Grey Wolf Optimization-Based Hyperparameter Tuning for Madura Tobacco Leaf Sorting

Rosida Vivin Nahari[1, a)], Adi K. Saputro[1,b)], Riza Alfita[1,c)], Achmad F. Ibadillah[1,d)]

Author Affiliations
[1]*Faculty of Engineering, Universitas Trunojoyo, Bangkalan, Indonesia*
[a)]Corresponding author: rosida.nahari@trunojoyo.ac.id
[b)]adi.kurniawan@trunojoyo.ac.id
[c)]riza.alfita@trunojoyo.ac.id
[d)]fiqhi.ibadillah@trunojoyo.ac.id

Abstract. Tobacco leaf sorting is an important part of the Madura tobacco industry. This process selects quality tobacco leaves by considering several parameters, including leaf shape and color. The use of a machine learning-based model can aid the tobacco leaf sorting process through tobacco leaf image processing. Tobacco leaf images are sorted into two categories: low-quality leaves and high-quality leaves. A common challenge in machine learning is the selection of the appropriate learning rate value. A small learning rate requires more iterations to converge and can potentially yield high accuracy. Conversely, a high learning rate can accelerate training but may miss the optimal solution. This study implements hyperparameter tuning, specifically focusing on adjusting the learning rate using Grey Wolf Optimization (GWO), to optimize the computation time and accuracy of the tobacco leaf sorting machine learning model. Computational testing of tobacco leaf sorting using GWO took 61.46 seconds. In comparison, Grid Search and Random Search took 106.72 seconds and 87.28 seconds, respectively. Additionally, GWO achieved a higher accuracy of 82% compared to Grid Search (67%) and Random Search (67%).

INTRODUCTION

The tobacco industry has long been an important sector in the economies of many countries, including Indonesia [1]. One of the tobacco industry's biggest contributions is as a source of state income through excise revenues [2]. Excise tax on tobacco products is a fairly large source of state income. Madurese tobacco has long been known as one of the leading commodities in Indonesia, especially on Madura Island. Its unique quality and high market demand give Madura tobacco enormous potential for development. [3,4].

The quality of Madurese tobacco is a crucial factor in determining the selling value and competitiveness of local tobacco products in the international market [5]. The selling price of tobacco is greatly influenced by quality so the right tobacco leaf sorting system will have an impact on the selling value. Sorting tobacco leaves is an important step in post-harvest to obtain good quality final products. This process aims to separate tobacco leaves based on their quality (separating good quality leaves from damaged or defective ones). The process of sorting tobacco leaves can help and facilitate the drying, fermentation, and subsequent processing of tobacco [6,7].

Unclear quality standards and lack of strict supervision can result in the circulation of low-quality tobacco on the market. The traditional sorting method that has been carried out uses the visual skills of experienced students in sorting dried tobacco leaves [5,8]. It is hoped that the development of deep learning-based information technology can help farmers sort dry tobacco leaves without the help of a grader [9,10].

Several studies have shown that deep learning-based computer vision technology can be directly applied in the agricultural sector to increase crop production. The deep learning model, Convolutional Neural Network (CNN), has been widely applied as a very effective technological solution in processing visual data [11]. The physical characteristics of high-quality tobacco leaves can be seen from the golden yellow color that is evenly distributed on the tobacco leaves with the size, and texture of the leaves being thick and more prominent. This visual vision is difficult to implement if only rely on simple image processing methods due to differences in light when capturing tobacco leaf image data [12,13]. The CNN architectural model can automatically extract important features from data without requiring complex feature extraction engineering [14,15].

Classification research with developments in the changes to the CNN architecture, namely the number of layers used and the use of data augmentation techniques to increase the size of the data, succeeded in obtaining better accuracy compared to a simple machine learning-based model approach. in the past 5 years, the researchers has succeeded in building a CNN architecture with high-accuracy results. However, what needs to be emphasized is that the number of parameters and computing time required by the system is also very high. Hyperparameter tuning is a crucial stage in optimizing CNN performance. The techniques commonly used today only explore hyperparameter combinations and look for the best CNN parameter combination configuration [16-19]. Artificial intelligence-based optimization that imitates the behavior of a pack of gray wolves has provided many solutions for finding optimal values. The characteristics of the search model in Gray Wolf Optimizer are used as the basis for solutions to hyperparameter optimization problems in tobacco sorting systems. [20-23].

MATERIAL AND METHOD

This chapter describes the methodology that will be used in this research. Such as input data, method design, testing, and evaluation scenarios. The stages carried out in Figure 1 include literature study, method design and implementation, testing, performance measurement, and results analysis.

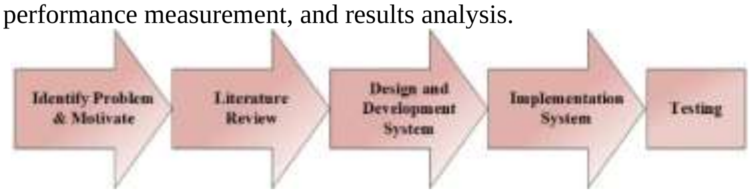

FIGURE 1. Display of high-quality tobacco leaf image dataset

1. Identify Problem & Motivate
   Hyperparameter tuning learning rate problem mapping for sorting tobacco leaves through a classification process using the CNN method optimized with the Gray Wolf Algorithm [24]. Determining the correct Learning Rate greatly influences the performance of the classification system. This is because if the Learning rate is too high: The model may "jump around" around the optimal solution and never converge well, resulting in overfitting. On the other hand, if the learning rate is too low: The model will be very slow to learn and may never reach the optimal solution, resulting in underfitting
2. Literature Review
   Literature Review is the initial stage of research to study the basic theoretical literature used to support this research. This includes initial data requirements, processing, and classification methods used
3. Design and Development System
   In the initial data collection process, the prototype was designed so that a webcam camera from above could capture images of dried tobacco leaves placed on an observation board. Then the installed camera automatically saves the image in still image form. All leaf image files that have been saved on the computer are then labeled into 2 categories, namely High Quality and Low Quality.
4. Implementation System
   Tobacco leaves with quality leaf characteristics seen based on leaf thickness, leaf texture, aroma, and leaf nicotine content. Leaf thickness and texture are extracted through texture features and shape features. Meanwhile, high leaf nicotine levels are inversely proportional to leaf chlorophyll levels. So leaves with low chlorophyll after harvest is characterized by golden yellowing evenly on the leaf surface, then the quality of the nicotine is the best. The type of tobacco used in this research is the prancak variety which comes from agricultural areas around the Kehi hills of Pamekasan. This research uses stand-holder equipment with a webcam camera to capture images of dry tobacco leaves. Meanwhile, roboflow is used for labeling and editing images used for datasets as well as preprocessing tobacco leaf images so that the dataset that is built has leaf image data that matches the input system specifications.

Dataset

The data collection process (taking pictures of tobacco leaves) was carried out using a room lighting scenario and adding lights directly above the object and camera so as to eliminate shadow mode. The results of data collection are then grouped and labeled according to class.

The size of the resulting image is 300x300 and preprocessing is carried out in the form of resizing the image to 224x224 to suit the size of the CNN mode input. The next preprocessing stage for tobacco images is the conversion from RGB to gray level and the noise filtering process. Image datasets that have passed the preprocessing stage are then separated into 2 groups, namely high-quality and low-quality. The dataset of tobacco leaves consists of 328 tobacco images.

FIGURE 2. Display of high-quality tobacco leaf image dataset

FIGURE 3. Display of low-quality tobacco leaf image dataset

The CNN model consists of several layers, namely the input layer, convolution layer, pooling layer, flatten layer, and fully connected layer. The neural network architecture built in this research is a Residual Network (ResNet-50) which has a shortcut connection concept that can prevent information from being lost during the training process. The CNN architectural design can be seen in Figure 4. Training is the process of the model learning patterns from the tobacco leaf dataset. At the training stage, there are two main processes, namely feed-forward and backpropagation. Feed-forward is a stage that processes input data through each existing layer until it becomes the desired output. Backpropagation is the opposite stage of feed-forward, namely the process of calculating and adjusting the weight and bias for each neuron based on errors in the output validation results. Because each output from each layer will be interconnected with the input of another layer, the results of the backpropagation process for each neuron in the output layer will represent the overall gradient that can be combined from all layers. Through these two processes, the model will adjust internal parameters to produce more accurate output.

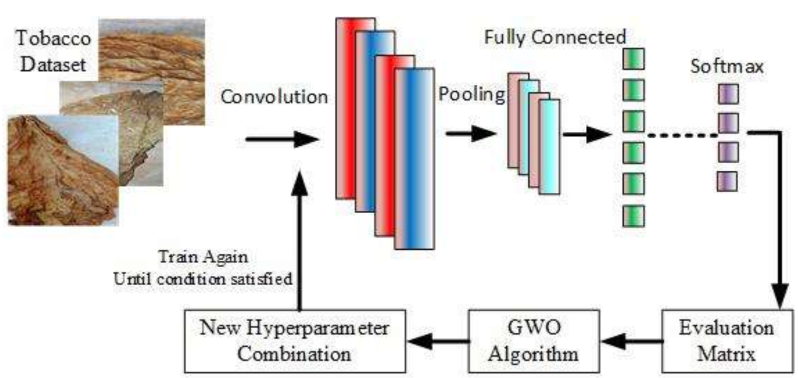

FIGURE 4. CNN training architecture

The training process and accuracy of predicting tobacco quality in neural network models are influenced by hyperparameter tuning. Hyperparameters are parameters used to control the training process of the model. The correct Hyperparameter Tuning configuration will also greatly influence the accuracy of results and computing time. Comparative observations of accuracy results and computing time were carried out according to the learning rate produced through the optimization stage using the gray wolf algorithm.

Grey Wolf Optimization (GWO)

All classes are trained using the CNN architecture and Hyperparameter tuning with Grey Wolf Optimization proposed in this research. The GWO algorithm is a metaheuristic method inspired by the hunting behavior of gray wolves. Fig. 5 shows the training process where the results of the best model will be used for the testing process so that it can identify high-quality dry tobacco or vice versa if it falls into the low-quality class.

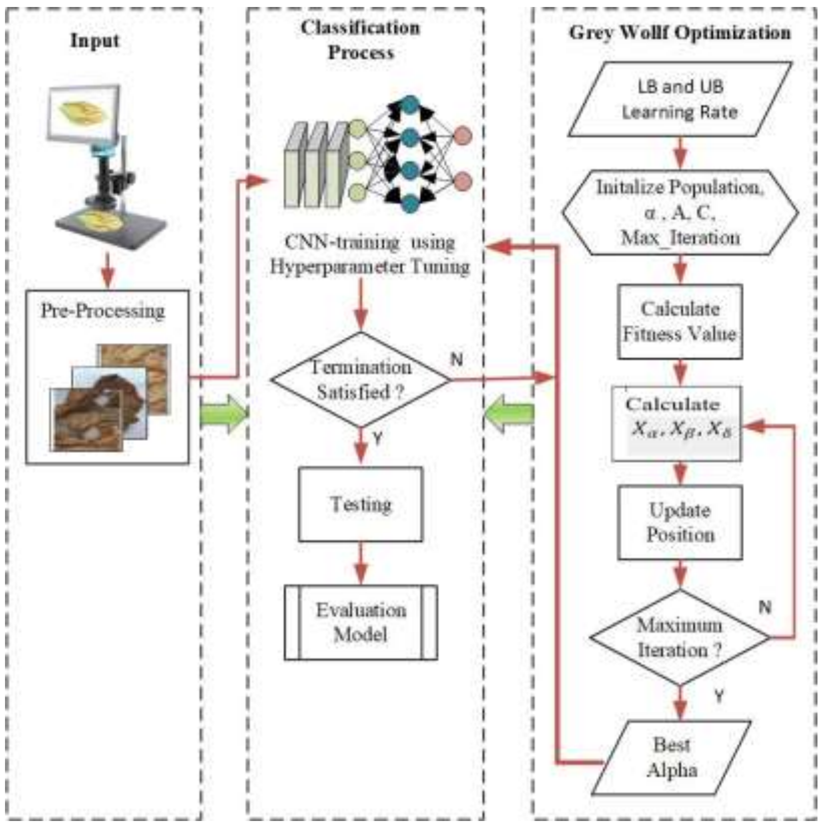

FIGURE 5. Tobacco leaf sorting system model design with hyperparameter tuning optimization

The implementation of GWO to find the optimal learning rate value is as follows:

1. Initialize the initial population

   Input a random learning rate value with a lower limit of 0.001 and an upper limit of 0.1. Next, initialize the population using the 5 random learning rate values that have been input. This population represents a candidate solution (as wolves). The maximum iteration used for population testing is 100. For each combination of parameter values, the average accuracy and average computing time are calculated every 10 rounds. The mathematical model to get the updated position of the wolf is as follows:

$$\vec{D} = \left|\vec{C}.\overrightarrow{X_p}(t) - \vec{X}(t)\right|$$
$$\vec{X}(t+1) = \overrightarrow{X_p}(t) - \vec{A}.\vec{D}$$

   Apart from that, the constant value α is also determined which linearly decreases from 2 to 0 during the loop with random values r_1,r_2 [0,1] to determine the value of the coefficient vectors A and C using the equation formula:

$$A = 2\alpha.r_1 - \alpha$$
$$C = 2\alpha.r_2$$

2. Fitness Function
   The stage of measuring the level of good and bad value of each search agent through accuracy results using a simple neural network.
3. Calculating Fitness Alpha (X_α), Beta (X_β) and Delta (X_δ) values.
   Sort the fitness values obtained from the previous step. The largest fitness value is chosen as X_α the second order as X_β and the third order as X_δ.
4. Update Wolf Position.
   Calculates the direction the search agent moves and updates the agent's position if the value is better than the previous position.
5. Check Stopping criteria
   If the iteration has reached a predetermined maximum value, the latest Alpha fitness value is used as the learning rate parameter index that will be used in CNN modeling
6. The CNN model with optimized hyperparameter tuning is saved for later use in the tobacco quality testing process.
7. The final results of the model are evaluated using the function of the confusion matrix

RESULTS AND DISCUSSION

This research conducted experiments with 2 test scenarios where the data was divided into 3 parts, namely training data, validation data, and testing data. Next, a comparison of hyperparameter tuning learning rate optimizers was carried out between the random search, grid search, and GWO methods. Testing was carried out to see the performance of the three methods in terms of accuracy and computing time.

Testing Scenario 1

The first trial was carried out by looking at the effect of the five learning rates of GWO results on accuracy. The five learning rate values are random numbers through a uniform distribution function with a lower limit of 0.001 and an upper limit of 0.1. The maximum iteration used is 10 iterations

TABLE 1. Accuracy values based on the implementation of learning rate candidates

Iterate	Candidate Solution	Accuracy	Iterate	Candidate Solution	Accuracy
1	0,00173	0,6748	6	0,00935	0,5261
	0,00368	0,7535		0,00852	0,7865
	0,03096	0,5706		0,00436	0,6774
	0,03722	0,5895		0,00798	0,7123
	0,01413	0,4704		0,00529	0,4512
2	0,05067	0,7521	7	0,00582	0,5415
	0,05215	0,6131		0,00612	0,8265
	0,00250	0,7485		0,00148	0,7662
	0,00370	0,4176		0,00261	0,5472
	0,00216	0,4319		0,00702	0,6891
3	0,00155	0,4229	8	0,00318	0,5275
	0,00868	0,5986		0,00342	0,5277
	0,00231	0,4133		0,00962	0,7149
	0,00225	0,4503		0,00151	0,4325
	0,00448	0,6600		0,00139	0,4734
4	0,00556	0,5347	9	0,00851	0,4865
	0,00402	0,5167		0,00960	0,4058
	0,00695	0,5162		0,00146	0,5627
	0,00625	0,7183		0,00983	0,6399
	0,00839	0,6477		0,00596	0,4586

5	0,01994	0,5245	10	0,00897	0,6221
	0,08692	0,4209		0,04016	0,6944
	0,02507	0,5267		0,08630	0,5324
	0,02639	0,4848		0,07096	0,6912
	0,02448	0,7130		0,07311	0,6036

Testing Scenario 2

Analyze the comparison of hyperparameter tuning learning rate using the GWO algorithm and 2 other conventional optimization algorithms, namely Random Search and Grid Search when applied to a tobacco leaf sorting system. The maximum iteration for each algorithm is the same, namely 10 iterations.

TABLE 2. Comparison of accuracy values and computing time

Method	Best Hyperparameter (Learning Rate)	Best Accuracy	Time Taken (seconds)
GWO Algorithm	0,0061	0,8265	61,46
Grid Search	0,001	0,6753	106,72
Random Search	0,0955	0,6752	87,28

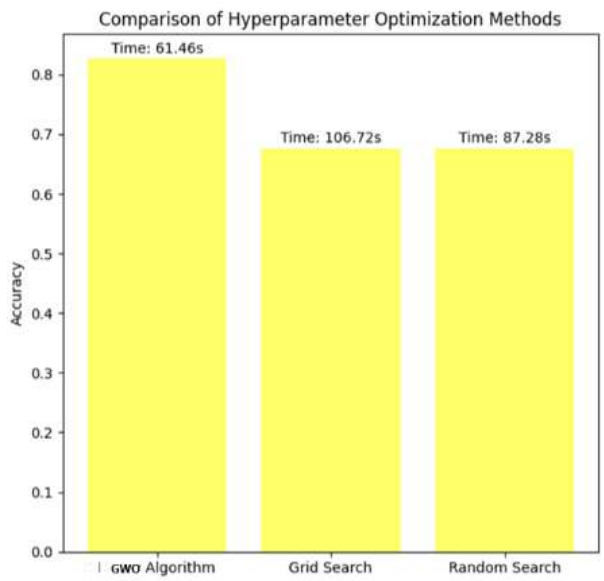

FIGURE 6. Comparison graph of GWO, grid search and random search computing times

CONCLUSION

- The GWO achieved a perfect accuracy of 82,65%, demonstrating its ability to effectively navigate the hyperparameter space and identify optimal settings rapidly. The total time taken for optimization was significantly lower than both Grid and Random Search.
- Grid Search and Random Search have almost the same accuracy value, namely 67,53% and 67,52% but Random search has a shorter computing time, namely 87,28s compared to Grid search with a time of 106,7s
  .

ACKNOWLEDGMENTS.

The research team would like to thank Universitas Trunojoyo Madura for giving us the opportunity to take part in science development activities in the field of computer vision

REFERENCES

1. Joko Riyono, Pujiastuti CE, Putri ALR. Clustering Negara Berdasarkan Skor Pengendalian Konsumsi Tembakau Menggunakan Algoritma DBSCAN. JTIK (Jurnal Tek Inform Kaputama). 2024;8(1).
2. Siregar GT. Pengyidikan Tindak Pidana Pemalsuan Pita Cukai Hasil Tembakau Pada Kantor Pengawasan Dan Pelayanan Bea Dan Cukai Belawan). J Ilm METADATA. 2023;5(3).
3. Renaldi VP, Bakhtiar A, Mufriantie F. Motivasi Petani Dalam Usahatani Tembakau Di Desa Bulangan Barat Kecamatan Pegantenan Kabupaten Pamekasan. J KIRANA. 2022;3(2).
4. Tiyasha, Tung TM, Yaseen ZM. A survey on river water quality modelling using artificial intelligence models: 2000–2020. Vol. 585, Journal of Hydrology. 2020.
5. Setyanti SWLH, Nilasari FS, Prajitiasari ED, Sudarsih S. Berbagi Pengetahuan Sebagai Mediator Pengaruh Persepsi Dukungan Organisasi terhadap Perilaku Inovasi UMKM Batik Jember. J Nusant Apl Manaj BISNIS. 2023;8(1).
6. Ning Y, Zhang LY, Mai J, Su JE, Cai JY, Chen Y, et al. Tobacco microbial screening and application in improving the quality of tobacco in different physical states. Bioresour Bioprocess. 2023;10(1).
7. Kim V Van, Nguyen HM, Greiman SE, Nguyen H Van, Nguyen CN, Vu MD, et al. Molecular and morphological characterization of Dollfustrema bagarii (Digenea: Bucephalidae) metacercariae from aquaculture channel catfish (Ictalurus punctatus) in northern Vietnam. J Fish Dis. 2022;45(8).
8. Arifianto K, Wulandari A. Pola Komunikasi Tata Niaga Tembakau (Studi Kasus Grader Tembakau dan Petani Tembakau di Kecamatan Parakan). J Ilm Manaj Inf dan Komun. 2023;7(1).
9. Wang W, Zhang L, Wang J, Long Z. Locating Defects and Image Preprocessing: Deep Learning in Automated Tobacco Production. J Sensors. 2022;2022.
10. Dey B, Masum Ul Haque M, Khatun R, Ahmed R. Comparative performance of four CNN-based deep learning variants in detecting Hispa pest, two fungal diseases, and NPK deficiency symptoms of rice (Oryza sativa). Comput Electron Agric. 2022;202.
11. Setiawan B, Siradjuddin I, Sumari ADW, Widjanarko, Mandyatma E, Putradi DF. Identification OF CNN Hyper-Parameters For Tobacco Leaf Quality Classification On Nvidia Jetson Nano. Eastern-European J Enterp Technol. 2023;6(2(126)).
12. Budhi Pramono T, Ika Qothrunnada N, Asadi F, Wawan Cenggoro T, Pardamean B. Water quality monitoring system for aquaponic technology using the internet of things (IoT). scik.org [Internet]. 2023 [cited 2024 Jul 28]; Available from: https://scik.org/index.php/cmbn/article/view/8221
13. Lu M, Chen D, Zhou Q, Wang Z, Chen T, Jiang S. Tobacco leaf grading and feature visualization based on deep learning. Tob Sci Technol. 2023;56(6).
14. Marzan CS, Ruiz CR. Automated tobacco grading using image processing techniques and a convolutional neural network. Int J Mach Learn Comput. 2019;9(6).
15. Chen Y, Bin J, Zou C, Ding M. Discrimination of Fresh Tobacco Leaves with Different Maturity Levels by Near-Infrared (NIR) Spectroscopy and Deep Learning. J Anal Methods Chem. 2021;2021.
16. Maulani AA, Winarno S, Zeniarja J, Putri RTE, Cahyani AN. Comparison of Hyperparameter Optimization Techniques in Hybrid CNN-LSTM Model for Heart Disease Classification. Sinkron. 2024;9(1).
17. Yi L. Improving University Second-year Students' Problem-Solving Skills in Computer Science Course (Digital Logic course) through Engineering Design Process in Inner Mongolia. J Educ Educ Res. 2024;7(2).
18. Hanifi S, Cammarono A, Zare-Behtash H. Advanced hyperparameter optimization of deep learning models for wind power prediction. Renew Energy. 2024;221.
19. Ayan E. Genetic Algorithm-Based Hyperparameter Optimization for Convolutional Neural Networks in the Classification of Crop Pests. Arab J Sci Eng. 2024;49(3).
20. Geng YR, Shen HC, Ni HF, Chen Y, Liu XS. Support Vector Machine Optimized by Near-Infrared Spectroscopic Technique Combined with Grey Wolf Optimizer Algorithm to Realize Rapid Identification of Tobacco Origin. Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy Spectr Anal. 2022;42(9).
21. Premkumar M, Sinha G, Ramasamy MD, Sahu S, Subramanyam CB, Sowmya R, et al. Augmented weighted K-means grey wolf optimizer: An enhanced metaheuristic algorithm for data clustering problems. Sci Rep. 2024;14(1).
22. Wang Y, Ran S, Wang GG. Role-oriented binary grey wolf optimizer using foraging-following and Lévy flight for feature selection. Appl Math Model. 2024;126.
23. Yang Z. Competing leaders grey wolf optimizer and its application for training multi-layer perceptron classifier. Expert Syst Appl. 2024;239.
24. Yu X, Duan Y, Cai Z, Luo W. An adaptive learning grey wolf optimizer for coverage optimization in WSNs.

Expert Syst Appl. 2024;238.