Ann. For. Res. 66(2): 301-315, 2023
ISSN: 18448135, 20652445
ANNALS OF FOREST RESEARCH
www.e-afr.org

MACHINE LEARNING MODEL FOR MULTIPARAMETRIC CROP TYPE AND YIELD FORECASTING

Ms. Deepali Jawale[1], Dr. Sandip Malik[2]

[1]Research Scholar, Department Of CSE, Oriental University Indore MP INDIA

[2]Professor, Department Of CSE, Oriental University Indore MP INDIA

ABSTRACT
The multidisciplinary process of "smart farming" entails forecasting the kind and quality of plant that will produce the highest yield in a specific geographic area. This entails accounting for elements like soil, climate, and other environmental elements. Deep learning models are used by the great majority of intelligent farming systems to achieve this, however their accuracy, scalability, and real-time deployment capabilities are all lacking. This is due to the fact that most deep learning models need a lot of training data, which lengthens the time needed before they can be put into use. Furthermore, a model that has been trained on one kind of crop may not be as applicable to other types of crops. To improve the usability and practicality of deep learning models, this paper suggests an improved machine learning model for multiparametric crop type and yield prediction. The proposed model fuses the 'you look only once' (YoLo) model with the VGGNet-19 model, the Inception Net model, the Xception Net model, and the GoogleNet model.This makes processing at a high speed possible. The model uses multiparametric data to predict the maximum yield that can be produced from a crop and the kind of crop that can be grown under a specific set of environmental conditions. This includes the use of plant imagery, soil parameters, weather data, temporal geographical data, and nutrient information. The suggested model has a high degree of accuracy when tested against a variety of crop and soil types; it received scores of 98.7% and 97.6%, respectively, for crop-type prediction and yield prediction. The suggested model performs, on average, 8% better in terms of accuracy, 6% better in terms of precision, 3% better in terms of recall, and 6.5% better in terms of area under the curve (AUC) when compared to other cutting-edge models. The suggested model also demonstrates a 9% decrease in latency, which qualifies it for high-speed real-time deployments. Additionally, this text provides a number of case studies that were conducted in order to validate the model's performance and makes some recommendations for future research in order to achieve even higher application-specific performance levels.
Keywords: Smart farming, machine learning, ensemble, augmented, yield, crop-type

1. INTRODUCTION

Predicting yield and crop type from parametric non-image data necessitates precise feature extraction, forecasting, post-processing model designs, and data collection. These models combine a number of yield- and crop-specific parameters, such as

• Information about the crop, such as the leaf area index and growth process.

• Soil information, including pH levels, soil type, cation exchange capacity (CEC), and crop area on that soil.
• Information on relative humidity, which covers humidity levels, precipitation, and rainfall.
• Soil nutrient information, including details on fertilizers, nutrients added to the soil, nutrients already present in the soil, how rain affects nutrient levels, etc.
• Field management, encompassing details on fertilization and irrigation.
• Solar data, including temperature, length of sun rise, short wave radiation, and gamma radiation.
• You can also use other parameters for this, such as vegetation indices, wind speed, pressure, etc.
A machine learning model is given all these parameters, as shown in figure 1, along with tagged data regarding crop type and yield.

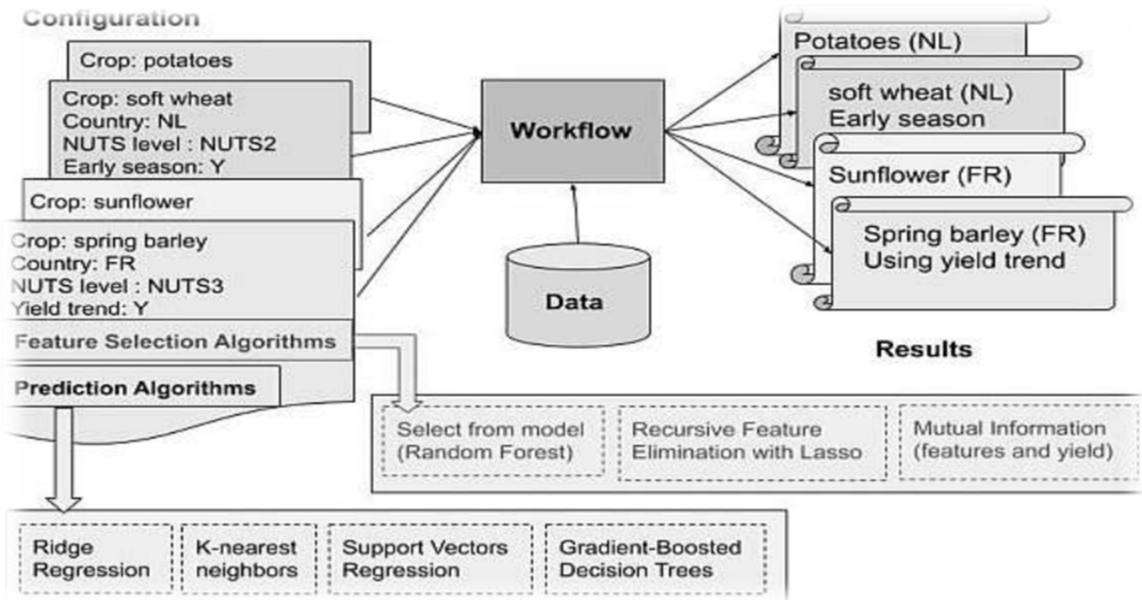

Figure 1. Multivariate machine learning model for predicting crop type and yield [42]

The model illustrates how the crop configuration—which could include the kind of crop, geographic data, wind data, and other details of that kind—is supplied for training. Numerous machine learning algorithms, including random forest (RF), recursive feature elimination, mutual information, gradient boosted trees, and others, are used to predict the crop types and yield values. These algorithms are predicated on the previously mentioned features. An overview of these algorithms and an analysis of their approximate performance characteristics are presented to readers in the section that follows. The survey's findings indicate that, when it comes to general-purpose accuracy or the classification of crop type and yield, machine learning models like convolutional neural networks (CNN) are generally thought to outperform linear classification models. The design of the suggested augmented machine learning model for multiparametric crop type and yield prediction comes after the review. The observation that was just made is used in this model. The performance of the suggested model is assessed on a range of crop types after the design section, and this performance is then compared with several other approaches that are thought to be state-of-the-art. This article concludes with some insightful observations about the suggested model and some recommendations for how to improve its functionality even further for various use cases.

2. literature review :

Researchers are creating a wide range of systems to precisely predict crop type and yield using machine learning. These systems work by mapping different crop, field, and environmental parameters to the final crop type or yield category after first analyzing various crop, field, and environment parameters. To achieve this goal, for example, the research presented in [1, 2, 3, 4] suggests using different types of convolutional neural networks (CNN), vegetation indices, support vector machines (SVM), logistic regression (LR), and random forests (RF). These algorithms are appropriate for use in real-time analysis because they consistently achieve an accuracy of between 80% and 94% on the dataset, regardless of the evaluation conditions. In order to increase the accuracy of yield prediction, the research presented in [5, 6, and 7] furthers the improvement of performance via deep learning models such as recurrent neural networks (RNN), extreme gradient boosting (XGBoost), least absolute shrinkage and selection operators (LASSO), and 'k' nearest neighbor (kNN) classifiers. Although these models have been trained with limited data, their accuracy is high, but their applicability to a wide variety of crop types is limited. It is possible to enhance this performance by incorporating the research conducted in [8, 9, 10, 11]. The study employs a range of feature extraction models, including cosine transforms, wavelet transforms, and others, in conjunction with Naive Bayes (NB), recurrent convolutional neural networks (RCNN), decision trees (DT), gated recurrent units (GRU), long-short-term memory (LSTM), and variations of support vector machine models to estimate the final By offering deep incremental learning frameworks that continuously learn from output evaluations, these models help to increase accuracy. These frameworks have a cycle of learning. Similar models using hardware-based designs, hybrid CNN with LR, Auto Regressive Integrated Moving Average (ARIMA) with kNN, and hybrid RNN can be found in [12, 13, 14, 15]. which employ hardware-based designs: Auto Regressive Integrated Moving Average (ARIMA) with kNN, hybrid RNN & LSTM, and hybrid CNN with LR. These hybrid models leverage multiple learners to enhance the use of multiple learners, which helps to improve accuracy and scalability performance for crop type and yield prediction. To further enhance the accuracy performance, techniques like deep reinforcement learning [16], gradient descent (GD) with deep neural networks (DNN) [17], auto encoder with SegNet [18], convolutional LSTM [19], and growth stage normalization [20] are applied. In [16] DNN stands for deep neural networks, and GD stands for gradient descent. Compared to linear SVM and RF models, these models exhibit superior accuracy and increased versatility. This is because feedback learning allows these models to process features more quickly and with fewer errors. The error rates can be further decreased by using CNN to analyze genomic data [21], SVM to analyze soil nutrients [22], an ensemble of artificial neural networks (ANN) with SVM and maximum likelihood (ML) classifiers [23], and an ensemble of linear regression (LR) with classification and regression trees (CART) and gradient Naive bayes (GNB) for improved prediction. The overall performance of these models is very high due to the integration of a wide range of classification and prediction algorithms. On the other hand, because these models employ a multitude of learning algorithms, they have a low delay. The research described in [24, 25, 26, 27] proposes the use of particular deep learning models, like support vector regression (SVR), RCNN, and sophisticated clustering techniques, to get around this issue. These techniques help lower error rates overall by progressively learning from test classifications. Through the work in [28, 29, 30, 31], the performance of these deep learning models is further modeled and enhanced. These papers propose a variety of hybrid classification models, including recursive feature elimination with adaptive bagging classification (RFEABC), random forest with fuzzy C means (RF2CM), apriori Naive Bayes (ANB), and SVM with kNN. The purpose of these models is to model and enhance the deep learning models' performance. The objective of these models is to minimize the number of features through variance normalization, which will decrease classification time and increase accuracy. Additionally, a wide range of distinct algorithms are proposed by researchers for crop data management. These suggested algorithms, which support common agricultural policy and employ random forest for band-based classification, variable rate technology (VRT), and semantically enriched crop classification, are highlighted by the research reported in [32, 33, 34]. By offering an accurate estimation of crop types, soil types, and approximate yield crop levels on the soil with the least amount of delay and error rates, these algorithms help manage entire farm fields. Similar models are also proposed in [35, 36, 37, 38, and 39]. These models include Support Vector Regression with Radial Basis Function kernel (SVR-RBF), farm land optimization techniques, ridge regression models, deep transfer learning techniques, and ensemble classification techniques. Using these techniques to estimate crop type and yield values yields very accurate results with little lag time. The results of this review show that when it comes to crop type and yield prediction, ensemble architectures and deep learning techniques outperform linear classification models by a significant margin. In order to create an improved machine learning model for multiparametric crop type and yield prediction, the work that is being proposed uses these approaches. The next paragraph of this text will discuss the construction of the suggested model.

Improved machine learning model recently created for multiparametric yield and crop type prediction

GoogLeNet, Xception Net, VGGNet-19, and Inception Net To augment features, the proposed augmented machine learning model for crop type & yield prediction combines CNN models.. For the purpose of training these networks, these models gather multiparametric data such as crop, soil, relative humidity, soil nutrient, field management, solar, and other random parameters. For high-speed processing, the trained networks are combined with the "you look only once" (YoLo) model. A super-feature vector is created as a result of using numerous CNN models for feature extraction. With more than 50k feature values, this super-feature vector is given to the recurrent CNN YoLo model for final classification after being transformed into a 2D vector. As a result, the suggested model design deploys using these procedures. Acquiring data and enhancing feature representation through the use of an ensemble CNN model Feature translation into twodimensional vectors for RCNN assessment and training. Through incremental learning, crop type classification performance and yield are improved. All of these processes are described in detail in different sections of this text in order to streamline the design process. Figure 2 shows the proposed model and defines the entire process flow. Field data is collected and sent to a pre-processing layer for organization. After that, an ensemble of CNN layers is given the structured data for augmented feature extraction, which results in the creation of a super feature vector.

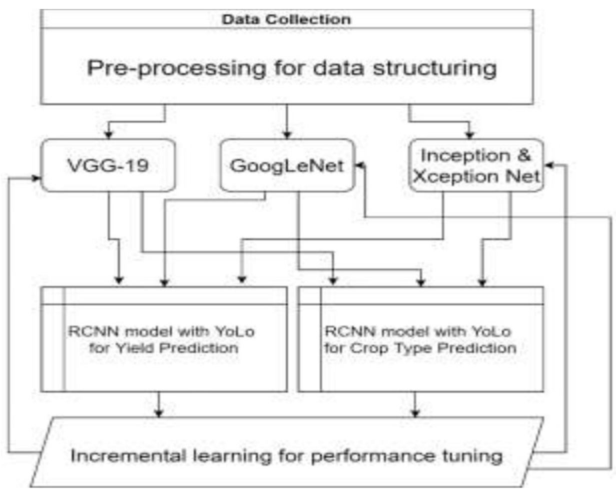

Figure 2. A proposed augmented machine learning model for multiparametric crop type and yield prediction [43]

Two distinct RCNN models that make use of YoLo for high-speed operation are given the feature vector. An incremental learning model receives the yield and crop prediction results, and correlation values are used as confidence levels to adjust CNN performance.

The internal design of each of these blocks is explained in the text's following subsections.

3.1. Using an ensemble CNN model for data collection and enhanced feature extraction

- Sensors are placed in the field to collect data, and the following physiological parameters are extracted:
- Time-related data regarding the crops that were previously planted in the region (Icrop) crop area (Carea), cation exchange capacity (Scec), pH level (Sph), and soil type (Stype) on that soil.
- Soil humidity levels (Hsoil), water level (Wlevel), precipitation (P), relative humidity (Hrel), and temporal rainfall (Rtemp).
- Nutrients added to the soil (Nadded), nutrients present in the soil (Npresent), nutrients temporally affected by rain (NRains), and nutrients existing in the soil (Snut).
- Information on historical fertilization (Finfo) and irrigation level (Hirr).
- Temperature (T), sun rise duration (Srise), short-wave radiation (SWR), and temporal gamma radiation (Gtemp).

- Temporal vegetation indices (VItemp), environmental pressure (Pe), and temporal wind speed (Wt).
- The area's temporal yield per crop information (YCroptemp) and temporal crop type information (Croptemp) are both available.

To eliminate any unwanted outlier data, median filtering is used as part of the pre-processing step for all of this data. Equation 1 is used to process each temporal piece of data in order to eliminate unnecessary information from the input set.

$$T(out) = \mathrm{Tin}, \mathrm{else}\ \mathrm{T(out)} = Yin \sum_{n=1}^{N} (\mathrm{I/N}) \quad (1)$$

In this case, Tin, Tout, and N stand for the number of records in the input data, the temporal output data, and the input temporal data. The ensemble feature extraction layer receives the processed data. A layer with 3x3 sized convolutions and a 64x64 window is utilized for feature augmentation in order to deploy this CNN. By now, nearly 150k features have been extracted, and each feature vector is sent to a subsequent 3x3 convolutional layer with a 64x64 window size.

$$M(POUT) = \mathrm{MAX} + \sum_{n=1}^{\infty} \left(\frac{1}{xK} * \frac{XK}{I} \right) I/LR \quad (2)$$

The Max Pooling outputs are represented by Mk, the inputs supplied to the layers with k = 224 features are represented by Xk, and the rate at which layers learn cognitively is represented by lr. The outputs of the maximum pooling layer are then sent to additional convolutional layers, which have the following sizes: 128 x 128; 256 x 256; and 512 x 512. These layers extract 800 k features, 400 k features, and 200 k features, in that order.

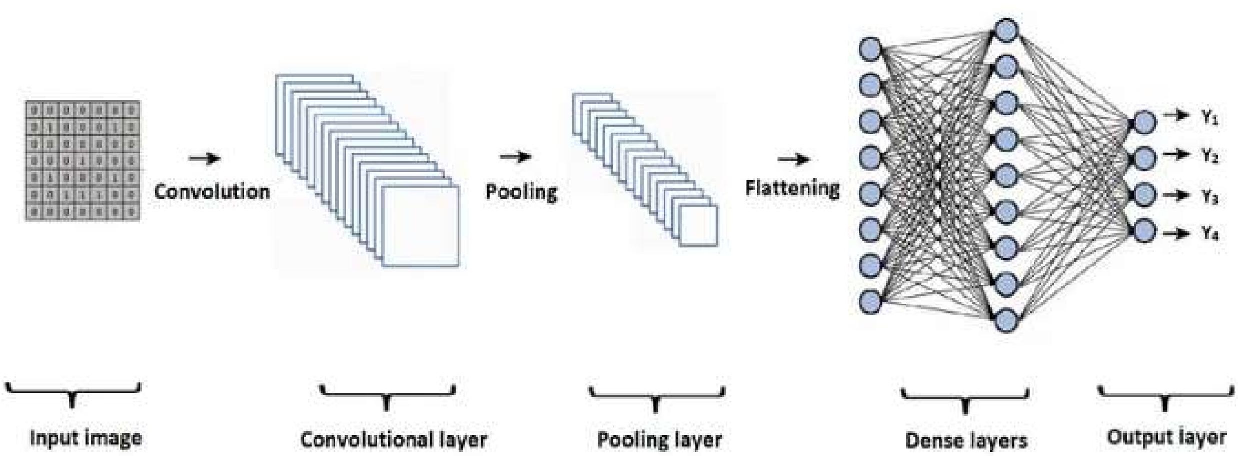

Figure 3.The VGGNet-19model [41]

A layer of max pooling and filtering is used to standardize the variation in the features. This layer lowers the feature variance for features that are similar to each other and increases it for characteristics that belong to distinct classes. The next layer is called a max-pooling layer, and it is controlled by equation 2. In this layer, features are reduced by identifying the characteristics with the highest levels of variance. Consequently, 1.5 million distinct characteristics have been generated. The same process is repeated multiple times for the GoogLeNet, Inception Net, and Xception Net models. For every entry, a total of six million features are extracted. As will be explained in more detail in the following section of this text, these features are then sent to a YoLo RCNN model and a feature variance reduction layer.

3.2. Conversion of features into 2D vectors for RCNN training and evaluation

The extracted features are given to a YoLo RCN model, where variance of each class is extracted using equation 3,

$$v_{Avg} = \left[\sum_{a=1}^{m} \frac{1}{m-1}\left[x_a - \sum_{i=1}^{m}\frac{1}{m}\left[\sum_{j=1}^{n}\left(x_j * \sum_{k=1}^{n} x_k\right)^{2/n}\right]^{1/2}\right]\right]^{1/2} \tag{3}$$

where 'x' denotes the value of a single sample,'m' denotes the total number of samples in this class, and 'n' denotes the total number of samples in the other class. A super-feature vector with almost 50,000 features is then extracted after eliminating from the set any features with variances less than V avg. After that, these features are arranged into a 224x224 2D vector and supplied to an RCNN YoLo network in the manner depicted in figure 4:

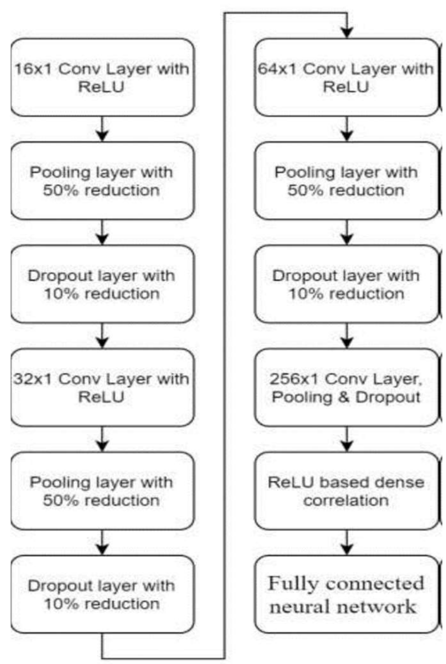

Figure 4. The RCNN YoLo layer design for super-feature vector classification [43]

The represented feature sets are then fed into a model of the RCNN architecture, which incorporates multiple optimizers, leaky rectilinear units (ReLUs), and level 2 regularization. Put another way, it's a combination of these three components. The first convolutional layer of the model uses 1024 features from each of the feature sets that were previously extracted by the process to better represent the data.A three-by-three window for maximum pooling is used to alter the stride, which facilitates the extraction of longer feature sets. A leaky ReLU layer receives the features that were taken out of the layer below and uses them to maximize the variance of the features. The results of this layer are determined by Equation 6, which indicates that feature variance levels can be used to find output first.

$$\text{When } var(xin) > 0,\ xout = alpha * xin;\ \text{otherwise},\ xout = 0 \ \ldots\ldots(4)$$

where input data, output data, and attenuation constant are represented by the variables xin, xout, and alpha. The retrieved features are scaled down by a factor of two times two after the highest feature variance value is subtracted. This implies that half as many features are produced when employing this method. Consequently, 512 features are generated for every stride and subsequently routed to a dropout layer. To enhance the feature selection procedure, roughly 25% of the features in the dropout layer are eliminated. From the retrieved feature sets, the next layer is responsible for extracting 4096 features per stride. The extraction of longer feature sets is made easier and maximum pooling is enabled by a three-by-three window with a variable stride. The features that were taken out of the layer below are then moved to a leaky ReLU layer in order to decrease the feature variance. Equation 4, which states that output may be discovered by first assessing feature variance, determines the outcomes of this layer. Once the highest value of feature variance is subtracted from the retrieved features, they are scaled down by a factor of two times two. This indicates that half as many features are produced when employing this method. About 25% of the 1024 features that are produced for each iteration are then removed from feature sets in order to improve the quality of the feature selection process. This process is repeated for various input layer sizes, and the final feature extraction process uses a 256x1 layer combination. Lastly, a fully connected layer for crop yield classification into various categories of crop types.Several performance metrics, including accuracy, fMeasure, precision, and recall, are taken into consideration when analyzing the performance of this network. An incremental learning model, which will be discussed in the section that follows, will help to further optimize these parameters.

3.3. Using incremental learning to improve crop type classification performance and yield

A correlation model is trained on the testing set data to help with incremental learning and continuous accuracy improvement. The steps listed below are how this yield accuracy improvement model operates. By comparing the final obtained yield with the test set yield, test accuracy is estimated.The chosen feature vectors are compared with accurately classified examples of the suggested model for every new input. Equation 5 is used to estimate the correlation between the original yield and the estimated yield derived from our model.

$$Corr = \frac{\sum_{i=1}^{NF_{Test}} F_{Testi} - F_{Newi}}{\left[\sum_{i=1}^{NF_{Test}} (F_{Testi} - F_{newi})^2\right]^{1/2}} \quad [5]$$

where Nftest is the total number of features chosen by the feature extraction model, and Ftesti&Fnewi are the ith test set and new input features, respectively. If the correlation merit of the new input data is greater than 0.999, it is added to the training set, indicating that the input closely matches the training and testing sequences that have already been stored. As a result, the total classification accuracy rises with the number of testing sequences. This accuracy is evaluated using various input sequences and contrasted with other cutting-edge techniques. The evaluation's findings are shown in the following section, where these values are tabulated for various testing sample counts, helping to assess the proposed model's overall accuracy.

\RESULTS AND STATISTICAL COMPARISON

A significant amount of data was contributed by a number of sources, including the Food and Agricultural Organization of the United States of America (FAOSTAT), the Indian Meteorological Department (IMD), the International Soil Reference and Information Center (ISRIC), Agriculture Crop Production in India (ACPI), and the European Centre for Medium-Range Weather Forecasts (ECMWF). Precision, recall, fMeasure, and accuracy values were compared to those predicted by the proposed model after the efficacy of the work detailed in [6] and [15] was evaluated. The methodology used to derive the evaluation's findings for the various testing samples indicated in tables 1, 2, 3, and 4 is explained below.

Testing samples	Avg. P ([6])	Avg. P ([15])	Avg. P (Proposed)
100	0.784	0.797	0.881
200	0.799	0.810	0.895

Table 1. Precision values for different models

It has been found that adding multiple CNN-based techniques for feature set re-presentation and YoLo for classification leads to improvements of over 10%.Additional validation of this performance gain comes from a recall value analysis, as presented in table 2 below.

Testing samples	Avg. R ([6])	Avg. R ([15])	Avg. R (Proposed)
100	0.615	0.651	0.705
200	0.626	0.661	0.714

Table 2. Recall values for different models

There is a 7% increase in recall ability when compared to reference models. The true accuracy of the models that have been provided is represented by the value of fMeasure, which is created by adding the values of precision and recall. Table 3, which is formatted as follows, can be used to observe this performance statistic.

Testing samples	Avg. F ([6])	Avg. F ([15])	Avg. F (Proposed)
100	0.690	0.716	0.783
200	0.702	0.728	0.794

Table 3. fMeasure values for different models

The suggested model performs well and can be used for the real-time detection of crop damage, as evidenced by the improvements in fMeasure values that parallel those in accuracy and recall. It was shown that there was an 8% improvement in fMeasure values. This performance is further taken into account for the optimization process through an assessment of its accuracy for general applications, which is displayed in table 4 and explained in the evaluation sets that follow.

Testing samples	Avg. A ([6])	Avg. A ([15])	Avg. A (Proposed)
100	0.836	0.865	0.877
200	0.851	0.880	0.890

Table 4. Accuracy values for different models

Accuracy values have increased by 4%, in line with fMeasure. This illustrates the usefulness of the suggested model and how it can be used to detect crop damage in real time. In addition to this performance, a classification delay evaluation is also taken into account. The outcomes are shown in table 5 as follows:

Testing samples	Avg. D ([6])	Avg. D ([15])	Avg. D (Proposed)
100	3.38	3.50	1.77
200	3.44	3.56	1.80

Table 5. Delay values for different models

YoLo was discovered to offer a 45% efficiency boost in the delay performance, suggesting that the suggested model is quick and may be used for a variety of real-time crop damage detection scenarios.

CONCLUSION & FUTURE WORK

Using ensemble convolutional neural network (CNN) models improves the performance of feature extraction for crop type and yield prediction. This performance has been enhanced by the use of dropout layers, max pooling layers, ReLU layers, and high efficiency convolutional layers. Each of these layers contributes differently to the effective feature extraction processAn ensemble of these layers is utilized in the provided model to extract features from the input. This guarantees the most effective processing of the data. In order to further enhance the accuracy performance of the dense layer, this ensemble design aims to integrate several CNN models, including VGGNet19, Xception Net, and Inception Net in addition to GoogleNet. Accuracy, precision, recall, and fMeasure values can be progressively raised by combining an ensemble CNN classifier with an incremental learning model for accuracy tweaking. Because each of these metrics can be enhanced separately, this is made possible.This impact can be better understood by looking at tables 1, 2, 3, and 4, each of which evaluates relevant parameters, and comparing those evaluations with those of high-efficiency models that are currently in use. Alternatively, this impact can be shown using the graph below. By using the YoLo RCNN Model, which allows for the single-iteration evaluation of the accumulated data, the required amount of time is reducedConclusion: The suggested model outperforms existing deep learning models in terms of parametric performance and can be highly effectively used for crop type recognition and yield estimation. Taking into account that the suggested model outperforms other deep learning models currently in use in terms of parametric performance, this conclusion seems reasonable. Applying long-short-term picture data can help this model perform better by allowing for more in-depth research on the gated recurrent unit (GRU) and long short-term memory (LSTM) models. These two models are available right here. It may be possible to improve the performance of both picture and parametric data to predict crop subtypes and yield quality for more recent soil sets by looking into the use of Generative Adversarial Networks (GANs) in more detail.

REFERENCES

[1] Udendhran, R., Balamurugan, M. Towards secure deep learning architecture for smart farming-based applications. Complex Intell. Syst. 7, 659–666 (2021). https://doi.org/10.1007/s40747-020-00225-5

[2] Vallentin, C., Harfenmeister, K., Itzerott, S. et al. Suitability of satellite remote sensing data for yield estimation in northeast Germany. Precision Agric (2021). https://doi.org/10.1007/s11119-021-09827-6

[3] Amitabha Chakrabarty, Nafees Mansoor, Muhammad Irfan Uddin, Mosleh Hmoud Al-adaileh, Nizar Alsharif, Fawaz Waselallah Alsaade, "Prediction Approaches for Smart Cultivation: A Comparative Study", Complexity, vol. 2021, Article ID 5534379, 16 pages, 2021. https://doi.org/10.1155/2021/5534379

[4] A. Sharma, A. Jain, P. Gupta and V. Chowdary, "Machine Learning Applications for Precision Agriculture: A Comprehensive Review," in IEEE Access, vol. 9, pp. 4843-4873, 2021, doi: 10.1109/ACCESS.2020.3048415.

[5] Klompenburg, Thomas & Kassahun, Ayalew & Catal, Cagatay. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture. 177. 105709. 10.1016/j.compag.2020.105709.

[6] Shahhosseini, Mohsen & Hu, Guiping & Huber, Isaiah & Archontoulis, Sotirios. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Scientific Reports. 11. 10.1038/s41598-020-80820-1.

[7] K. N. Bhanu, H. J. Jasmine and H. S. Mahadevaswamy, "Machine learning Implementation in IoT based Intelligent System for Agriculture," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9153978.

[8] S. G, N. M. Davasam and N. M. Krishna, "Smart Farming Using IoT and Machine Learning Techniques," 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC), 2020, pp. 1-5, doi: 10.1109/B-HTC50970.2020.9297916.

[9] Fenu, Gianni & Malloci, Francesca. (2021). Forecasting Plant and Crop Disease: An Explorative Study on Current Algorithms. Big Data and Cognitive Computing. 5. 2. 10.3390/bdcc5010002.

[10] K. S. Pratyush Reddy, Y. M. Roopa, K. Rajeev L.N. and N. S. Nandan, "IoT based Smart Agriculture using Machine Learning," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020, pp. 130-134, doi: 10.1109/ICIRCA48905.2020.9183373.

[11] R. Reshma, V. Sathiyavathi, T. Sindhu, K. Selvakumar and L. SaiRamesh, "IoT based Classification Techniques for Soil Content Analysis and Crop Yield Prediction," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 156-160, doi: 10.1109/I-SMAC49090.2020.9243600.

[12] J. Kanimozhi, G. Pavithra, K. Pooja and S. Saranya, "A Study of Smart Farming Based on IOT," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), 2020, pp. 1-6, doi: 10.1109/ICSCAN49426.2020.9262277.

[13] J. G. N. Zannou and V. R. Houndji, "Sorghum Yield Prediction using Machine Learning," 2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART), 2019, pp. 1-4, doi: 10.1109/BIOSMART.2019.8734219.

[14] S. Y. Chaganti, P. Ainapur, M. Singh, Sangamesh and S. O. R., "Prediction Based Smart Farming," 2019 2nd International Conference of Computer and Informatics Engineering (IC2IE), 2019, pp. 204-209, doi: 10.1109/IC2IE47452.2019.8940834.

[15] M. R. S. Muthusinghe, P. S.T., W. A. N. D. Weerakkody, A. M. H. Saranga and W. H. Rankothge, "Towards Smart Farming: Accurate Prediction of Paddy Harvest and Rice Demand," 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), 2018, pp. 1-6, doi: 10.1109/R10-HTC.2018.8629843.

[16] D. Elavarasan and P. M. D. Vincent, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications," in IEEE Access, vol. 8, pp. 86886-86901, 2020, doi: 10.1109/ACCESS.2020.2992480.

[17] P. Kanaga Priya and N. Yuvaraj 2019 J. Phys.: Conf. Ser. 1362 012038

[18] Z. Sun, L. Di, H. Fang and A. Burgess, "Deep Learning Classification for Crop Types in North Dakota," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 2200-2213, 2020, doi: 10.1109/JSTARS.2020.2990104.

[19] M. M. G. de Macedo, A. B. Mattos and D. A. B. Oliveira, "Generalization of Convolutional LSTM Models for Crop Area Estimation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 1134-1142, 2020, doi: 10.1109/JSTARS.2020.2973602.

[20] Kerner, Hannah & Sahajpal, Ritvik & Skakun, Sergii & Becker-Reshef, Inbal & Barker, Brian & Hosseini, Mehdi & Puricelli, Estefania & Gray, Patrick. (2020). Resilient In-Season Crop Type Classification in Multispectral Satellite Observations using Growth Stage Normalization.

[21] Plaat, Artur & Treuren, Rob & Hintum, Theo. (2021). Reliable genomic strategies for species classification of plant genetic resources. BMC Bioinformatics. 22. 10.1186/s12859-021-04018-6.

[22] Kulkarni, Nachiket & Thakur, Amit & Rajwal, Tanay & Tornekar, Rupali & Patil, Saraswati. (2019). Smart soil nutrients analysis and prediction of the level of nutrients using a bot. 663-668. 10.1109/RDCAPE47089.2019.8979007.

[23] Aicha Moumni, Abderrahman Lahrouni, "Machine Learning-Based Classification for Crop-Type Mapping Using the Fusion of High-Resolution Satellite Imagery in a Semiarid Area", Scientifica, vol. 2021, Article
ID 8810279, 20 pages, 2021. https://doi.org/10.1155/2021/8810279

[24] Mupangwa, W., Chipindu, L., Nyagumbo, I. et al. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa. SN Appl. Sci. 2, 952 (2020). https://doi.org/10.1007/s42452-020-2711-6

[25] Campos-Taberner, M., García-Haro, F.J., Martínez, B. et al. Understanding deep learning in land use classification based on Sentinel-2 time series. Sci Rep 10, 17188 (2020). https://doi.org/10.1038/s41598-020-74215-5

[26] Ruß G. (2009) Data Mining of Agricultural Yield Data: A Comparison of Regression Models. In: Perner P. (eds) Advances in Data Mining. Applications and Theoretical Aspects. ICDM 2009. Lecture Notes in Computer Science, vol 5633. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-03067-3_3

[27] Majumdar, P. and Mitra, S. (2021). IoT and Machine Learning-Based Approaches for Real Time Environment Parameters Monitoring in Agriculture: An Empirical Review. In Agricultural Informatics (eds A. Choudhury, A. Biswas, M. Prateek and A. Chakrabarti). https://doi.org/10.1002/9781119769231.ch5

[28] S. V. Bhosale, R. A. Thombare, P. G. Dhemey and A. N. Chaudhari, "Crop Yield Prediction Using Data Analytics and Hybrid Approach," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697806.

[29] Yi Z, Jia L, Chen Q. Crop Classification Using Multi-Temporal Sentinel-2 Data in the Shiyang River Basin of China. Remote Sensing. 2020; 12(24):4052. https://doi.org/10.3390/rs12244052

[30] Sinwar, Deepak & Dhaka, Vijaypal & Sharma, Manoj Kumar & Rani, Geeta. (2019). AI-Based Yield Prediction and Smart Irrigation. 10.1007/978-981-15-0663-5_8.

[31] A. Suruliandi, G. Mariammal & S.P. Raja (2021) Crop prediction based on soil and environmental characteristics using feature selection techniques, Mathematical and Computer Modelling of Dynamical Systems, 27:1, 117-140, DOI: 10.1080/13873954.2021.1882505

[32] Saiz-Rubio, Veronica & Rovira-Más, Francisco. (2020). From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management. Agronomy. 10. 10.3390/agronomy10020207.

[33] Xie, Qinghua & Lai, Kunyu & Wang, Jinfei & Lopez-Sanchez, Juan & Shang, Jiali & Liao, Chunhua & Jian-Jun, Zhu & Fu, Haiqiang & Xing, Peng. (2021). Crop Monitoring and Classification Using Polarimetric RADARSAT-2 Time-Series Data Across Growing Season: A Case Study in Southwestern Ontario, Canada. Remote Sensing. 13. 1394. 10.3390/rs13071394.

[34] Rousi, Maria & Sitokonstantinou, Vasileios & Meditskos, Georgios & Papoutsis, Ioannis & Gialampoukidis, Ilias & Koukos, Alkis & Karathanassi, Vassilia & Drivas, Thanassis & Vrochidis, Stefanos & Charalabos, Kontoes & Kompatsiaris, Ioannis. (2020). Semantically Enriched Crop Type Classification and Linked Earth Observation Data to Support the Common Agricultural Policy Monitoring. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. PP. 1-1. 10.1109/JSTARS.2020.3038152.

[35] Khan, Nawab & Ray, Ram & Sargani, Raza & M., Ihtisham & Khayyam, Muhammad & Ismail, Sohaib. (2021). Current Progress and Future Prospects of Agriculture Technology: Gateway to Sustainable Agriculture. Sustainability. 13. 10.3390/su13094883.

[36] A. Suruliandi, G. Mariammal & S.P. Raja (2021) Crop prediction based on soil and environmental characteristics using feature selection techniques, Mathematical and Computer Modelling of Dynamical Systems, 27:1, 117-140, DOI: 10.1080/13873954.2021.1882505

[37] D. J. Reddy and M. R. Kumar, "Crop Yield Prediction using Machine Learning Algorithm," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1466-1470, doi: 10.1109/ICICCS51141.2021.9432236.

[38] Marion Olubunmi Adebiyi, Roseline Oluwaseun Ogundokun, Aneoghena Amarachi Abokhai, "Machine Learning–Based Predictive Farmland Optimization and Crop Monitoring System", Scientifica, vol. 2020, Article
ID 9428281, 12 pages, 2020. https://doi.org/10.1155/2020/9428281

[39] Crop Yield Prediction Integrating Genotype and Weather Variables Using Deep Learning, https://arxiv.org/abs/2006.13847

[40] Ms Deepali Jawale and Dr Sandeep Malik, "Design of a high efficiency augmented machine learning model for multiparametric crop type & yield prediction" Telematique, Vol 21, No 1 (2022)

[41] https://medium.com/mlearning-ai/image-detection-using-convolutional-neural-networks-89c9e21fffa3

[42] Paudel, Dilli, Hendrik Boogaard, Allard de Wit, Sander Janssen, Sjoukje Osinga, Christos Pylianidis, and Ioannis N. Athanasiadis. "Machine learning for large-scale crop yield forecasting." Agricultural Systems 187 (2021): 103016.

[43] Jawale, Ms Deepali, and Sandip Malik. "Design of A High Efficiency Augmented Machine Learning Model For Multiparametric Crop Type And Yield Prediction." Telematique (2022): 5551-5568.