Crop Yield Estimation Using CNN

Apoorva Hegde, Phalguni S P
School of Computer Science of Engineering
RV University
Bengaluru, Karnataka
apoorvah.bsc22@rvu.edu.in,
phalgunisp.bsc22@rvu.edu.in

***Abstract*—This paper presents the development and evaluation of a Convolutional Neural Network (CNN) to estimate crop yields based on historical data and various predictive features. Utilizing the dataset from the paper A CNN-RNN Framework for Crop Yield Prediction, the proposed CNN model aims to capture complex patterns and improve predictive accuracy compared to traditional machine learning models. The CNN achieved competitive R-squared (R^2) scores, demonstrating its potential in agricultural yield prediction.**

Keywords— regression models, crop yield prediction, features, CNN, Random Forest

I. Introduction

Accurate crop yield prediction is crucial for effective agricultural planning and resource management. Traditional methods often fail to capture the complexity of agricultural systems. Machine learning (ML) and deep learning (DL) have shown promise in various predictive tasks, including agriculture. This study aims to implement a CNN to predict crop yields using historical data, comparing its performance with Decision Tree and Random Forest regressors. Initial attempts using Linear Regression showed suboptimal performance, prompting the exploration of CNN for its superior ability to model complex data patterns.

II. Related work

Previous research has explored various ML techniques for crop yield prediction. Regression models, such as Decision Trees and Random Forests, have been widely used. More recent studies have employed DL techniques, such as CNNs, to capture spatial and temporal dependencies in agricultural data, showing improved performance in yield prediction tasks.

The paper A CNN-RNN Framework for Crop Yield Prediction [1] provided a robust dataset that was utilized in this study. This paper integrates both Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to enhance crop yield prediction by capturing spatial and temporal patterns. The results demonstrated significant improvements in prediction accuracy compared to traditional methods. Additionally, the Syngenta Crop Challenge 2018 [2] has encouraged the development of innovative predictive models, and insights from this competition have been referenced to guide the methodology in this study.

III. Dataset and Features

The dataset used in this study is derived from the paper A CNN-RNN Framework for Crop Yield Prediction. It includes a variety of factors influencing crop yield, such as weather conditions, soil properties, and management practices. The dataset was preprocessed by standardizing the features and splitting it into training and testing sets (80%-20% split). The features were reshaped to fit the input requirements of the CNN.

IV. Methods

A. Data Pre-processing:

- Standardization of features
- Splitting data into training and testing sets
- Reshaping data for CNN input

B. Model Architecture:

- Multiple convolutional layers to extract features
- Dense layers for regression tasks

Initial experiments involved implementing a Linear Regression model. However, the performance was inadequate due to the model's inability to capture non-linear relationships and complex patterns in the data. This limitation led to the adoption of a CNN, which is better suited for such tasks.
The model was compiled with the Adam optimizer and mean squared error (MSE) as the loss function. The model was trained for 30 epochs, and its performance was evaluated on the test data using R-squared (R^2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

V. Experiments, Results and Discussion

CNN Performance

- R-squared (R^2): 0.92
- Mean Squared Error (MSE): 121.77
- Root Mean Squared Error (RMSE): 11.03
- Mean Absolute Error (MAE): 8.81

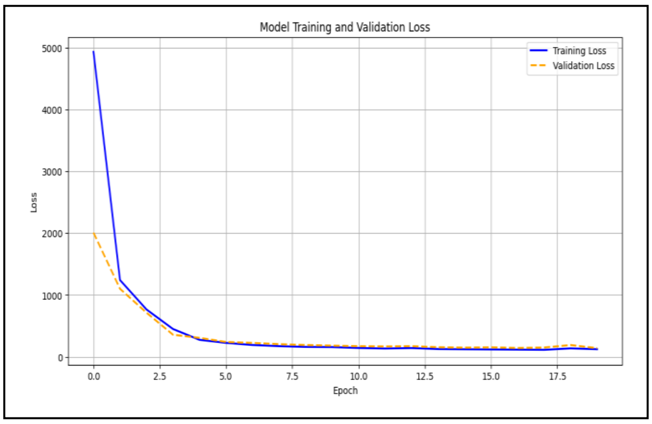
For comparison, a Decision Tree Regressor was trained and evaluated with the following results:

- R-squared (R^2): 0.74
- Mean Squared Error (MSE): 390.24
- Root Mean Squared Error (RMSE): 19.75
- Mean Absolute Error (MAE): 14.71

Similarly, a Random Forest Regressor was implemented, yielding the following metrics:

- R-squared (R^2): 0.87
- Mean Squared Error (MSE): 187.51
- Root Mean Squared Error (RMSE): 13.69
- Mean Absolute Error (MAE): 10.04

The CNN model outperformed both the Decision Tree and Random Forest Regressors in terms of R-squared, MSE, RMSE, and MAE. This demonstrates the capability of CNNs to capture complex patterns in the data, leading to more accurate yield predictions.



MODEL TRAINING AND VALIDATION LOSS

VI. CONCLUSION

This study demonstrates the efficacy of CNNs in crop yield prediction. The proposed model achieved competitive performance, underscoring the potential of DL in agricultural applications. These findings can aid in better decision-making and resource management in agriculture.

VII. FUTURE WORK

Future work could explore the integration of Recurrent Neural Networks (RNNs) with CNNs to further capture temporal dependencies in the data. Additionally, the application of data augmentation techniques and hybrid models could enhance predictive accuracy. Expanding the dataset and incorporating more diverse features may also improve the model's generalization capabilities.

VIII. ACKNOWLEDGMENTS

We would like to express our gratitude to the authors of the paper A CNN-RNN Framework for Crop Yield Prediction for providing the dataset used in this study. Their work laid the foundation for our research and significantly contributed to the advancement of crop yield prediction models. Additionally, we acknowledge the insights gained from the Syngenta Crop Challenge, which helped guide our methodology and approach. Lastly, we thank the developers of TensorFlow and Keras for their excellent tools that facilitated the implementation of our Convolutional Neural Network model.

IX. REFERENCES

[1] L. W. a. S. A. Saeed Khaki, "CNN-RNN-Yield-Prediction," 2019.

[2] L. W. Saeed Khaki, "Crop Yield Prediction Using Deep Neural Networks," 2019.