Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi

Afyon Kocatepe University Journal of Science and Engineering

AKÜ FEMÜBİD **23** (2023) 035101 (648-660) AKU J. Sci. Eng. **23** (2023) 035101 (648-660)

DOI: 10.35414/akufemubid.1116187

Araştırma Makalesi / Research Article

Hybrid Deep Learning Implementation for Crop Yield Prediction

Halit ÇETİNER[1]

[1] *Isparta University of Applied Sciences, Vocational School of Technical Sciences, Isparta*

e-posta: halitcetiner@isparta.edu.tr *ORCID ID: http://orcid.org/0000-0001-7794-2555*

Geliş Tarihi: 13.05. 2022 Kabul Tarihi: 13.06.2023

Keywords
Crop; Yield estimation; CNN; LSTM; Hybrid model

Abstract

Agriculture producers should be supported technologically in order to continue production in a way that meets the worldwide food supply and demand. Automatic realization of crop yield estimation calculation is a desired need of farmers. Automatic yield estimation also facilitates the work of agricultural producers with different goals such as imports and exports. To achieve the stated objectives, deep learning models have been developed that estimated yield using parameters such as the amount of water per hectare, the average amount of sunlight received by the hectare, the amount of fertilization per hectare, the number of pesticides used per hectare, and the area of cultivation. With the hybrid model created by combining the strengths of the LSTM and CNN models developed within the scope of this article, the success rate of data prediction has increased with fine adjustments. Success rates of 89.71 R^2, 0.0035 MSE, 0.0248 RMSE, 0.0461 MAE, and 10.10 MAPE have been achieved with the Proposed hybrid model. This model is competitive with similar studies with the stated values.

Mahsul Verim Tahmini için Hibrit Derin Öğrenme Gerçekleştirimi

Anahtar kelimeler
Mahsul; Verim tahmini; CNN; LSTM; Hibrit model

Öz

Tarım üreticilerinin dünya çapındaki gıda arz ve talebini karşılayacak şekilde üretime devam edebilmesi için teknolojik olarak desteklenmesi gerekmektedir. Mahsul verim tahmini hesaplamasının otomatik olarak gerçekleştirilmesi, çiftçilerin arzu ettiği bir ihtiyaçtır. Otomatik olarak verim tahmini gerçekleştirilmesi ithalat ve ihracat gibi farklı hedefleri olan tarım üreticisinin işlerini de kolaylaştırmaktadır. Belirtilen amaçlara ulaşabilmek için hektar başına su miktarı, hektar tarafından alınan ortalama güneş ışığı miktarı, hektar başına verilen gübreleme miktarı, hektar başına kullanılan pestisit miktarı, ekim yapılan alan bölgesi parametrelerini kullanarak verim tahmini gerçekleştiren derin öğrenme modelleri geliştirilmiştir. Bu makale kapsamında geliştirilen LSTM ve CNN modellerinin güçlü yanları birleştirilerek oluşturulan hibrit modelde ile veri tahmin başarı oranının ince ayarlamalar ile artırılmıştır. Önerilen hibrit model ile 89.71 R^2, 0.0035 MSE, 0.0248 RMSE, 0.0461 MAE, ve 10.10 MAPE başarı oranlarına ulaşılmıştır. Bu model, belirtilen değerlerle benzer çalışmalarla rekabet edebilir seviyededir.

1. Introduction

Developing countries, their income sources need to be fed and developed. At this point, agriculture is one of the important sources of income for most countries around the world. Developments in agriculture not only improve the sustainability of nutrition and the food supply chain but also increase development.

For the sustainability of the food supply and supply chain to continue, agricultural activities must continue uninterruptedly in a certain order and stability. Although this necessity is possible in many non-agricultural areas, it is not possible due to environmental and natural factors affecting production in agriculture. In addition to these, many factors affect the production of crops in a wide range, from the flat or handicapped area of cultivated land to the hot or cold planting weather.

It is seen that scientists doing academic research in this field around the world are researching different methods of crop productivity prediction (Asseng *et al.* 2017, Cao *et al.* 2021, Jeong *et al.* 2016, Vanli, Ustundag, Ahmad, Hernandez-Ochoa, and Hoogenboom 2019).

Food insecurity is increasing day by day in the world population (FAO 2017). In the current situation, it is expected that the population whose food supply and supply continues to increase by another two billion people in approximately thirty years (Cao *et al.* 2021, Dodds and Bartram 2016). With the increase in population, there will be an increase in food demand in terms of the sustainability of the food supply (Gorelick *et al.* 2017). A large proportion of the world's food demand is provided by wheat (Vanli, Ahmad, and Ustundag 2020). Wheat productivity must be calculated correctly to meet the food supply and need according to the growing population. There is a decrease in wheat cultivation areas due to different reasons (Deutsch *et al.* 2018). As a result of the decrease in cultivation, it is necessary to increase the production that will provide the increased consumption supply. It is seen that the product warehouses of different countries are empty (Chen, Zhang, Tao, Wang, and Wei 2017). For the reasons stated, crop productivity must be achieved with the least error to protect the profits of crop producers, maintain the global food supply, and protect the interests of the producers.

Environmental difficulties can cause irregularities in temperature due to climate change. It is necessary to reduce environmental irregularities and uncertainties for the continuation of crop production without affecting food safety (Ahmad *et al.* 2018, Nasim *et al.* 2018). Global temperatures are expected to increase by a few degrees in the next quarter-century. It is predicted that this temperature increase will adversely affect crop production (Ben-Asher, Yano, Aydın, and Garcia y Garcia 2019, Nasim *et al.* 2018). The fluctuation between seasons is considered a harbinger of this situation (Ahmad, Wajid, Ahmad, Cheema, and Judge 2019, Asseng *et al.* 2017). The increase in temperature, which is predicted to adversely affect crop production, will increase the risk to the sustainability of the food supply (Dogan and Karakas 2018). As a continuation of this information, some researchers expect a decrease in precipitation rates and an increase in temperature in the next half-century (Cline 2007). Increased yield losses are expected if no effective adaptation or genetic-based adaptation is carried out on the crop (Zhao *et al.* 2017). The wheat crop is a crop that grows in rainy conditions rather than in extreme hot and cold weather (Dudu and Cakmak 2018). If the temperature increases above the growing conditions of normal wheat, there is a decrease in wheat grains (Ahmed *et al.* 2019, Asseng *et al.* 2015). As a result of this situation, efficiency decreases.

Farmers who produce for different reasons such as population growth, temperature fluctuations, and extreme temperatures should be supported for crop production planning. For this stated purpose, accurate estimation of wheat productivity is a must in the production planning that will continue the food supply.

Annual insurance of the crop, yield forecasting, and market planning is of great importance for the sustainability of the food supply and supply of the countries. Efficiency estimation for topics of specified importance has been analyzed with deep learning-based approaches within the scope of this article. Considering the growing crop data piles, complex deep learning models based on the Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) are needed to extract meaningful content from these data. The input parameters in the crop yield estimate of the study to be carried out for this purpose are as follows;

- amount of water per hectare,
- the average amount of sunlight received per hectare,
- fertilization amount per hectare,
- amount of pesticide used per hectare,
- Sowing area.

Analyzes were performed based on the above-mentioned items. The prediction accuracy of deep learning models is improved by extracting features from the inputs given above (Lago, De Brabandere, De Ridder, and De Schutter 2018, Qing and Niu 2018, Srivastava and Lessmann 2018, Ye, Cao, and Xiao 2017). In this context, three different deep learning models based on CNN and LSTM, which use the specified parameters as inputs for crop yield estimation, are proposed. In one of the proposed models, a CNN-based model was developed. With this model developed, it is possible to capture temporal information. CNN-based models provide superior results in obtaining spatial correlations (Ye *et al.* 2017). Second, an LSTM-based model has been developed that offers good performance in estimating the data from the past time series (Qing and Niu 2018).

Depending on water, sunlight, fertilization, pesticides, and field area, support for continued production is needed to ensure the sustainability of agriculture around the world. For this purpose, the main additives provided to the literature of the study, which were carried out to timely and accurately predict food efficiency to maintain production planning by maintaining food supply, are as follows.

- Data should be normalized so that the data can be processed quickly.
- A hybrid deep learning model consisting of CNN, LSTM, and a combination of both useful models in normalized crop yield estimation is presented.
- The results of similar studies were compared according to the R^2 score, MSE, RMSE, MAE, and MAPE measurement metric results to evaluate the performance of the presented yield estimation model.
- Success rates of 0.8624, 0.8834, and 0.8971 were obtained from the R^2 performance metric of the proposed CNN, LSTM, and hybrid models, respectively.

The following sections of the article consist of Material and Methods, Experimental Results, Conclusion and Discussion sections. In the Material and Methods section, detailed information about the data set used in the regression analysis is given. The methods used and recommended are mentioned.

The performance results obtained from the proposed method are presented in the Experimental Results section. In the last section, the study concludes with controversial analyzes.

2. Material and Methods

In this section of the article, detailed information is given about the data set used to evaluate the performance results of CNN, LSTM, and the hybrid model. At the same time, three different deep learning models are presented for crop yield prediction using the data set.

2.1 Material

The data set used in the regression analyses carried out within the scope of this article has the parameters of the amount of water per hectare, the average amount of sunlight received by the hectare, the amount of fertilization per hectare, the number of pesticides used per hectare, the area of cultivation, and the actual yield of the crop.

The values in the data set were obtained from the World Bank, the Food and Agriculture Organization (FAO) web pages (FAO 2023; WorldBank 2023). The values in the dataset consist of data between 1990 and 2022 years. The dataset consists of 2000 data for the specified years. The scales of the parameter inputs and outputs used in the article are shown in Figure 1.

The data whose distributions are shown in Figure 1 have been reduced to the 0-1 range to increase the processing speed. A built-in function was not used in the normalization process. Normalization was performed by dividing all variables by the maximum of the values in them. All other parameters except the yield parameter were used as input for regression analysis. The values of the yield parameters in the data set are used as the output value.

The data shown in Table 1 are the detailed parameters used in the analysis of the study. In order not to encounter different analysis results in each run of three different deep learning models created within the scope of this article, it is divided into two separate parts training and testing according to the K-fold 4 value.

Table 1. Features used in wheat yield estimation

Raw inputs					Outputs
Water	**UV**	**Area**	**Fertilizer**	**Pesticides**	**Yield**
5.615	65.281	3.23	0	8.969	7.977
7.044	73.319	9.081	0	23.009	7.197
….	…	…	…	….	…
5.607	60.038	2.864	2	23.019	7.424
….	…	…	…	….	…
9.346	64.719	2.797	2	28.066	1.256
….	…	…	…	….	…
6.11	89.28	7.367	1	37.244	0.321
….	…	…	…	….	…
5.92	78.735	5.245	2	29.507	1.136
….	…	…	…	….	…
9.07	71.769	4.13	2	29.673	2.075

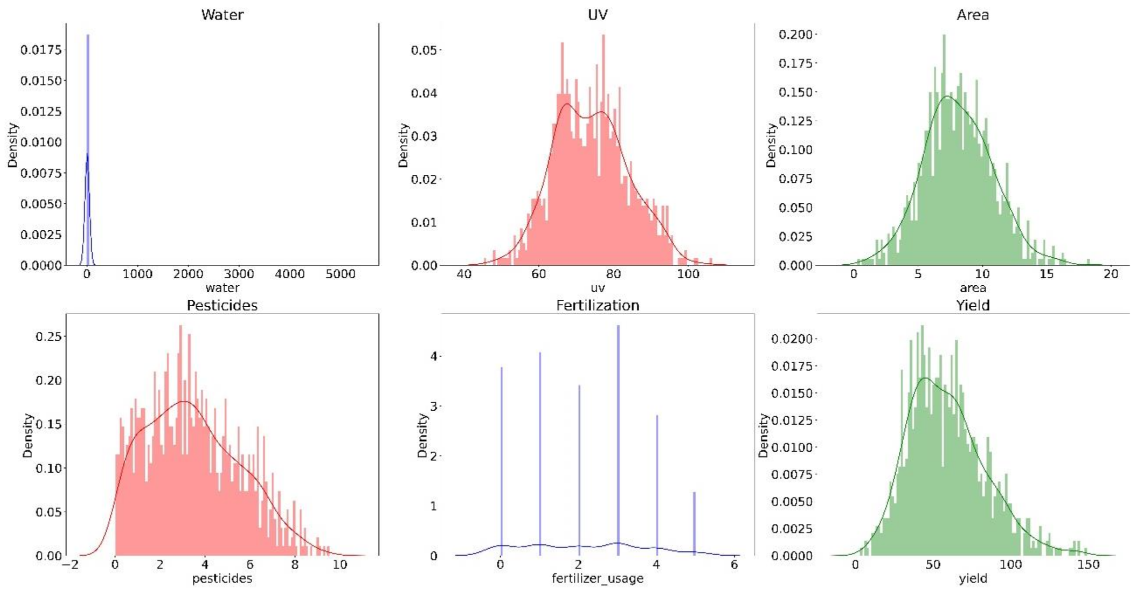

Figure 1. Distribution of parameters in the data set

2.2 Methods

Within the scope of this article, LSTM and CNN-based methods that perform active analysis of temporal information and serial data are used. The model was created from LSTM structures, which removed the gradient boosting problems in the RNN structure.

A model has also been created from CNN structures that provide detailed and distinctive features to the data with convolution layers. Subsequently, a hybrid model is proposed to include strong features of both structures. The LSTM structure, which eliminates the problem of keeping long-range data in RNN structures in memory, has been utilized in efficiency estimation (Jayaraman, Murugappan, Trueman, and Cambria 2021).

2.2.1 LSTM

In Recurrent Neural Networks, the training process takes a lot of time depending on the length of the input parameters in the data set. These networks have loss functions with variable precision. There may be different gradient values depending on the loss variation of the layers that make up the model (Aggarwal 2018).

These different gradient values can often cause gradient boosting in RNN structures (Liu and Guo 2019). The specified boosting usually occurs in the backpropagation stage of the RNN structure by successive multiplication of the weight matrices. Although RNN structures are good for keeping data in short-term memory, RNN-based structures that have the feature of remembering and not forgetting in long-term processes are LSTM structures (Srinivasu *et al.* 2021, Wang, Du, and Wang 2020).

$$i_t = \sigma(W_i X_t + U_i h_{t-1} + b_t) \quad (1)$$

$$f_t = \sigma\left(W_f X_t + U_f h_{t-1} + b_f\right) \quad (2)$$

$$o_t = \sigma(W_o X_t + U_o h_{t-1} + b_o) \quad (3)$$

$$g_t = tanh\left(W_g x_t + R_g h_{t-1} + b_g\right) \quad (4)$$

In LSTM structures with three gates and one layer, the input data at time t is shown as $\mathrm{x_t}$, and the structure showing the hidden state is $\mathrm{h_t}$ (Çetiner and Çetiner 2021). The hidden state structure at time $t-1$ is represented as $\mathrm{h_{t-1}}$. When time t, forgetting, entrance and output gates are shown in Figure 2 with terms $\mathrm{i_t}$, $\mathrm{f_t}$ and $\mathrm{o_t}$.The state layer at time t is represented by the symbol $\mathrm{g_t}$. R, b, and W used in Equations 1-4 are used as a recurrent weight, bias, and weight values.

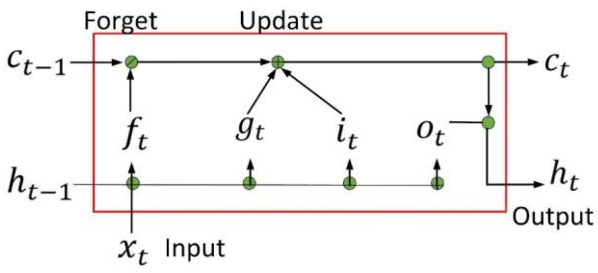

Figure 2. LSTM cell structure (Çetiner and Kara 2022)

2.2.1 CNN

In classical neural networks, it is necessary to create a fully connected structure between the previous layer and the next layer. In this case, it causes the model to work slowly and ineffectively. Instead, CNN methods connect the previous layer to a small point of the next layer by interlayer connections.

With the mentioned approach, a great gain is obtained in model cost calculations. In LSTM structures, it is prevented from forgetting the words in the time series by keeping them in memory for a long time. In CNN models, automatic detection of words in space is provided (LeCun, Bengio, and Hinton 2015).

The convolution layer, which is the basic layer of the CNN architecture, moves filters on the data to obtain distinctive features in the time series. During filtering, the values in the convolution kernel are multiplied by the corresponding values in the hovering window, and the convolution operation is performed. The specified process is represented by the I data matrix in Equation 5. A filter matrix K of size [i,j] is circulated in this matrix.

$$(I * K)_{xy} = \sum_{i=1}^{h} \sum_{j=1}^{w} K_{ij} \cdot I_{x+i-1,y+j-1} \quad (5)$$

2.2.1 Hybrid Model

A hybrid model was created as a result of combining the LSTM cell of the LSTM architecture and the convolution layers of the CNN architecture with fine adjustments in a certain order and plane, the theoretical definitions of which were made in

Sections 2.2.1 and 2.2.2. Information was given about the number of layers, structure, number of steps, and batch size values of the hybrid model, which was obtained as a result of a detailed study with fine adjustments. The Adam optimization method was used to run all models suggested in the article. In Figure 3, the number of layers and structure of the hybrid CNN-LSTM deep learning model are given.

In the first layer, the amount of water per hectare of the crop, the average amount of sunlight received by the hectare, the amount of fertilization per hectare, and the amount of pesticide used per hectare are given as inputs.

In the second and third layers, a one-dimensional convolution operation was carried out by using 64 filters with a window size of 3x3. Capturing distinctive features is provided. In the fourth layer, there is an LSTM structure with 200 units of hidden neurons. With this structure, it is ensured that the data in the time series are kept in memory for a long time. In the fifth layer, the maximum pooling was performed using a 1x1 filter. The maximum numbers in the filtered windows have been captured. In the sixth layer, the fluctuations between the layers were eliminated by mass normalization. In the seventh layer, there is a Dense layer with 512 neurons. It realizes the full connection with the previous layers. In the eighth layer, there is a dropout layer that releases 0.2 neurons.

The ninth layer contains a layer similar to the dense layer in the seventh layer. In the tenth and eleventh layers, the fully connected layer with the linear activation function is connected with the features from the previous layers. At the end of the twelfth step, the yield value obtained as a result of the specified steps is obtained as output. This value is the output value. It is the output value of the proposed model. How close the value obtained as a result of the output process is to the real value is compared with the performance metrics.

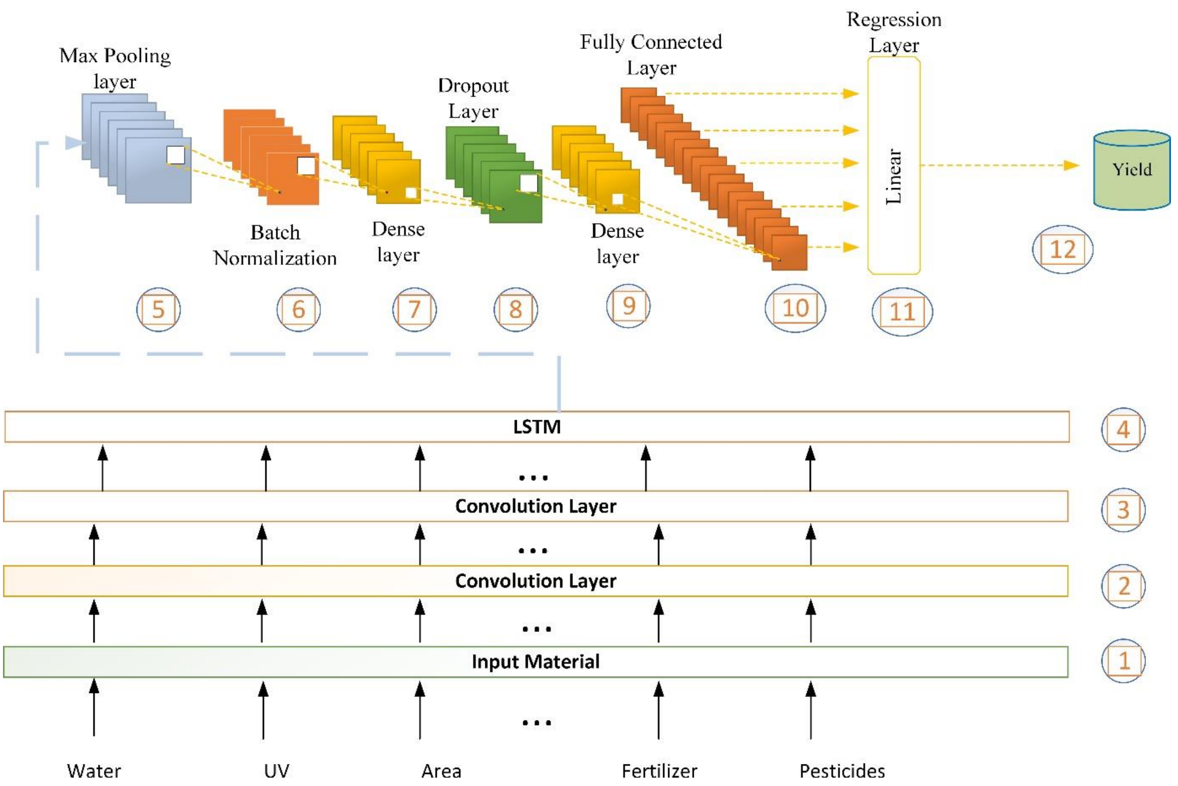

Figure 3. The proposed hybrid deep learning model

3. Experimental Results

The study was carried out on a 64-bit Windows 10 operating system on a computer with an Nvidia GeForce RTX 3060 graphics card. In the experimental analyzes carried out within the scope of the article, the data were first subjected to the normalization process. In the normalization process, the data is reduced to the 0-1 range, thus accelerating the model training process. The data, whose data are normalized to a certain range value, are divided into two training and testing according to the K-fold 4 value. The separated training data were trained separately in 3 different models.

The performance results of the trained models were compared with the performance measurement metrics given in Equations 6-10. Equations 6-10 used MAE, MSE, RMSE, MAPE, and R^2 measurement metrics (Çetiner and Çetiner 2021). The first of the models used in crop forecasting is based on the LSTM structure detailed in Figure 2. The parameters of the LSTM model used in training data are given in Table 2.

Table 2. LSTM model parameters were used.

Parameter	Value
Layers	3, 5, 6
Loss	Mean squared error
Optimizer	Adam
Epochs	100
Batch size	16, 32, 64
Activation name	ReLU

$$MAE = \frac{100}{m}\sum_{i=1}^{m}\left[\frac{Y_i - \hat{Y}_i}{Y_i}\right] \tag{6}$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left(Y_i - \hat{Y}_i\right)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(Y_i - \hat{Y}_i\right)^2} \tag{8}$$

$$MAPE = \frac{100}{m}\sum_{i=1}^{m}\left[\frac{Y_i - \hat{Y}_i}{Y_i}\right] \tag{9}$$

$$R^2 = 1 - \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{\sum(Y_i - \bar{Y})^2} \tag{10}$$

The LSTM model was created according to the number of basic layers and the setting parameters given in Table 2. In the LSTM model created, the best results were obtained in the 5-layer structure with 125 neurons. In the analyses made, it was seen that the best result was obtained in 100 iteration steps. The MAE, MSE, RMSE, MAPE, and R^2 measurement metrics given in Equations 6-10, which are widely used in the literature, were used to measure the performance of the proposed crop productivity.

The performance results obtained in the training and testing of the LSTM model are given. According to these performance results, the training and test performance results of the proposed LSTM model are shown in Table 3.

Table 3. Performance results of the LSTM model

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
LSTM Model with Adam (Testing)	0.8834	0.0043	0.0369	0.0512	11.22
LSTM Model with Adam (Training)	0.8727	0.0045	0.0385	0.0518	11.30

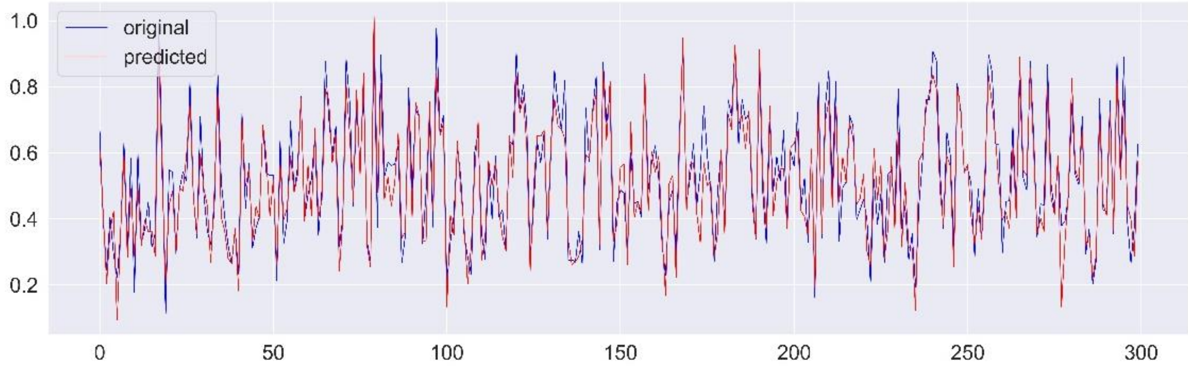

Figure 4. The proposed LSTM model performance output

The test results gave a slightly better result than the training results. As the R^2 value approaches 1, the success rate increases, while the other MSE, RMSE, MAE, and MAPE values have fewer error values, indicating a higher success rate.

For all model outputs, the average of K-fold 4 values is given. In Figure 4, the estimated yield estimation graphs are drawn with the actual crop yield estimation. For the structure in this graph to be seen more closely, the first 40 indexed forms shown in Figure 5 were drawn. Deviations appear on the bottom, top, and sides of this drawing.

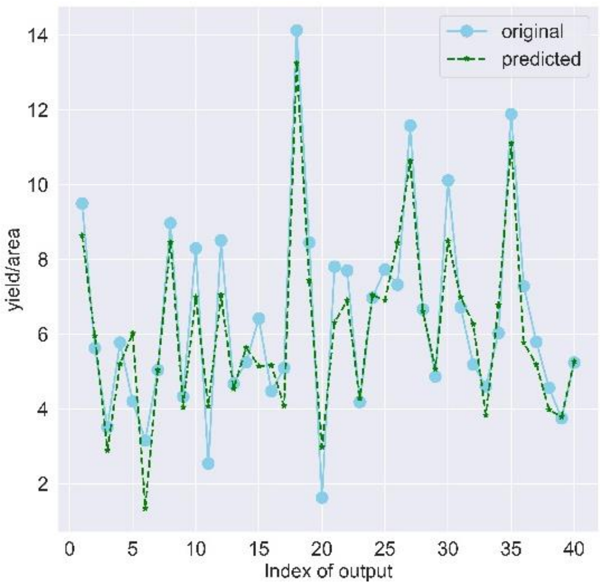

Figure 5. Prediction output of 40 scale LSTM model

In Figure 6, training and test loss rates are plotted. While the test loss rate is less than the training loss rate, both curves converge towards the 100th iteration.

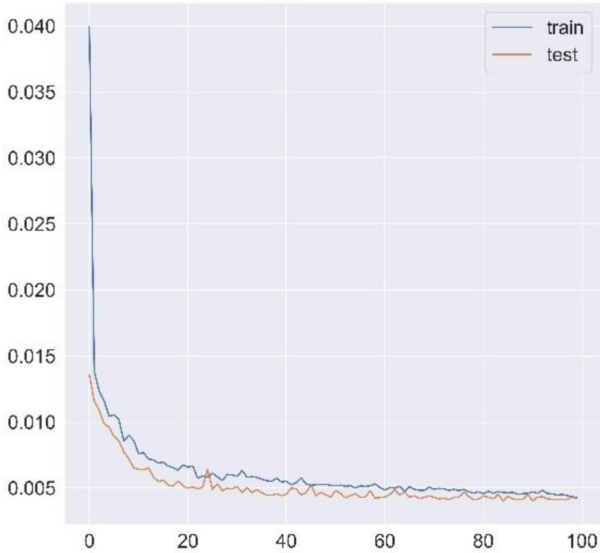

Figure 6. LSTM model training and test loss graph

The CNN model was created according to the number of basic layers and the setting parameters given in Table 4. In the proposed CNN model, the model was created using the 1-dimensional convolution layer. In the proposed CNN model, the best results were obtained in the 6-layer structure using 64 filters in 3x3 size. In general, convolution, maximum pooling, dense layer, dropout, batch normalization, fully connected layer, and regression layer are used. The linear activation function is used in the fully connected layer. Filtering outputs were subjected to maximum pooling, and the highest values were selected. In this model, MAE, MSE, RMSE, MAPE, and R^2 measurement metrics given in Equations 6-10, which are widely used in the literature, were used to measure the performance of the proposed crop productivity.

Table 4. Used CNN model parameters

Parameter	Value
Layers	6
Loss	Mean squared error
Optimizer	Adam
Epochs	100
Batch size	32, 128
Activation name	Linear

Table 5. Performance results from the CNN model

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
CNN Model with Adam (Testing)	0.8624	0.0051	0.0386	0.0573	12.94
CNN Model with Adam (Training)	0.8537	0.0062	0.0398	0.0618	13.98

The performance results obtained in the training and testing of the CNN model are given. According to these performance results, the training and test performance results of the proposed CNN model are shown in Table 5. The test results gave a slightly better result than the training results.

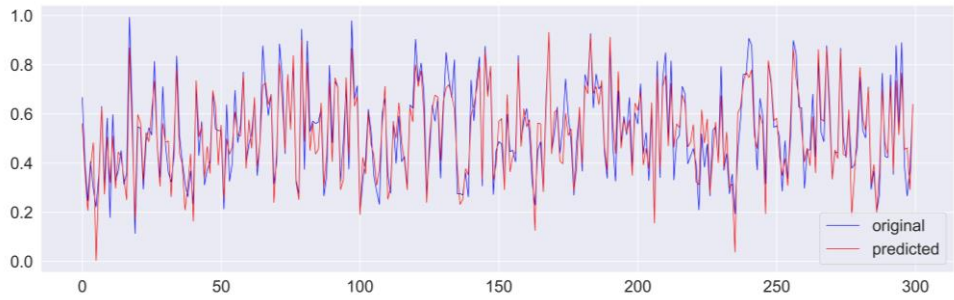

Figure 7. The proposed CNN model performance output

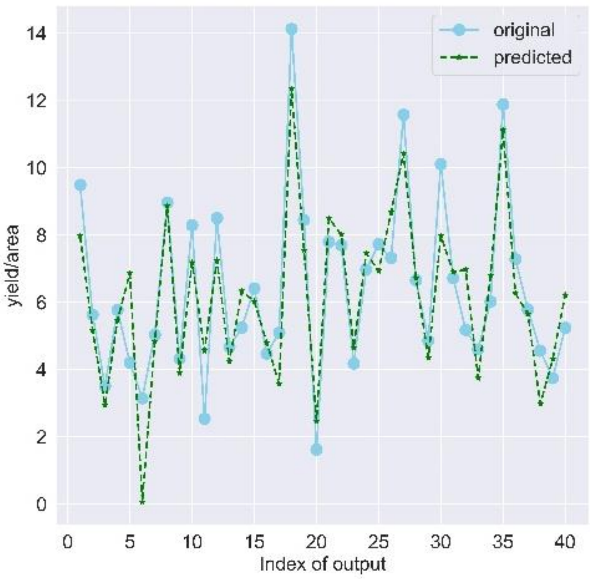

Figure 8. Prediction output of 40 scales CNN models

In Figure 7, the estimated yield estimation graphs are drawn with the actual crop yield estimation. For the structure in this graph to be seen more closely, the first 40 indexed forms shown in Figure 8 have been drawn. There are deviations on the bottom, top, and sides of this drawing. It is seen that the deviations on the sides increases between 35-40 scale. It is seen that the deviations at the peaks are more than the LSTM model. In Figure 9, the training and test loss rates of the CNN model are plotted. The test loss rate is less than the training loss rate. Towards the 100th iteration, both curves move in parallel with each other.

The remainder of this section of the article focuses on the hybrid model, which is one of the important points of the article.

In the hybrid model, which was created by combining the strengths of the LSTM and CNN models, fine adjustments were made to the model to increase the success rate. An LSTM cell layer has been added to ensure that the distinctive features remain in memory for a long time, by enabling the identification of distinctive features with convolution layers. By performing maximum pooling on this cell output, the most remarkable values are separated. The performance results obtained in the whole of the operations carried out by the specified flow chart are given in Figure 10.

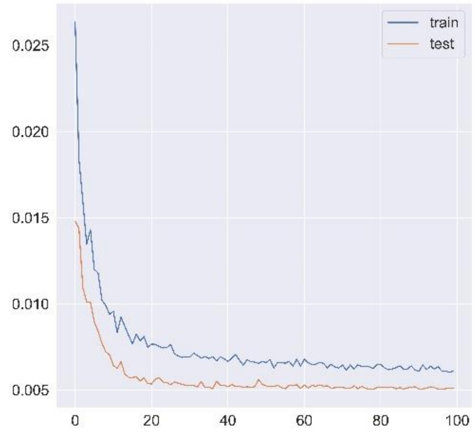

Figure 9. CNN model training and test loss graph

Table 6. Used hybrid model parameters

Parameter	Value
Layers	12
Loss	Mean squared error
Optimizer	Adam
Epochs	100
Batch size	32
Activation name	Linear

The tuning parameters given in Table 6 were used to train the hybrid model. The model, consisting of 12 layers, was run at 100 epochs. The specified parameters were used to obtain the obtained R^2 Score, MSE, RMSE, MAE, MAPE values. While the training time of the hybrid model was 25.56 minutes, the training times of the CNN and LSTM models took 10.91 and 19.38 minutes, respectively. Although the hybrid model is better than the CNN and LSTM models in terms of performance, the training time is high.

Table 7. Performance results of the hybrid model

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
Hybrid Model with Adam (Testing)	0.8971	0.0035	0.0248	0.0461	10.10
Hybrid Model with Adam (Training)	0.8945	0.0037	0.0276	0.0483	10.89

The performance results obtained in the training and testing of the hybrid model are given. According to these performance results, the training and test performance results of the proposed hybrid model are shown in Table 7. When the results of the hybrid model in Table 7 are compared with the results of the LSTM model in Table 3 and the results of the CNN model in Table 7, the hybrid model has the highest R^2 score. When the MSE, RMSE, MAE, and MAPE values are examined, it is seen that the lowest error values are obtained in the hybrid model. It is seen that the proposed hybrid model can be used effectively in crop yield estimation.

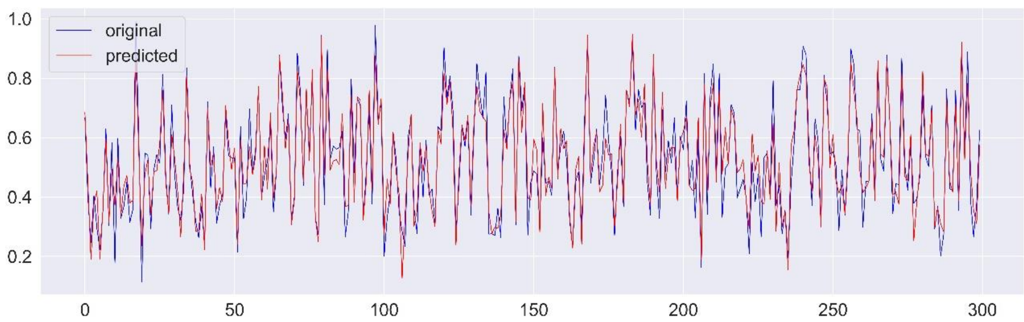

Figure 10. The proposed hybrid model performance output

Although the test results give a result close to the training results, a higher result was obtained than the success rates of other models.

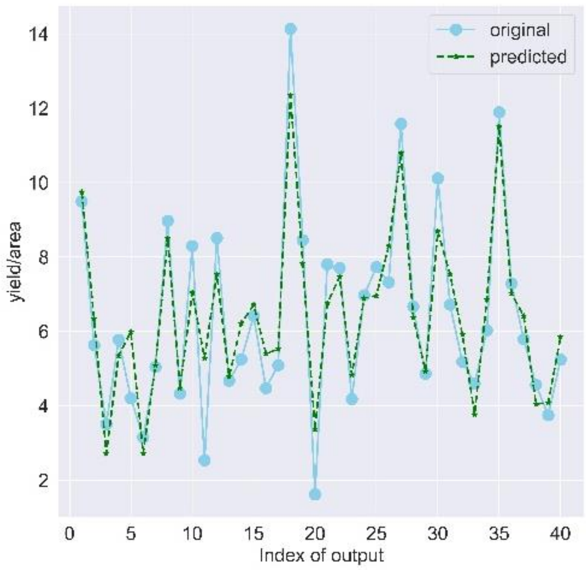

Figure 11. The first 40 index outputs of the hybrid model

In Figure 10, the estimated yield estimation graphs are drawn with the actual crop yield estimate. For the structure in this graph to be seen more closely, the first 40 indexed forms shown in Figure 11 have been drawn.

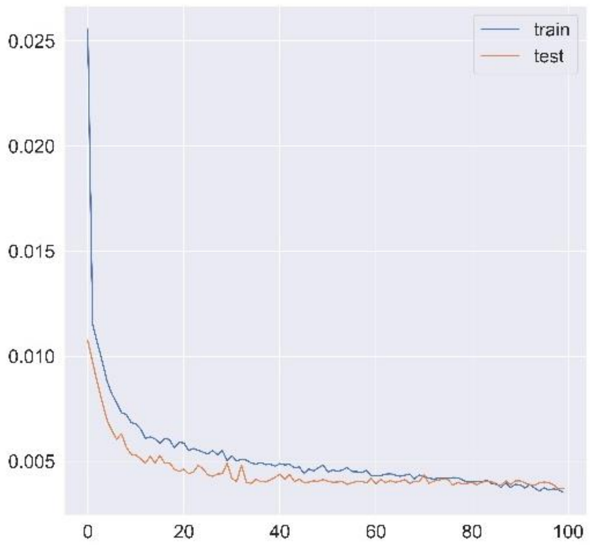

Figure 12. Proposed hybrid model training and test loss graph

From this drawing, it is seen that the deviations in Figure 5 and Figure 8 are slightly reduced. In particular, it is seen that the lateral deviations in the 35-40 scale have decreased. In Figure 12, the training and test loss rates of the hybrid model are plotted. The test loss rate is less than the training loss rate. Towards the 100th iteration, both curves move in parallel with each other.

Table 8. Comparison results of the hybrid model with similar studies

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
CNN+GP (Gavahi *et al.* 2021)	0.803	-	0.5755	-	-
CNN+LSTM (Gavahi *et al.* 2021)	0.786	-	0.5844	-	-
DT (Gavahi *et al.* 2021)	0.774	-	0.7441	-	-
DeepYield (Gavahi *et al.* 2021)	0.864	-	0.4803	-	-
Hybrid Model with Adam (Testing)	0.8971	0.0035	0.0248	0.0461	10.10
Hybrid Model with Adam (Training)	0.8945	0.0037	0.0276	0.0483	10.89

Since there is no study using the same dataset, a comparison has been made with a model developed using CNN and LSTM architectures in the literature. The comparison results shown in Table 8 show that the models benefiting from the strengths of CNN and LSTM architectures can compete with the studies in the literature. The comparison results shown in Table 8 show that the models benefiting from the strengths of CNN and LSTM architectures can compete with the studies in the literature. The GP Gaussian Processor is denoted by GP, while the Decision Tree is abbreviated as DT.

4. Conclusion and Discussion

Deep learning models have been proposed to perform yield estimation, which depends on many different parameters to meet food supply and demand. Among the models proposed in this article, the CNN model took 10.91 minutes, the LSTM model took 19.38 minutes, and the hybrid model took 25.56 minutes.

The hybrid model R^2, MSE, RMSE, MAE, and MAPE performance metrics were 0.8971, 0.0035, 0.0248, 0.0461, 10.10, respectively. From the proposed CNN model, 0.8624, 0.0051, 0.0386, 0.0573, and 12.94 results were obtained for the R^2, MSE, RMSE, MAE, and MAPE performance metrics, respectively. From the proposed LSTM model, the results for 0.8834, 0.0043, 0.0369, 0.0512, and 11.22 were obtained for the R^2, MSE, RMSE, MAE, and MAPE performance metrics, respectively.

When the results obtained are evaluated, the hybrid model is more successful than both the LSTM and CNN models. In future studies, the aim is to develop different studies that will perform an in-depth analysis of a data set that affects the entire crop yield.

5. References

Aggarwal, C. C. 2018. Neural networks and deep learning. Cham: Springer International Publishing, 1-493.

Ahmad, I., Saeed, U., Fahad, M., Ullah, A., Habib ur Rahman, M., Ahmad, A., and Judge, J. 2018. Yield forecasting of spring maize using remote sensing and crop modeling in faisalabad-punjab Pakistan. *Journal of the Indian Society of Remote Sensing*, **46(10)**, 1701–1711.

Ahmad, I., Wajid, S. A., Ahmad, A., Cheema, M. J. M., and Judge, J. 2019. Optimizing irrigation and nitrogen requirements for maize through empirical modeling in semi-arid environment. *Environmental Science and Pollution Research*, **26(2)**, 1227–1237.

Ahmed, I., Ullah, A., Rahman, M. H. ur, Ahmad, B., Wajid, S. A., Ahmad, A., and Ahmed, S. 2019. Climate change impacts and adaptation strategies for agronomic crops. *In Climate Change and Agriculture*, 1–14. IntechOpen London, UK.

Asseng, S., Cammarano, D., Basso, B., Chung, U., Alderman, P. D., Sonder, K., … Lobell, D. B. 2017. Hot spots of wheat yield decline with rising temperatures. *Global Change Biology*, **23(6)**, 2464–2472.

Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., … White, J. W. 2015. Rising temperatures reduce global wheat production. *Nature Climate Change*, **5(2)**, 143–147.

Ben-Asher, J., Yano, T., Aydın, M., and Garcia y Garcia, A. 2019. Enhanced growth rate and reduced water demand of crop due to climate change in the eastern mediterranean region, *Climate Change Impacts on Basin Agro-ecosystems,* **18,** 269–293.

Cao, J., Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., and Tao, F. 2021. Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine. *European Journal of Agronomy*, **123**, 126204.

Çetiner, H., and Çetiner, İ. 2021. Analysis of different regression algorithms for the estimate of energy consumption. *European Journal of Science and Technology*, **31**, 23–33.

Çetiner, H., and Kara, B. 2022. Recurrent neural network based model development for wheat yield forecasting. *Journal of Engineering Sciences of Adiyaman University*, **9(16)**, 204–218.

Chen, Y., Zhang, Z., Tao, F., Wang, P., and Wei, X. 2017. Spatio-temporal patterns of winter wheat yield potential and yield gap during the past three decades In North China. *Field Crops Research*, **206**, 11–20.

Cline, W. R. 2007. Global warming and agriculture: end-of-century estimates by country. Peterson Institute.

Deutsch, C. A., Tewksbury, J. J., Tigchelaar, M., Battisti, D. S., Merrill, S. C., Huey, R. B., and Naylor, R. L. 2018. Increase in crop losses to insect pests in a warming climate. *Science*, **361(6405)**, 916– 919.

Dodds, F., and Bartram, J. 2016. The water, food, energy and climate nexus: challenges and an agenda for action. *Routledge*.

Dogan, H. G., and Karakas, G. 2018. The effect of climatic factors on wheat yield in Turkey: a panel dols approach. *Fresenius Environ Bull.*, **27**, 4162–4168.

Dudu, H., and Cakmak, E. H. 2018. Climate change and agriculture: an integrated approach to evaluate economy-wide effects for Turkey. *Climate and*

Development, **10(3)**, 275–288.

FAO, I. 2017. The State of food insecurity in the world, meeting the 2015 international hunger targets: taking stock of uneven progress. *Rome*, FAO.

Gavahi, K., Abbaszadeh, P., and Moradkhani, H. 2021. DeepYield: a combined convolutional neural network with long short-term memory for crop yield forecasting. *Expert Systems with Applications*, **184**, 115511.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, **202**, 18–27.

Jayaraman, A. K., Murugappan, A., Trueman, T. E., and Cambria, E. 2021. Comment toxicity detection via a multichannel convolutional bidirectional gated recurrent unit. *Neurocomputing*, **441**, 272–278.

Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., … Kim, S.-H. 2016. Random forests for global and regional crop yield predictions. *Plos One*, **11(6)**.

Lago, J., De Brabandere, K., De Ridder, F., and De Schutter, B. 2018. Short-term forecasting of solar irradiance without local telemetry: a generalized model using satellite data. *Solar Energy*, **173**, 566–577.

LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. *Nature*, **521(7553)**, 436–444.

Liu, G., and Guo, J. 2019. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, **337**, 325–338.

Nasim, W., Amin, A., Fahad, S., Awais, M., Khan, N., Mubeen, M., … Jamal, Y. 2018. Future risk assessment by estimating historical heat wave trends with projected heat accumulation using simclim climate model in Pakistan. *Atmospheric Research*, **205**, 118–133.

Qing, X., and Niu, Y. 2018. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, **148**, 461–468.

Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., and Kang, J. J. 2021. Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors*, **21(8)**.

Srivastava, S., and Lessmann, S. 2018. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Solar Energy*, **162**, 232–247.

Vanli, Ö., Ahmad, I., and Ustundag, B. B. 2020. Area estimation and yield forecasting of wheat in southeastern Turkey using a machine learning approach. *Journal of the Indian Society of Remote Sensing*, **48(12)**, 1757–1766.

Vanli, Ö., Ustundag, B. B., Ahmad, I., Hernandez-Ochoa, I. M., and Hoogenboom, G. 2019. Using crop modeling to evaluate the impacts of climate change on wheat in southeastern Turkey. *Environmental Science and Pollution Research*, **26(28)**, 29397–29408.

Wang, J. Q., Du, Y., and Wang, J. 2020. LSTM based long-term energy consumption prediction with periodicity. *Energy*, **197**, 117197.

Ye, L., Cao, Z., and Xiao, Y. 2017. DeepCloud: ground-based cloud image categorization using deep convolutional features. *IEEE Transactions on Geoscience and Remote Sensing*, **55(10)**, 5729–5740.

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., … Ciais, P. 2017. Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences*, **114(35)**, 9326–9331.

İnternet kaynakları

1- https://www.fao.org/home/en/, (07.12.2022).

2- https://data.worldbank.org/, (07.07.2022).