International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Sugarcane Yield Classification and Prediction Using Light weight Deep Network

Mrs. M. Deepanayaki[1*], Dr. Vidyaathulasiraman[2]

Submitted: 17/09/2023 **Revised**: 18/11/2023 **Accepted**: 29/11/2023

Abstract: Sugarcane is the most important renewable commercial crops in India. The sugarcane farming and sugar industry are essential to the socio-economic development of rural communities by generating greater income and employment opportunities. The ability of decision-makers and planners to choose import or export strategies is based on the early detection and control of issues related to sugarcane yield indicators. In this manuscript, Sugarcane Yield Classification and Prediction Using Light weight Deep Network (SY-CP-LWDN) proposed. At first, the data are gathered via field and mill. Afterward, the data are fed to pre-processing using Triangle Filter (TF). Here the noise present in the data's are reduced. Then these data's are undergoes segmentation for segmenting canopy, leaf size and color. The segmented data's are given for Adaptive Fuzzy Segmentation Algorithm (DAFSA) for segmenting canopy, leaf size and color of Sugarcane. Then the segmented data's are fed to the Feature extraction using Swin Transformer (ST) for extracting the features such as Standardizing and Imputation. Finally classification is done using light weight deep network (LWDN). The classification results are grade1, grade2 and grade 3 of sugarcane yield. Simulation of the model done using python and performance metrics also examined. The performance metrics like accuracy, ROC used to analyze performance of proposed technique. The Performance of the proposed SY-CP-LWDN approach attains 24.11%, 27.12% and 32.73% high accuracy compared with existing methods such as Sugarcane yield prediction with data mining and crop simulation models (SYP-DM-CSM), Integration of RNN with GARCH refined by whale optimization algorithm for yield forecasting: a hybrid machine learning approach (IRNN-GARCH-WOA-HMA) and An improved multilayer perceptron approach for detecting sugarcane yield production in IoT based smart agriculture(AIMP-DSYR-IoTSA), respectively.

Keywords: *Adaptive Fuzzy Segmentation Algorithm, Cat Swarm Optimization Algorithm, Light Weight Deep Network, Sugarcane yield prediction, Triangle Filter.*

1. Introduction

Predicting yield is a difficult task in every agriculture crops. It is important to build a hybrid machine-learning technique using the available data for yield predictions. [1]. The major goal of this study is to create unique hybrid technique for predicting sugarcane yield using non linear time series data. Recurrent neural networks frequently have large memories, allowing for adequate forecasts with fewer parameters [2]. To increase the effectiveness of neural network produce correct outcomes, the whale optimization method optimizes the weights and thresholds of the recurrent neural network [3]. It is difficult to improve performance when predicting the volatility of time series. An estimated 12 billion shillings are generated by sugarcane production each year, providing employment for around 6 million people [4]. As a result, yield rarely has access to enough opportunity to profit from sugarcane sales. This might have contributed to their lack of desire and cause of division's falling sugarcane growing and less yield for more than a decade. The estimation of crop output issues is difficult across different Indian states since sugarcane is grown by sowing in January–February, July–August, or October–November, with a maturation period of 12–18 months. [5]. Today's society has developed a range of terms to elicit and retain motivation such as usage of loyalty cards, bonuses to endure engagement.

The key contributions of this paper include;

- In this manuscript, An Optimization Light weight Deep Network to implement for the Sugarcane Yield classification (SY-CP-LWDN) is proposed.
- Initially sugarcane yield data is taken from farmer plot by collecting the data at the mill.
- Then the collected data's are preprocessed using Triangle Filter for reducing the noise in the data. After this preprocessed result is given for segmentation.
- Segmentation is done using Adaptive Fuzzy Segmentation Algorithm (DAFSA) for segmenting canopy, leaf size and color of Sugarcane. And these features are extracted using Swin Transformer (ST)for Standardizing and Imputation features.
- Finally by using light weight deep network(LWDN) the classification process is done as grade1, grade2 and grade 3 of sugarcane yield.
- Where LWDN do not show any use of optimization methods for getting best accuracy in sugarcane yield.

[1]Research Scholar, Department of Computer Science, Periyar University, Salem, India.*
[2]Associate Professor and Head, Department of Computer Science, Government Arts & Science College (W), Barugur, India.
[1] Corresponding Author Email: sureshdeepanayaki@gmail.com*

So, Sand Cat Swarm Optimization Algorithm (SCSOA) used to enhance LWDN weight parameter.

- The proposed SY-CP-LWDN method is implemented in python and performance of suggested technique is compared with existing methods such as SYP-DM-CSM [11], (IRNN-GARCH-WOA-HMA [12] and, (AIMP-DSYR-IoTSA) [13], respectively.

The remaining manuscripts are organized as follows: part 2 analyzes Literature survey, part 3 defines proposed approach; part 4 illustrates results; and part 5 presents conclusion.

2. Literature Survey

Numerous research studies presented in the literature to categorize Sugarcane Yield Analysis and Prediction; a few current works are described below.

Hammer et al., [6], have presented Sugarcane yield estimation through data mining and crop simulation techniques. The purpose of this work was to build mathematical models for predicting sugarcane production using data mining (DM) techniques as well as to identify and ordinate the primary variables that condition sugarcane output, according to their relative importance. The findings made it possible to draw the conclusion that, among all the variables examined, the number of cuts was the aspect that all DM approaches took into account the most. It provides high accuracy with lower ROC.

Murali et al., [7] have presented a hybrid machine learning strategy for yield forecasting that integrates RNN with GARCH optimized by whale optimization technique. The primary objective of the presented study was to develop brand-new hybrid model for forecasting sugarcane yield utilizing non-linear time series data. By combining statistical technique such generalized autoregressive conditional recurrent neural network augmented by whale optimization approach enabled the novel capacity to anticipate the yield with volatility in time series analysis. The presented method offers high ROC with lower accuracy.

Wang et al., [8] have presented, An improved multiple layer perceptron method for identifying sugarcane production in Internet of things basis smart agriculture. The Internet of Things was the potent technologies; provide managers and manufacturing facilities with precision management and intelligent navigation. Utilizing intelligent systems such as machine learning presented as an alternate remedy in such circumstances. In order to forecast how much sugar was produced as yield in Internet of Things agriculture, the presented paper paper developed an enhanced Multilayer Perceptron technique. It provides high ROC with lower accuracy.

3. Proposed Methodology

In this research An Optimization Light weight Deep Network to implement for the Sugarcane Yield classification (SY-CP-LWDN) is proposed. Data collection, Pre-processing, segmentation and classification are done in this method. The block diagram of the proposed SY-CP-LWDN methodology is presented in Figure 1.

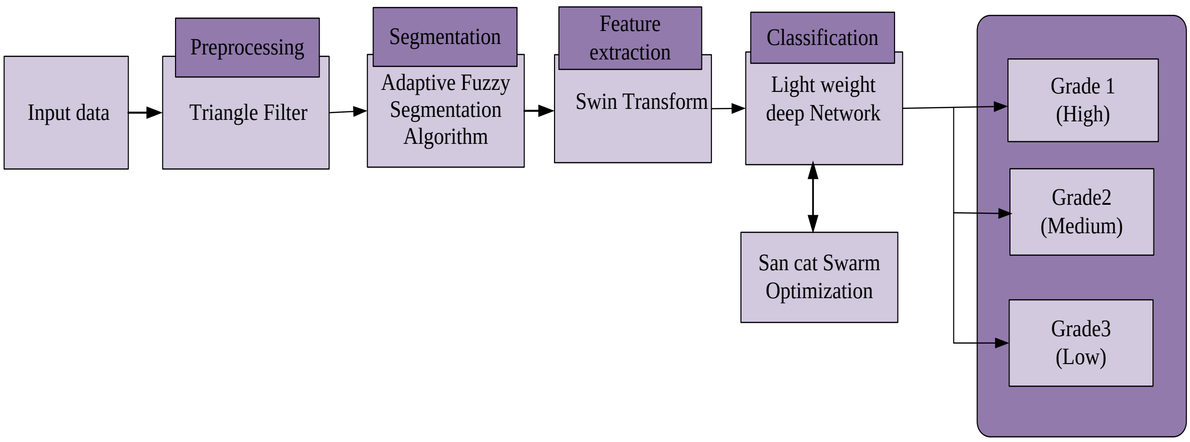

Fig 1: Block diagram for the proposed SY-CP-LWDN method

3.1 Data Acquisition

In this step, sugarcane yield data is taken from farmer plot and collecting the data at the mill to classify the grade. There are 12,521 records, out of this 80% data are taken for testing, remaining 20% for training. Table 1 tabulates the list of features in the dataset.

Table 1: List of featuresat the dataset

(Cat. = Category, Cont. = Continual data type)

S.No	Name	Category	S.No	Name	Type	S.No	Name	Type
1	Class_cane	Cat.	5	Fertilizer	Cat.	9	Farmer Contract Grad	Cat.
2	Type_Cane	Cat.	6	Type Soil	Cat.	10	Area_ Remain	Cont.
3	Water Type	Cat.	7	Groove Wide	Cat.	11	Rain_Avg	Cont.
4	Action Water	Cat.	8	Yield Old Grade	Cat.	12	Contracts Area	Cont.

3.2 Pre-processing using Triangle Filter (TF)

In this step, Triangle Filter (TF) [9] performs data pre-processing which is used for reducing noise in the data. When dealing with irregular data, Triangle Filter employs interpolation to approximate the original continuous function. It shows a reconstructed data that was obtained using linear interpolation of the filtered data. The complete dataset is analyzed for anomaly detection during the preprocessing step, which eliminates the need for data preprocessing. Because header accounts for such a small proportion of total network data, they require fewer resources to process the whole packet payloads. Filtering carried out with the continuous convolution it is expressed in equation (1),

$$Z^Y = \int_{\lambda_x} Z_w^V(r) U(y(\hat{w}), r) dr \tag{1}$$

where U signifies normalized continuous function, Z_w^V signifies reconstructed data, λ_x denotes linear interpolation. The input data contains label noise and interference. The interpretation of Triangle Filter is appealing: it is a linear diffusion process that operates on data; it is expressed in equation (2),

$$U(y(\hat{w}), r) = \gamma \{ |y(\hat{w}) - r| \le t \} / 2t \tag{2}$$

where t signifies radius, while substituting above two equations, There are considerable odds that the forecast will not be recognized if pre-processing stages are not carried out correctly. There is no link between the incoming data and the noise, Z^Y found by equation (3),

$$Z^Y = \int_{y(\hat{w})-t}^{y(\hat{w})+t} Z_w^V(r) dr \tag{3}$$

where Z_w^V is the linearly interpolated data, Equation (3) can be used to accomplish the data transfer procedure, where Z^Y is the Triangle Filtering input, it is an input data with large-scale structures. Triangle Filter as previously indicated capture small scale information in large-scale buildings. V is iterated Y times. By using this Triangle Filter the noise is removed from the data. Finally, pre-processed data fed to segmentation phase.

3.3 Segmentation using Deep Adaptive Fuzzy Segmentation Algorithm (DAFSA)

After complete the preprocessing the collected data’s are given to Deep Adaptive Fuzzy Segmentation Algorithm (DAFSA) [10] for segmentation. The segmentation results as canopy, leaf size and color of Sugarcane. DAFSA segmented the parameter of mapping loss function of reconstruction formulated as equation (4),

$$A_{\mathrm{Re}\,c} = \min_{\omega,\theta} \sum_{I=1}^{N} \frac{1}{N} \ell \left[y\theta(h\omega(v_i)), u_i \right] \tag{4}$$

where $A_{\mathrm{Re}\,c}$ denoted as loss function, $(h\omega(v_i))$ denoted as frequency of particulate component, u_i is denoted as angular moment, N denoted as varience , y and ℓ denoted as squared element in loss function. The similarity function is represented using equation (5)

$$y_\mu(u_i, v_i) = D_{i,j}(u_i) \cdot v_i \tag{5}$$

where $y_\mu(u_i, v_i)$ denoted as similarity function, u_i is denoted as angular moment, $D_{i,j}$ is denoted as cluster map function and v_i is denoted as objective function. The goal of the fuzzy clustering model is to collect related or identical patterns from a large number of diverse data points and it is segment the canopy, leaf size and color by utilizing Deep Adaptive Fuzzy Segmentation Algorithm. Then the segmented results are given for feature extraction.

3.4 Feature Extraction by Swin transformer

Following the segmentation, a method called feature extraction used to extract features of Standardizing, imputation. After segmentation, various types of features are extracted from the segmented output utilizing Dual-Domain Feature Extraction (DDFE) [11]. It helps to extract the Standardizing and Imputation features. Substantial features are present below preprocessing illuminated through support of Swin transforms. Preprocessing output has significant appearances are extracted with the help of Swin transform. When compared to other vision models, the Swin transformer, a type of Vision transformer, achieves a better speed-accuracy trade-off. Hierarchical feature maps are produced by fusing image patches in deeper levels. Its linear computing complexity for input image size results from the fact that it computes self-attention only within each local window. The whole process of the STL is given by the equation (6) and (7),

$$Y' = R_{(P)V-QPL}(R_{GF}(Y)) + Y \quad (6)$$

$$Y'' = R_{QNS}(R_{GF}(Y')) + Y' \quad (7)$$

where Y' and Y" denotes input, output of Swin transformer layer. R_{QNS} And R_{GF} denote multilayer perceptron layer. $R_{(P)V-QPL}$ Are utilized consecutive STL. Spatial restraints added in Swin transformer layer equated original transformers. Thus it is given by the equation (8)

$$\Psi(R_{QPL}) = 4RVB^2 + 2(RV)^2 B \quad (8)$$

where Ψ denotes multiple head self-attenuation and R_{QPL} signifies learnable relative position, RVB^2, signifies input and output of the variations. Here RV denotes Multilayer perceptron and B is the computational complexity. Thus given by the equation (9)

$$S = YM_S, T = YM_T, W = YM_W \quad (9)$$

where they M_S, M_T, M_W signifies shared projection matrices over every windows. Query S key T, value W, learnable, relation location encoding used to calculate self-attenuation mechanism in local window. The process of categorization uses adaptive improved permutation entropy. In the selection below, the classification method is described in depth.Then these extracted features are given into LWDNto effectively classify the skin cancer diseases. Finally the features extraction is supplied to the classification.

3.5 Classification using Light Weight Deep Network

In this section, sugarcane yield classification using Light Weight Deep Network(LWDN) [12] is discussed. A popular technique for sugarcane yield classification is the LWDN method, which may be enhanced to become a two classification recognition method. The label information from the authentic and fake samples is integrated into a single vector that feeds the LWDN discriminator. By changing the label on the generator with conditional information, LWDN is able to adjust the output type. Two portions make up the overall loss of the LWDN. By using this Light Weight Deep Network the classified results is arrived by grade1 (High), grade2 (Medium) and grade 3 (Low) of sugarcane yield. For finding this classification results equations are needed. First for finding loss function it is expressed using equation (10)

$$\gamma = \beta[win_c] \rightarrow_{lag} \beta[win_d] \quad (10)$$

where α, β signifies represents the frequent with respective windows, $[win_c]$ denoted as minimized loss function , win_d denoted as network hyper parameter.Then user sets of the window width since the environment determines the length of time that will be an interesting time measure. Association rules are still extracted from sequence data, but in order to make predictions based on the currently known rules, a decision tree is constructed over the association rules obtained. The confidence measure employed involves more variables than just how confident one is given to the equation (11),

$$\gamma_i = (L \rightarrow M) = \frac{|c_i \in C_i | L \cup M \subseteq c_i|}{|C_i|} \quad (11)$$

where C_i being the set of transactions for time period of L, M the use of clustering to find segments that exhibit same trend over temporal sequence is a complement to trend prediction with classification algorithms.

The association rule algorithm's adaption makes use of the following definition of confidence and it is given to the equation (12),

$$\gamma_i = (L \rightarrow M) = \frac{H(L, M, P)}{H(L)} \quad (12)$$

where F denotes frequency of networks, P denoted as learning rate of the network, H denoted as initial parameter, L denotes vector parameter in division factor , M denoted as parameters in batch normalization. The label information from the authentic and fake samples is integrated into single vector that feeds Light Weight Deep Network (LWDN)discriminator. By changing the label on the generator with conditional information, LWDN is able to adjust the output type. Two portions make up the overall loss of the LWDN. By using this Light Weight Deep Network the classified results is arrived by grade1 (High), grade2 (Medium) and grade 3 (Low) of sugarcane yield. But Light Weight Deep Network does not reveal optimization techniques adoption for computing ideal parameters and for assuring precise Sugarcane Yield. So these results are given for SCSOA to optimize weight parameter η of LWDN.

3.5.1 Sand Cat Swarm Optimization Algorithm (SCSOA) for optimizing weight parameters of the Light Weight Deep Network

Here Sand Cat Swarm Optimization Algorithm (SCSOA) [13] used to enhance weight parameter η of Light Weight Deep Network (LWDN) for getting better accuracy.The sand cat (Felis margarita) belongs to the Felis family of mammals and is known for its ability to thrive in harsh desert environments, including central Asian Sahara, African Sahara, Arabian Peninsula. Despite its small, nimble, unassuming appearance, the sand cat exhibits distinct behaviors in both hunting and living.Unlike domestic cats, the sand cat does not adopt a group-living lifestyle. Its unique adaptations include the presence of greater density of sandy to light grey fur on the palms and soles of its feet, which serve as insulation against the extreme hot and cold conditions of the desert. Additionally, fur characteristics of sand cat make difficult to detect and track.Here, step by step procedure is defined to get ideal value of LWDN based on SCSOA. Initially, SCSOAmakes the equally distributing populace to optimize the optimum parameter η of LWDN.

Step 1: Initialization

The defined bounds are used to initialize the search space at random. Each current search agent's position is updated during searching phase depend on a random position. The population of a swarms is divided as c big males, one female, and d tiny males are expressed in equation (13),

$$c = [(b-1)a] \tag{13}$$

where c signifies number of big male swarms, d signifies small male swarms. However, a percentage b is too tiny or too large result in zero for c or d. As a result, a straightforward strategy is utilized to entice both large and tiny males to interact with at least two swarms.

Step 2: Random generation

Input parameters produced at random after initialization. Best fitness value selection is depending onobvious hyper parameter condition.

Step 3: Fitness Function Estimation

The initialized evaluations are used to generate a random solution. Fitness function is assessed with parameter optimization value for enhancing weight parameter η of classifier. This is given in eqn (14),

$$fitness\ \ function = optimizing\ \ (\eta) \tag{14}$$

Step 4: Exploration phase

A female, as a middle-class individual, is built to undertake exploration at the same time. It is carried out with a preset probability of 0.5. In exploration, the female broadens the feasible region surround global ideal solution by breeding with winner big male and giving birth 2 offspring. Parthenogenesis, on the other hand, broadens the search field by diversifying the solution. The best-generated offspring eventually update the female. Furthermore, the mating (of two swarms is described as equation (15) and (16),

$$E'_{fg} = h_g . E_{fg} + (1 - h_g) - z \tag{15}$$

$$E'_{kg} = h_g . E_{kg} + (1 - h_g) - E_{fg} \tag{16}$$

where E_{fg} implies g^{th} dimension of f^{th}, E_{kg} implies g^{th} dimension of k^{th}, E'_{fg} and E'_{kg} implies E^{th} dimensions of two young ones.

Step 5: Exploitation phase for optimizing η

Tiny swarms are expected to hunt for solutions in a broad area while also engaging in some exploitation as low-quality persons. By randomly selecting a portion of the large men's dimensions with a predetermined probability rate, each tiny guy backs away from them. The movement of the u^{th} swarms to track the b^{th} one is stated with a Milpir rate l in interval (0, 1) given in equation (17)

$$m_{ub} = \begin{cases} \sum_{g=1}^{n} h_1 (E_{kg} - E_{ug}), & if\ h_{2<l} \\ 0 & otherwise \end{cases} \tag{17}$$

where h_1, h_2 denotes random numbers, n denotes dimension, g implies randomly chosen dimension, c denotes number of big males. Exploration is certain to be more likely than exploitation in this sand cat movement.

Step 6: Termination

The weight parameter values of generator η from Light Weight Deep Network are optimized with the help of SCSOA, will iteratively repeat step 3 until fulfill halting criteria $KM = KM + 1$. Then SY-CP-LWDNis improved with higher accuracy and higher ROC with no error.

4. Result and Discussion

In this Research work, An Optimization Light weight Deep Network to implement for the Sugarcane Yield classification (SY-CP-LWDN) is proposed. The simulations are done in Python and the models are run on a PC with a 2.50 GHz central processor unit, an Intel Core i5, 8GB RAM, Windows 7. The performance metrics such as accuracy, ROC are evaluated. The proposed SY-CP-LWDN compared with existing SYP-DM-CSM, IRNN-GARCH-WOA-HMA and AIMP-DSYR-IoTSA methods.

4.1 Performance measures

Performance metrics like accuracy, and ROC are examined to measure performance. confusion matrix is used to measure performance parameters. The True Negative, True Positive, False Negative, False Positive numbers necessary to measure confusion matrix.

- True Positive (TN): Non-Defective class properly categorized into Defective class.
- True Negative (TN): Non-Defective class properly categorized into Non-Defective class.
- False Positive (FP): Non-Defective class in exactly categorized into Defective class.
- False Negative (FN): Defective class inexactly categorized into Non-Defective class.

4.1.1 Accuracy

Accuracy measures the proportion of samples (positives and negatives) besides total samples and it is calculated by equation (18),

$$accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{18}$$

4.1.2 ROC

The ratio of false negative to true positive area, it is calculated by equation (19),

$$ROC = 0.5 \times \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \tag{19}$$

4.2 Performance Analysis

Figure 2 to 3 depicts the simulation outcomes of proposed SY-CP-LWDN method. Then, the suggested SY-CP-LWDN method compared with existing SYP-DM-CSM, IRNN-GARCH-WOA-HMA and AIMP-DSYR-IoTSA methods.

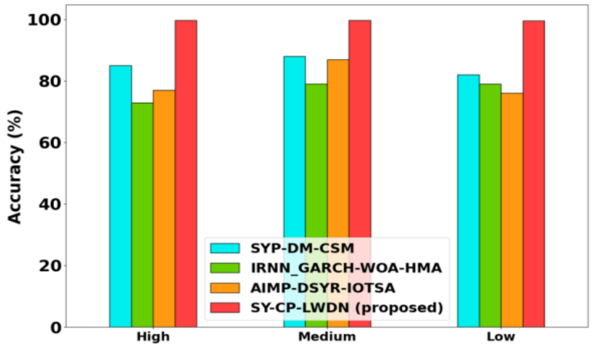

Fig 2: Accuracy Analysis

Figure 2 depicts Accuracy analysis. The proposed WSN-AD-ADRGN method attains 40.56%, 20.76%, and 20.67% greater Accuracy in high (grade 1) , 35.50%, 23.65%, and 20.64% greater accuracy in medium(grade 2), 45.65%, 32.54% and 22.76% greater accuracy in low (grade 3), when evaluated to existing SYP-DM-CSM, IRNN-GARCH-WOA-HMA and AIMP-DSYR-IoTSA methods.

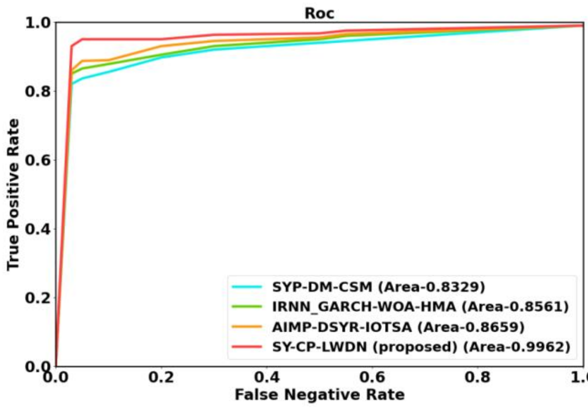

Fig 3: ROC Analysis

Figure 3 depicts ROC Analysis. The proposed WSN-AD-ADRGN techniqueachieves56.56%, 57.76%, and 65.67% greater ROC when evaluated to the existing methods such as SYP-DM-CSM, IRNN-GARCH-WOA-HMA and AIMP-DSYR-IoTSA methods.

4.3 Discussion

Sugarcane yield classification and prediction are essential aspects of modern agriculture, and the application of lightweight deep learning networks can significantly enhance our ability to accurately forecast sugarcane yields. Accurate predictions also assist the sugar and bioenergy industries in managing their supply chains and optimizing production processes. Light Weight Deep Network have demonstrated capacity to handle complex, high-dimensional data. In the context of sugarcane yield prediction, these networks analyze a wide range of input data sources, including satellite imagery, weather data, soil information, and historical yield data. These networks often involve techniques like model pruning, quantization, architecture design to reduce their size and computational requirements while maintaining performance. To predict sugarcane yield accurately, data from various sources must be integrated. Remote sensing data from satellites and drones provide real-time information on crop health, while weather stations offer crucial climate data. Soil sensors and historical yield records add valuable context. Preprocessing and feature engineering play a vital role in preparing data for deep learning models. Feature extraction techniques help identify relevant patterns and relationships within the data. By leveraging a wide range of data sources and optimizing model architectures, these models can help increase yield and resource efficiency while contributing to the sustainability of sugarcane agriculture.

5. Conclusion

This paper proposes An Optimization Light weight Deep Network to implement for the Sugarcane Yield classification (SY-CP-LWDN) for classifying grade1, grade2 and grade 3 of sugarcane yield. The proposed SY-CP-LWDN method is activated in python with the help of Sand Cat Swarm Optimization Algorithm. The proposed technique attains attains 56.56%, 57.76%, and 65.67% higher ROC when evaluated to the existing methods such as SYP-DM-CSM, IRNN-GARCH-WOA-HMA and AIMP-DSYR-IoTSA methods respectively.

Reference

[1] N. A. Husaini, R. Ghazali, N. Arbaiy, N. A. Hamid, and L. H. Ismail, 2020. A modified weight optimization for artificial higher order neural networks in physical time series. *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3.

[2] A. Rajagopal, S. Jha, M. Khari, S. Ahmad, B. Alouffi, and A. Alharbi 2021. A novel approach in prediction of crop production using recurrent cuckoo search optimization neural networks. *Applied Sciences*, vol. 11, no. 21, p.9816.

[3] C. E. Tujah, R. A. Ali, and N. N. L. N. Ibrahim, 2023. Optimization of Sugarcane Bagasse Conversion Technologies Using Process Network Synthesis Coupled with Machine Learning. *Pertanika Journal of Science & Technology*, vol. 31, no. 4.

[4] T. H. Thai, R. A. Omari, D. Barkusky, and S. D. Bellingrath-Kimura, 2020. Statistical analysis versus the m5p machine learning algorithm to analyze the yield of winter wheat in a long-term fertilizer experiment. *Agronomy*, vol. 10, no. 11, p.1779.

[5] L. Zhu, X. Liu, Z. Wang, and L. Tian, 2023. High-precision sugarcane yield prediction by integrating 10-m Sentinel-1 VOD and Sentinel-2 GRVI indexes. *European Journal of Agronomy*, vol. 149, p.126889.

[6] R. G. Hammer, P. C. Sentelhas, and J. C. Mariano, 2020. Sugarcane yield prediction through data mining and crop simulation models. *Sugar Tech*, vol. 22, no. 2, pp.216-225.

[7] P. Murali, R. Revathy, S. Balamurali, and A. S. Tayade, 2020. Integration of RNN with GARCH refined by whale optimization algorithm for yield forecasting: a hybrid machine learning approach. *Journal of Ambient Intelligence and Humanized Computing*, pp.1-13.

[8] P. Wang, B. A. Hafshejani, and D. Wang, 2021. An improved multilayer perceptron approach for detecting sugarcane yield production in IoT based smart agriculture. *Microprocessors and microsystems*, vol. 82, p.103822.

[9] Z. Li, J. Wang, J. Huang, and M. Ding, 2023. Development and research of triangle-filter convolution neural network for fuel reloading optimization of block-type HTGRs. *Applied Soft Computing*, vol. 136, p.110126.

[10] https://www.kaggle.com/abhinand05/crop-production-in-india

[11] J. Zhang, W. Wang, C. Lu, J. Wang, and A. K. Sangaiah, 2020. Lightweight deep network for traffic sign classification. *Annals of Telecommunications*, vol. 75, pp.369-379.

[12] A. Seyyedabbasi, and F. Kiani, 2023. Sand Cat swarm optimization: A nature-inspired algorithm to solve global optimization problems. *Engineering with Computers*, vol. 39, no. 4, pp.2627-2651.

[13] J. Huang, Y. Fang, Y. Wu, H. Wu, Z. Gao, Y. Li, J. Del Ser, J. Xia, and G. Yang 2022. Swin transformer for fast MRI. *Neurocomputing*, vol. 493, pp.281-304.