Nitrogen Fertilizer Recommendation for Paddies through Automating the Leaf Color Chart (LCC)

Torikul Islam[1], Rafsan Uddin Beg Rizan[2], Yeasir Arefin Tusher[3], Md Shafiuzzaman[4],
Md. Alam Hossain[5] and Syed Galib[6*]

Department of Computer Science and Engineering, Jashore University of Science and Technology, Jashore, Bangladesh

Abstract—Nitrogen fertilizer is inevitable for rice production to ensure that the crop's nitrogen need is adequately supplied, during the growing season. International Rice Research Institute (IRRI) has proposed Leaf Color Chart (LCC) to detect the exact nitrogen need of paddy. Farmers generally monitor the plant's growth (which is also an indicator of the nitrogen concentration of leaves) by comparing the leaf color with the corresponding color of the LCC. Currently, in most cases, LCC is used manually to determine the fertilizer need and thus, there is a chance of either overestimating or underestimating the amount of fertilizer. To avoid this problem, a smart fertilizer recommendation system is proposed in this paper. The proposed method is able to automate the manual acquisition and interpretation of leaf color for classification through LCC. The experimentation considers a sample of 6000 Aman paddy leaf images. The data acquisition process was performed according to IRRI's guidance of taking the paddy leaf images within the body shade by our developed application. The data/images have already been made public in Kaggle - a well-known dataset website. The semantic segmentation of the dataset was performed by a powerful Convolutional Neural Network (CNN) backbone architecture - DeepLabV3+. Color classification into 4 categories of the LCC was performed by CNN architecture which consists of seven layers. Information gain based evaluation was performed in the Decision Tree (DT) approach to select features and with the selected features DT classified images into 4 categories. Color classification by our two proposed methods achieved 94.22% accuracy in CNN Model and 91.22% accuracy in the DT classifier.

Keywords—*Leaf Color Chart (LCC); Convolutional Neural Network (CNN); fertilizer recommendation system; color classification; Decision Tree (DT)*

I. INTRODUCTION

Paddy (*Oryza sativa L.*) is one of the principal food crops and also is consumed by one third of the world population residing in developing countries [1]. Nitrogen is a principle nutrient for paddy. The crop yields in the world are improved significantly due to the use of Nitrogen fertilizer [2]. However, farmers around the world are facing a problem to detect the nitrogen need of paddies for proper cultivation. Farmers attempt to estimate the required amount of nitrogen fertilizers through simply looking at the color of the crop's leaves. They generally apply nitrogen fertilizer too much (little Phosphorus (P) and Potassium (K) and other nutrients) that results in high pest and disease incidence [3]. Therefore, overestimation can be resolved by carefully matching leaf color with the LCC [4]. The optimum use of nitrogen fertilizer can be achieved by matching nitrogen supply with crop demand.

International Rice Research Institute (IRRI) has proposed Leaf Color Chart (LCC) to detect the exact nitrogen level of paddy [5]. LCC is used in the agricultural areas for recommending accurate amount of nitrogen fertilizer. In practice, leaf color is compared with its corresponding color in LCC inside body shade with proper lighting conditions. An exact color calibration process is necessary in the digital dimension for interpreting leaf colors. The calibration process evaluates the performance with the operational lighting conditions and determines whether the crops need fertilizers.

The LCC is being made of the plastic body having four green color levels with the standard suggestion of the amount of nitrogen fertilizers for different species such as Aman and Boro. It is arranged with such a shape where the panels are shown horizontally from yellowish green to dark green. IRRI's standard version ensures that the colors of the paddy leaves can be matched with corresponding LCC colors. The standard version having four green color variations as represented from two for yellowish green to five for the dark green is shown in Fig. 1.
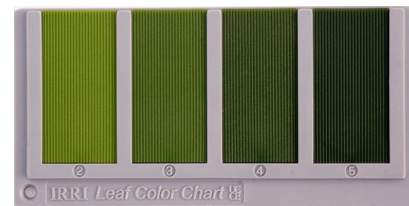

Fig. 1. Standard Leaf Color Chart [5]

Currently, the leaf color comparison process with LCC is manual i.e. farmers have to place paddy leaves on the shades of LCC and match it by simply visualizing. Thus, accurate measurement process is difficult for them. The farmers are also required to compare 6 to 10 leaves with the LCC level and find the average values. If the average value of Aman paddy leaf color level is less than or equal to 3.5 LCC level, nitrogen fertilizer is required of 7.5 kg per 0.133 hectares of land [6]. Usually, this process lacks the ease of detecting the accurate amount of fertilizers due to assumption. With the use of LCC for paddy, there is a possibility of saving fertilizers which may cause a positive environmental effect. The estimated annual saving of urea is 261.6 tons for Bangladesh if 50 per cent of the farmers use LCC in the irrigated rice area of 3488 million hectares of land [4].

Similar research has been proposed for nitrogen fertilizers of Soybean crops [7] by using the Fuzzy Logic method. The nitrogen fertilizer demand for Broccoli Plants was estimated

by Simone Graeff et al. [8]. The authors used the L*a*b* color system, also known as CIELAB color space (more on L*a*b* color system can be found in Malacara, 2003 [9]). The parameter, b*, was used to determine the nitrogen fertilizer demand. Mercado-Luna et al. [10] analysed color images to determine the nitrogen need for tomato seedlings. Wang et al. [11] measured the leaf nitrogen concentration as well as rice chlorophyll content where they took pictures under natural lights using a digital still color camera. Need based nitrogen fertilizer management was performed for different maize (Zea mays L.) genotypes by calibrating the leaf color chart [12]. Sing et al. estimated the need based nitrogen management for Rice and Wheat by using Chlorophyll Meter along with the Leaf Color Chart [13]. However, an automated need based nitrogen fertilizer recommendation system for paddies is required based on standard LCC rules.

Classification of paddy leaves into four categories of LCC color is the principal step of this research. The exact color classification method can bring a solution to this problem. Color classification can be performed by Convolutional Neural Network as well as other Machine Learning (ML) algorithms. These ML techniques have been applied to color classification for crops, fruits or even other objects. Son, Kim and Kim [14] compared the performances of a Gaussian mixture model, an Artificial Neural Network (ANN) model, and a Support Vector Machine (SVM) with two color spaces to detect concrete structural component and found that SVM with HSI (Hue-Saturation-Intensity) color model performed best. The SVM and K nearest neighbors (K-NN) were developed for hair color classification [15]. Automated mango ripening stages detection by color classification was performed using DT classifier [16]. Without using CNN based structure, SVM machine learning algorithm was also used for vehicle color recognition [17]. However, Convolutional Neural Network based backbone architecture can recognize vehicle color more efficiently [18]. Those color classification backbone architectures and ML algorithms motivated us to design our model for paddy leaf color classification.

Therefore, we have considered manually comparing LCC as our research problem and wanted to address it by using CNN and DT classifier as our two proposed automated methods. Those two methods can be integrated through a mobile app and thus a smart farming system can be introduced. Hence, by automating the fertilizer recommendation system can help farmers use fertilizers efficiently. This can make a great change over the total economy of a country. Therefore, our proposed solution will economically empower farmers through automating the determination of fertilizers in paddy fields.

The rest of paper is organized as follows. In the Section II, we have discussed the related background of our work. The data collection and validation has been discussed in section III. The overall system model of our proposed scheme, image segmentation and color classification are discussed in Section IV. Experimental analysis is represented in Section V. Its tell us about the fruitfulness of our model. The conclusion is presented in the Section VI. Finally in the Section VII, the acknowledgment is represented.

II. BACKGROUND

Need based nitrogen fertilizers implementation in economical crops like paddy is a concerning issue across the world. The LCC is basically a guideline to supply the necessary amount of nitrogen fertilizer which is optimal for achieving maximum yield. Researchers around the world proposed some systems to recommend nitrogen based fertilizer for paddies as well as other crops. This section reviews some related works on fertilizer recommendation system using image processing and/or LCC.

The suggestion of standard amount of nitrogen fertilizers is essential for rice production. In the critical growth stage, need base nitrogen management is suggested by using the leaf color chart (LCC) [13]. Modeling of demand of on-farm soil nitrogen supply and nitrogen nutrition in rice and wheat is conducted also [19]. Chlorophyll meter, also known as soil plant analysis development (SPAD), is used for determining the chlorophyll concentration on the leaf. SPAD technology is used for real time nitrogen management in rice [20]. In season, need based nitrogen management have demonstrated for four maize genotypes by calibrating leaf color chart [12].

Data acquisition technique from the field is one of the momentous phases of this research. There are some constraints for capturing images which is proposed by IRRI. The farmers need to capture images inside the body shade with proper lighting conditions [5]. Some researchers have used white paper as a background for better segmentation and prepossessing [7], [21]. Accurate segmentation depends on the proper data acquisition process. Some researchers have embedded Otsu's method in mobile application to obtain threshold values where the leaf will be segmented as object pixels from background pixels [7]. The background pixels are the area with white paper and object pixels are the area of the leaf portion. Reducing the limitation of removing background, we have used the DeepLabV3+ segmentation model which is invented by Google. Moreover, it can segment Regions of Interest (ROI) pixels from the leaf whatever the background is [22].

Accurate color level prediction of paddy leaf into 4 categories of LCC can be a fruitful way of this research. Automatic prediction of leaf color level by using digital image processing techniques has been proposed by Singh and Singh [21]. Here, test images are compared with database generated LCC value and then the color level is predicted. For need based nitrogen fertilizer recommendation on Soyabean crops, the fuzzy logic method has been proposed for automatic leaf color level determination [7]. All object pixel values are transferred to Citrus Color Index (CCI) values which is used for the fuzzy logic inference method. Color histogram analysis and pixels Bitwise operations have been proposed for paddy leaf color perception with standard LCC color level [23]. However, the machine learning and the Deep Neural Network based algorithm may bring more accurate results.

Color classification by deep convolutional neural network and machine learning algorithm have been proposed by some researchers for other color classification problems. The CNN based backbone architecture has been proposed for recognition of vehicle color [18]. The support vector machines (SVM) and Kth Nearest Neighbors (KNN) was developed for hair color classification [15] and vehicle color classification [17].

However, from those color classification architectures and ML algorithms, we have designed our own optimal model structure for paddy leaf color classification with the best result.

As it was seen, a number of analyses have been performed for other color classification. However, there are some limitations. For example, the approaches with the application of fuzzy logic suffers from the problem of being computationally intensive i.e. requires much computation power and takes long time. Moreover, in the color histogram analysis and pixels' bitwise operation, the greenness value of input images is compared with the standard greenness value of LCC each time. This creates repetition of comparison which may be eliminated easily through ML algorithms. Therefore, different ML algorithms or color classification neural network methods can learn and solve the limitations more efficiently. Furthermore, the limitation of background removal and segmentation of a leaf can be performed with Google's segmentation model - DeepLabV3+ - which can achieve better accuracy. More on DeepLabV3+ is discussed in Section IV-A.

III. DATA COLLECTION AND VALIDATION

The entire data set is collected from the paddy fields in the Aman season by our own developed application with the tremendous help of Bangladesh Agricultural Research Institute (BARI), Jashore, Bangladesh and the office of the Upazila Agriculture officer near Jashore. We have used our application in Nokia 3 (8MP Camera) and Samsung S8 (12MP Camera) devices for collecting data. In Fig. 2, we have demonstrated the developed application data acquisition process.
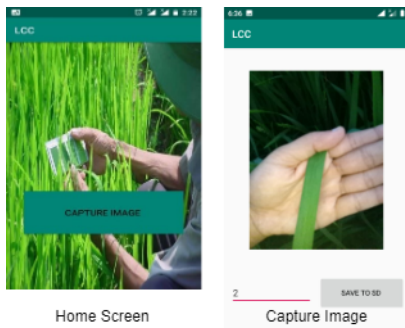

Fig. 2. Developed Data Collection Application

The color of the paddy leaf depends on the maturity as well as Nitrogen concentration. However, Nitrogen concentration can determine by analysing only the top most leaves. At first, a total of 560 top most paddy leaf images were taken. The resolution of the images were 500*500 pixels and we used our developed mobile application where each of 2 to 5 LCC level has 158, 162, 124 and 116 number of images respectively. Moreover, the color of the paddy leaf depends on the ambient light. Therefore, those images are captured in the natural daylight condition inside the body shade which condition is proposed by IRRI [4].

Data validation has been carried out by the two principal scientific officers at Bangladesh Agricultural Research Institute (BARI) Jashore. At the time of data collection, we have maintained the LCC data acquisition procedure as well as classified each images comparing to LCC stage. We have

carefully labeled our total data set corresponding to the LCC level with the help of domain experts. We have made our data set public in Kaggle for the academic, educational,and research community [24].
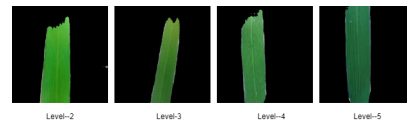

Fig. 3. Four LCC Level of Paddy Leaf

If the color of a paddy leaf is yellowish green, this color corresponds to LCC level 2. On the other hand, if the color of a paddy leaf is dark green, this color corresponds to LCC level 5. The middle 2 levels of the LCC is between the colors yellowish green and dark green and all these four types of paddy leaves are shown in Figure 3. Those four paddy leaf color levels reflect the nitrogen concentration of crops. The leaf color level 2 and 3 represent that the crops need fertilizers. The fertilizer is not needed for leaf color level 4 and level 5 as the nitrogen concentration is maximum in those two color level.

IV. METHODOLOGY

The main aim of this study is to propose a suitable method for automatic categorization of paddy leaves using their digital images. For that at first images of paddy leaves are taken as inputs. After that the DeepLabV3+ semantic segmentation is performed for background removing. The segmented images are used for both the CNN method and the DT method. In the DT method three features are extracted from segmented images. In both methods, the output is the LCC level. The entire flow diagram of our work is depicted in the Fig. 4.
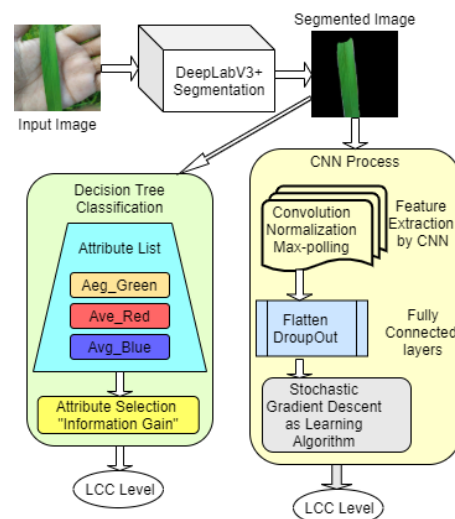

Fig. 4. Proposed Two Methods

A. DeepLabV3+ Segmentation

Object region segmentation from paddy images is implemented by DeepLabV3+ in TensorFlow [22]. This model generates semantic labels to every pixel in the input image and semantic segmentation masks for a specific Region of

Interest (ROI) in paddy leaf. Assigning these semantic labels involves specifying the outline of artifacts. Hence, this imposes much more stringent localization accuracy criteria than other visual entity recognition tasks, such as classifying image levels or recognizing boundary box rates. It explicitly controls the resolution using atrous convolution process. After that feature responses are computed within Deep Convolutional Neural Networks. This model has been driven by Fully Convolutional Network (FCN) frameworks for object detection and semantic segmentation, respectively.
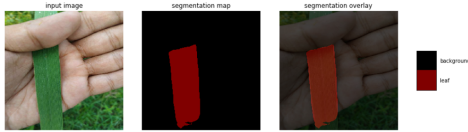

Fig. 5. Semantic Segmentation Process

In this paper, we propose a comparably enabling framework for semantic segmentation by DeepLabV3+ procedure training with our own data set of paddy leafs which is depicted in Fig. 5. As the backbone networks for implementation of this model, we have used MobileNetv2 architecture [25] rather than using other backbone networks such as Xception (server-side deployment) [26]. MobileNetv2 is a fast network structure designed for mobile devices implemented in real field. This model is conceptually intuitive with fast training and inference time for paddy leaf segmentation. Moreover, it offers flexibility and robustness [25].

B. Classification using Decision Tree

As discussed earlier, DT was applied to fruit categorization through color classification and achieved 96% accuracy [16]. Therefore, we apply this classification technique to leaf color classification.

Decision Tree is a supervised machine learning classification method where the data is relentlessly fragmented by a certain parameter. In this approach, a set of training examples are divided into smaller and smaller subsets while at the same time a corresponding decision tree is formed iteratively.

In this classification approach, we have fitted our data set by selecting three features from segmented images. The features are the average values of the three color channels - average red, average green and average blue. The average of each color is calculated by taking the average of the non-zero values of the image pixels. We have augmented initially our 560 paddy leaf images in total 6000 where each LCC level has 1500 images. The augmentation parameters width shift, height shift, shear, zoom and horizontal flip are used which sustain the color pixel information. After that, we have prepared data set based on 6000 images extracting three features. This data set is made public in the Kaggle data set [24] for the academic, educational,and research community.

Classification using a Decision Tree from a data set can be performed by using some properties or attributes of the data. Attributes may be selected based on many criteria and a selection measure is necessary here in our case. An attribute selection measure is a heuristic for selecting the splitting criterion. We have used Information Gain or Entropy as our attribute selection measure criterion which is also known as ID3 approaches for learning process. Entropy is a degree of randomness of elements or in other words it is a measure of impurity. Mathematically, it can be calculated with the help of the probability of the items as,

$$H = -\sum p(x) log p(x) \qquad (1)$$

Where $p(x)$ is the probability of the item, x. It is a negative summation of probability times the log of the probability of the item, x [27]. Percentage split technique was used to specify the training and testing datasets. 70% data of the whole dataset were used as training data. Testing was done on the remaining 30% data. Before going to train-test split function, we have shuffled our entire dataset that has reduced the bias error of accuracy. Classification using Decision Tree still have some problems especially when the sunlight condition is different.

Sunlight illumination problem is one of the concerning issues of this method. Different day light conditions will change the average values of the image pixels. As a result, miss classification rate may occur and farmers would not be able to detect the exact LCC level. However, to address this problem, we have moved to the convolutional neural network (CNN) model for better accuracy. Deep convolutional neural network works better than normal machine learning algorithms for feature extraction as well as classification.

C. Color Classification Using CNN

Leaf color information contributes significantly to need based nitrogen management. Extracting optimal features depend on the convolutional layers in the CNN model. We have demonstrated the CNN can not only learn for classifications based on shape information but also learn in color distribution. We have designed an optimal CNN model for paddy leaf color classification.

Our CNN architecture consists of seven layers where five layers are for convolution and 2 layers are fully connected layers. In our CNN architecture, the first two layers are convolution layers. Normalization and pooling processes are performed after the convolution process. The architecture of our CNN model is depicted in the Fig. 6.
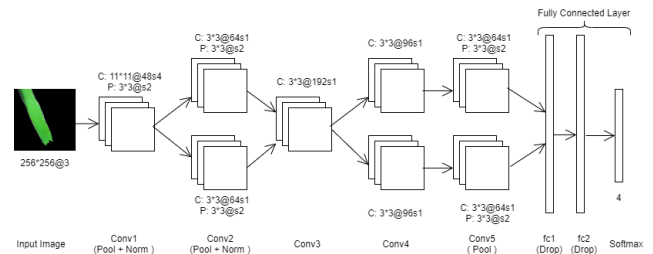

Fig. 6. CNN Architecture of Our System

Features from the images are extracted in the convolution layer. This process is similar to the convolution process in image processing algorithms.If the input image is I_i and the convolution kernel is h, then after the convolution process the output image I_o can be written as,

$$I_o[m, n] = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} I_i[i, j].h[m, n] \qquad (2)$$

where $I[m, n]$ is the value of the pixel of the image I at the coordinate (m, n). In neural networks, activation factions are used to obtain the output of a node. It determines the resulting value of a node between 0 and 1 or -1 and 1. Moreover, in the convolutional layer, the choice of a specific activation function has vital consequences for the networks. There are different activation functions including $tanh$ and $ReLU$ (Rectified Linear Unit). Activation function $ReLU$ is used for all layers in our CNN networks. The ReLU activation function is used as it does not saturate. When the neuron activates, the gradient is still high (equal to 1). The normalization process is done by equation 3,

$$l_{x,y}^i = K_{x,y}^i / (1 + \alpha/n \sum_{j=i-n/2}^{i+n/2} (K_{x,y}^i)^2)^\beta \qquad (3)$$

Here, $l_{x,y}^i$ is the output of the normalization process and $K_{x,y}^i$ is the result of the layer activation function for convolution at coordinate (x, y). Pooling is performed as the last process in the first two layers. Pooling operation can be of two types: max pooling and mean pooling. The maximum result from the convolutionary process formed by sharp edges will be taken by the max pooling process. On the other hand, the average result of the convolutionary process that sums up the shape in the neighbourhood will be taken by mean pooling. For overlapping pooling, we have used max pooling with size 3x3 and stride 2 in our CNN architecture. The second, fourth and fifth layers are grouped into two independent groups. Convolutional process is also performed in the third and fourth layers without the process of pooling and normalization. The outputs of these two layers are same as input. This is because a 3x3 kernel with pad 1 added has been used for each border. In the fifth layer, only convolution and pooling process is performed without the normalization process.

One long vector is formed by concatenating and flattening the pooling output of the fifth layer from the base networks. This converted long vector is passed to the fully connected layer after the fifth layer. The sixth and seventh layers are two fully connected layers where dropout regularization approach is used to reduce overfitting [28]. Dropout is a strategy that eliminates randomly selected neurons. In the last layer softmax activation function is employed which can be defined in equation 4.

$$p(y^{(i)} = j | x^{(i)}; \theta) = e^{\theta_j^T x^{(i)}} / \sum_{l=1}^k e^{\theta_l^T x^{(i)}} \qquad (4)$$

Where, $x^{(i)}$ is the given input with weight parameter θ and $p(y^{(i)} = j | x^{(i)}; \theta)$ is the probability of $y^{(i)}$ being class j.

Moreover, CNN architecture consists of 7 layers. 11x11@3 kernels for a total of 48 kernels are used in the first layer. The second layer uses 3x3@48 kernels for a total of 128 kernels. On the other hand, the third layer uses 3x3@128 kernels for a total of 192 kernels while 3x3@192 kernels for a total of 192 kernels are used in the fourth layer. Finally, the fifth layer uses 3x3@192 for a total of 128 kernels. In the first, second and fifth layers, the pooling process is employed with the same parameter. The pooling size of 3x3 with 2 pixel stride is used in those layers. The sixth and seventh layers are fully-connected layers. Moreover, in the sixth, seventh and output final layer the

number of neuron is 4096-4096-4 in each layers respectively with dropout regularization method.

We have trained the CNN model by stochastic gradient descent (SGD) as the learning algorithm. We have used 50 images as batch size, three channel image with 256x256@3 resolution, momentum of 0.9 and weight decay of 10^{-6}. The optimization method, SGD, wants to find the minimum or maximum value of some function. SGD will work for all functions with a gradient or first derivative. Moreover, SGD is used by the system for minimizing the error or loss function. The update of the weight parameters based on the function is performed according to equation 5.

$$w_{i+1} = w_i - \alpha \nabla L(z, w_i) \qquad (5)$$

Where the current weight parameters is w_i, the learning rate is α and $\nabla L(z, w_i)$ is the gradient of loss function L with respect to input examples z. To achieve better model convergence, the momentum and the weight decay are employed to the update equation. The final equation of update function in the SGD method is described in equations 6 and 7.

$$\Delta w_{i+1} = \gamma \Delta w_i | (1 - \gamma)(-\alpha \nabla L(z, w_i)) \qquad (6)$$

$$w_{i+1} = w_i - \alpha \nabla L(z, w_i) - \alpha \zeta w_i \qquad (7)$$

where γ is momentum variable and ζ is weight decay. Changing momentum and weight decay will speed up the training process.

The prepossessed dataset is used for training in the CNN model. Data augmentation is a powerful technique to reduce overfitting of the model [29], [30]. We have augmented initially our 560 paddy leaf images to 6000 where each LCC level has 1500 paddy leaf images after augmentation These images are also made public in kaggle dataset. Moreover, The percentage split method is used to specify the training and testing sets. Through this method, 70% of the data were used for training and the remaining 30% data were used for testing. Before going to train-test split function, we have shuffled our entire data set that reduces the bias error.

V. EXPERIMENTAL ANALYSIS

Many simulations are conducted to validate our result using two proposed methods. In this section, we describe how we conducted the experiments and demonstrated the simulation outcomes of the proposed two methods. Following metrics are used to measure the performance of our proposed model:

- *Accuracy:* Accuracy is one metric for evaluating classification models. Here, we present accuracy for both CNN method and decision tree method.

- *Loss:* Overall miss classification rate in decision tree method and CNN method.

- *Precision:* Precision can be thought of as a measure of exactness of the model.

- *Sensitivity:* Sensitivity is a measure of completeness of the proposed model.

A. Evaluation of the Decision Tree Approach

The decision tree classifier is significant in machine learning. Evaluation of a model represents how accurately a model classifies the test data set. In the machine learning model, the evaluation belongs to the accuracy score for testing set, miss classification rate, sensitivity and precision scores. We have split our entire 6000 images data set into 70% and 30%. Moreover, before splitting, we shuffled our entire data set which helps us reduce the biasness. We have trained our model by 4200 images data set and tested on the 1800 images data set. Using our standard data set, the decision tree is built with ID3 or information gain algorithm. In this method, the attributes with higher information gain are ranked higher.

In this section, we show, how accurately our model learns and correctly classifies our test data. By using the DT model, we have achieved the test accuracy of 91.22%. The confusion matrix of test data set demonstrates the prediction of each LCC level in Table I.

TABLE I.   CONFUSION MATRIX OF DECISION TREE

Actual Class		Predicted Class			
		Level-2	Level-3	Level-4	Level-5
	Level-2	425 & 26	0	0	
	Level-3	20 & 405	30	0	
	Level-4	0 & 18	405	32	
	Level-5	0 & 0	32	407	

In Table I, we can see the LCC level-2 425 images are correctly classified and only 26 images are misclassifed. Level-2 misclassification rate is 5.76%. In LCC level-3 and level-4, their misclassification rate is 10.99 % in both two levels. Those level misclassification rate is higher than others because of their color level are very close to each other. A little bit daylight change will increase the misclassification rate on these two levels. In LCC level-5, 407 are correctly classified and 32 images are misclassifed. The misclassification rate level-5 is 7.29%. A model's fruitfulness can be found by testing on a totally new data set and accurate class prediction on a new data set. Therefore, We also have checked our model performance by testing 60 totally new images for which each LCC level has 15 paddy images. The DT model achieved 83.33% accuracy in this approach.

There are different evaluation criteria to measure our proposed model's completeness. Here, we have also evaluated our decision tree model by sensitivity and precision score. Precision can be thought of as a measure of exactness of the model. Sensitivity is a measure of completeness of a proposed model. We have calculated the accuracy, sensitivity and precision score of our model by the equations 8, 9 and 10,

$$Accuracy = (TP + TN)/(P + N) \qquad (8)$$

$$Sensitivity = TP/P \qquad (9)$$

$$Precision = TP/(TP + FP) \qquad (10)$$

Where, TP= True Positives, TN= True Negatives, FP= False Positives, P= Totals positive and N= Totals negatives samples. DT model miss classification rate is 8.78%, Precision score is 91.24% and sensitivity score is 91.28%.

B. Evaluation of the CNN Model

In this paper, our aim is to automate the manual LCC system through processing digital images of paddy leaves. To fulfill that expectation, we have proposed a CNN model. In this approach, we have used our 6000 paddy leaf images. We have split the data set into 70% and 30% using train test split function. Therefore, we used 4200 images as our training data and 1800 as our test data set. Then we train our model with this new data set and successfully achieved 94.22% accuracy on test data set.

Model evaluation involves completeness of our proposed model, exactness of our proposed model and accurate classification rate. In this section, we are going to evaluate our CNN model by confusion matrix, precision, and recall score. The confusion matrix is generated from our 1800 test images. From the confusion matrix, we have proved our model miss classification rate is very low in the CNN model. The confusion matrix is shown in Table II,

TABLE II.   CONFUSION MATRIX OF CNN MODEL

Actual Class		Predicted Class			
		Level-2	Level-3	Level-4	Level-5
	Level-2	415	41	0	0
	Level-3	19	433	14	1
	Level-4	0	7	417	11
	Level-5	0	0	11	431

The sensitivity and precision score of our CNN model is 94.27% and 94.31% respectively. We have also checked our model performance by testing 60 new images data set where each LCC level have 15 paddy images. The CNN model have achieved 91.66% accuracy in this approach.

C. Result Analysis

This section presents an analysis of the results for the two proposed methods and the accuracy between two proposed methods. In experimental analysis section, we have shown that the CNN model achieved 94.22% and the DT model achieved 91.22% accuracy by using the same data set. CNN model achieved better accuracy than the DT model. This is due to the fact that sunlight has vital consequences on the paddy leaf images and it changes the pixels' information. As a result, misclassification rate was increased in the DT method. However, in the CNN model our proposed filter and kernel size extract feature in a coherent way that reduces the chances of misclassification. Our proposed CNN model learned in the better way and classified our paddy leaf images with better accuracy.

VI.   CONCLUSION

In this work, we mainly give attention to transform the manual reading of LCC level to an automated method that could be easily deployed in smartphones. Entire work is summarized by three main steps: (1) segmentation, (2) feature extraction, and (3) training and classification. These approaches are performed here using DeepLabV3+ and Convolutional Neural Network. Reducing the sunlight illumination problem while data acquisition can bring out a revolutionary change in this filed. The DeepLabV3+ process can be integrated in Mobile Phone by using the MobileNetV2 backbone network

with TensorFlow. This fruitful solution can also be implemented for other crops like soybean and wheat for nitrogen fertilizer recommendation. Moreover, we have implemented two improved algorithms that accurately classify the Aman paddy leaf according to LCC level with the best 94.22% accuracy in the CNN method and 91.22% in the Decision Tree method.

This automated method of LCC will help the farmer and recommend need based nitrogen fertilizers. The implementation of this system in smart devices will monitor the paddy field automatically that is fruitful for rural area farmers. The farmer will not need manual process to detect the nitrogen level and measuring out the exact amount of fertilizers. Hopefully, Our research will bring an economically fruitful way for Bangladesh as well as ASIA Pacific countries.

VII. Acknowledgment

References

[1] D. B. Tari, A. Gazanchian, H. A. Pirdashti, and M. Nasiri, "Flag leaf Morphophysiological Response to Different Agronomical Treatments in a Promising Line of Rice (Oryza sativa L .)," *Natural Resources Research*, vol. 5, no. 3, pp. 403–408, 2009.

[2] S. Peng, R. J. Buresh, J. Huang, X. Zhong, Y. Zou, J. Yang, G. Wang, Y. Liu, R. Hu, Q. Tang, K. Cui, F. Zhang, and A. Dobermann, "Improving nitrogen fertilization in rice by site-specific N management. A review," *Agronomy for Sustainable Development*, vol. 30, no. 3, pp. 649–656, 2010.

[3] S. Yosef Tabar, "Evaluation use leaf color chart in rice for nitrogen management," *J Sci Agric*, vol. 3, pp. 66–9, 2013.

[4] IRRI, "About leaf color chart," https://bit.ly/3eU6dt3, 2019, access date: February, 10, 2020.

[5] I. R. R. Institute, "Leaf Color Chart - IRRI Rice Knowledge Bank," https://bit.ly/2SbIqLs, access date: February, 08, 2020.

[6] B. R. K. Bank, "Leaf Color Chart (LCC) for Fertilizer N Management in Rice," http://knowledgebank-brri.org/how-to-use-lcc.php, access date: February, 09, 2020.

[7] K. R. Prilianti, S. P. Yuwono, M. A. Adhiwibawa, M. N. Prihastyanti, L. Limantara, and T. H. Brotosudarmo, "Automatic leaf color level determination for need based fertilizer using fuzzy logic on mobile application: A model for soybean leaves," *Proceedings - 2014 6th International Conference on Information Technology and Electrical Engineering: Leveraging Research and Technology Through University-Industry Collaboration, ICITEE 2014*, 2015.

[8] S. Graeff, J. Pfenning, W. Claupein, and H. P. Liebig, "Evaluation of image analysis to determine the N-fertilizer demand of broccoli plants (Brassica oleracea convar. botrytis var. Italica)," *Advances in Optical Technologies*, vol. 2008, 2008.

[9] D. Malacara, "Color vision and Colorimetry: Theory and applications," *Color Research & Application*, vol. 28, no. 1, pp. 77–78, feb 2003. [Online]. Available: https://doi.org/10.1002/col.10118

[10] A. Mercado-Luna, E. Rico-García, A. Lara-Herrera, G. Soto-Zarazúa, R. Ocampo-Velázquez, R. Guevara-González, G. Herrera-Ruiz, and I. Torres-Pacheco, "Nitrogen determination on tomato (Lycopersicon esculentum Mill.) seedlings by color image analysis (RGB)," *African Journal of Biotechnology*, vol. 9, no. 33, pp. 5326–5332, 2002. [Online]. Available: https://www.ajol.info/index.php/ajb/article/view/92074/81517

[11] Y. Wang, D. Wang, P. Shi, and K. Omasa, "Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still color camera under natural light," *Plant methods*, vol. 10, no. 1, p. 36, 2014.

[12] H. Thind, A. Kumar, M. Vashistha *et al.*, "Calibrating the leaf colour chart for need based fertilizer nitrogen management in different maize (zea mays l.) genotypes," *Field Crops Research*, vol. 120, no. 2, pp. 276–282, 2011.

[13] B. Singh, Y. Singh, J. K. Ladha, K. F. Bronson, V. Balasubramanian, J. Singh, and C. S. Khind, "Chlorophyll meter–and leaf color chart–based nitrogen management for rice and wheat in northwestern india," *Agronomy Journal*, vol. 94, no. 4, pp. 821–829, 2002.

[14] H. Son, C. Kim, and C. Kim, "Automated color model–based concrete detection in construction-site images by using machine learning algorithms," *Journal of Computing in Civil Engineering*, vol. 26, no. 3, pp. 421–433, 2012.

[15] S. Sarraf *et al.*, "Hair color classification in face recognition using machine learning algorithms," *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, vol. 26, no. 3, pp. 317–334, 2016.

[16] F. S. Mim, S. M. Galib, M. F. Hasan, and S. A. Jerin, "Automatic detection of mango ripening stages–an application of information technology to botany," *Scientia Horticulturae*, vol. 237, pp. 156–163, 2018.

[17] N. Baek, S.-M. Park, K.-J. Kim, and S.-B. Park, "Vehicle color classification based on the support vector machine method," in *International conference on intelligent computing*. Springer, 2007, pp. 1133–1139.

[18] A. Dehghan, S. Z. Masood, G. Shu, E. Ortiz *et al.*, "View independent vehicle make, model and color recognition using convolutional neural network," *arXiv preprint arXiv:1702.01721*, 2017.

[19] C. Adhikari, K. Bronson, G. Panuallah, A. Regmi, P. Saha, A. Dobermann, D. Olk, P. Hobbs, and E. Pasuquin, "On-farm soil n supply and n nutrition in the rice–wheat system of nepal and bangladesh," *Field Crops Research*, vol. 64, no. 3, pp. 273–286, 1999.

[20] V. Balasubramanian, A. Morales, R. Cruz, T. Thiyagarajan, R. Nagarajan, M. Babu, S. Abdulrachman, L. Hai *et al.*, "Adaptation of the chlorophyll meter (spad) technology for real-time n management in rice: a review." *International Rice Research Notes*, vol. 25, no. 1, pp. 4–8, 2000.

[21] A. Singh and M. L. Singh, "Automated color prediction of paddy crop leaf using image processing," *Proceedings - 2015 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development, TIAR 2015*, no. Tiar, pp. 24–32, 2015.

[22] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *ECCV*, 2018.

[23] R. John, B. C. Ii, B. S. Ece, S. R. Soltes, and A. V. Tatel, "Android-based image processing application for rice nitrogen management," no. March, 2012.

[24] Torikul, "Kaggle dataset," https://bit.ly/34rE4EM, 2019.

[25] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *CVPR*, 2018.

[26] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[27] Chen Jin, Luo De-lin, and Mu Fen-xiang, "An improved id3 decision tree algorithm," in *2009 4th International Conference on Computer Science Education*, July 2009, pp. 127–130.

[28] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[29] D. A. Van Dyk and X.-L. Meng, "The art of data augmentation," *Journal of Computational and Graphical Statistics*, vol. 10, no. 1, pp. 1–50, 2001.

[30] C. Zhang, P. Zhou, C. Li, and L. Liu, "A convolutional neural network for leaves recognition using data augmentation," in *2015 IEEE International Conference on Computer and Information Technology;* *Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing.* IEEE, 2015, pp. 2143–2150.