ACTA SCIENTIFIC AGRICULTURE (ISSN: 2581-365X)

Volume 3 Issue 5 May 2019

Mini Review

A Short Review on Agriculture Based on Machine Learning and Image Processing

Karisma Mohiuddin[1]* and Mirza Mohtashim Alam[2]

[1]*Computer Science, BRAC University, Bangladesh*

[2]*Department of Multimedia and Creative Technology, Daffodil International University, Bangladesh*

***Corresponding Author:** Karishma Mohiuddin, Post-graduate Student, Computer Science, BRAC University, Bangladesh.

Received: March 28, 2019; **Published:** April 16, 2019

Abstract

In this paper, some research papers are highlighted on the basis of the holistic approach and the novelty of works. These papers are mainly focused the on the approaches of the art of machine learning and image processing. Some of the authors on selective papers used pictures as dataset and some others used the chemical substances as raw dataset to analyze their models by applying the state of the art of machine learning. Most authors are comfortable to use machine learning algorithm such as Support Vector Machine (SVM), and Artificial Neural Network (ANN) etc. to conduct research on the crop science. Again, the authors are applying image processing method on the pictures of crop, seed, flower, leaf etc to collect the visual information. Hence, these reviewed papers signify the implementation of modern technology, especially the artificial intelligence on the agricultural aspect to enhance the food growth system, to create a subtle environment of crop, and to modify the plant cultivation to produce more crops from less resource.

Keywords: SVM; ANN; Machine Learning; Image Processing

Introduction

Agriculture field is a prominent area to implement scientific methods and to conduct research on various aspects of it. Recently, researchers are conducting and adapting technology to implement on crop cultivation, weather prediction, yield prediction, insect detection, fertilizer detection, distinguish weed from the crops, detecting the absolute environment to grow plants perfectly. Computer science and its distinctive subjective knowledge especially machine learning, artificial intelligence, deep learning, image processing, computer vision, robotics, automated machineries contribute a major role in the field of agricultural research. Moreover, they play a massive role in agriculture economy. Researchers are using Machine Learning and deep learning approach in agriculture to predict the growth of harvest, to detect the disease beforehand, to suggest the fertilizer for specific crop and so on. The remarkable and promised works are captured in this short review where possibilities of new inventions highlighted prominently with the help of machine learning and image processing techniques.

Notable works

Crop yield is important part of agriculture output, as it refers to the growth of crops per unit area of land cultivation and the seed generation of the plant itself. Before using prediction method via artificial neural network it was merely impossible to determine beforehand that how much crop would grow per unit of land. Crop yield depends on natural phenomena such as climate and weather changes, soil quality, rainfall, sun exposure, water, fertilizer etc. The authors in their paper [1] predicted crop yield on the basis of fertilizer. By using the artificial neural network, they have predicted the best fertilizer and suggested best fitted crop according to soil features for crop yield purposes. Their system model is given below-

Figure 1: Design flow of the model [1].

To prepare the ANN model they used 9 types of crops such as cotton, sugarcane, jowar, bajra, soy beans, corn, rice wheat, groundnut for training the algorithm. There are seven parameters for this algorithm, they are -pH, N (Nitrogen), P (Phosphorus), K (Potassium), depth, rainfall, and temperature. The authors used MATLAB to build the prediction model. By using their prediction model they have come to two different decisions. They could predict the suggested crop according the level of parameters. In this case, they demonstrated sugarcane as suggested crop on the basis of those distinct seven parameters.

Figure 2: Result of predicted crop- Sugarcane [1].

Again, the authors also predicted the amount of fertilizers required for cultivation of the crop. In this scenario, if parameters N. P. and K were in higher value then no fertilizers was needed but if the values were lower than the actual measurement then it showcased the necessary fertilizers for crop cultivation.

Figure 3: The prediction of using fertilizers [1].

This system is very important for the farmers who are unable to test their soil for cultivation. Yet, the technique is really handy but the authors did not showcase the accuracy of the model. Moreover, it only applied on only 9 crops; they did not apply on the big data to learn the accuracy. Also they did not mention the time frame analysis.

Another research focuses on using the clustering algorithm to identify management zones [2]. For this purpose, k-means clustering has been used to recognize optimal number of management zones for cotton field. They have considered Elements of Precision Agriculture which involves sensing, control, calculating prescription through induction and management zones. For study purpose they have used data from a cotton field named “Helena Fertility Trial”. The dependent variables include measures of yield, bales of cotton per acre and biomass flow for the first class representation. Second class represents geo- referenced field topological features. To collect data, the authors used Global Positioning System (GPS). Besides, other physical variables are slope, soil series type, and some operational variables that are part of treatment structure. Their main constraint is they could not do the time series analysis as they did not have temporal information because they had only one year of data collection. Like previous researcher, they did not consider NDVI as an independent variable because it results in low Pearson Coefficient Correlation with yield. On the other hand, yield, bales of cotton per acre and biomass flow are variables since they are informative. In their set of attribute they included field topography. They normalized each variables or attribute between 0 and 1 for removing any kinds of scale biasness. They have further augmented the dataset with weighting factor which is less than 1. Furthermore, K-means was used in augmented data. Segmentation was done for identification of large number of clusters where dissimilarity among attributes was given higher weight than spatial proximity. A second set of data was then created using the clusters which was obtained in the previous step. Finally, they have implemented K-means on the obtained data with smaller number of clusters which is equivalent to 15. The aim was to merge the clusters from previously created small clusters.

In paper [3], authors emphasis on color images of plant leaf to detect early stage of the disease. For this purpose, the authors used Support Vector Machine (SVM) on the visual symptoms of plants where plant diseased areas are showed as spots, stains or strikes. These marks were identified, segmented, pre- processed and features are extracted from each distinctive region of picture. The most added image features are added as set of features for further process and little or no feature pictures were omitted from the feature selection process. A set of 117 cotton crop images was used for implementing the SVM algorithm.

Figure 4: Image of cotton crops showing the visual symptoms of damages caused by: (a) Southern green stink bug (Nezaraviridula); (b) Bacterial angular (Xanthomonascampestris); (c) Ascochyta blight (Ascochytagossypii) [3].

They extracted features by marking the diseased region of leaf's shape, size, dimension, color texture etc. Their categorical feature's table is given below.

Categories	Features	Description	Number of measurements
1	Shape	Solidity Extent Major axis length Minor axis length Eccentricity Centroid Diameter Area	8
2	Texture	d =1, 3, 5	90
3	Fractal dimension	Box's size -2, 4, 6, 8	4
4	Lacunarity	Box' size-2, 4, 6, 8	4
5	Dispersion	RGB	3
6	Grey levels	For RGB and HSV channels	6
7	Grey histogram discrimination	For RGB and HSV channels	6
7	Fourier descriptor		20

Table 1: Candidate features discriminated by categories [3].

To implement SVM algorithm, they separated features into three separate classes beforehand. In 1st class, there would be one feature which was used as input into classifier, in 2nd class, groups of features were used as input in the classifier and lastly, in 3rd class, all features were used by excluding one feature and used as input for classifier. The dataset was divided into seven subclasses from where six sub classes are used as training sets and one was used as testing set. In their result part, they explained that area, diameters, perimeter, Fractal Dimension were not feasible features to detect the disease of the plant as they were giving accuracy less than 50%. For, texture and Lacunarity are giving accuracy below 50%, for some special cases such as distance pixel from one to three and dimension from six to eight respectively for single input, the accuracy is below 50%. When bag of features were applied, the accuracy picked 93.1% where 45 features were combined efficiently. From this paper, the author could apply other algorithms to classify the crop disease and make comparison of the best techniques.

In this research paper [4], the authors worked on plant and weed distinguishing technique via Machine Learning and image processing technique to successfully manage weed separation from the crops. It is a vital research as weed management is heavily important in the field of crop harvesting. As this method aids to reduce the cost of weed management by enhancing crop growth and yield, the authors applied this technique on 3 types of plant. They are canola, corn and radish. They used algorithm called Local Binary Pattern (LBP) for extracting textural features of crop leaf and Support Vector Machine (SVM) to classify the multiclass plant. The dataset is referred as "bccr-segset" which is published online. In this dataset there are four subclasses which refer to the four stage of the growth of crops. The researchers here used 24,000 images for training set and 6000 pictures for validation.

The combination of LBP and SVM demonstrated accuracy of 91% to discriminate plants from the background. It is real time detection as LBP takes short period of time for detecting the crops. Hence, LBP is a promising algorithm to detect weed from the background where other crops are present. In future work the authors wanted to introduce color features and non-uniform patterns method to identify weed more accurately. In this paper, they only introduced grey scaled images for detection. Therefore, to improve their technique the detection of broad leaf is essential and it requires further investigation.

In their paper [5], the author used Machine Learning technique on beans color to determine ten local beans of Iran. The method was accumulated by 3 consecutive steps to detect the bean's varieties. They applied multi-layer perception artificial neural network algorithm on beans' images to identify the exact class of the bean. The three steps of the procedure are imaging, processing and information display unit. Their proposed model is given below.

Figure 5: The Proposed Model (Algorithm of the Coded Program) [5].

In image processing part, the authors used image acquisition, smoothing, image segmentation, feature extraction techniques. They applied Gaussian filter to remove the noise from the image and image segmentation method to capture the ROI (Region of Interest) of the bean. They also applied binary images to mask on color image from which they could extract the color features.

Figure 6: IRGB (a), R-B (b), binary of R-B (c), final image (d) [5].

They applied MLP-ANN on 1000 datasets. From the data set 70% was used for training, 15% for validating and another 15% for testing purposes. In their research, they had used 12 color features of 10 verities of bean and the general sensibility results into 100%, 97.33%, 96% and the specificity of 100%, 97.9%, and 97.1% for training. Overall their technique showcases the accuracy of sensitivity and specificity over 96% and 97% respectively. This is an effective way to identify bean variants yet the author had not described the future scope and the probable use in industry or to detect other variants of crops.

Conclusion

In this paper, a few of the recent works are marked to indicate the advancement of science in the area of agriculture. Besides, the authors are still working on their respective scope to improve their proposed model. They could still compare their model with others on the same area of interest to determine whether their works are more efficient or not. Moreover, most of the papers work on specific plant or the same class of plant. They could expand their dataset and work on big data which is still missing. Again, the researchers only work with single machine learning algorithm to demonstrate their researches. They did not apply multiple machine learning algorithms on their proposed model to see the outcome. This point should be noted to conduct research further on this area. Moreover, some researchers in their paper [6] conducted Machine Learning technique on water stress of plants by predicting the soil, environment and plant conditions. This precision technique aids to the management process of plant cultivation. Besides, there are some new kind of research is conducted to detect plant by using vein morphology. The authors in their paper [7] applied Convolutional Neural Network (CNN) on plant vein pattern to identify plants. This new technique would certainly improvise better identification of plant detection in the field of agriculture. Even researchers in their paper [8], used MLP on satellite based image data as remote sensing data to detect and classify plants which is certainly a new approach in research field. As new type of deep learning algorithms are using in other field of sciences, those similar type of algorithms could be implemented on agricultural science to get better result of crop cultivation and plant protection.

Bibliography

1. Dahikar Snehal S., *et al*. "An Artificial Neural Network Approach for Agricultural Crop Yield Prediction Based on Various Parameters". *International Journal of Advanced Research in Electronics and Communication Engineering* 4.1 (2015): 94-98.
2. Schuster Edmund W., *et al*. "Infrastructure for data-driven agriculture: identifying management zones for cotton using statistical modeling and machine learning techniques". 2011 8th International Conference & Expo on Emerging Technologies for a Smarter World. IEEE, (2011): 1-6.
3. Camargo A and JS Smith. "Image pattern classification for the identification of disease causing agents in plants". *Computers and Electronics in Agriculture* 66.2 (2009): 121-125.
4. Le., *et al*. "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods". *Information processing in agriculture* 6.1 (2019): 116-131.
5. Nasirahmadi Abozar and N Behroozi-Khazaei. "Identification of bean varieties according to color features using artificial neural network". *Spanish Journal of Agricultural Research* 11.3 (2013): 670- 677.
6. Dimitriadis Savvas and Christos Goumopoulos. "Applying machine learning to extract new knowledge in precision agriculture applications". 2008 Panhellenic Conference on Informatics. IEEE, (2008): 100-104.
7. Grinblat Guillermo L., *et al*. "Deep learning for plant identification using vein morphological patterns". *Computers and Electronics in Agriculture* 127 (2016): 418-424.
8. Kussul Nataliia., *et al*. "Deep learning classification of land cover and crop types using remote sensing data". *IEEE Geoscience and Remote Sensing Letters* 14.5 (2017): 778-782.

Volume 3 Issue 5 May 2019