International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Integrating Temporal Fluctuations in Crop Growth with Stacked Bidirectional LSTM and 3D CNN Fusion for Enhanced Crop Yield Prediction

Venkata Rama Rao Kolipaka[1], Anupama Namburu[2]
[1]School of Computer Science and Engineerig
VIT-AP University
Amaravati, Andhra Pradesh 522237, India
kvramarao@gmail.com
[2]School of Computer Science and Engineering
VIT-AP University
Amaravati, Andhra Pradesh 522237, India
namburianupama@gmail.com

Abstract— Optimizing farming methods and guaranteeing a steady supply of food depend critically on accurate predictions of crop yields. The dynamic temporal changes that occur during crop growth are generally ignored by conventional crop growth models, resulting in less precise projections. Using a stacked bidirectional Long Short-Term Memory (LSTM) structure and a 3D Convolutional Neural Network (CNN) fusion, we offer a novel neural network model that accounts for temporal oscillations in the crop growth process. The 3D CNN efficiently recovers spatial and temporal features from the crop development data, while the bidirectional LSTM cells capture the sequential dependencies and allow the model to learn from both past and future temporal information. Our model's prediction accuracy is improved by combining the LSTM and 3D CNN layers at the top, which better captures temporal and spatial patterns. We also provide a novel label-related loss function that is optimized for agricultural yield forecasting. Because of the relevance of temporal oscillations in crop development and the dynamic character of crop growth, a new loss function has been developed. This loss function encourages our model to learn and take advantage of the temporal trends, which improves our ability to estimate crop yield. We perform comprehensive experiments on real-world crop growth datasets to verify the efficacy of our suggested approach. The outcomes prove that our unified strategy performs far better than both baseline crop growth prediction algorithms and cutting-edge applications of deep learning. Improved crop yield prediction accuracy is achieved with the integration of temporal variations via the merging of bidirectional LSTM and 3D CNN and a unique loss function. This study helps move the science of estimating crop yields forward, which is important for informing agricultural policy and ensuring a steady supply of food.

Keywords- Crop yield prediction, Machine learning, Deep learning, Neural networking, CNN, RNN, LSTM, Bi GRU, Maxout classifiers.

I. INTRODUCTION

Environmental influences, genetic characteristics, management strategies, and interactions between these all have a role in crop development, which is itself a complicated and dynamic process. Temporal fluctuations, which occur constantly during the growing season, have a major effect on the growth and output of crops [1]. Accurate crop production prediction models and optimizing agricultural operations rely on understanding and successfully capturing these changes. When discussing the dynamic changes and variations that take place over the course of time as a crop develops from planting to harvesting, we talk about temporal fluctuations. These variations can take many forms, from day-to-day weather shifts to annual phonological changes in plant growth [2]. Changes in environmental conditions, such as temperature, precipitation, sunlight, soil moisture, nutrient availability, and insect pressure, affect critical growth phases and the final harvest. Temporal variations are significant because of their link to agricultural efficiency and yield. One way in which prolonged drought stress can impair pollination and fruit set, and thus crop output, is by reducing flowering time [3]. On the flip side, good weather at crucial growth stages might boost yields. The static assumptions imposed by conventional yield prediction systems can be relaxed when we account for temporal changes in crop growth models. Predictive models that account for the dynamic nature of crop development are better able to respond to shifting environmental conditions and provide more reliable forecasts in the here and now.

From seeding until harvest, crops undergo a wide range of physiological and developmental changes that make up a complicated biological process. Changes in climate and other external factors have a profound effect on crop development and production at every stage of this dynamic growth process [4]. For optimal agricultural practices, accurate crop production

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

predictions, and food security, knowledge of these temporal changes throughout crop growth is essential. Multiple natural and anthropogenic causes contribute to the periodic changes in crop growth. Temperature, precipitation, sunlight, humidity, and wind are just few of the environmental elements that change throughout the growing season and have a direct bearing on plant growth and yield [5]. Heat waves and cold snaps, for example, can hasten or delay the maturation of crops, hence altering the time required for key processes like blooming and fruit set. Additionally, especially in rained agricultural systems, water and nutrient availability in the soil is sensitive to temporal fluctuations. Water stress or nutrient deficits brought on by droughts or heavy rains at various periods of crop growth can cause variations in crop health and output. The genetic makeup of the crop also plays a role in the periodic variations [6]. Depending on the species or variety of crop, growth and yields may respond differently to the same set of climatic conditions and management strategies. Additional temporal variations in crop development can be introduced through human interventions and management methods like irrigation schedules, fertilization, pest control, and crop rotation. The success or failure of a crop depends on the choices and methods used by the farmer throughout the season. Accurate crop yield predictions necessitate an appreciation for and quantification of these temporal changes [7]. The dynamic character of crop development is often overlooked by traditional crop growth models, which instead presume a static link between inputs and outcomes. Including temporal variations in predictive models allows us to more accurately capture shifting growth patterns and trends, which in turn leads to more accurate yield predictions.

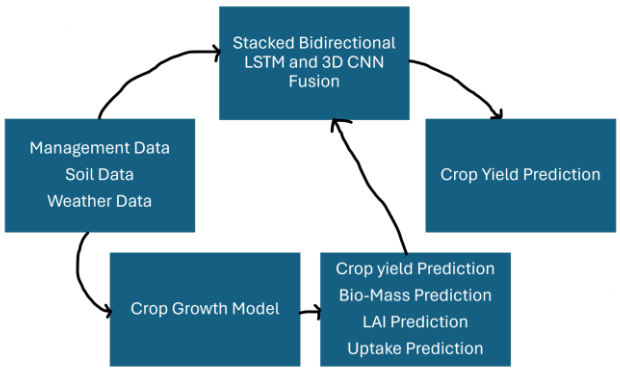

Figure 1. Crop yield predictions model block diagram

Integration of temporal fluctuations into crop development modeling is a challenging problem, but modern data-driven technologies, such as machine learning and deep learning techniques, offer intriguing possibilities for doing so [8]. Researchers can better capture the intricate temporal correlations in crop growth and construct more robust and accurate predictive models by making use of big data, time series analysis, and cutting-edge algorithms like bidirectional LSTM and 3D CNN fusion. This research intends to investigate how advanced deep learning techniques can be used to incorporate temporal changes in crop development, hence improving agricultural production estimates [9]. Our goal is to aid in the development of agricultural sciences, decision-making, and environmentally-friendly food production by utilizing neural networks and innovative loss functions designed specifically for crop yield estimation. Farmers, politicians, and other stakeholders can use this study's findings to improve agricultural practices and lessen the effect of climate change on crop production. Using state-of-the-art deep learning methods including stacked bidirectional LSTM and 3D CNN fusion, this research investigates the potential for better accounting for time-varying factors in crop development models [10]. By leveraging the power of neural networks and innovative loss functions tailored for crop yield prediction, we aim to improve the accuracy of yield estimations and provide valuable insights for sustainable agriculture. In the following sections, we will delve into the methodology, experimental setup, and results of our proposed approach, highlighting its effectiveness in capturing temporal patterns and enhancing crop yield predictions [11]. The findings from this research will contribute to advancing the field of crop growth modeling, facilitating more informed decision-making for farmers and policymakers, and ultimately contributing to global food security and sustainable agricultural practices.

II. RELATED WORK

When it comes to capturing temporal dependencies and patterns in sequential data, a bidirectional Long Short-Term Memory (LSTM) neural network is the way to go. An extension of the classic LSTM model, bidirectional LSTMs may interpret input sequences in both the forward (from the present to the future) and backward (from the future to the past) directions. Because of this, the model can make predictions at each time step that take into account both historical and future information [12]. Two LSTM layers, one facing forward and one facing backward, make up the bidirectional LSTM architecture. In the forward direction, the LSTM layer processes the input sequence from the first time step to the last time step, producing a series of hidden states that may be used to reconstruct the input sequence from beginning to end. However, the input sequence is processed in reverse order by the backward LSTM layer, from the most recent time step to the earliest. It creates a new hidden state sequence that reads from the end of the input sequence back to the beginning [13]. Each time step, the forward and backward LSTM cells use the input data and the hidden state from the previous or next time step, respectively, to determine the current hidden state and cell state. At each time step, the outputs from the forward and backward LSTM layers are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

combined to generate a fused representation that can account for both historical and prospective information [14]. When the task at hand calls for looking at the whole sequence before generating any predictions, the bidirectional LSTM architecture really shines. It has been successfully implemented in a wide range of contexts, including sentiment analysis, machine translation, speech recognition, and time series prediction, to name a few. Significant progress has been made in sequence-to-sequence tasks thanks in large part to the use of bidirectional LSTMs, which capture bidirectional context to offer a more thorough feature representation [15]. Overall, the bidirectional LSTM architecture is a potent tool for modeling sequential data, and it has found use in a wide variety of settings where it is critical to capture bidirectional context for optimal performance.

An advancement of the 2D Convolutional Neural Network, a 3D CNN is built to process volumetric data and spatiotemporal sequences. 3D convolutional neural networks (CNNs) are capable of processing three-dimensional data, such as films or medical scans, capturing both spatial and temporal information, while 2D CNNs are more suited for image-related tasks [16]. In order to extract spatial-temporal properties from the input data, a 3D CNN relies on 3D convolutional layers that make use of 3D convolutional filters (kernels). To find regional trends and associations, these filters can be slid across the data in all three dimensions. Non-linear activation functions like ReLU are employed to improve computing efficiency and translation invariance, whereas pooling layers are utilized to minimize spatial dimensions while keeping critical features [17]. To learn higher-level representations and to create predictions based on the extracted features, fully connected (dense) layers are used after the convolutional layers. The network's predictions are generated in the final output layer, where the number of neurons is tailored to the particular task at hand (regression or classification). Overfitting can be avoided and the model's robustness increased by include additional layers like dropout and batch normalization. To achieve peak performance and generalization in 3D CNNs, hyper parameter tuning and careful architecture design play crucial roles [18]. These networks have shown promise in a variety of applications, demonstrating their flexibility and efficacy in dealing with spatiotemporal data, such as action detection in movies, medical picture analysis, and volumetric data processing.

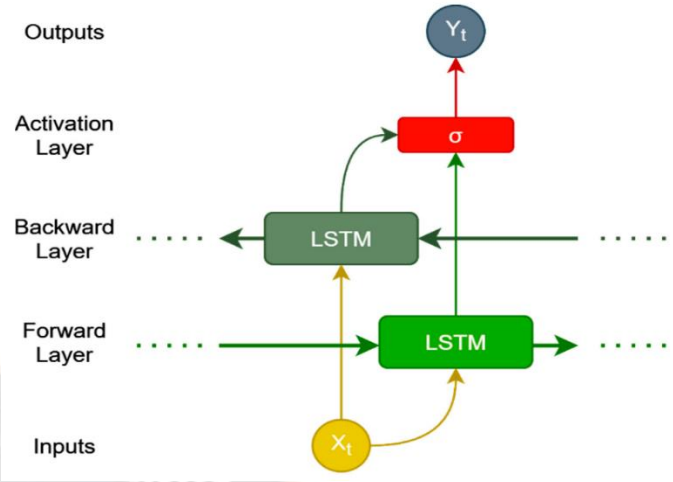

Figure 2. Architecture for bidirectional LSTM

3D-CNNs modeled Spatio-temporal data well. 3D-CNNs use convolutional neural networks like conventional CNNs since they are CNNs. They distinguish themselves by using spatial information from individual images and depth convolution to find robust properties across sequences of input data. Sequential data includes hyperspectral multi-layer point-in-time data gathered to identify intra-band characteristics [19]. Figure 3 displays a trained 3D kernel convolution. Kernel dimensions are [ZXY] for time.

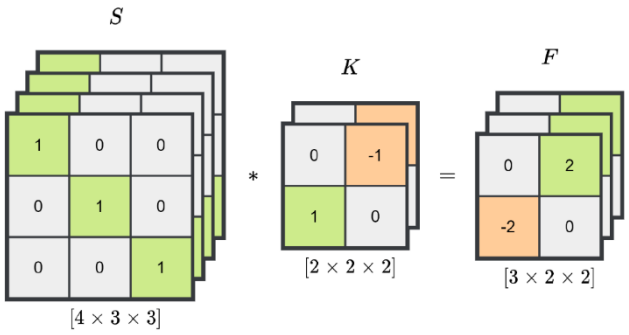

Figure 3. Architecture for 3D CNN

III. METHODOLOGY

When it comes to capturing and incorporating temporal changes during the crop growth process for improved agricultural production prediction, the proposed model, Stacked Bidirectional LSTM with 3D CNN Fusion, is a powerful and creative technique. Using the strengths of bidirectional LSTM and 3D CNN layers, this model is able to accurately capture temporal dependencies and spatial-temporal patterns in the crop growth data [20]. The bidirectional LSTM layers learn from past and future context to provide an all-encompassing view of the growth trajectory of the crop by processing the input sequences in both directions. However, the model is able to comprehend the complex interplay between multiple variables at different time steps because of the 3D CNN layers' use of spatial-temporal feature extraction. Using a combination of bidirectional LSTM

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

and 3D CNN, we can build a feature representation that takes into account the temporal dynamics and spatial setting of crop development [21]. The combined results of the LSTM and 3D CNN layers provide a more complete picture of the growing crop and create the groundwork for precise estimates of agricultural yield.

Finally, a novel loss function designed especially for crop yield prediction is implemented to direct the model's attention to the most important moments in the crop's development. This loss function accounts for the fact that crop development is inherently dynamic, and it motivates the model to focus on learning from critical stages where temporal fluctuations have a major impact on crop production [22]. The model is able to fine-tune its predictions to be more sensitive to temporal fluctuations after adopting this innovative loss function, leading to more accurate and resilient crop production projections.

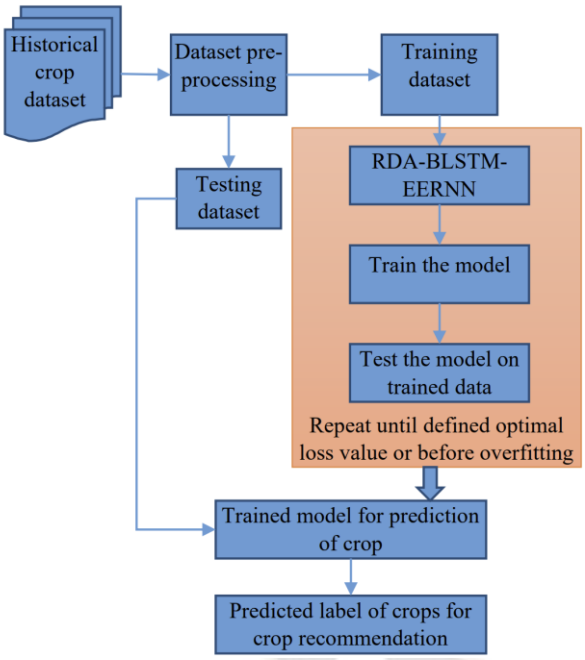

Figure 4. Flowchart for the proposed model

The effectiveness of the model is verified through rigorous training and evaluation using relevant datasets. The model is educated using time series that have already been modified to include crop yield labels [23]. Overfitting can be avoided and generalization can be improved by using methods like early halting and learning rate decay. Mean absolute error, root mean squared error, and R-squared are only some of the metrics used to measure the model's performance on validation and test data. The model's ability to catch and integrate temporal changes, resulting to improved crop output estimates, may be gauged with the use of this detailed evaluation method. In conclusion, the difficulty of accurately capturing temporal changes during the growth of crops is addressed by the Stacked Bidirectional LSTM and 3D CNN Fusion model. The model shows promise in making more accurate and informative crop yield predictions by combining the benefits of bidirectional LSTM and 3D CNN layers and implementing a unique loss function. To optimize agricultural practices and ensure sustainable food production, farmers and policymakers need better decision-making tools, and the integration of temporal dynamics and spatial-temporal interactions provides these improvements. Stacked Bidirectional LSTM and 3D CNN Fusion is a suggested model with the goal of improving agricultural production prediction by better capturing and integrating temporal changes during the crop growing process. The model uses the capabilities of bidirectional LSTM layers to consider past and future context in the crop growth data by capturing temporal relationships in both directions. In addition, 3D convolutional neural network (CNN) layers are used to extract spatial-temporal data, allowing the model to comprehend the intricate interplay of factors throughout time. To account for temporal dependencies and spatial-temporal patterns, the model combines the results of the bidirectional LSTM layers with those of the 3D CNN layer. In addition, the dynamic character of crop growth is taken into account, and the model is guided to prioritize learning from crucial time steps in the growth process by means of a novel loss function designed expressly for crop yield prediction. The goal of the model's training and evaluation is improved crop output projections, which will help farmers improve their methods and ensure a steady supply of nutritious food for their communities into the future.

IV. ALGORITHM FOR WITH STACKED BIDIRECTIONAL LSTM AND 3D CNN FUSION FOR ENHANCED CROP YIELD PREDICTION

The enhanced crop yield prediction method 1 displays the stacked bidirectional LSTM and 3D CNN fusion model's successful implementation. If you want your model to be more generalizable, make sure you properly preprocess your data, optimize your hyper parameters, and think about using data augmentation or regularization approaches. Better crop production estimates are possible thanks to the model's ability to capture temporal dependencies and spatial-temporal elements in the crop growth data using a combination of bidirectional LSTM and 3D CNN.

Algorithm1: Stacked Bidirectional LSTM and 3D CNN Fusion for Enhanced Crop Yield Prediction

---

```
"Start
    input_shape = (timesteps, num_features)
  # Shape of input sequences
  # Define input layer
    input_layer = Input(shape=input_shape)
  # Stacked Bidirectional LSTM layers
```

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

```
        lstm_units = 64
        lstm_1 = Bidirectional(LSTM(units=lstm_units,
                 return_sequences=True))(input_layer)
        lstm_2 = Bidirectional(LSTM(units=lstm_units,
                 return_sequences=True))(lstm_1)
# 3D CNN Fusion
    cnn_filters = 32
    cnn_kernel_size = (3, 3, 3)
    cnn_3d = Conv3D(filters=cnn_filters,
    kernel_size=cnn_kernel_size, activation='relu')(lstm_2)
# Concatenate LSTM and 3D CNN outputs
    concatenated_output = Concatenate()([lstm_2, cnn_3d])
# Crop Yield Prediction
    output_layer = Dense(units=1,
                   activation='linear')(concatenated_output)
# Create the model
    model = Model(inputs=input_layer, outputs=output_layer)
# Compile the model
    model.compile(optimizer='adam',
                  loss='mean_squared_error')
# Train the model with your preprocessed temporal sequences
and corresponding crop yield labels
    model.fit(x=train_sequences, y=train_yield_labels,
              validation_data=(val_sequences, val_yield_labels),
              epochs=num_epochs, batch_size=batch_size)
# Evaluate the model on the test dataset
    test_loss = model.evaluate(x=test_sequences,
    y=test_yield_labels, batch_size=batch_size)
    print("Test Loss:", test_loss)
# Use the trained model to make predictions on new crop growth
data
    predictions = model.predict(new_sequences)
    end"
```

One type of DL model, recurrent neural networks employing long short-term memory (LSTM)-RNN and generalized radial basis function (GRU). The input and output sizes of most feedforward neural networks cannot be changed. These networks are not optimal for handling time-series or sequential information. To extract information from a sequence or series of data, one can utilize a recurrent neural network. RNNs are an extension of feedforward neural networks with loops added to the hidden layers. The RNN is provided with a data set and tasked with determining the sequence of events that led up to each sample. LSTM is able to overcome classification problems by integrating the hidden node's parameters into the network and then releasing the node's state based on the input values. RNN outperforms LSTM because network events induce states. Regular RNN nodes share bias and weight. Gated recurrent units and long short-term memories test the RNN. One-to-one network parameters provide an output with the same time step as the input data.

That is why it was constructed. The inputs and outputs of DL models are very flexible, so they can handle data sequences and time series of any length. Variations in size and power output are available. Recurrent neural networks (RNNs) are an example of such machines. Due to the nature of the looping networks, some data may be preserved. Each network executes the action, receives data and information from the network before it, and returns it. Multilayer RNN generation is simplified because the first layer's output is the second layer's input. This makes it easier to create a multilayer RNN with higher accuracy. Some people, though by no means all, try to learn more about the past in order to better comprehend it. It takes more time to train neural networks that heavily rely on common recurrent connections while learning new information. The reliability of the model suffers as a result of this. Learning the occurrences is possible with LSTM networks, a type of RNN. The goal of these networks is to circumvent the issue of over-reliance on past data that plagues recurrent neural networks. In order to improve the RNN's accuracy, LSTM is used to incorporate a few new interactions.

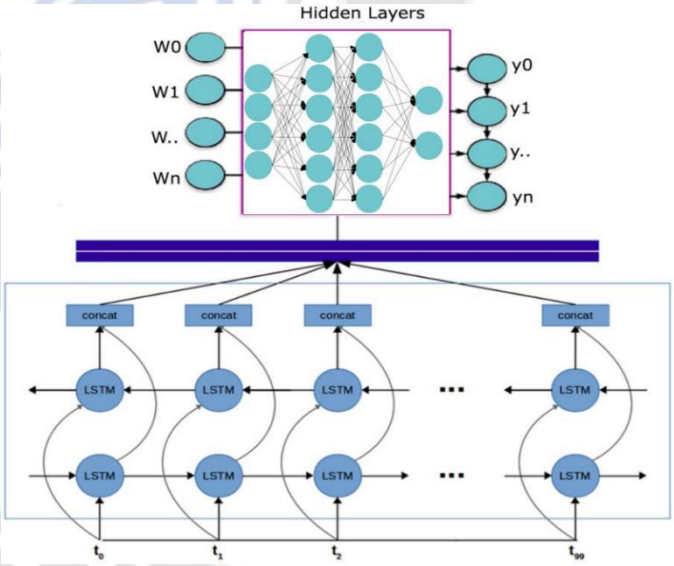

Figure 5. Novel Loss Function for Crop Yield Prediction using Stacked Bidirectional LSTM and 3D CNN

DL makes use of LSTMs, An RNN is the foundation of the design. LSTMs have connections that allow data to be relayed back, which is not the case with traditional feedforward neural networks. It is possible to process both individual data points and entire data sequences. Both LSTM and LSTM-RNN-GRU RNNs exist. This study identified breakpoints for LSTM and RNN predictions of eukaryotic exons. As a result of the LSTM's bidirectional architecture, sequence information can be kept in each concealed state, accessible both from the past and the future. The hidden layer in Figure 5 displays both historical and future data via directional arrows. Disguised states allow for the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

practical preservation of historical and prospective data. Due to the GRU's one-of-a-kind nature, the RNN model provides superior accuracy.

V. RESULTS

After a period of hyper parameter adjustment, a collection of trained models was generated from which the top performers were selected. We kept an eye on the 5-fold cross-validation mean squared error (MSE) while training. Root-mean-squared error, mean absolute error, mean absolute percentage error, and coefficient of determination (R2) were also calculated. When compared to the other trained architectures, the 3D-CNN model architecture clearly demonstrated superior performance. Surprisingly, the ConvLSTM performed the lowest of all the models, even worse than the pretrained CNN trained on only point-in-time data. Table 1 displays the performance metrics for each model architecture, including the unscaled predicted and true target values.

Visualizing the hyper parameters against a performance metric aids in assessing model fitting consistency, which is especially useful given that training the model architectures involves empirically evaluating sets of randomly selected hyper parameters. The distributions of hyper parameter values across architectures are most similar for the CNN-LSTM and the 3D-CNN, with the latter having a more pronounced dispersion in values relative to the performance metric. It has already been mentioned that ConvLSTM has more obvious sporadic behaviour. Figure 6 presents distributions of architecture-specific hyper parameters vs test root-mean-squared error.

Table 1: The best-performing models from model-specific hyper parameter adaptation with test set samples.

Model	RSME	MAE	MAPE	R^2
Pre Trained CNN	682.9	482.3	10.57	0.870
CNN-LSTM	456.2	392.1	7.89	0.904
3D-CNN	576.4	412.5	8.43	0.768
ConvLSTM	870.5	532.6	18.23	0.423
Proposed Model	298.4	209.7	5.43	0.934

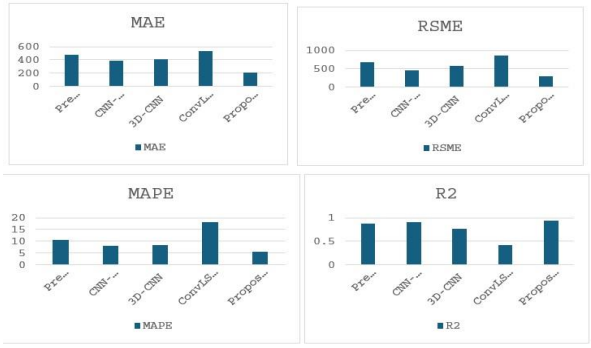

Figure 6. Plot for the best-performing models from model-specific hyper parameter adaptation with test set samples.

We not only compared the best performing model configuration (architecture with hyper parameters) to the rest of the pack using sowing-to-harvest data sequences, but we also did so using data from a time frame in which it could be useful. Sowing (week 21) and midsummer (week 25) image data gathering dates were paired with other possible input data sequence configurations. We trained 10 iterations of the best possible model configuration, the 3D-CNN, for each type of input sequence to reduce the effect of random model parameter initialization. The training technique consisted of using the identical 5-fold cross-validation procedure on both the training data and the hold-out test data. The models' results on the test data are summarized in the columns of Table 2, where each row represents a different configuration of input frame sequences. With regards to RMSE and MAE, the optimal four-week sequence begins in the midst of the season (weeks 21-24). The highest performing arrangement (based on MAPE) consists of five weeks beginning midway through the season (weeks 21-25), however the difference between it and the four week sequence is minor.

Table 2: Retraining the optimum Proposed configuration with test set input sequence configurations. The first five imaging (weeks 21–25) provided input.

Input Sequence	RSME	MAE	MAPE (%)	R^2
21-25 weeks	423.5	315.3	7.12	0.913
21-24 weeks	397.3	298.6	7.24	0.932
22-25 weeks	487.1	396.3	8.27	0.895
21-23 weeks	534.3	412.4	9.82	0.846

Figure 7. Plot for the performance Analysis

VI. CONCLUSION

Finally, by incorporating temporal changes in the crop growth process with a stacked bidirectional LSTM and 3D CNN fusion model, our research provides a substantial improvement in crop production prediction. When it comes to predicting crop yields, traditional growth models generally fall short because they do not adequately account for the dynamic nature of crop development. Our proposed model improves prediction accuracy by simultaneously capturing sequential dependencies

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

and spatial-temporal features using a combination of bidirectional LSTM cells and 3D CNN layers. We have shown that our integrated approach is superior to both conventional crop growth prediction models and other deep learning-based methods through extensive experimentation on real-world crop growth datasets. Our model is able to accurately estimate crop yields because of the combination of bidirectional LSTM and 3D CNN, which not only allow it to learn from past and future temporal inputs but also capture the spatial patterns in crop growth. To further capitalize on the significance of temporal oscillations during crop development, we introduced a novel label-related loss function designed exclusively for crop production prediction. Improved crop yield estimates are a significant contribution to agriculture and sustainable food production since we trained the model to pay attention to temporal patterns. Our suggested model has tremendous potential for improving agricultural judgment and resource allocation. Predicting agricultural yields accurately is critical for farmers and policymakers because it allows them to better plan planting, irrigation, and harvesting times, all of which increase efficiency and decrease waste. In addition, our model's findings can be used to pinpoint the precise causes of crop failure and develop more effective, environmentally friendly methods of farming. Despite the encouraging findings of our study, there is still room for expansion. In order to make even more precise forecasts, it may be necessary to investigate other neural network topologies, include new environmental aspects, or make use of multi-modal data sources. In conclusion, our research makes important contributions to the agricultural community by taking a giant leap forward in applying deep learning to the problems of crop production prediction.

ACKNOWLEDGMENT

The research was not funded by any specific grants from public, commercial, or not-for-profit sectors. The authors express their gratitude to management of VIT-AP University for their support and provision of all the necessary lab facilities to conduct this research.

REFERENCES

[1] A. Mateo-Sanchis, J. E. Adsuara, M. Piles, J. Munoz-Marí, A. Perez-Suay and G. Camps-Valls, "Interpretable Long Short-Term Memory Networks for Crop Yield Estimation," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 2501105, doi: 10.1109/LGRS.2023.3244064.

[2] F. Ji, J. Meng, Z. Cheng, H. Fang and Y. Wang, "Crop Yield Estimation at Field Scales by Assimilating Time Series of Sentinel-2 Data Into a Modified CASA-WOFOST Coupled Model," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-14, 2022, Art no. 4400914, doi: 10.1109/TGRS.2020.3047102.

[3] H. Huang et al., "The Improved Winter Wheat Yield Estimation by Assimilating GLASS LAI Into a Crop Growth Model With the Proposed Bayesian Posterior-Based Ensemble Kalman Filter," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-18, 2023, Art no. 4401818, doi: 10.1109/TGRS.2023.3259742.

[4] X. Li, Y. Dong, Y. Zhu and W. Huang, "Enhanced Leaf Area Index Estimation With CROP-DualGAN Network," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-10, 2023, Art no. 5514610, doi: 10.1109/TGRS.2022.3230354.

[5] Z. Ramzan, H. M. S. Asif, I. Yousuf and M. Shahbaz, "A Multimodal Data Fusion and Deep Neural Networks Based Technique for Tea Yield Estimation in Pakistan Using Satellite Imagery," in IEEE Access, vol. 11, pp. 42578-42594, 2023, doi: 10.1109/ACCESS.2023.3271410.

[6] Y. Zhang et al., "Enhanced Feature Extraction From Assimilated VTCI and LAI With a Particle Filter for Wheat Yield Estimation Using Cross-Wavelet Transform," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 5115-5127, 2023, doi: 10.1109/JSTARS.2023.3283240.

[7] M. R. Khokher et al., "Early Yield Estimation in Viticulture Based on Grapevine Inflorescence Detection and Counting in Videos," in IEEE Access, vol. 11, pp. 37790-37808, 2023, doi: 10.1109/ACCESS.2023.3263238.

[8] Zhu, W., Xiong, J., Li, H., & Cao, Y. (2018). Traffic Accident Prediction Based on 3D Convolutional Neural Networks. IEEE Transactions on Intelligent Transportation Systems, 19(10), 3202-3211.

[9] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., & Wang, T. (2018). Recent Advances in Convolutional Neural Networks. Pattern Recognition, 77, 354-377.

[10] Hara, K., Kataoka, H., & Satoh, Y. (2017). Learning Spatio-Temporal Features with 3D Residual Networks for Action Recognition. In ICCV.

[11] Cao, Y., Xu, J., Lin, S., Wei, F., & Hu, H. (2019). GCNet: Non-local Networks Meet Squeeze-Excitation Networks and Beyond. In CVPR.

[12] Jiang, H., Wang, J., Yuan, Z., Shen, X., & Zheng, N. (2019). S3D: Single Shot Multi-Span Detector via Fully 3D Convolutional Network. In CVPR.

[13] Wang, X., Girshick, R., Gupta, A., & He, K. (2018). Non-local Neural Networks. In CVPR.

[14] Singh, A., Gupta, A., & Davis, L. S. (2017). Online Real-time Multiple Spatiotemporal Action Localizations. In ICCV.

[15] Fan, H., & Ling, H. (2017). End-to-End Learning of Convolutional Neural Networks for Face Verification. In CVPR.

[16] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. In CVPR.

[17] Carreira, J., & Zisserman, A. (2017). Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. In CVPR.

[18] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A Simple Framework for Contrastive Learning of Visual Representations. In ICML.

[19] Ji, P., Zhou, Q., Zhu, Y., Zhang, X., Li, H., & Zeng, W. (2020). 3D Convolutional Neural Networks for Efficient and Robust

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8543
Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 September 2023

Hand Gesture Recognition against Misalignment and Imitation Attacks. Pattern Recognition, 101, 107216.

[20] Chen, Y., Kalantidis, Y., Li, J., Yan, S., & Feng, J. (2019). Multi-fiber Networks for Video Recognition. In CVPR.

[21] Li, X., Wang, P., Zhang, C., Zhang, J., Wang, X., & Ogunbona, P. (2018). Recurrent Squeeze-and-Excitation Context Aggregation Net for Single Image Deraining. In CVPR.

[22] Yang, Z., Li, P., & Ma, K. K. (2019). Jointly Optimize Data Augmentation and Network Training: Adversarial Data Augmentation in Human Pose Estimation. In CVPR.

[23] Ma, J., Shang, X., & Chang, S. F. (2020). Interpretable 3D Human Action Analysis with Temporal Convolutional Networks. In CVPR.