Convolution and Recurrent Hybrid Neural Network for Hevea Yield Prediction

Lince Rachel Varghese[1,*] & Vanitha Kandasamy [2]

[1]Department of Computer Science, Dr G R Damodaran College of Science, Coimbatore, Tamilnadu, India
[2]Department of Computer Science, Dr G R Damodaran College of Science, Coimbatore, Tamilnadu, India
*E-mail: lincerachel@gmail.com.

Abstract. Deep learning techniques have been used effectively for rubber crop yield prediction. A hybrid of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) is the best technique for crop yield prediction because it can effectively handle uncertainty of features. Hence, in this paper, a hybrid CNN-RNN method is proposed to forecast Hevea yields based on environmental data in Kerala state, India. The proposed hybrid CNN-RNN method reduces the internal covariate shift of CNN by batch normalization and solves the gradient vanishing or exploding problem of RNN using LSTM with a cell activation mechanism. The proposed method has three essential characteristics: (i) it captures the time dependency of environmental factors and improves the inherent computational time; (ii) it is capable of generalizing the yield prediction under uncertain conditions without loss of prediction accuracy; (iii) combined with the back propagation and feed forward method it can reveal the extent to which samples of weather conditions and soil data conditions are suitable to provide a clear boundary between rubber yield variations.

Keywords: *convolutional neural networks; feature selection; recurrent neural networks; rubber yield prediction; weather prediction.*

1 Introduction

The rubber tree (Hevea brasiliensis) is an important agricultural crop that occupies more than 10 million hectares of the earth's terrestrial area, as described by Rivano, *et al.* [1]. Latex harvesting from rubber trees is often done repeatedly on the same trees over a period of time. Data collection on the same measuring units (rubber trees) over time is unavoidable when investigating the latex yield from each rubber harvest. As a result, the resulting longitudinal data will almost certainly be serially clustered, necessitating the use of mixed models to allow for autocorrelation, as defined by Golbon, *et al.* [2].

Machine learning algorithms have been utilized in a number of researches to estimate crop yields. Jeong, *et al.* [3] employed a machine learning system and

Received April 29th, 2021, Revised July 12th, 2021, Accepted for publication August 6th, 2021.
 ISSN: 2337-5787, DOI: 10.5614/itbj.ict.res.appl.2021.15.2.6

found that Random Forest was the best algorithm for crop yield prediction. Romero, *et al.* [4] tested several machine learning algorithms and found that the association rule mining approach produced the best results for a wide range of crops. However, the quality of the output of machine learning algorithms is mostly determined by the quality of the input data.

Deep learning approaches are representation learning methods that have several levels of representation, each progressed with nonlinear modules that convert the representation at the current level (beginning with the raw input) to a likely more abstract level [5]. Deep neural networks also provide a basis for uniform approximation methods, which ensures to reach a quality solution regardless of the concept of the function. Deep learning approaches do not require handcrafted features. Instead, these approaches can learn them from the input, resulting in higher performance accuracy. Deep learning techniques often make use of cutting-edge modeling and solution techniques to forecast crop yields based on environmental data and management methods.

The deep neural networks constructed by Khaki & Wang [6] were able to find nonlinear and complicated correlations between environmental factors as well as their interactions from historical data and create reasonably accurate yield estimates for various crops. The model's performance was found to be relatively sensitive to the quality of the weather forecast, implying the necessity of weather prediction methodologies. The guided back propagation approach was used to locate input variables that maximize the activation of interested neurons by back propagating positive gradients. The complicated structure of this model makes it difficult to generate testable hypotheses.

A convolutional neural network is a specialized neural network architecture that is well-suited for processing multidimensional data. CNNs are typically composed of the following layers: a convolution layer, an activation function layer, a pooling layer, a fully connected layer, and a loss layer. They have certain specification criteria or parameters, such as the number of filters, filter size, inserting form, and stride. Filters are weighted matrixes in CNNs and they are used to convolve the input results. The CNN architecture does not consider temporal dependency between features. Nevavuori, *et al.* [7] specifically developed CNNs for crop yield prediction.

Recurrent Neural Networks (RNNs) are designed primarily for processing sequential data with inter-sample correlations. They have the advantage of interpreting sequential data and can be programmed to model both long-term and short-term data associations. RNNs are used to collect time dependencies in tasks containing sequential data. RNNs store the background of all previous elements of a sequence in their hidden units, known as state vectors, and use this information as they process the input sequence one element at a time. RNNs are very effective in sequence modeling. Long short-term memory (LSTM) is a type of RNN proposed by Sharma, *et al.* [8] for crop yield prediction. RNNs do not solve multi-dimensional data efficiently.

Recently, CNN and RNN based structures have been combined efficiently for various applications. A hybrid CNN-RNN model was proposed by Khaki, *et al.* [9] for crop yield prediction with different environmental data. The CNN layers learn strong middle-level features and assist the RNN layer in learning successful spatial relations between image region features. Meanwhile, the RNN learns temporal dependency between features to improve crop yield by considering seasonally varying weather parameters. The back propagation method is used with RNN for selecting only the more relevant features to obtain high-level prediction accuracy.

The primary goal of this study was to create an efficient hybrid CNN-RNN model for predicting rubber crop yield. The proposed method not only predicts the yield of Hevea brasiliensis (rubber crop), but it also overcomes challenges in hybrid CNN-RNNs. The internal covariate shift problem arises in CNN as a result of variance in the sharing of network activations, as changes in the network parameters create variation in the network activations. Batch normalization is used to solve CNN's internal covariate shift, fixing the means and variances of the layer inputs. It also improves gradient flow through the network by lowering the gradient's reliance on the scale of the parameters or their initial values.

The gradient vanishing or exploding problem affects RNN. In each training iteration, each of the neural network's weights receives an update proportionate to the partial derivative of the error function with respect to the current weight. The issue is that the gradient may be vanishingly small in some situations, thereby preventing the weight from changing its value. Exploding gradients are a problem that occurs when huge error gradients build and cause very large modifications to the neural network model weights during training. LSTM cells, which are precisely constructed recurrent neurons that provide excellent performance in a wide range of sequence modeling applications, improve RNNs. To avoid the vanishing gradient and exploding problem, LSTM cells utilize a special unit called the memory cell to recall inputs for a long time period.

Thus, the proposed hybrid CNN-RNN structure eliminates the problems of internal covariate shift and gradient vanishing or exploding, reduces the classifier's computational cost while also improving the accuracy of the rubber yield prediction. In the results and discussion section, the performance of the proposed structure is briefly explained.

2 Literature Survey

CNN and RNN were improved by various techniques. Finally, CNN and RNN were hybridized for achieving better prediction results. The improved methods for CNN, RNN and the hybridized CNN-RNN are discussed in this section.

The Deep Recurrent Q-Network model was built by Elavarasan & Vincent [10] to forecast crop yields over the Q-Learning reinforcement learning algorithm. The data parameters are initially fed into the sequentially stacked layers of RNN. The Q-learning network was developed on a crop yield prediction environment, based on the input parameters. A linear layer converts the RNN output values to Q-values. The Reinforcement Learning agent combines parametric characteristics with the threshold to aid in the prediction of the crop yield. Finally, the agent is given an overall score for the measuring performances in order to minimize the error and maximize forecast precision. This strategy, however, does not address prediction uncertainty.

You, *et al.* [11] used an improved CNN and LSTM to develop a deep learning system for crop yield prediction. Even though the labeled training data were sparse, the mean field approximation dimensionality reduction strategy was employed to train the CNN or LSTM network. Finally, a Deep Gaussian Process structure was employed to reduce correlated errors and make the spatio-temporal data structure model clear. On large datasets, however, this model is slow.

Hybrid neural networks were proposed by Lin, *et al.* [12] to provide a method for learning both local and global contextual variables to identify trends in time series. In order to extract the most prominent features from the local raw time series data, a CNN is utilized. By successfully including LSTM into the system, long-term dependencies in the history of prior trends are tracked. Then, a feature fusion layer is constructed to learn joint representation for trend prediction. However, the structure fails to accurately describe abruptly changing variables, which could affect trend progression.

Deep learning prediction methods like CNN, Bi-LSTM, and Gated Recurrent Unit (GRU) were developed by Milad, *et al.* [13] to predict asphalt pavement temperature (APT). The characteristics of asphalt change over air temperature level at different depths and times. The temperature level of the changing material is known as APT. GRU was developed by enhancing the structure of the LSTM. Additional structures were implemented with gates to modify and reset earlier information. However, the relationship between parameters is not considered in this method.

A hybrid CNN and RNN was proposed Li, *et al.* [14] for more in-depth hippocampal study, utilizing structural MRI for the purpose of evaluating the status of Alzheimer's disease. With this strategy, multi-level and multi-type features are learned simultaneously and their complementing information and illness categorization are enhanced. The local patches from the hippocampal region are used to form a deep DenseNet, which makes the network structure much simpler and helps training to proceed rapidly. At that point, bidirectional GRU (which is used to learn features that relate to the left and right hippocampus) is employed to make the disease classification more accurate. Nonetheless, it could not be proved that the identified features provide the ability to interpret brain areas involved in the sickness.

A new spatiotemporal optimization model was developed by Kousik, *et al*. [15] by combining CNN and RNN for video saliency detection. This model was designed in such a way that it reduces the computational load in analyzing video frames by aggregating layers from both CNN and RNN. The presented model computes dynamic video saliency model maps by analyzing information from temporal and spatial video frames. This method does not consider the gradient vanishing problem in RNN.

A new deep learning framework called YieldNet was proposed Khaki, *et al.* [16] by utilizing a novel CNN architecture. This model uses a transfer learning method between corn and soybean yield predictions by sharing the weights of the backbone feature extractor. The convolution operation in the YieldNet model captures both the temporal effect of data collected over the growing season and the spatial information of bins in histograms. A new loss function was developed to handle multiple response variables because of the uniqueness associated with simultaneous yield predictions for both crops. The shared backbone feature extractor in the YieldNet model substantially decreases the number of model parameters and subsequently helps the training process despite the limited labeled data.

In the above survey, a detailed review of crop yield prediction based on various improved techniques of CNN, RNN and a working structure for a hybridized CNN-RNN model was presented, including their limitations. According to this analysis, it is known that the hybrid CNN-RNN model is efficient for any type of crop yield prediction. The drawbacks of the CNN-RNN model identified from the survey were rectified in the proposed hybrid CNN-RNN model, which helps to achieve accurate predictions of rubber crop yield with different environmental factors.

3 Data Acquisition

The primary work of this research was to increase the accuracy of predicting the quality of yields of rubber crops planted in Kerala state. Therefore, the different parametric conditions for predicting rubber crop yields such as soil, rainfall, humidity, temperature and average wind speed were acquired from the India Metrological Department [17] and published data from the Rubber Institute Of India (RRII) [18], Kottayam. For the experimental analysis, the collected parametric conditions from the years 2018 to 2021 were evaluated. Datasets of these parameters were collected from different blocks in the state of Kerala district, such as Alappuzha, Ernakulam, Kannur, Idukki, Kasargode, Kollam, Kottayam, Kozhikode, Palakkad, Pathanamthitta Thiruvananthapuram, Thrissur, Wayanad for rubber crop yield prediction, optimizing the time spent on image dataset creation.

4 Proposed Methodology

In this section, the proposed methodology for rubber crop yield prediction using the hybrid CNN-RNN model is described in detail. The CNN network works with fixed-size inputs and produces fixed-size outputs, while RNN can deal with arbitrary input or output lengths. In this work, the hybrid CNN-RNN model enhances the prediction accuracy and reduces the computational complexity.

The proposed structure for rubber crop yield prediction includes a fusion element that combines CNNs, fully connected layers, and RNNs, as illustrated in Figure 1. The model's descriptions represented by the input variables in relation to the assigned weather parametric conditions are $X_1, X_2 \dots X_t$ at time phase t, and k represents the duration of time dependencies. The hybridization model was developed like the model proposed by Hu, *et al.* [19].

The proposed CNN was enhanced with different features like soil, rainfall, humidity, temperature and average wind speed, which are represented by the terms (s-CNN), (r-CNN), (h-CNN), (t-CNN), (a-CNN) respectively. The suggested CNN segments are used in a one-dimensional CNN to capture the temporal and spatial dependencies of the weather environment in various locations. Significantly, these CNN models are used in various rubber crop field applications for effective accuracy prediction.

A fully connected layer (FC) is used to integrate the high-level features extracted by (s-CNN), (r-CNN), (h-CNN), (t-CNN), and (a-CNN), which also reduces the dimension of the CNN model's performance.

The RNN model is made up of LSTM cells and is used to estimate the rubber crop yield of a specific state for year t using datasets from years $t-k$ to t. Input into the cell consists of average yield data (across all states in the same year), management data, and the performance of the FC sheet, which derives essential features processed by the (s-CNN), (r-CNN), (h-CNN), (t-CNN), and (a-CNN) segments using weather parametric condition data.

CNN is used as a feature selection method to convert conditional datasets $\{X_1, X_2, \ldots, X_t\}$ into feature vectors $\{F_1,\ F_2, \ldots, F_t\}$. The CNN platform is made up of seven layers. Following two locally connected layers, the first two layers are convolutional layers with $64 \times 3 \times 3$ kernels. The locally bound layer with $64 \times 1 \times 1$ kernels is used to obtain local features of the rubber yield.

For each of the layers, batch normalization by Ioffe & Szegedy [20] is used to reduce the internal covariate transfer shift. The final three layers are all fully connected layers with batch normalization added to the first two fully connected layers. Each RNN unit has a dropout with a probability of 0.5 and 512 hidden units for the sequence modeling level, followed by a G-way fully connected layer and a Softmax classifier. G denotes the number of rubber yield details that must be remembered. The final label is determined by averaging the Softmax outputs.

The RNN has a feedback loops and encodes temporal sequence contextual information. Given an input sequence of $\{F1, F2, \ldots, FT\}$, i.e. function vectors derived from a CNN model, the hidden states and h_t and y_t outputs can be computed as follows:

$$h_t = H(W_{ih}F_t + W_{hh}h_{t-1} + b_h) \tag{1}$$

$$y_t = W_{ho}h_t + b_o \tag{2}$$

where $W_{ih},\ W_{hh}, W_{ho}$ are weight matrices between the input, hidden and output layers respectively. As standardized RNN suffers from the gradient vanishing or exploding problem, so long short-term memory (LSTM) by Lipton, *et al.* [21] was used to alleviate this issue. Each LSTM unit is made up of an input gate, output gate, forget gate and cell, with the following measuring relationships among them:

$$i_t = \delta\ (W_i[\ h_{t-1}, F_t] +\ b_i) \tag{3}$$

$$f_t = \delta\ (W_f[\ h_{t-1}, F_t] +\ b_f) \tag{4}$$

$$o_t = \delta\ (W_o[\ h_{t-1}, F_t] +\ b_o) \tag{5}$$

$$\widehat{c_t} = \tan(W_c[\ h_{t-1}, F_c] +\ b_c) \tag{6}$$

$$c_t = f_t\ \odot\ c_{t-1} +\ i_t\ \odot c_t \tag{7}$$

$$h_t = o_t \odot \tanh(c_t) \tag{8}$$

where δ is a logistic sigmoid function and i, f, o and c are input, forget, output gate and cell activation, h_t is the output of the $t-th$ hidden unit of the RNN module, W_h and W_t are the weighted matrices, and r is the output of the hybrid module.

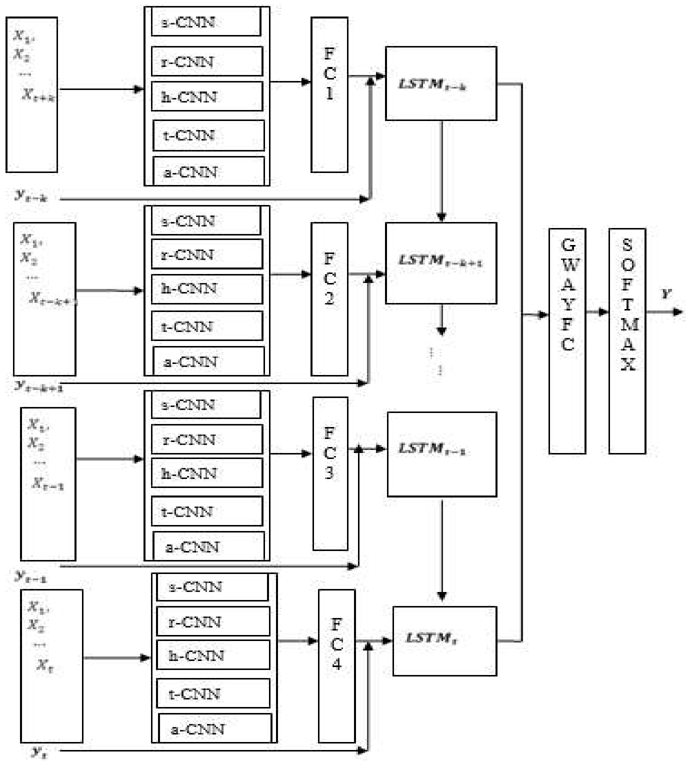

Figure 1 The structure of the proposed hybrid CNN-RNN model.

The given output ***r*** is followed by a G-way fully connected layer and a Softmax classifier.

$$r = \sum_{t=1}^{T} h_t \tag{9}$$

4.1 Loss Function

The attention-based hybrid CNN-RNN architecture in the loss function is defined as:

$$loss = \beta \,.\, loss_{target} + \lambda \,.\, \|w\|^2 \tag{10}$$

where the target replication loss is considered as the first label and the last label is the regularization term. The considered weight parameters $\beta \ and \ \lambda$ are:

$$loss_{target} = \frac{1}{T} \sum_{t=1}^{T} l(g_2(X_t), y) \tag{11}$$

$$l(g_2(X_t), y) = -\sum_{t=1}^{G} l_t \ (y) log(g_2(X_t)^t \tag{12}$$

$$g_2(X_t) = f_s\big(f_h(X_t)\big) \tag{13}$$

where X is the given dataset, y is the ground-truth label, and T is the number of time steps in RNN. G is the number of gestures, X_t is the th-th sub-segment of X, $g_2(X_t)$ and the indicator function is defined as l_t , $g_2(X_t)^{\ i}$ is the i-th dimension of $g_2(X_t)$, and f_h and f_s stand for the hybrid CNN-RNN architecture and the last Softmax layer.

5 Result and Discussion

The proposed fusion of a convolutional neural network (CNN) and a recurrent neural network (RNN), i.e. the proposed hybrid CNN-RNN for rubber crop yield prediction, was analyzed in terms of accuracy, precision, recall, and F-measure. The experimental performance showed that the proposed method produced better result compared with an existing method, i.e. an artificial neural network (ANN).

The classification accuracy was evaluated by comparing the actual and the predicted yield of rubber in the given dataset. Matches and mismatches between the predicted and the actual rubber crop yield in the given datasets were defined as follows:

1. *TP (True Positive)* – The number of samples with the yield quality label high, low, medium was predicted as high, low and medium, respectively.
2. *FP (False Positive)* – The number of samples with the yield quality label medium was predicted as high/low.
3. *FN (False Negative)* – The number of samples with the yield quality label low was predicted as high/medium
4. *TN (True Negative)* – The number of samples with the yield quality label high was predicted as medium/low.

5.1 Accuracy

The accuracy of the classifiers was defined by measuring how many instances are correctly predicted (crop yield as low, medium and high) among all instances. It was computed as follows:

$$Accuracy = \frac{TP + FP}{TP+TN+FP+FN}$$

Table 1 shows a comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of accuracy for respective years.

Table 1 Comparison of ANN and CNN-RNN in terms of accuracy.

Method/ Year	ANN	CNN-RNN
2018	75	77
2019	78	81
2020	81	85
2021	83	88

Figure 2 shows a comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of accuracy. The considered prediction of rubber crops for respective years are on the x-axis and the accuracy ranges are on the y-axis.

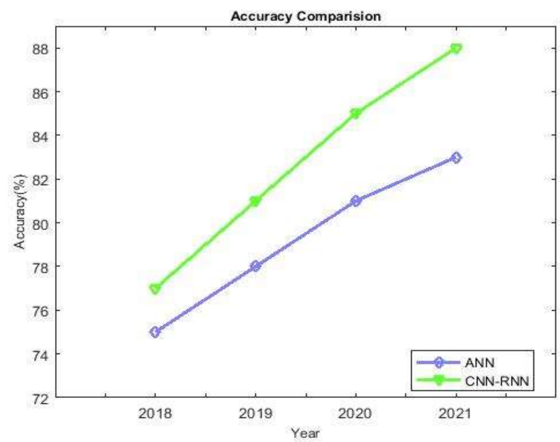

Figure 2 Comparison of ANN and CNN-RNN in terms of accuracy.

The accuracy of the proposed hybrid CNN-RNN structure was 2.67%, 3.84%, 4.93%, and 6.02% greater than that of the existing ANN model for prediction of rubber crop yields in the years 2018, 2019, 2020 and 2021, respectively. The result shows that the performance of the proposed CNN-RNN structure produced higher accuracy values compared to the existing method.

5.2 Precision

Precision is used to determine a classifier's ability to forecast only the appropriate rubber crop yield in a set of data. It is expressed as the percentage of accurately estimated rubber yields at TP and FP rates, or the proportion of actual positives and predicted rubber yields.

$$precision = \frac{TP}{TP+FP}$$

Table 2 shows the comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of precision with respect to respective years.

Table 2 Comparison of ANN and CNN-RNN in terms of precision.

Method/ Year	ANN	CNN-RNN
2018	80	82
2019	83	84
2020	86	88
2021	88	89

Figure 3 shows a comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of precision. The considered rubber crop yield predictions for respective years are on the x-axis and the precision ranges are on the y-axis. The precision of the proposed CNN-RNN structure was 2.5%, 1.20%, 2.32%, and 1.14% greater than that of the existing ANN model for predicting rubber crop yields in the years 2018, 2019, 2020 and 2021, respectively. The result shows that the performance of the proposed CNN-RNN structure produced higher precision values compared to the existing method.

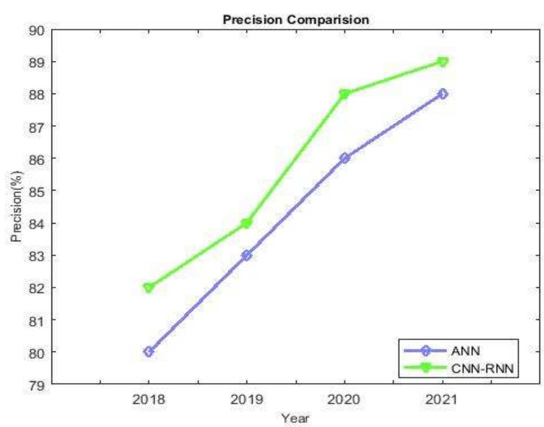

Figure 3 Comparison of ANN and CNN-RNN in terms of precision.

5.3 Recall

Recall can be used to assess a model's ability to identify each of the feature vectors of interest in a set of data. It is expressed as the ratio of accurately observed rubber yields at TP and FN rates, or the proportion of observed positive cases.

$$Recall = \frac{TP}{TP+FN}$$

Table 3 shows a comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of recall values for respective years.

Table 3 Comparison of ANN and CNN-RNN in terms of recall.

Method/ Year	ANN	CNN-RNN
2018	66	69
2019	69	73
2020	71	77
2021	75	85

Figure 4 shows a comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of recall. The considered prediction of years for rubber crops are on the x-axis and the recall ranges are on the y-axis. The recall of the proposed CNN-RNN structure was 4.54%, 5.97%, 8.45%, and 13.34% greater than that of the existing ANN model in predicting rubber crop yields in the years 2018, 2019, 2020 and 2021, respectively. The result shows that the performance of the proposed CNN-RNN structure produced higher recall values compared to the existing method.

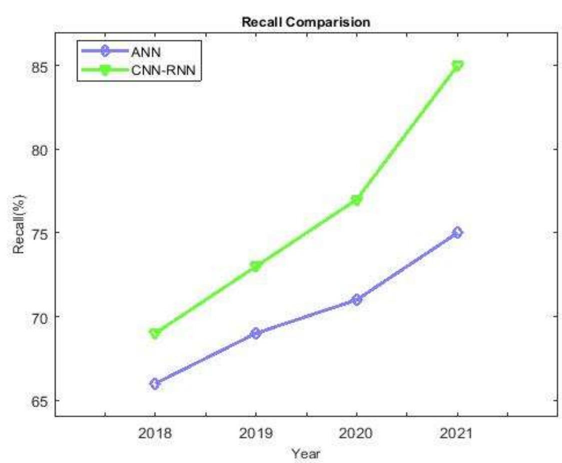

Figure 4 Comparison of ANN and CNN-RNN in terms of recall.

5.4 F-Measure

F-measure is the harmonic average mean of precision and recall when calculating the approximate prediction of rubber crop yields.

$$\text{F-measure}=\left(\frac{Precision \cdot Recall}{Precision+Rec}\right) \times 2$$

Table 4 shows the comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of F-Measure values for respective years.

Table 4 Comparison of ANN and CNN-RNN in terms of F-measure.

Method/ Year	ANN	CNN-RNN
2018	72	75
2019	76	78
2020	78	82
2021	80	86

Figure 5 shows a comparison of the existing ANN and the proposed hybrid CNN-RNN in terms of F-measure. The considered predictions for the yearly rubber crops are on the x-axis and the F-measure ranges are on the y-axis. The F-measure of the proposed CNN-RNN structure was 4.17%, 2.63%, 5.13%, and 7.5% greater than the existing ANN model for the prediction of rubber crop yields in the years 2018, 2019, 2020 and 2021, respectively. The proposed CNN-RNN structure produced higher F-measure values compared to the existing method.

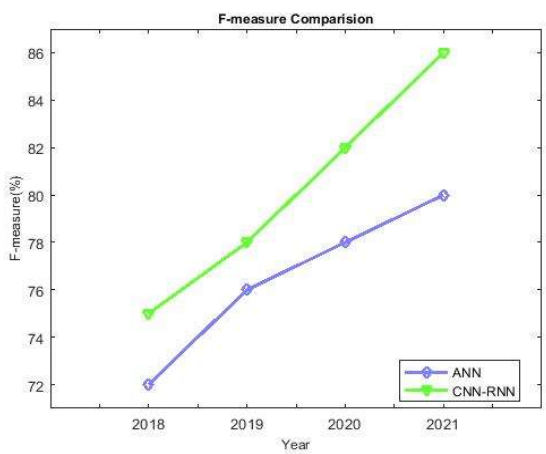

Figure 5 Comparison of the existing ANN and the proposed CNN-RNN in terms of F-measure.

6 Conclusion

Crop yield prediction is a critical task for decision-makers at the national and regional levels for rapid decision-making. A reliable crop yield prediction model can assist farmers in deciding what to grow and when to grow it. In recent years, various methods have been developed for the prediction of crop yield and its factors. Thus, deep learning methods play an important role in the prediction of crops. A rubber crop yield prediction method was developed by combining the structures of CNN and RNN (CNN-RNN) for various soil and weather conditions. Drawbacks in CNN and RNN were rectified to improve the prediction of Hevea (rubber). The CNN is capable of learning weight values by considering spatially dependent features from large numbers of data, solving the problem of internal covariate shift. Thus, the learned weight model is more accurate than that of a normal CNN. The temporal dependency for long sequences is maintained by the RNN after solving the gradient vanishing or exploding problem. Thus, the number of learning iterations is increased without any stuck until the minimum number of errors is obtained. Because of the higher learning accuracy, the prediction accuracy is also improved. The experimental results proved that the proposed method always performed better than the existing ANN in terms of all performance metrics.

In the future, deep transfer learning will be utilized for learning from different datasets collected from various states across India for rubber yield prediction. The growth rate of rubber trees can be estimated along with yield prediction by considering factors such as pesticides and fertilizers used.

References

[1] Rivano, F., Mattos, C.R., Cardoso, S.E., Martinez, M., Cevallos, V., Le Guen V. & Garcia, D., *Breeding Hevea Brasiliensis for Yield, Growth and SALB Resistance for High Disease Environments*, Industrial Crops and Products, **44**, pp. 659-670, 2013.

[2] Golbon, R., Ogutu, J.O., Cotter, M. & Sauerborn, J., *Rubber Yield Prediction by Meteorological Conditions Using Mixed Models and Multi-model Inference Techniques*, International Journal of Biometeorology, **59**(12), pp. 1747-1759, 2015.

[3] Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K. & Butler, E.E., *Random Forests for Global and Regional Crop Yield Predictions*, PloS One, **11**(6), 2016.

[4] Romero, J.R., Roncallo, P. F., Akkiraju, P.C., Ponzoni, I., Echenique, V. C. & Carballido, J.A., *Using Classification Algorithms for Predicting Durum Wheat Yield in the Province of Buenos Aires*, Comput. Electron. in Agric. **96,** pp. 173-179, 2013.

[5] LeCun, Y., Bengio, Y. & Hinton, G., *Deep Learning*, Nature, **521**(7553), pp. 436-444, 2015.
[6] Khaki, S. & Wang, L., *Crop Yield Prediction Using Deep Neural Networks,* Frontiers in Plant Science, **10,** 621, 2019.
[7] Nevavuori, P., Narra, N. & Lipping, T., *Crop Yield Prediction with Deep Convolutional Neural Networks*, Computers and Electronics in Agriculture, **163**, p. 104859, 2019.
[8] Sharma, S., Rai, S., & Krishnan, N.C., *Wheat Crop Yield Prediction Using Deep LSTM Model*, arXiv preprint arXiv:2011.01498, 2020.
[9] Khaki, S., Wang, L. & Archontoulis, S.V., *A CNN-RNN Framework for Crop Yield Prediction*, Frontiers in Plant Science, **10**, 1750, 2020.
[10] Elavarasan, D. & Vincent, P.D., *Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications*, IEEE Access, **8**, pp. 86886-86901, 2020.
[11] You, J., Li, X., Low, M., Lobell, D. & Ermon, S., *Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data,* in Proceedings of the AAAI Conference on Artificial Intelligence, **31**(1), pp. 455-4565, 2017.
[12] Lin, T., Guo, T. & Aberer, K., *Hybrid Neural Networks for Learning the Trend in Time Series*, Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp. 2273-2279, 2017.
[13] Milad, A., Adwan, I., Majeed, S.A., Yusoff, N.I.M., Al-Ansari, N. & Yaseen, Z.M., *Emerging Technologies of Deep Learning Models Development for Pavement Temperature Prediction*, IEEE Access, **9**, pp. 23840-23849, 2021.
[14] Li, F., Liu, M. & Alzheimer's Disease Neuroimaging Initiative, *A Hybrid Convolutional and Recurrent Neural Network for Hippocampus Analysis in Alzheimer's Disease*, Journal of Neuroscience Methods, **323**, pp. 108-118, 2019.
[15] Kousik, N., Natarajan, Y., Raja, R.A., Kallam, S., Patan, R. & Gandomi, A. H., *Improved Salient Object Detection Using Hybrid Convolution Recurrent Neural Network*, Expert Systems with Applications, **166**, 114064, 2021.
[16] Khaki, S., Pham, H. & Wang, L., *Yieldnet: A Convolutional Neural Network for Simultaneous Corn and Soybean Yield Prediction Based on Remote Sensing Data*, Arxiv Preprint Arxiv:2012.03129, 2020.
[17] India Meteorological Department of Kerala, mausam.imd.gov.in/Thiruvananthapuram (10 July 2021).
[18] For the Rubber Research Institute of India (Rubber-board), http://rubsis.rubberboard.org.in/app/index.html?lang=en (10 July 2021).
[19] Hu, Y., Wong, Y., Wei, W., Du, Y., Kankanhalli, M. & Geng, W.A., *Novel Attention-Based Hybrid CNN-RNN Architecture for SEMG-Based Gesture Recognition*, PloS One, **13**(10), e0206049, 2018.

[20] Ioffe, S., & Szegedy, C., *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, International Conference on Machine Learning, pp. 448-456, 2015.

[21] Lipton, Z.C., Kale, D.C., Elkan, C. & Wetzell, R., *Learning to Diagnose with LSTM Recurrent Neural Networks*, International Conference on Learning Representations, 2016.