNN used in this project was implemented in Google's TensorFlow using MobileNet version 1.0, which provided a high level of abstraction over TensorFlow. The CNN was trained on an ASUS PBH67-V computer that ran in Core i7-2600 CPU 3401 Mhz that took 10-20 minutes. The following script was used in the training process. Since increased samples were used for input, they were highly correlated even if they increased the size of the dataset. This could have led to overfitting of the data. The overfitting problem has been solved by modulating the network's entropic capacity - the amount of information stored in the model that can store much information can be more accurate, but it also saves unnecessary features. In our case, we used a very trivial CNN with few layers and few filters per layer together with data increase and 0.5 dropouts. Dropout also helped reduce overfitting by preventing twice the same pattern from being seen by a layer. The model that stored less information therefore only saved the most important features found in the data and is likely to become more relevant and more widespread.

3 RESULTS AND DISCUSSIONS

TABLE2
PARAMETERS

Neural Network Parameters	*Values*
Total Training Samples	*800*
Total Validation Sample	*200*
Number of Hidden Neuron	*1*
Learning Rate	*0.5*
Training Time (sec)	*38*
Epoch	*100*
Mean Square Error	*0.033*

3.1 Dataset and Neural Network Parameters
Dataset acquired from the Department of Agriculture Cagayan Valley, Philippines was divided into two using random split.
Eighty (80) percent was used to train the neural network and 20% for validating the testing the model.
Learning rate of the neural network was set to 0.5 and 100 epochs were also used to train the network. The network took 38 seconds to be trained and a mean square error of 0.033 was observed.

3.2 Program Interface

***Fig. 4**. Screenshot of the Android app.*

Figure 4 shows a sample output of the developed Android application for soil classification.

3.3 Evaluation

TABLE 3
CLASSIFICATION RESULTS

Soil Series	*Number of Samples*	*Correctly Classified*	*Mis Classified*	*Accuracy*
Annam	*9*	*8*	*1*	*88.89%*
Alaminos	*8*	*8*	*0*	*100%*
Bago	*5*	*5*	*0*	*100%*
Bantog	*9*	*9*	*0*	*100%*
Bigaa	*4*	*2*	*2*	*50%*
Cuayan	*8*	*7*	*1*	*87%*
Faraon	*5*	*5*	*0*	*100%*
Guimbalaon	*4*	*4*	*0*	*100%*
Ilagan	*9*	*8*	*1*	*88%*
Quinga	*8*	*8*	*0*	*100%*
Rugao	*5*	*5*	*0*	*100%*
San Juan	*6*	*6*	*0*	*100%*
San Manuel	*3*	*3*	*0*	*100%*
Sibul	*8*	*8*	*0*	*100%*
Sta. Rita	*3*	*3*	*0*	*100%*
Tagulod	*6*	*6*	*0*	*100%*

In order to test and validate the accuracy of the network model, actual soil data were fed into the program and an Accuracy Matrix was constructed for the results (Figure 4). Basing on classified soil from 100 sites in Cagayan Valley, Philippines it was discovered that there were only five cases of misclassifications. This points to an overall accuracy rate of 91%. Close analysis of the soil data point to the highly comparative values between the attributes of the soil types being mistaken as the reason for the confusion and misclassifications.

4 CONCLUSION

This paper has shown the application of an Artificial Neural Network for soil classification using a program developed for this purpose. Using actual soil data, it provided an accurate way to classify soils according to series with accuracy level of 95%. Such high accuracy together with the program developed may provide an important tool for farmers and government agriculture personnel for soil classification.

5 ACKNOWLEDGMENT

The authors wish to thank Isabela State University for the support of this study.

6 REFERENCES

[1] And, N. C., & Deshmukh H R. (2012). ANDROID OPERATING SYSTEM. Software Engineering. https://doi.org/http://en.wikipedia.org/wiki/Android_(operating_system)

[2] Azizi, A., Abbaspour-Gilandeh, Y., Nooshyar, M., & Afkari-Sayah, A. (2015). Identifying Potato Varieties Using Machine Vision and Artificial Neural Networks. International Journal of Food Properties, 19(3), 618–635. https://doi.org/10.1080/10942912.2015.1038834

[3] Bennett, J. K., Wheatley, J. K., & Walton, K. N. (1984). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning arXiv:1602.07261v2. Journal of Urology. https://doi.org/10.1007/s10236-015-0809-y

[4] Beynon-Davies, P., Carne, C., Mackay, H., & Tudhope, D. (1999). Rapid application development (RAD): An empirical review. European Journal of Information Systems, 8(3), 211–223. https://doi.org/10.1057/palgrave.ejis.3000325

[5] Developers, A. (2014). Android, the world's most popular mobile platform. Android, the World's Most Popular Mobile Platform | Android Developers.

[6] Eggert, C., Brehm, S., Winschel, A., Zecha, D., & Lienhart, R. (2017). A closer look: Small object detection in faster R-CNN. In Proceedings - IEEE International Conference on Multimedia and Expo. https://doi.org/10.1109/ICME.2017.8019550

[7] Fei-Fei Li, J. J. and S. Y. (2016). CS 231N Course Slides and Notes", Stanford University. Retrieved from http://cs231n.stanford.edu/2017/syllabus

[8] Felt, A., Chin, E., & Hanna, S. (2011). Android permissions demystified. In Proceedings of the 18th ACM conference on Computer and communications security - CCS '11 (2011). https://doi.org/10.1145/2046707.2046779

[9] Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision. https://doi.org/10.1109/ICCV.2015.169

[10] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/TPAMI.2016.2577031

[11] Sakr, G. E., Mokbel, M., Darwich, A., Khneisser, M. N., & Hadi, A. (2016). Comparing deep learning and support vector machines for autonomous waste sorting. In 2016 IEEE International Multidisciplinary Conference on Engineering Technology, IMCET 2016. https://doi.org/10.1109/IMCET.2016.7777453

[12] Sermanet, P., & LeCun, Y. (2011). Traffic sign recognition with multi-scale Convolutional Networks. In Neural Networks (IJCNN), The 2011 International Joint Conference on. https://doi.org/10.1109/IJCNN.2011.6033589

[13] T.Vizhanyo, J. F. (2000). Enhancing color differences in image of diseased mushrooms,". Computer and Electronics in Agriculture, 26, 187–198.

[14] Tan, X., & Triggs, B. (2010). Recognition Under Difficult Lighting Conditions. IEEE Transactions on Image Processing(TIP). https://doi.org/10.1109/TIP.2010.2042645

[15] Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations. Health Care and Life Sciences. https://doi.org/10.1080/14759551.2011.530745

[16] Bhargavi, P. and Jyothi, S. 2011. Soil Classification Using Data Mining Techniques : A Comparative Study.

[17] Mutalib, S. and Jamian, S. 2010. Soil Classification : An Application of Self Organising Map and k -means, pp. 439–444.

[18] Neto, A. B., Bonini, C. S., Bisi, B. S., Colleta, L. F. S., and Reis, A. R. 2017. Artificial Neural Network for Classification and Analysis of Degraded Soils, vol. 15, no. 3, pp. 503–509.

[19] Han, P., Dong, D., Zhao, X., Jiao, L., and Lang, Y. 2016. A smartphone-based soil color sensor: For soil type classification. Computer Electronic Agriculture.

vol. 123, pp. 232–241.

[20] Beucher, A., Moller, A. B., and Greve, M. H. 2017. Artificial neural networks and decision tree classification for predicting soil drainage classes in Denmark. Geoderma, no. May, pp. 0–1.

[21] Bondarenko, A., Aleksejeva, L., Jumutc, V., and Borisov, A. 2016. Classification Tree Extraction from Trained Artificial Neural Networks. Procedia Comput. Sci., vol. 104, pp. 556–563.

[22] Reale, C., Gavin, K., Libric, L., and Jurić-Kaćunić, D. 2018. Automatic classification of fine grained soils using CPT measurements and Artificial Neural Networks. Advanced Engineering Informatics. pp. 207–215.

[23] Sitton, J. D., Zeinali, Y., and Story, B. A. 2017. Rapid soil classification using artificial neural networks for use in constructing compressed earth blocks. vol. 138, pp. 214–221.