Journal of Engineering Research and Reports

Volume 24, Issue 12, Page 29-44, 2023; Article no.JERR.98594
ISSN: 2582-2926

Forecasting of Crop Yield using Remote Sensing Data, Agrarian Factors and Machine Learning Approaches

Jasmin Praful Bharadiya [a*], Nikolaos Tzenios [b] and Manjunath Reddy [c]

[a] *University of the Cumberlands, United States.*
[b] *Charisma University, United Kingdom.*
[c] *Qualcomm , San diego ,CA, USA.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JERR/2023/v24i12858

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/98594

Original Research Article

Received: 09/02/2023
Accepted: 11/04/2023
Published: 15/04/2023

ABSTRACT

The art of predicting crop production is done before the crop is harvested. Crop output forecasts will help people make timely judgments concerning food policy, prices in markets, import and export laws, and acceptable warehousing. It is possible to reduce the socioeconomic effects of crop loss brought on by a natural disaster, such as a flood or a drought, and to organize humanitarian food assistance. It has been suggested that deep learning, which lets the model to automatically extricate features and learn from the datasets, could be useful for predicting agricultural yields. This review helps to understand that how vegetation indices and environmental variables affect agricultural output by revealing gaps in our understanding of deep learning methodologies and remote sensing data in a specific area. Literature review of 2011-2022 has been collected from different databases and sites and analyzed to meet the aims of this review. The study mainly focused on the benefits of machine learning, agrarian factors and remote sensing for forecasting

---

**Corresponding author: Email: jasminbharadiya92@gmail.com;*

crop yield. The most often employed form of remote sensing is satellite technology, namely the usage of the Moderate-Resolution Imaging Spectro radiometer. Vegetation indices referred to as the most often employed attribute for forecasting crop yield, according to the results. This review compares all these techniques and pros and cons related to them.

Keywords: Machine learning; artificial intelligence; yield prediction; algorithms; agriculture; remote sensing.

1. INTRODUCTION

Predicting crop yields is one of the most difficult problems in agriculture. It is crucial to decision-making at the international, regional, and local levels. Crop, soil, climatic, environmental, and other characteristics are used to predict agricultural yield. Machine learning, a branch of artificial intelligence, allows computers to learn from data without having to be explicitly programmed. The use of machine learning has improved thanks to big data technology. A considerable amount of data that is produced quickly from multiple sources is simply referred to as big data [1-4]. For the production of food on a worldwide scale, crop yield prediction is crucial. Policymakers rely on accurate projections to quickly decide which foods to buy and export in order to improve national food security. To breed for superior types, seed firms must forecast how new hybrids will function in diverse conditions. The ability to estimate production helps farmers and growers make wise management and financial decisions [5-9]. The temperature, the soil, the crop, the use of fertilizer, and the type of seeds used are some of the factors that affect crop production [10,11]. For accurate agricultural yield estimation findings, a variety of crop simulation and yield estimation methods have been applied [11,12]. Researchers frequently use Deep Learning techniques to estimate agricultural yields based on the aforementioned variables [11].

Crop output forecasting is becoming increasingly important due to growing worries about food security. Early predictions of crop production, which foretell the availability of food for the expanding world population, can considerably reduce famine [13]. One of the most serious issues of our day, ending world hunger, may be accomplished by increasing food yields. Despite recent progress, the World Health Organization estimates that 820 million people still lack access to enough food globally [14]. By 2030, the Sustainable Development Goals of the UN with a focus on agriculture seek to eliminate hunger issues, achieve security of food, and promote more sustainable agriculture [15]. Crop output forecasts may provide essential data for creating a viable plan to attain the goal of ending hunger [16-18]. Many factors need be taken into account while forecasting agriculture yield, which makes difficult to create a reliable predicting model using conventional methods [19]. Yet, recent developments in computer technology have opened up the prospect of developing and honing a fresh approach to forecasting agricultural productivity. Since deep learning makes use of a wide variety of data technologies and can handle large amounts of information quickly, it finds widespread application in the agriculture industry [20]. The term "deep learning" refers to a kind of machine learning that makes use of numerous layers of neural networks and is capable of learning from inputs that are both unstructured and unlabeled. Depending on the learning environment, the learning may be supervised, semi-supervised, or unsupervised [21,22]. Sarker [23] pointed out that, in contrast to typical machine learning techniques, deep learning models focus on learning abstract properties from large datasets. It is essential to have a full grasp of the interactions that exist between functional qualities and interacting variables in order to accurately predict crop yield. Large datasets and high-efficiency algorithms are needed for such correlations; deep learning can be used to achieve both of these goals [24,25]. Since machine learning has been widely studied over the past 10 years, it is currently being applied globally to forecast and boost agricultural produce outputs [11,18,26]. Multivariate regression, decision trees, association rule mining, and artificial neural networks are only some of the machine learning methods that have been used to forecast agricultural yields in recent years. One defining feature of machine learning models is their implicit treatment of the output (crop yield) as a function of the input variables (genes and environmental factors), which may be a very intricate and non-linear function [27,28].

Another important technique to predict yield in agriculture is remote sensing. The science of remote sensing is the non-intrusive method of learning about an object or phenomena without coming into close touch with it. It is employed in agriculture to keep an eye on the moisture, soil, and crop conditions. Remote sensing makes use of electromagnetic radiation emissions such radio waves, microwaves, infrared, visible light, and ultraviolet light. Crop growth conditions can be tracked over time via remote sensing of crops. Additionally, it offers details on the condition of crops at particular junctures in time and space. Crop yields may be calculated using this data, and it can also predict when the harvest will take place. Remote sensing data can be used to track changes in land usage, track agricultural production and growth, evaluate salinity and moisture levels in the soil, assess pest infestation levels, and more [29,30]. The study provides a review of the literature on various remote sensing methods, deep learning models and various techniques that are utilized with satellite data. Many models are created, and calculated results are contrasted with benchmark models that are also supplied.

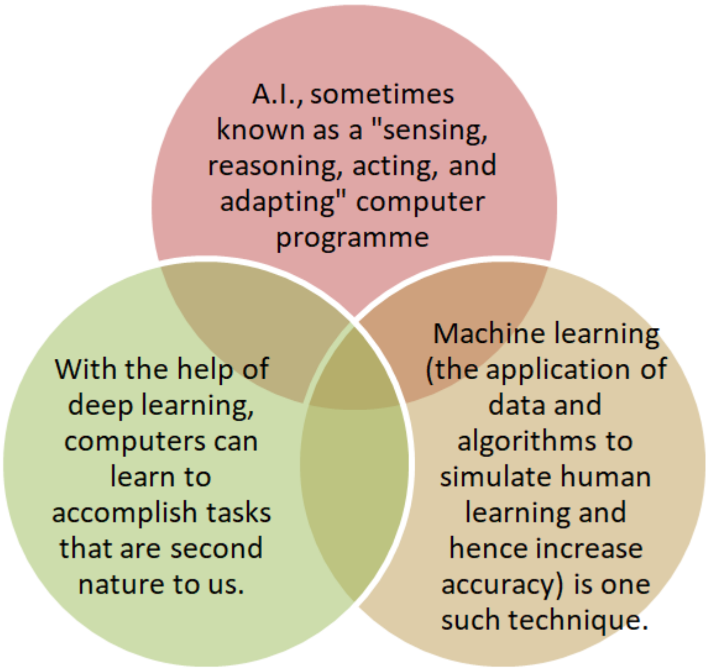

Fig. 1. How machine learning, deep learning and artificial intelligence are interlinked

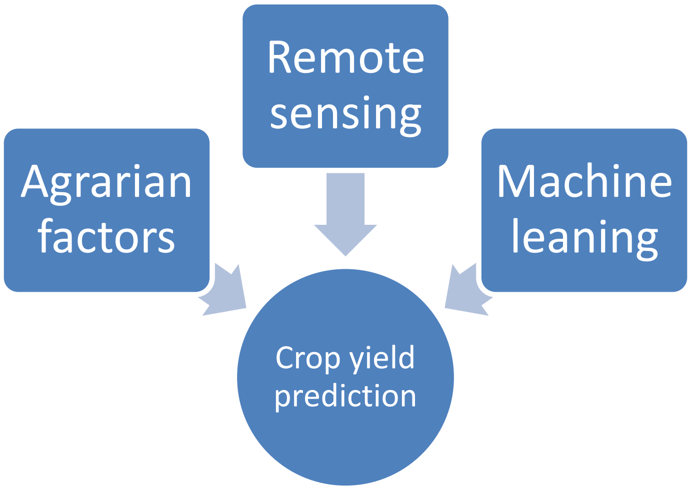

Fig. 2. Crop yield forecasting methods

2. Existing AL Techniques IN AGRICULTURE SECTOR

2.1 Deep Learning

Due to their limited applicability and unpredictability, traditional methodologies like the static regression approach and the mechanistic approach make it difficult to develop a crop production forecast model that is accurate [16,24,31]. For the prediction of crop yield, many researchers have employed ML approaches such as regression trees, random forests, multivariate regression, association rule mining, and artificial neural networks [12,32-35]. Machine learning models view the output, or crop production, as an implicit function of the input variables, which can include things like weather and soil conditions. Furthermore, the nonlinear link between input and output variables is lost on supervised learning methods employed in machine learning [36-39]. Yet, recent technological developments have made it possible to create a sophisticated model for predicting agricultural yields using deep learning [40]. Since deep learning can examine enormous datasets, discover correlations between different variables, and employ nonlinear functions, it is widely applied in the agriculture sector. In an unsupervised setting, these techniques can extract features for big datasets. Deep learning approaches outperform conventional machine learning methods in feature extraction [41-43]. Deep learning has a significant ability to extract features from the existing data because an effective agricultural yield prediction depends on the variables controlling crop growth.

Each layer of a deep neural network's nonlinear processing transforms unseen input data into a usable form [44]. Finding the nonlinear association between the input and response variables requires the use of deep neural networks with a variety of hidden layers [45]. However, they are challenging to train and require cutting-edge technology and optimization techniques [46]. So, adding more hidden layers can be useful but comes with some limitations that can be overcome by using certain strategies. Deeper neural networks can avoid the vanishing gradient problem by making use of the network's remaining skip connections [47]. Furthermore, by implementing several techniques that includes stochastic gradient descent (SGD), batch normalization, and dropout, the performance of deep learning systems has been enhanced. The following list contains a few deep learning techniques.

2.1.1 Deep Neural Networks (DNN)

The DNN techniques are relatively comparable to the traditional artificial neural networks ANN procedures, with the number of hidden layers being the only difference. DNN networks feature many hidden layers that are virtually always fully connected, just like ANN techniques [48].

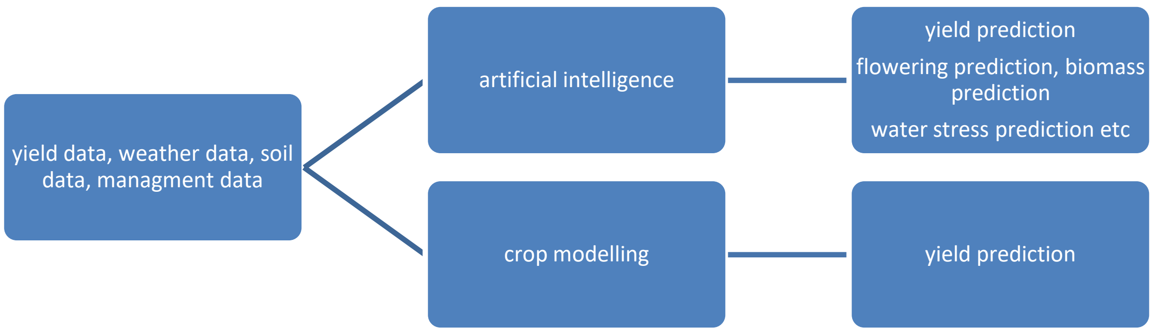

Fig. 3. Artificial Intelligence role in crop yielding

2.1.2 Convolutional neural network

Compared to a network with all connections, CNN has fewer parameters to learn. Three different kinds of layers—convolutional, pooling, and fully connected—combine to form a CNN model. Convolutional layers are made up of filters and feature maps. The neurons of the layer are filters, which generate a value from weighted inputs. The output of a filter is occasionally referred to as a feature map. Pooling layers are used to down sample the feature map of the preceding layers, generalize feature representations, and reduce over fitting. At the network's edge, predictions are often performed using fully connected layers. In CNN models, a pooling layer is often followed by one or more convolutional layers, and this structure is repeatedly used. In most cases, a pooling layer comes after one or more convolutional layers, and before fully linked layers are employed in CNN models, this pattern is repeated several times [49,50].

2.1.3 Long short term memory

For problems with sequence prediction, LSTM networks were developed specifically. The stacked LSTM, CNN-LSTM, encoder-decoder LSTM, bidirectional LSTM, and generative LSTM architectures are only a few examples of the numerous LSTM designs. Statelessness, insensitivity to temporal structure, messy scaling, fixed sized inputs, and fixed sized outputs are only a few of the shortcomings of Multi-Layer Perceptron (MLP) feedforward ANN methods. LSTM can be viewed as the network's loop addition when compared to the MLP network. The LSTM is a distinct variant of the Recurrent Neural Network (RNN) method. In addition to having an internal state, being aware of the temporal structure of the inputs, being able to simulate parallel input series, and processing variable-length input to produce variable-length output, LSTMs differ significantly from MLP networks in several respects. The memory cell serves as the LSTM's computational unit. These cells are made up of gates and weights (such as internal state, input weights, and output weights) (i.e., forget gate, input gate, and output gate) [51].

2.1.4 3D-CNN

The kernels in this network's variant of the CNN model travel through depth, height, and width. It consequently generates 3D activation maps. This kind of model was created to enhance the recognition of moving objects, such as in the case of security cameras and x-rays. In CNN's convolutional layers, 3D convolutions are conducted [52].

2.1.5 CNN-LSTM

The strength of various deep learning algorithms can be combined. As a result, researchers integrate several algorithms in various ways. Chu and Yu [53] developed a model for predicting crop productivity by combining Back-Propagation Neural Networks (BPNNs) and Independently Recurrent Neural Networks (IndRNN). Convolutional Neural Networks and Long-Short Term Memory Networks (CNN-LSTM) were integrated by Sun et al. [54] to predict soybean yield. Convolutional and recurrent neural networks were integrated (CNN-RNN) by Khaki et al. [42] to forecast yield. Wang et al. [55] merged CNN and LSTM networks (CNN-LSTM) to solve the challenge of predicting wheat yield.

2.1.6 Multi-Task Learning (MTL)

To enhance the performance of our models created for various tasks, we share representations between tasks in multi-task learning. It has been used in a variety of fields, including speech recognition, drug development, and natural language processing. Instead of focusing on enhancing the performance of just one task, the goal is to increase performance across the board. In their evaluation of several multi-task learning strategies for supervised learning tasks, Zhang and Yang also provided an explanation of how multi-task learning can be combined with other learning types, such as semi-supervised learning and reinforcement learning. The supervised MTL approaches were split into the following groups: decomposition approach, task relation learning approach, task clustering approach, feature learning approach, and low-rank approach [56,57].

2.1.7 Deep recurrent Q-Network (DQN)

In reinforcement learning, agents examine their surroundings and take appropriate action in accordance with the rules and information at hand. Agents work to maximize their benefits, which might be favorable or bad depending on their activities. Environment and agents are always interacting with one another. Researchers at Deep Mind, which Google purchased in 2014, created the DQN algorithm in 2015. In 2015, multiple Atari games were solved using the DQN technique, which combines the strength of reinforcement learning and deep neural networks. Deep neural networks were added to the traditional Q-learning algorithm, and the experience replay method was also included [58,59].

3. USING REMOTE SENSING TO PREDICT CROP YIELD

Agricultural output can change depending on the environment, weather, disease, and other factors. These aforementioned elements have an impact on crop growth at various stages, which has an impact on crop production. With the use of a range of instruments and methods, such as on-site surveying, ground observation, remote sensing, and global positioning systems, it is feasible to keep an eye on environmental conditions, other features, and crop growth. It is difficult to manually collect data for a big area using ground observation and other conventional methods, and the results will be less precise and unreliable. Remote sensing is currently being used more often for crop monitoring to overcome this constraint [64-66]. Remote sensing methods use spectral signatures to provide information on the status of crops at different growth levels that is comparable to thorough on-field surveying. Remote sensing technology is the non-contact, instrumental collection and analysis of data about the physical environment and its objects using a satellite or device put in the atmosphere. When compared to other methods of data acquisition, such as field surveying, remote sensing has the capacity to create a sufficient amount of data. It is the technique of seeing and identifying locations on Earth by utilizing sensors to measure the radiation that is emitted and reflected [67-72].

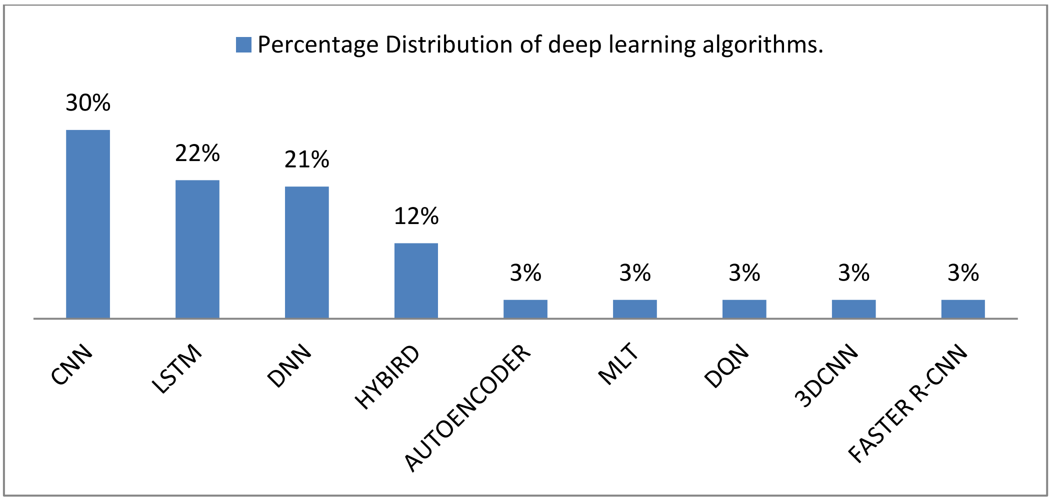

Fig. 4. Percentage distribution of DL techniques

Table 1. Examples of some AL/ML techniques used in crop yield prediction

Machine Learning techniques used	Year	Crops that were predicted	Reference
Long Short-Term Memory + Recurrent Neural Network	2021	Wheat	[60]
Gradient Boosting + k-Nearest Neighbors + Support Vector Regression	2021	1. Sunflower 2. Sugar beet 3. Potatoes 4. wheat 5. barley	[61]
Recurrent Neural Network + Long Short-Term Memory	2021	Wheat	[60]
Support Vector Regression, k-Nearest Neighbor, linear regression and Elastic net	2020	Potato	[40]
RNN and CNN	2020	Corn	[62]
Deep Neural Network	2019	Corn Hybrids	[63]

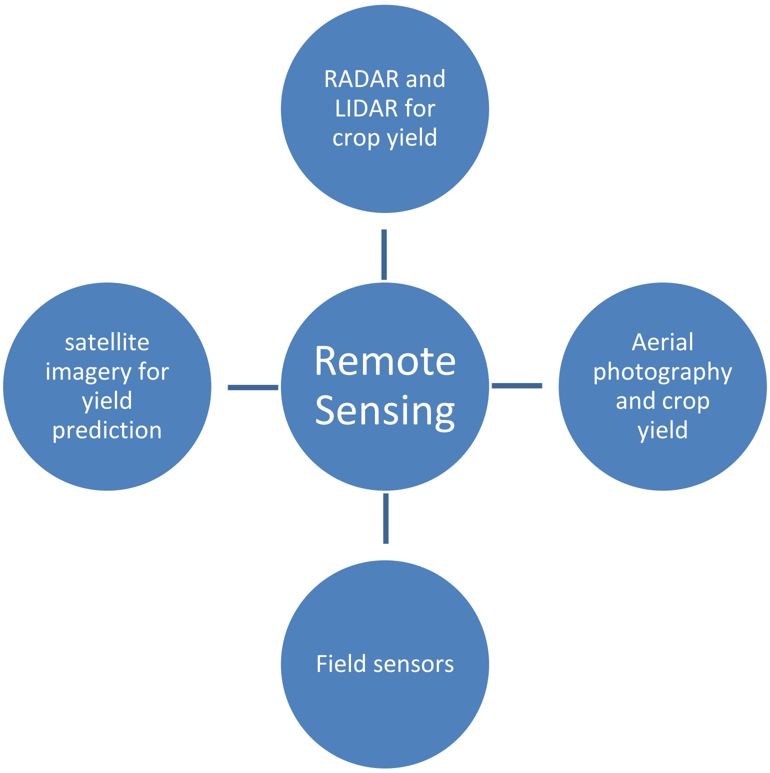

Fig. 5. Uses of Remote sensing technology

Table 2. List of remote index used in forecasting of crop yield [76]

Greenness index: enhanced vegetation index (EVI), enhanced vegetation index with two bands (EVI2), excessive green index Normalized difference vegetation index (NDVI), normalized difference red edge index, in-season estimated yield (INSEY), greenness index, and greenness normalized difference vegetation index (GNDVI) (NDRE), Broad dynamic range vegetation index, transformed soil adjusted vegetation index, and saturation-adjusted normalized difference vegetation index (SANDVI),
Chlorophyll index: Chlorophyll index red-edge, Fraction of absorbed photosynthetically active radiance (FAPAR)
Photochemical index: Photochemical reflectance index (PRI), Structural independent pigment index (SIPI)
Dryness index: Crop water stress index (CWSI), Normalized difference water index (NDWI), Normalized water index (NWI), Standardized crop-specific drought index (S-CSDI)
Temperature index: Land surface temperature (LST) , Vegetation condition index (VCI)
Ratio index: Area weighted radiance ratio, Green ratio vegetation index, Inverse simple ratio red-edge.
Reflectance/backscatter: Optical sensors: reflectance from green, NIR and Red, etc, Synthetic aperture radar sensors: backscatters from C, X, and L band

An important justification for using optical remote sensing to obtain agricultural data is the measurement of vegetation indices. Combinations of spectral observations at multiple wavelengths are known as "spectral indices." They are used to calculate biophysical parameters and derive vegetation phenology [73]. Of the several spectral indices, vegetation indices are the ideal indices that are frequently utilized in crop yield prediction. Healthy crops exhibit high red and near-infrared band absorption and reflection [74]. Many quantitative measures of the vegetative environment can be built using the stark contrast between the red and near-infrared bands' levels of absorption and reflection. Vegetation indices (VI) are the names for these linear or nonlinear combinational processes [75]. Green vegetation index (GVI), chlorophyll absorption ratio index (CARI), and normalized difference vegetation index (NDVI) are a few examples of the VI.

4. AGRARIAN FACTORS AND YIELD FORECASTING

Various factors affect the yield of crops including biotic and Abiotic factors. The hybrid system used for agricultural dynamic monitoring illustrates the key elements influencing crop yield in below Fig. In general, crop output is forecast by keeping an eye on a variety of elements, including water usage, plantation size, soil quality, weather, disease incidences, and more.

There are several variables that affect crop productivity and the inherent risks of farming. All the factors mentioned in above Fig. are he most important while predicting crop yield. When these factors are not adequately evaluated and managed, they can pose a significant risk to farmers. Also, it is crucial to understand precisely what effects crop productivity and the liabilities involved in order to increase crop yield and reduce risk [77].

Ml and remote sensing helps to access these factors more effectively.

5. HOW ML HELPS AGRARIAN FACTORS AND REMOTE SENSING IN CROP YIELD FORECASTING?

While remotely sensed photos typically offer more spatio-temporal-spectral information that may be exploited, more subtle and diverse patterns, and more complex patterns, there are stricter limitations on how these images can be processed than for natural photographs. The incorporation of DL into environmental remote sensing has allowed for its use in a wide variety of applications, such as land cover mapping, environmental parameter retrieval, data fusion and downscaling, information production and prediction, and so on, all thanks to DL's superior ability in feature representation [78].

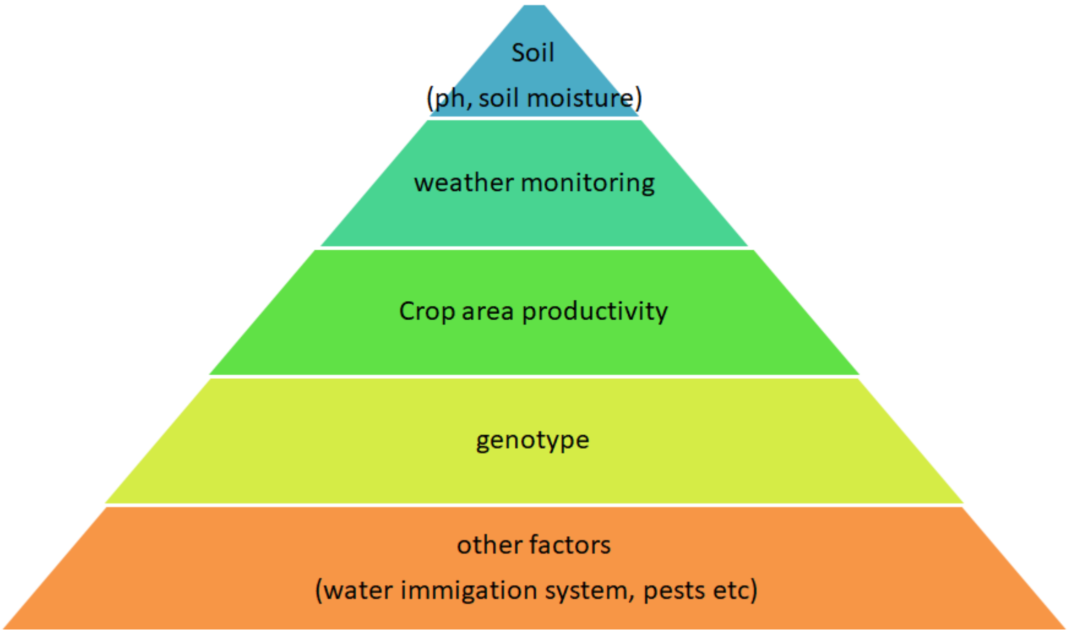

Fig. 6. Agrarian factorsaffecting crop yield

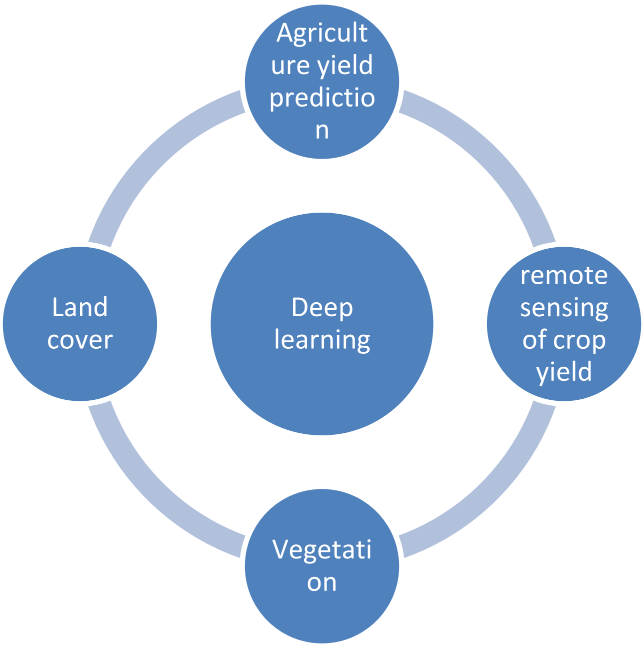

Fig. 7. Applications of deep learning

5.1 Mapping of Land Cover

Image categorization is required for the mapping of land cover from remote sensing data. According to various spatial units, such as moving windows, objects, and scenes as well as pixels, traditional classification algorithms identify photographs [79]. Unfortunately, it is usually difficult to distinguish between the complex terrain structures or patterns by using a small number of rules because standard methods only use low-level data in the spectral and spatial domains for categorization. As a result, methods for classifying data that incorporate a lot of features at high levels are recommended. The best results were obtained when DL was recently used to land cover mapping due to its benefits in multiscale and multilayer feature extraction [80]. In complex urban settings, the deep learning-based classification strategy offers substantial advantages in terms of classification accuracy compared to the traditional rule-based and ML methods. Current applications have shown the promise of DL-based land-cover classification methods due to the necessity for land cover mapping from high-resolution and even very-high-resolution satellite imageries.

5.2 Environmental Parameter Retrieval

Physical models that are based on systematic environmental data are frequently used in remote sensing to retrieve environmental parameters. The physical processes, however, are quite intricate and include a large number of model factors. Additionally, several environmental phenomena still lack a reliable physical model. This makes it possible for deep learning or machine learning to recover environmental factors. To begin with, deep leaning can replicate or condense the physical models for retrieving environmental factors. Physical models require extremely complicated calculation, and DL can be used in the forward simulation of physical models due to its significant simulation capability. As a result, retrieving environmental parameters can be made simpler. Second, due to its ability to approximate complex relationships, The statistical link between remote sensing measurements and in-situ environmental parameters can be determined using deep learning [81,82]. This can achieve a comparable performance without using complex physical models. Maybe more crucially, DL can offer an alternate and workable method for retrieving environmental parameters in particular environmental phenomena where there are no reliable physical models available.

5.3 Agricultural Yield Prediction by Remote Sensing, Agrarian Factors and Machine Learning

Large-area agricultural yield projections can assist policymakers and grain marketing organisations in making export and import plans

[90]. By building models connecting yields and influencing (like weather, soil conditions, terrain, disease, and vegetation growth conditions) and human (like irrigation and fertilizer management) elements, the majority of available methods to forecast agriculture yield few months before harvesting. With remote sensing data collected over huge areas, some parameters can be calculated. An articulation controller NN model for the cerebellum was created by Desachy and Simpson in 1994. They discovered that the addition of remote sensing data, such as Landsat Thematic Mapper (TM) observations based on agricultural data and climatic factors, will increase the prediction model's accuracy. Moreover, by utilizing remote sensing vegetation indexes and other parameters, NNs surpassed the conventional linear regression approaches in the prediction of crop yield [38,91-93]. NDVI is the most widely used index. Using historical yield data, MODIS, and AVHRR NDVI, Ju et al developed the shuffled complex evolution technique (SCE-UA) optimization NN approach to estimate corn and soybean yields [94]. In one study, sugarcane yields were forecasted using MODIS NDVI and an ensemble model of NN. The initial data set's redundant and unnecessary characteristics were eliminated using a sequential backward elimination NN wrapper. Similar research has also been done on other unique vegetation indices [95]. Using crop yield data, Johnson and his colleagues built Bayesian NNs in each hierarchically grouped region to assess the MODIS NDVI, MODIS EVI, and AVHRR NDVI. For all three crops, they discovered that MODIS NDVI was a reliable prediction, and MODIS EVI was an improved predictor [96]. NDVI, green vegetation index, soil-adjusted vegetation index, and perpendicular vegetation indices were used in one research to create the BPNN model. The outcomes showed that the grid images of perpendicular vegetation index were accurate in predicting the corn production [97]. In one study, the effect of irrigation on lettuce output was investigated by building a neural network model with the use of the NDVI, chlorophyll green, simple ratio, and red-edge chlorophyll. The scientists discovered that a drop in irrigation water caused a fall in lettuce yield. To create prediction models between auxiliary factors and agricultural yields, some more types of neural networks are also utilized [98]. Researchers used a fuzzy neural networks (FNN) or granular neural networks (GNN) to forecast crop yields using simulation parameters from the Crop Growth Monitoring System and SPOT NDVI [99]. In comparison to the conventional approaches, the use of neural networks and deep learning to predict agricultural output is significantly improved with the addition of remote sensing data based on meteorological data. There are numerous different retrieval models available right now. The forecast model's robustness, however, is limited as a result of

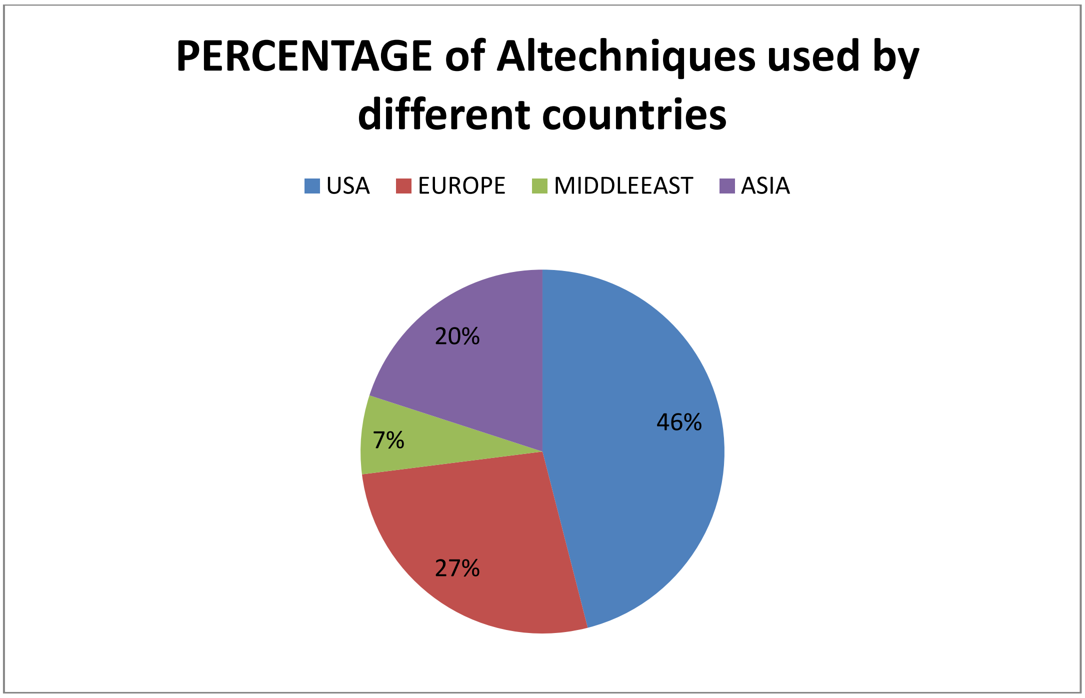

Fig. 8. Estimation of usage of AL in different countries to predict crop yield

Table 3. Some studies of varios crop yield pridticng techniques

Technique	Results	Reference
Neural network (Back propagation)	The suggested and trained models total four. Benchmark MLR model was outperformed by the fourth model.	[83]
Model named SRS (model of simulation remote sensing)	There were three different input types, and then calculation of LAI was conducted. When given AVHRR GAC input, the model produces good results.	[84]
Model named Monteith	The accuracy of model declines as crop heterogeneity increases.	[85, 86]
SVR Model (support vector regression)	Calculated MAPE & MAE were compared to other commercially available models. The proposed model's MAPE is higher but still within acceptable bounds.	[86]
Model RS-P-YEC (yield estimation for crop)	Data from meteorology as well as remote sensing were utilized. The outcomes of this model are contrasted with meteorological station observations, where R2 hits 0.817.	[87]
GPR , RFR, SVR and BRT are used.	The performance of machine learning approaches is superior to that of traditional regression techniques.	[88]
Prediction based on weather	The results show sensitivity of 89.36% + specificity of 91.72%.+ accuracy of 94.5%.	[89]

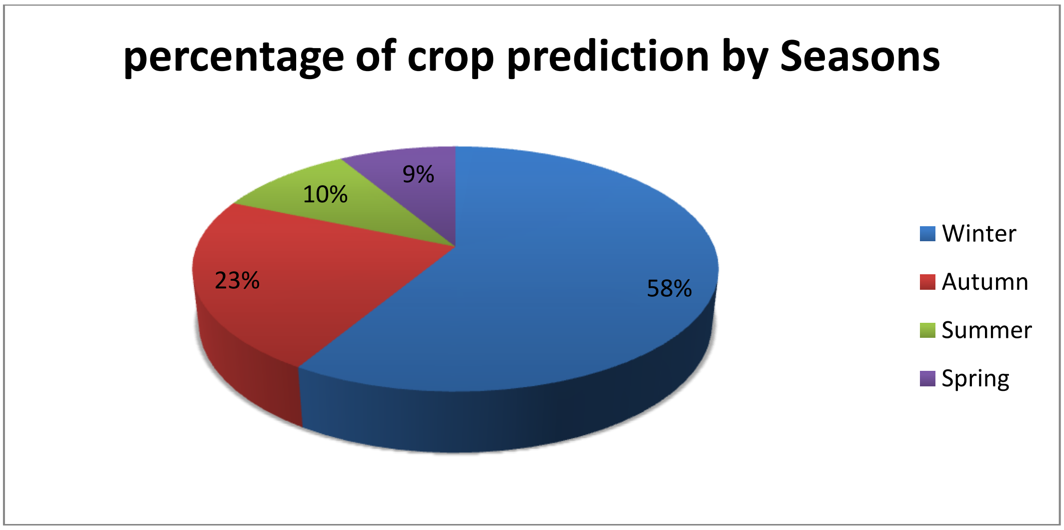

Fig. 9. Crop yield prediction in different seasons

specific circumstances, including various crop types, topography, and climate. Remote sensing may be used throughout the entire agricultural production cycle, from soil preparation to harvesting. Due to the advancement of low-cost unmanned aerial vehicles, high spatial and temporal resolution satellite data, and field hyper spectral measurements, remote sensing agricultural applications have undergone a significant transformation. Satellite data continue to be the most efficient remote sensing technique for scanning large areas and monitoring changes in national and regional agriculture [100]. In addition, high-precision forecasts are usually only applicable to the study area. So, increasing the universality and migration of the crop yield forecast model is a popular but tough field of future research.

6. CONCLUSIONS

The expected crop production is an important piece of data. This may be accomplished via surveys, statistical models as well as machine learning. Agricultural output is affected by many factors, including climate, soil type, soil nutrients, crop nutrients, crop canopy volume and biomass, water content, disease, weeds, insects, and cultivar and variety. The effects of the aforementioned factors may be observed in the crop's spectroscopic characteristics, which can be assessed by a variety of remote sensors. Crop yield may be tracked, assessed, and estimated quickly, affordably, and effectively using remote sensing. In this study, a detailed assessment of the use of DL techniques for agricultural production forecasting by using remote sensing data has been conducted. The objectives of this were to give useful information on how vegetation indices and environmental variables influence crop production forecast and to highlight the research gaps that still needed to be addressed in a specific field of deep learning methods. This comprehensive study of the literature has shown the various deep learning techniques, remote data sensing and agrarian parameters utilized for agricultural output forecasting. All deep learning algorithms may forecast crop output based on the variables and parameters included in the various models.. Based on the findings of this review, it is determined that the vegetation indices and meteorological data, which define the characteristics of the crops and help in monitoring the climatic conditions that directly influence crop yield forecast, are the most often utilized aspects. Furthermore, it is evident that the factors that affect crop yield forecasting are influenced by the crop yield and how it relates to other variables. Still further research is needed to examine the pro and cons of various techniques.

COMPETING INTERESTS

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. VanderPlas J. Python data science handbook: Essential tools for working with data.: "O'Reilly Media, Inc"; 2016.
2. Abu Al-Haij Q, Krichen M, Abu Elhaija W. Machine-learning-based darknet traffic detection system for IoT applications. Electronics. 2022;11(4):556.
3. Mihoub A, et al. Predicting covid-19 spread level using socio-economic indicators and machine learning techniques. in 2020 first international conference of smart systems and emerging technologies (SMARTTECH). IEEE; 2020.
4. Srinivasan S, et al. Deep convolutional neural network based image spam classification. in 2020 6th conference on data science and machine learning applications (CDMA). IEEE; 2020.
5. Horie T, Yajima M, Nakagawa H. Yield forecasting. Agricultural systems, 1992; 40(1-3):211-236.
6. Archontoulis SV, et al. Predicting crop yields and soil-plant nitrogen dynamics in the US Corn Belt. Crop Science. 2020; 60(2):721-738.
7. Bogard M, et al. Linking genetic maps and simulation to optimize breeding for wheat flowering time in current and future climates. Crop Science. 2020;60(2):678-699.
8. Ersoz ES, Martin NF, Stapleton AE. On to the next chapter for crop breeding: convergence with data science. Crop Science. 2020;60(2):639-655.
9. Washburn JD, Burch MB, Franco JAV. Predictive breeding for maize: Making use of molecular phenotypes, machine learning, and physiological crop models. Crop Science. 2020;60(2):622-638.
10. Xu X, et al. Design of an integrated climatic assessment indicator (ICAI) for wheat production: A case study in Jiangsu Province, China. Ecological Indicators. 2019;101:943-953.
11. Van Klompenburg T, Kassahun A, Catal C. Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture. 2020;177:105709.
12. Filippi P, et al. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agriculture. 2019;20:1015-1029.
13. Breure MS, Kempen B, Hoffland E. Spatial predictions of maize yields using QUEFTS–A comparison of methods. Geoderma. 2022;425:116018.
14. Holzworth DP, et al. Agricultural production systems modelling and software: current status and future prospects. Environmental Modelling and Software. 2015;72:276-286.
15. Tsani S, Koundouri P, Akinsete E. Resource management and sustainable development: A review of the European water policies in accordance with the United Nations' Sustainable Development

Goals. Environmental Science and Policy. 2020;114:570-579.
16. Kheir AM, et al. Calibration and validation of AQUACROP and APSIM models to optimize wheat yield and water saving in arid regions. Land. 2021;10(12):1375.
17. You J, et al. Deep gaussian process for crop yield prediction based on remote sensing data. in Proceedings of the AAAI conference on artificial intelligence; 2017.
18. Wang AX, et al. Deep transfer learning for crop yield prediction with remote sensing data. in Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies; 2018.
19. Wassmann R, et al. Regional vulnerability of climate change impacts on Asian rice production and scope for adaptation. Advances in agronomy. 2009;102:91-133.
20. Guo WW, Xue H. An incorporative statistic and neural approach for crop yield modelling and forecasting. Neural Computing and Application. 2012;21:109-117.
21. Qian B, et al. Statistical spring wheat yield forecasting for the Canadian prairie provinces. Agricultural and forest meteorology. 2009;149(6-7):1022-1031.
22. Prasad AK, et al. Crop yield estimation model for Iowa using remote sensing and surface parameters. International Journal of Applied earth observation and geoinformation. 2006;8(1):26-33.
23. Sarker IH. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN Comput Sci. 2021;2(6):420.
24. Tranfield D, Denyer D, Smart P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management. 2003; 14(3):207-222.
25. Shao Y, Ren J, Campbell J. Multitemporal remote sensing data analysis for agricultural application; 2018.
26. Lobell DB, et al. A scalable satellite-based crop yield mapper. Remote Sensing of Environment. 2015;164:324-333.
27. Marko O, et al. Soybean varieties portfolio optimisation based on yield prediction. Computers and Electronics in Agriculture. 2016;127:467-474.
28. Romero JR, et al. Using classification algorithms for predicting durum wheat yield in the province of Buenos Aires. Computers and Electronics in Agriculture. 2013;96:173-179.
29. Ali AM, et al. Integrated method for rice cultivation monitoring using Sentinel-2 data and Leaf Area Index. The Egyptian Journal of Remote Sensing and Space Science. 2021;24(3,1):431-441.
30. Bendig J, et al. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. Remote Sensing. 2014;6(11):10395-10412.
31. Sarker IH. Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. SN Computer Science. 2021; 2(6):420.
32. Kogan F, et al. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. International Journal of Applied Earth Observation and Geoinformation. 2013;23:192-203.
33. Kogan F, et al. Space-based vegetation health for wheat yield modeling and prediction in Australia. Journal of Applied Remote Sensing. 2018;12(2):026002-026002.
34. Cai Y, et al. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. Agricultural and forest meteorology. 2019;274:144-159.
35. Bhojani SH, Bhatt N. Wheat crop yield prediction using new activation functions in neural network. Neural Computing and Applications. 2020;32:13941-13951.
36. Page MJ, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. International journal of surgery. 2021;88:105906.
37. Shen C. A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research. 2018;54(11):8558-8593.
38. Jeong JH, et al. Random forests for global and regional crop yield predictions. PloS one. 2016;11(6): e0156571.
39. Oguntunde PG, Lischeid G, Dietrich O. Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis. International Journal of Biometeorology. 2018;62(3):459-469.

40. Abbas F, et al. Crop yield prediction through proximal sensing and machine learning algorithms. Agronomy. 2020; 10(7):1046.
41. Islam N, et al. Machine learning based approach for Weed Detection in Chilli field using RGB images. Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery. 2021;1097-1105.
42. Khaki S, Wang L, Archontoulis SV. A cnn-rnn framework for crop yield prediction. Frontiers in Plant Science. 2020;10:1750.
43. Szegedy C, et al. Going deeper with convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015.
44. Johnson MD, et al. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agricultural and Forest Meteorology. 2016;218:74-84.
45. Yamashita R, et al. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018;9: 611-629.
46. Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture. 2019;163: 104859.
47. Fernandez-Beltran R, et al. Rice-yield prediction with multi-temporal sentinel-2 data and 3D CNN: A case study in Nepal. Remote Sensing. 2021;13(7):1391.
48. Schwalbert RA, et al, Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. Agricultural and Forest Meteorology. 2020;284:107886.
49. Brownlee J. Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras. Machine Learning Mastery; 2016.
50. Brownlee J. Deep learning for computer vision: image classification, object detection, and face recognition in python. 2019: Machine Learning Mastery.
51. Kang Y, et al. Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest. Environmental Research Letters. 2020; 15(6):064005.
52. Ji S, et al. 3D convolutional neural networks for human action recognition. IEEE transactions on pattern analysis and machine intelligence. 2012;35(1):221-231.
53. Chu Z, Yu J. An end-to-end model for rice yield prediction using deep learning fusion. Computers and Electronics in Agriculture. 2020;174:105471.
54. Sun J, et al. County-level soybean yield prediction using deep CNN-LSTM model. Sensors. 2019;19(20):4363.
55. Wang Y, et al. Combining multi-source data and machine learning approaches to predict winter wheat yield in the conterminous United States. Remote Sensing. 2020;12(8):1232.
56. Ruder S. An overview of multi-task learning in deep neural networks; 2017. arXiv preprint arXiv:1706.05098
57. Zhang Y, Yang Q. A survey on multi-task learning. IEEE Transactions on Knowledge and Data Engineering. 2021;34(12):5586-5609.
58. Elavarasan D, and Vincent PD. Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. IEEE access. 2020;8:86886-86901.
59. Mnih V, et al. Human-level control through deep reinforcement learning. nature, 2015;518(7540):529-533.
60. Bali N, Singla A. Deep learning based wheat crop yield prediction model in punjab region of north india. Applied Artificial Intelligence. 2021l;35(15):1304-1328.
61. Paudel D, et al. Machine learning for large-scale crop yield forecasting. Agricultural Systems. 2021;187:103016.
62. Sun J, et al. Multilevel deep learning network for county-level corn yield estimation in the us corn belt. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020;13:5048-5060.
63. Khaki S, Wang L. Crop yield prediction using deep neural networks. Frontiers in plant science. 2019;10:621.
64. Chen Y, et al. Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages. Remote Sensing. 2019;11(13):1584.
65. Tian H, et al. An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China. Agricultural and Forest Meteorology. 2021;310:108629.

66. Tian H, et al. A deep learning framework under attention mechanism for wheat yield estimation using remotely sensed indices in the Guanzhong Plain, PR China. International Journal of Applied Earth Observation and Geoinformation. 2021; 102:102375.
67. KIRKAYA A. Smart farming-precision agriculture technologies and practices. Journal of Scientific Perspectives. 2020; 4(2):123-136.
68. Kobayashi N, et al. Crop classification using spectral indices derived from Sentinel-2A imagery. Journal of Information and Telecommunication. 2020; 4(1):67-90.
69. Liang S, Li X, Wang J. Compositing, smoothing and gap-filling techniques. Advanced remote sensing. 2012;75-90.
70. Vallentin C, et al. Suitability of satellite remote sensing data for yield estimation in northeast Germany. Precision Agriculture. 2022;23(1):52-82.
71. Prey L, Hu Y, Schmidhalter U. High-throughput field phenotyping traits of grain yield formation and nitrogen use efficiency: optimizing the selection of vegetation indices and growth stages. Frontiers in Plant Science. 2020;10:1672.
72. Wang X, et al. Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. Remote Sensing. 2020;12(11):1744.
73. de Freitas Cunha RL, Silva B. Estimating crop yields with remote sensing and deep learning. In 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS). IEEE; 2020.
74. Gavah K, Abbaszadeh P, Moradkhani H. DeepYield: A combined convolutional neural network with long short-term memory for crop yield forecasting. Expert Systems with Applications. 2021;184: 115511.
75. Qiao M, et al. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. International Journal of Applied Earth Observation and Geoinformation. 2021; 102:102436.
76. Basso B, Liu L. Seasonal crop yield forecast: Methods, applications, and accuracies. Advances in Agronomy; 2019.
77. Elavarasan D, et al. Forecasting yield by integrating agrarian factors and machine learning models: A survey. Computers and Electronics in Agriculture. 2018;155:257-282.
78. Abbaszadeh P, et al. Bayesian multi-modeling of deep neural nets for probabilistic crop yield prediction. Agricultural and Forest Meteorology. 2022; 314:108773.
79. Cao J, et al. Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine. European Journal of Agronomy. 2021;123:126204.
80. Wolanin A, et al. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. Environmental Research Letters. 2020; 15(2):024019.
81. Sharma S, Rai S, Krishnan NC. Wheat crop yield prediction using deep LSTM model; 2020.
arXiv preprint arXiv:2011.01498
82. Ghazaryan G, et al. Crop yield estimation using multi-source satellite image series and deep learning. in IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium; 2020.
83. Ruß G, et al. Data mining with neural networks for wheat yield prediction.. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008.
84. Mostaza Colado D, Mauri P, Capuano A. Assessing the Yield of a Multi-varieties Crop of *Camelina sativa* (L.) Crantz through NDVI Remote Sensing. 2019;596-602.
85. Patel NR, et al. Remote sensing of regional yield assessment of wheat in Haryana, India. International Journal of Remote Sensing. 2006;27(19):4071-4090.
86. Paidipati KK, et al. Prediction of Rice Cultivation in India—Support Vector Regression Approach with Various Kernels for Non-Linear Patterns. AgriEngineering. 2021;3(2):182-198.
87. Chu L, et al. Spatial Heterogeneity of winter wheat yield and its determinants in the yellow river delta, China. Sustainability. 2020;12(1):135.
88. Azadbakht M, et al. Alfalfa yield estimation based on time series of Landsat 8 and PROBA-V images: An investigation of machine learning techniques and spectral-temporal features. Remote Sensing Applications: Society and Environment. 2022;25:100657.

89. Ahmed MS, et al. Yield Response of different rice ecotypes to meteorological, agro-chemical, and Soil Physiographic Factors for Interpretable. Precision Agriculture Using Extreme Gradient Boosting and Support Vector Regression. Complex. 2022;2022:20.
90. Jeong S, Ko J, Yeom JM. Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea. Science of The Total Environment. 2022; 802:149726.
91. Ahmad I, et al. Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab Pakistan. Journal of the Indian Society of Remote Sensing. 2018;46: 1701-1711.
92. Gandhi N, et al. Rice crop yield prediction in India using support vector machines. in 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE; 2016.
93. Matsumura K, et al. Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. The Journal of Agricultural Science. 2015; 153(3):399-410.
94. Ju S, Lim H, Heo J. Machine learning approaches for crop yield prediction with MODIS and weather data. in 40th Asian Conference on Remote Sensing: Progress of Remote Sensing Technology for Smart Future, ACRS 2019; 2020.
95. Buschjäger S, Morik K. Decision tree and random forest implementations for fast filtering of sensor data. IEEE Transactions On Circuits And Systems I: Regular Papers. 2018;65(1):209-222.
96. Ranjan AK, Parida BR. Paddy acreage mapping and yield prediction using sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India). Spatial Information Research. 2019;27(4): 399-410.
97. Charoen-Ung P, Mittrapiyanuruk P. Sugarcane yield grade prediction using random forest with forward feature selection and hyper-parameter tuning. in Recent Advances in Information and Communication Technology 2018: Proceedings of the 14th International Conference on Computing and Information Technology (IC2IT 2018) Springer; 2019.
98. Ku SB, Edwards GE, Tanner CB. Effects of light, carbon dioxide, and temperature on photosynthesis, oxygen inhibition of photosynthesis, and transpiration in Solanum tuberosum. Plant physiology. 1977;59(5):868-872.
99. Zemba A, et al. Growth and yield response of irish potato (Solanum Tuberosum) to climate in Jos-South, Plateau State, Nigeria; 2013.
100. Ali AM, et al. Crop yield prediction using multi sensors remote sensing (Review Article). The Egyptian Journal of Remote Sensing and Space Science. 2022;25(3): 711-716.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/98594