Fertilizer Recommendation for Rice Crop based on NPK Nutrient deficiency using Deep Neural Networks and Random Forest Algorithm

Tanneru Bhavya
Undergraduate Student
Electronics and Communication Engineering
National Institute of Technology Warangal
tb21ecb0b60@student.nitw.ac.in

Ramyalaxmi Seggam
Undergraduate Student
Electronics and Communication Engineering
National Institute of Technology Warangal
sr21ecb0b55@student.nitw.ac.in

Prof. Ravi Kumar Jatoth
Electronics and Communication Engineering
National Institute of Technology Warangal
ravikumar@nitw.ac.in

***Abstract*—Nutrient deficiency has a noteworthy impact on agriculture which results in reduced plant quality in turn reduced crop yield. A plant can have Multiple deficiencies at the same time so there is need of suggesting an appropriate Fertilizer considering all nutrient deficiencies in it. In this paper we suggested suitable fertilizer using random forest for detected nutrient deficiency from leaf images using different Neural Networks. We trained our dataset using four different Convolution Neural Networks and selected the network with highest accuracy and then the output of this network is sent to the trained model of Random Forest which suggests the suitable fertilizer. Since Potassium(K), Phosphorous(P) and Nitrogen(N) are three key nutrients required for Rice plant we considered the image dataset having leaves with N, P, K deficiencies.**

Keywords*— *Densenet, Resnet, Alexnet, Random Forest, Dense block, transition layer, Decision tree

INTRODUCTION

Rice is one of the Principal Food crop of India. India has the largest area under rice cultivation. But compared to other countries the average rice yield per acre in India is very low. One of the important causes for it is imbalanced nutrient management. Since the population is increasing rapidly the production of rice need to be increased. So, it is significant to analyze the nutrient deficiencies in Rice crop and suggest suitable fertilizer for it. Nitrogen (N) deficient crops have low yield. Potassium (K) deficiency affects the crop growth as it effects the canopy photosynthesis. Phosphorous (P) deficiency decreases photosynthetic efficiency of rice. The Main moto of our project is to determine the percentages of nutrient deficiencies and to suggest the fertilizer.

A. RELATED WORKS

There are many works done on identifying nutrient deficiencies in plants based on notable different patterns and colors on leaf for each nutrient. Aditi Shah et al. [1] of Savitri Bhai Phule Pune University used Basic Image-Processing and machine Learning Algorithms to find macro-nutrient deficiencies (Nitrogen, Potassium, Calcium, Phosphorous, Sulphur, Magnesium). Supervised Machined learning Algorithm used is decision tree using the attributes like edges, spots, prescribed deficiency, color index, average color value and band ratios. Amirtha T et al. [2] of Agni College of Technology, Chennai proposed a model using Convolution Neural network to identify Manganese, Nitrogen, Potassium, Sulphur, Zinc deficiencies and healthy plant. Pavit Noinongyao et al. [3] proposed an about image analysis method for identifying the five different nutrients i.e., Mg, Ca, K, Fe, and N deficiencies in plants based on the image of leaf using Convolutional Neural Networks with the set of black grams grown under nutrient controlled environment. These results indicate the superiority of the proposed method over trained humans in nutrient deficiency identification. Lisu Chen et al. [4] Deng used Static Scanning Technology and Hierarchical Identification Method to identify the potassium, phosphorous and nitrogen deficiencies in paddy leaf. R. Sathyavani et al. [5] used LSTM ResNet50, DenseNet-BC, ResNet50, VGG16 DenseNet-169 and LSTM-VGG16 neural networks and compared the results. The principle used is since specific color and shape may have many root cause problems it is hence necessary to carefully analyze the texture of leaf with proper training of a classifier. They designed an image acquisition-based classification model such that it utilizes Internet of Things (IoTs) for data acquisition and a variant of convolutional neural networks namely DenseNet-BC classification purpose.

METHODOLOGY

A. Work Flow

The proposed workflow for the paper is as shown in Fig.1

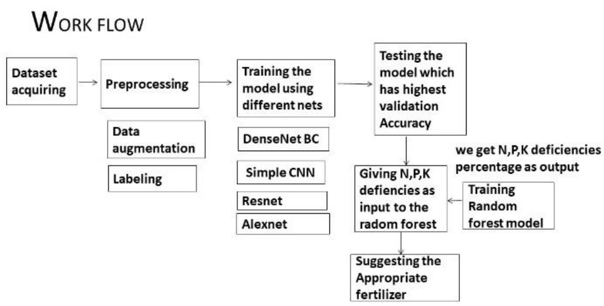


Fig.1-Work Flow

we acquire the data and then we do data augmentation for the data. Then we train our data using 4 different convolution neural networks and after finding the accuracy we select the network to be used. The output of selected network is given

2023 3rd International conference on Artificial Intelligence and Signal Processing (AISP) | 979-8-3503-2074-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/AISP57993.2023.10134912

as input to the random forest which suggests the suitable fertilizer.

B. Data Acquisition

We used two datasets one for Convolution Neural Network and other for Random Forest.

The first one is an image dataset which has 1156 images classified into Nitrogen (440 images), Potassium (383 images) and Phosphorous (333 images).

We divided the given dataset into three folders i.e. testing (10 percentage), training (80 percentage) and validation data (10 percentage). The image dataset is collected from Kaggle [6].
The second dataset is fertilizer recommendation from percentage of Deficiencies. This data is collected from Kaggle [7].

C. Description of Neural networks

Here we trained our dataset using 4 different Neural Networks

- *Convolution Neural Network*

  The CNN model we used consists of 4 convolution layers with striding, padding and with the activation function ReLU each followed by a max pooling layer. Then it is flattened and given to a Neural Network with 3 neurons as output. The convolution neural network used is as shown in Fig.2

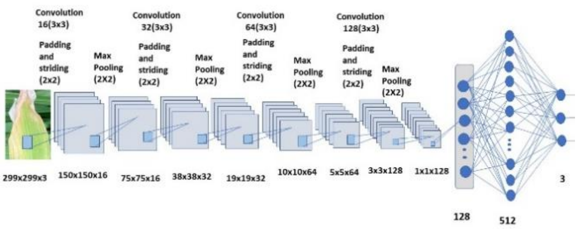


Fig.2-The proposed Convolution Neural network

- *Alexnet*

  The Alexnet consists of 8 layers, the first 5 are the convolution layers and the remaining 3 are the fully connected layers. The first, second and fifth layers are Max pooled before passing to the next layers. Except the last layer each of these layers are followed by the ReLU activation function. The last layer is followed by the softmax activation.

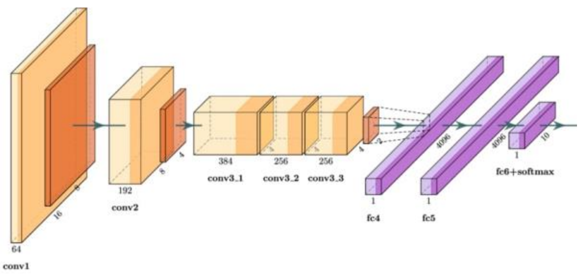


Fig.3-Architecture of Alexnet

- *Resnet*

  In this network, the skip connections are used to connect the outputs of a layer to its next layers by skipping few layers in between . This will form a residual block. Resnet is made by stacking these residual blocks together. Resnet50 Architecture is shown in Fig.4

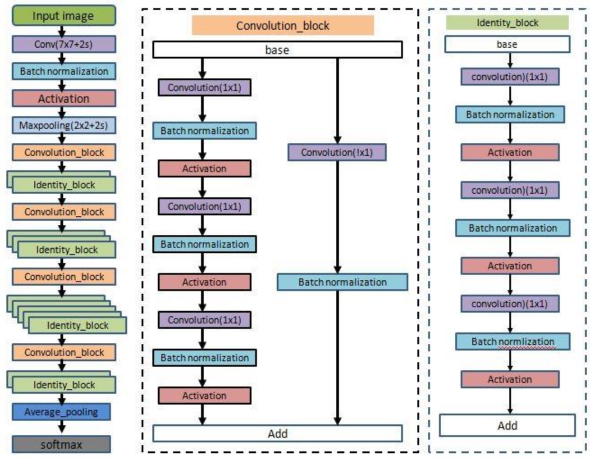


Fig.4 - Resnet 50 Architecture

- *DenseNet-121*

  DenseNet-121 have 4 dense blocks each having some convolution layers and a transition layer is added between two dense blocks. In dense block, every layer is connected to all other layers. Each layer gets feature-maps from previous layers. There are 6,12,24,16 convolution layers in 1st, 2nd, 3rd, 4th Dense Blocks respectively as shown in Fig.5

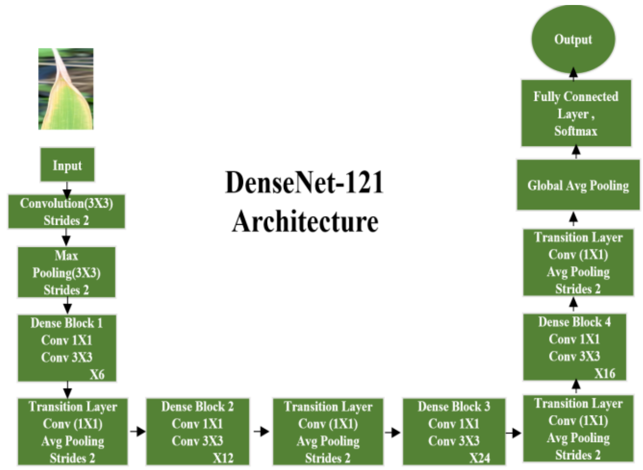


Fig.5- Architecture of DenseNet

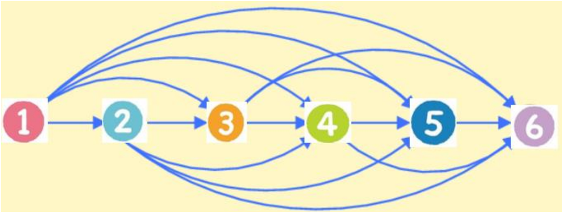


Fig.6 - Inside Dense Block

Inside Dense Block there are 6 convolution layers as shown in Fig.6

We have operations inside convolution layers as shown in Fig.7. Batch Normalization with subsequent activation function (ReLU) and then (3X3) Convolution and depending on the requirement Dropout is also added. Batch Normalization is nothing but another network layer which is placed between hidden layer and its next hidden layer. It takes output from one hidden layer and normalize it before sending it as an input to its next hidden layer.

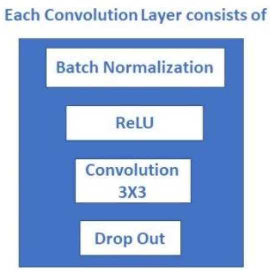


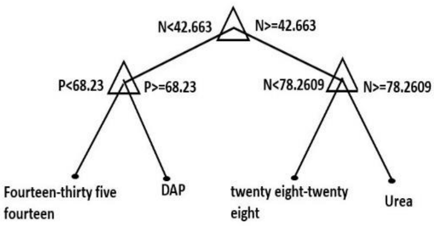
Fig.7- Inside Each Convolution Layer

D. Fertiliser Recommendation using Random Forest

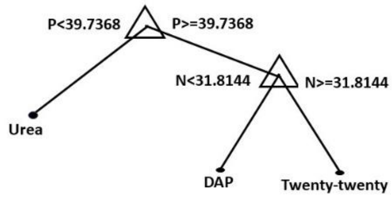
The output of the SoftMax function is the probability of the respective nutrition deficiency i.e. it indicates the percentage of the particular deficiency features in the leaf .Therefore we used this output to suggest the appropriate fertiliser depending on the percentage of the deficiencies in them. We gave this output (from neural networks) to a trained random forest model.

Random Forest is an ensemble learning algorithm in which the given dataset is bootstrapped that is row and feature sampled randomly and individual decision trees are trained. Each of these decision trees generates an output. Final output is considered based on Majority Voting for classification.

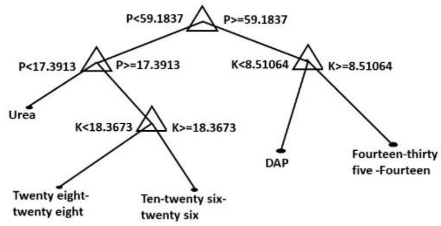
Here we used Information gain method for attribute selection of the Decision trees in the Random Forest.



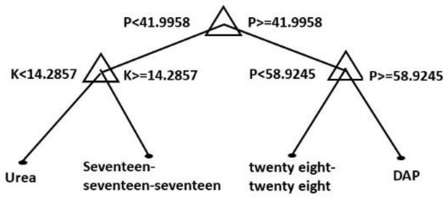
(a)



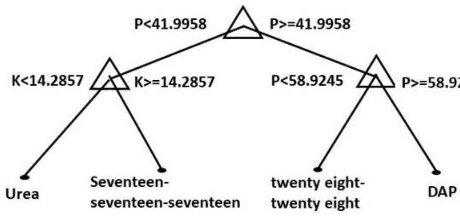
(b)



(c)



(d)



(e)

Fig.8(a-e)-Few individual decision trees trained for the randomly selected sample data sets

Since the dataset here is numerical i.e. the NPK relative percentages of the respective nutrients present in the respective fertilizer. Hence for the attribute selection we find the threshold such that the information gain for it is maximum and these continuous values are converted to the threshold based Boolean features.

Example calculations for the randomly selected decision trees are shown below in Table-1

Table-1

N%	P%	Fertilizer Name
100	0	Urea
26.53061224	73.46938776	DAP
52.63157895	47.36842105	Twenty-Twenty
31.57894737	31.57894737	Seventeen-Seventeen-Seventeen
22.91666667	77.08333333	DAP
100	0	Urea
56.52173913	43.47826087	Twenty-Twenty
53.48837209	46.51162791	Twenty Eight-Twenty Eight
19.14893617	61.70212766	Fourteen-Thirty Five-Fourteen
100	0	Urea
23.52941176	76.47058824	DAP
51.85185185	48.14814815	Twenty-Twenty
51.06382979	48.93617021	Twenty Eight-Twenty Eight
23.07692308	76.92307692	DAP
100	0	Urea
15.2173913	65.2173913	Fourteen-Thirty Five-Fourteen
54.76190476	45.23809524	Twenty Eight-Twenty Eight
100	0	Urea
16.32653061	63.26530612	Fourteen-Thirty Five-Fourteen
22.64150943	77.35849057	DAP
37.5	27.5	Seventeen-Seventeen-Seventeen
28.84615385	71.15384615	DAP
50	50	Twenty-Twenty
20.40816327	65.30612245	Fourteen-Thirty Five-Fourteen
47.82608696	52.17391304	Twenty Eight-Twenty Eight
100	0	Urea

Table-2

fertiliser	No of samples
urea	6
DAP	6
Twenty-Twenty	4
Seventeen-Seventeen-Seventeen	2
Twenty Eight-Twenty Eight	4
Fourteen-Thirty Five-Fourteen	4

Total Entropy(S) = -P(urea) $\times log_2^{\mathrm{P(urea)}}$ - P(DAP) $\times log_2^{\mathrm{P(DAP)}}$ - P(Twenty-Twenty) $\times log_2^{\mathrm{P(Twenty-Twenty)}}$ - P(seventeen-seventeen-seventeen) $\times log_2^{\mathrm{P(seventeen-seventeen-seventeen)}}$ - P(TwentyEight-TwentyEight) $\times log_2^{\mathrm{P\ (TwentyEight-TwentyEight)}}$ - P(Fourteen-ThirtyFive- Fourteen) $\times log_2^{\mathrm{P\ (Fourteen-ThirtyFive-\ Fourteen)}}$

= 2.507380102

Here for a random point N=52.63

The number of values less than 52.63 are 16

fertiliser	No of samples
DAP	6
Seventeen-seventeen-seventeen	2
Fourteen-Thirty Five-Fourteen	4
Twenty-Twenty	2
Twenty Eight-Twenty Eight	2

The Entropy (N<52.63) = -P(DAP)$\times log_2^{\mathrm{P(DAP)}}$ – P(seventeen-seventeen-seventeen) $\times log_2^{\mathrm{P(seventeen-seventeen-seventeen)}}$ – P (Fourteen-ThirtyFive- Fourteen) $\times log_2^{\mathrm{P\ (Fourteen-ThirtyFive-\ Fourteen)}}$ – P(Twenty-Twenty) $\times log_2^{\mathrm{P(Twenty-Twenty)}}$ – P (TwentyEight-TwentyEight)$\times log_2^{\mathrm{P\ (Twenty\ Eight-Twenty\ Eight)}}$

=2.155639

The number of values greater than or equal to 52.63 are 10

fertiliser	No of samples
urea	6
Twenty-Twenty	2
Twenty Eight-Twenty-Eight	2

The Entropy (N>=52.63) = -P(urea) $\times log_2^{\mathrm{P(urea)}}$ – P(Twenty-Twenty) $\times log_2^{\mathrm{P(Twenty-Twenty)}}$ – P (TwentyEight-TwentyEight) $\times log_2^{\mathrm{P\ (Twenty\ Eight-Twenty\ Eight)}}$

=1.370950594

Gain(S,52.63) = Entropy(S) - $\sum_{v \in <52.63,>52.63} \frac{|S_v|}{|S|} \boldsymbol{Entropy(S_v)}$

=0.653544297

Similarly, for other points considering some random step size information gain is calculated.

Let us consider step size = 10

Then we calculate same as above for 33.63,43.63,63.63,73.63 and so on. Then the attribute with maximum information gain is selected.

The attributes are split iteratively until the leaf nodes are reached. Similarly, many more decision trees are trained for the randomly selected sample datasets and the majority of the outputs is considered.

SIMULATIONS AND RESULTS

All the CNN architectures and algorithm are simulated in google colab environment using TensorFlow with Keras API. To fasten the simulations the python codes are executed on colab with GPU.

Comparison results of different neural networks

Neural network	No. of epochs	Train. loss	Train. accuracy	Valid. loss	Valid. accuracy
CNN	275	0.1012	0.9683	0.4123	0.8609
Alexnet	200	0.2171	0.9201	1.4099	0.8291
Resnet	100	0.4321	0.8522	1.4099	0.7725
Densenet121	1600	0.0872	0.9708	0.3437	0.8696

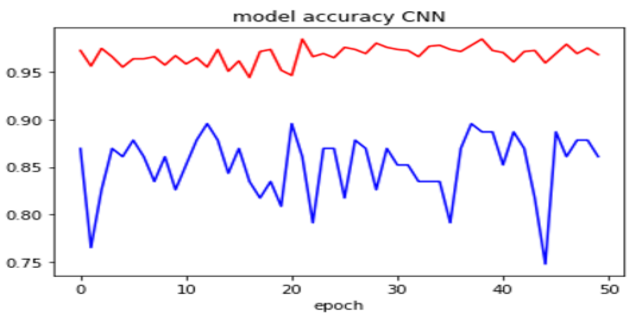


Fig.9-model accuracy CNN

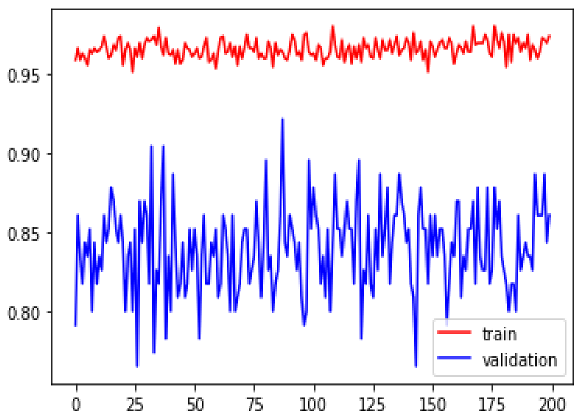


Fig.10-model Accuracy Denset

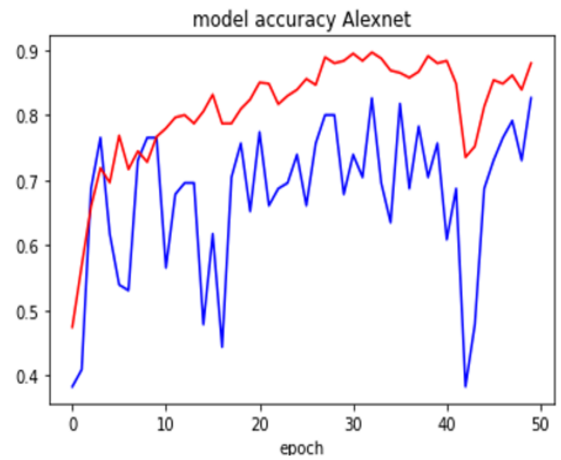


Fig.11-model accuracy Alexnet

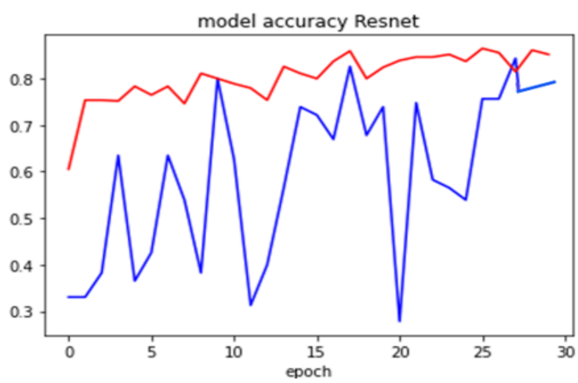


Fig.12-model accuracy Resnet

We trained our dataset using all 4 nets mentioned and we got highest accuracy for densnet so we gave the result of densenet (percentages of P,K,N) as input of Random forest and obtained the results . Few of the results are shown in Fig 13.

```
TEN-TWENTY SIX-TWENTY SIX
Fourteen-Thirty Five-Fourteen
TWENTY-TWENTY
TWENTY-TWENTY
TWENTY-TWENTY
TWENTY-TWENTY
```

```
Fourteen-Thirty Five-Fourteen
UREA
TWENTY-TWENTY
Seventeen-Seventeen-Seventeen
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
UREA
TEN-TWENTY SIX-TWENTY SIX
TWENTY-TWENTY
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
TEN-TWENTY SIX-TWENTY SIX
Fourteen-Thirty Five-Fourteen
TWENTY-TWENTY
```

Fig.13,14-Results of Predicting suitable Fertilizer

Conclusion

In this paper, we estimated the percentage of the deficiencies of the macro nutrients N P K in the rice leaf using four different networks and the network which gave highest accuracy is selected. Then this selected network output i.e. the output from the SoftMax is given to the random forest model. This suggests the appropriate fertilizer for the respective percentages of the nutrition deficiencies. The salinity of the soil should also be tested to ensure that the color of the leaf is mostly due to nutrition deficiency in the soil.

ACKNOWLEDGEMENT

This work is encouraged and developed with the support of Project EEQ/2021/001087, funded by the program "SERB" under the department of Science and Technology, Government of India.

References

[1] Shah, A., Gupta, P., & Ajgar, Y. M. (2018, April). Macro-nutrient deficiency identification in plants using image processing and machine learning. In *2018 3rd International conference for convergence in technology (I2CT)* (pp. 1-4). IEEE.

[2] Amirtha, T., Gokulalakshmi, T., Umamaheswari, P., & Tech, T. R. M. (2020). Machine Learning Based Nutrient Deficiency Detection in Crops. *Int. J. Recent Technol. Eng*, *8*(6), 5330-5333.

[3] Watchareeruetai, U., Noinongyao, P., Wattanapaiboonsuk, C., Khantiviriya, P., & Duangsrisai, S. (2018, March). Identification of plant nutrient deficiencies using convolutional neural networks. In *2018 International Electrical Engineering Congress (iEECON)* (pp. 1-4). IEEE.

[4] Chen, L., Lin, L., Cai, G., Sun, Y., Huang, T., Wang, K., & Deng, J. (2014). Identification of nitrogen, phosphorus, and potassium deficiencies in rice based on static scanning technology and hierarchical identification method. *PloS one*, *9*(11), e113200.

[5] Sathyavani, R., JaganMohan, K., & Kalaavathi, B. (2022). Classification of nutrient deficiencies in rice crop using denseNet-BC. *Materials Today: Proceedings*, *56*, 1783-1789.

[6] https://www.kaggle.com/datasets/guy007/nutrientdeficiencysymptom sinrice

[7] https://www.kaggle.com/datasets/mohitsingh1804/fertilizer-prediction-dataset