World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

WJARR

World Journal of Advanced Research and Reviews

(Review Article)

Analysis of SVM and KNN in the detection of leaf diseases

Sasmita Kumari Nayak *

Computer Science and Engineering, Centurion University of Technology and Management, Odisha, India.

World Journal of Advanced Research and Reviews, 2024, 21(01), 220–226

Publication history: Received on 23 November 2023; revised on 01 January 2024; accepted on 03 January 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.21.1.2716

Abstract

By utilizing the proposed automated strategy we can recognize the diseases without disappointment so we can lessen the use of manures by distinguishing the diseases. And along these lines we can lessen the wastage of cash and human exertion. The accuracy of the framework can be expanded by expanding the quantity of information base pictures. Here we got the greatest accuracy. It tends to be improved by expanding the pictures for every disease. Our efforts were focused on creating an automated method that could be used in agriculture to identify, count, categorize, and inspect mature or immature leaves. Over Leaf Diseases, the applied methodology was successfully implemented, and a satisfactory outcome was noted.

Keywords: Leaf Disease; Bacterial Blight; Multiclass SVM; Cercospora Leaf Spot.

1. Introduction

Leaf diseases continue to be extremely common in our nation, making it impossible for us to obtain leaves at the same rate as they are produced. Because the producer frequently encounters problems during packaging, occasionally plugs the leaf before it reaches maturity, etc., this is a result of their incorrect understanding of leaf size [1].

The estimation of organic product measure on-tree is useful in predicting harvest yield and can suggest buying pressing material (plate embed) and exhibiting strategies. Estimating the lifespan of a single natural product over time allows for the calculation of the rate at which organic products develop in physiological studies. Due to natural product covering way, its response to the disease and agronomic conditions error occurs during the season of estimating the extent of each organic product on-tree [2]. In some articles the classification can be done with the hybrid algorithms using any optimization method [3-17].

The color recognition technique is the primal recognition method that the human visual system has evolved to distinguish between leaf diseases that are green and magenta in color and to identify leaves against backgrounds of leaves. The process of identifying the leaf diseases using machine learning model is shown in Figure 1.

This article explains the initial attempts to use machine learning algorithms for dynamic task scheduling. Section 3 concludes by summarizing the difficulties in research for leaf disease identification and emphasizing the feasibility of machine learning algorithms in an effort to offer a better solution.

* Corresponding author: Sasmita Kumari Nayak

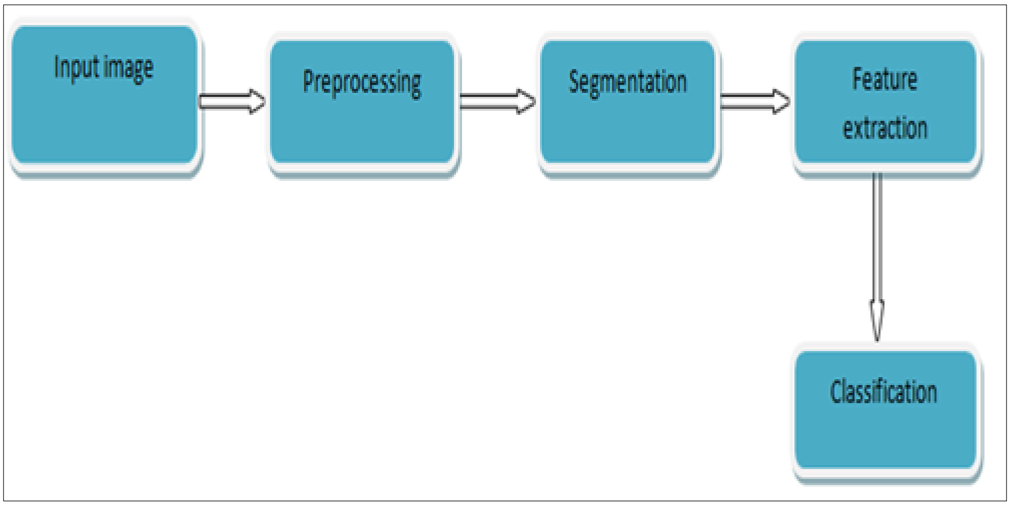

Figure 1 Process to classify the leaf disease

2. Machine Learning Algorithms for Leaf Disease Classification

The insights from earlier studies on leaf disease classification with machine learning algorithms are presented at this part. The section begins by explaining the underlying theory of machine learning algorithms [18-23]. The majority of machine learning Algorithms used for prediction as well as classification is then explored. Deep learning and optimization algorithms researches are also conducted by the image classification. The readers can get the idea of machine learning [10-15] and deep learning in details from the papers [24-30].

2.1. Machine Learning Algorithms

Machine learning is an artificial intelligence approach and a subfield of computer science. This method has the benefit of allowing a model to address issues that are not amenable to explicit algorithms, and it can be applied in multiple domains. A thorough analysis of various deterministic and machine learning techniques for predicting food, crop yield, weather, and hepatitis is provided in [31-34]. Even in situations where representation is not feasible, machine learning models identify relationships between inputs and outputs; this characteristic allow the use of machine learning models in many cases, for example in data mining and forecasting problems, spam filtering, classification problems, and pattern recognition. Because one must work with large datasets and machine learning models can handle pre-processing and data preparation, the classification and data mining aspects of this field are especially intriguing. Following this stage, forecasting issues can be solved using the machine learning models. This article provides the details of leaf disease classification using machine learning algorithms as shown in Figure 2.

2.1.1. Feature Extraction

It entails reducing the number of resources required to properly explain a sizable body of data. The primary disadvantage of analyzing complex data is that classification calls for a significant amount of memory and processing power. The classification algorithm performs poorly when applied to a new sample despite overfitting the trained sample. Thirteen number of texture are extracted from GLCM i.e. IDM, skewness, kurtosis, smoothness, variance, RMS, entropy, standard deviation, mean, homogeneity, energy, correlation and contrast.

2.1.2. Support Vector Machine

Support vector machines are learning algorithms that analyze data used for classification and are supervised learning models. A non-probabilistic binary linear classifier is created by an SVM training algorithm given a set of training examples that are each labeled as falling into one of two categories. The model then assigns new examples to the appropriate category. Different types of leaves are classified using multi-class SVM by taking into account the features of both the test sample image and the trained image [20–25].

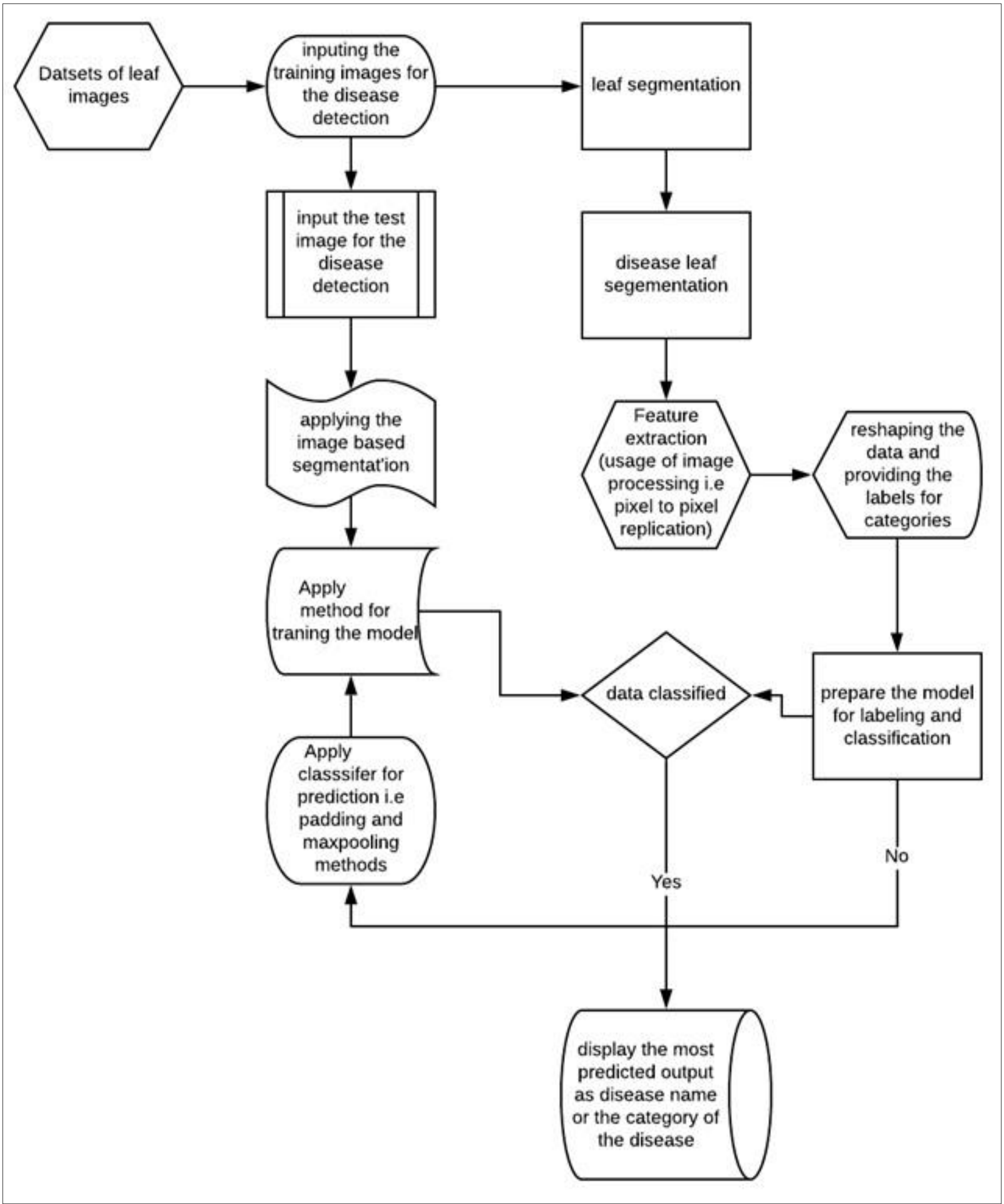

Figure 2 Step-by-step to classify the leaf disease using Machine Learning

2.1.3. KNN

An efficient machine learning technique for both regression and classification problems is K-nearest neighbors, or CNN. The idea behind KNN is to use the distance between an unknown sample and the K closest samples in the training set to classify it. The process of classification involves designating the most prevalent class among the K closest neighbors. Because KNN is a lazy learning algorithm, all that needs to be done for training is storing the training data. KNN is quick and memory-efficient because the real classification or regression of fresh samples is done at the prediction stage. KNN can handle both linear and nonlinear data, and it is simple to comprehend and apply. Nevertheless, the selection of K, the size of the features, and features that are not relevant can affect how well it performs. Unlabeled observations are classified using a KNN classifier by putting them in the same class as the labeled examples that are the closest to them. Both the training and test dataset's characteristics are gathered. For instance, the crunchiness and sweetness of fruit, vegetables, and grains can be used to identify them [35].

3. Results and Discussion

The experiment has done using MATLAB. The process of feature extraction has initially taken place. Thirteen numbers of features have been extracted during this process. These characteristics allow us to categorize the type of leaves. Only three of the thirty leaf diseases that we experimented with were incorrectly identified by the suggested algorithm. We found out of 26137 samples, there are more than 10000 events turned out to be Tropical storms, 6733 Hurricanes and so on. The Figure 3 shows the samples of spot leaf disease images as shown in below [36-41].

Figure 3 Represents Input dataset for Leafs-Brown Spot Disease

After that, detected the edges of all input images from the dataset. The sample of edge detection for leaf image is shown in Figure 4.

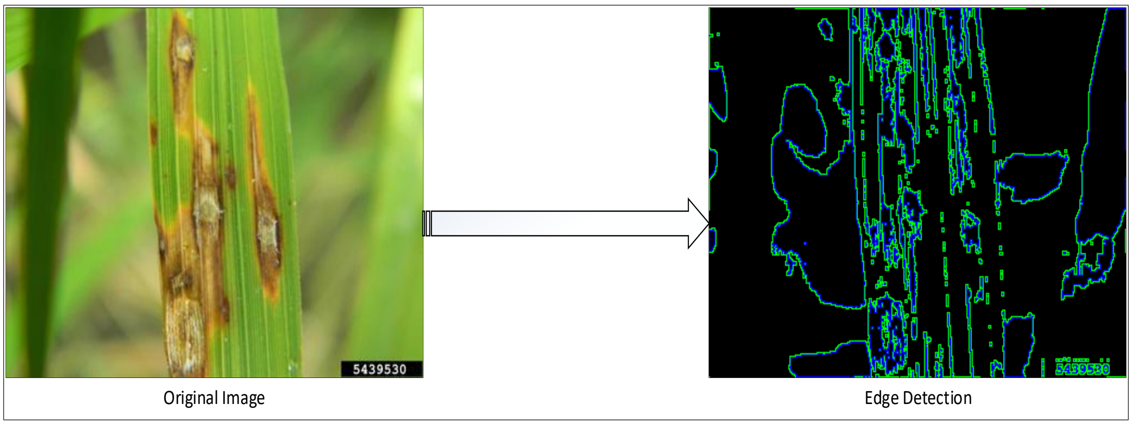

Figure 4 Edge detection results

After edge detection, classified the images. In this work, a K-Nearest Neighbor classifier is used for the classification purpose based on Euclidean distance measures. Here, Leaf diseases are categorized into 4 types: healthy, yellow leaf spot, Red rust and Powdery Mildew. We have performed KNN and SVM model for our collected data to found out which model has better accuracy and the results we found were given in the below table .

Table shows the classification results on features with Euclidean Distance Measure. Here we got the maximum accuracy of 78.1250% with K value of 11.

Table 1 Classification Results Using SVM and KNN

Classification Results Using SVM			**Classification Results Using KNN**			
Training Samples				**Training Samples**		
40%	**50%**	**60%**	**K values**	**40%**	**50%**	**60%**
68.7500	67.500	65.6250	1	72.9167	72.9167	68.7500
68.7500	65	65.6250	3	45.8333	60	62.5000
72.9167	70	71.8750	5	68.7500	67.5000	68.7500
68.7500	67.500	71.8750	7	70.8333	62.5000	71.8750
68.7500	67.500	75	9	60.4167	67.5000	68.7500
70.8333	72.500	75	11	68.7500	75	78.1250
70.8333	67.500	75	13	68.7500	67.5000	75

From the above table, we can easily conclude that KNN model has the best accuracy in comparison with SVM model. The accuracy score of KNN as well as accuracy score also found out to be maximum.

4. Conclusion

In first phase, we detect the Leaf on the tree and count its accuracy was Eighty percent, after that, we check matures immature where the accuracy was Ninety-Five percent. Due to some background problem, we can't count the Leaf properly every time so in the second stage, we add color thresholding process in that case we count the Leaf exactly with 78.125 percent accuracy. Using feature extraction, we categorize the various forms of leaf diseases in the third stage. There have occasionally been issues, particularly with size classification and measurement. It happens because of the classification and background, and some leaves are of the same quality. Our primary goal going forward is to address the shortcomings, enhance the algorithm, and provide better outcomes in size measurement section and try to measure the weight of each Leaf and analyses mature immature, which will be helpful for the farmers.

References

[1] Sachin D. Khirade, A.B Patil, "Plant Disease Detection Using Image Processing", International Conference on Computing Communication Control and Automation", 2015.

[2] Malvika Ranjan, Manasi Rajiv Weginwar, NehaJoshi, Prof.A.B. Ingole, "detection and classification of leaf disease using artificial neural network", International Journal of Technical Research and Applications, 2015.

[3] Nayak, S. K., Padhy, S. K., & Panigrahi, S. P. (2012). A novel algorithm for dynamic task scheduling. Future Generation Computer Systems, 28(5), 709-717.

[4] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Efficient multiprocessor scheduling using water cycle algorithm. Soft Computing Applications, 131-147.

[5] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Efficient multiprocessor scheduling using water cycle algorithm. In Soft Computing Applications (pp. 131-147). Springer, Singapore.

[6] Nayak S.K., Padhy S.K., Panda C.S. (2018) Efficient Multiprocessor Scheduling Using Water Cycle Algorithm. In: Pant M., Ray K., Sharma T., Rawat S., Bandyopadhyay A. (eds) Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing, vol 583. Springer, Singapore.

[7] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2019, February). Dynamic Task Scheduling Problem Based on Grey Wolf Optimization Algorithm. In 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP) (pp. 1-5). IEEE.

[8] S. K. Nayak, C. S. Panda and S. K. Padhy, "Dynamic Task Scheduling Problem Based on Grey Wolf Optimization Algorithm," 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), Gangtok, India, 2019, pp. 1-5, doi: 10.1109/ICACCP.2019.8882992.

[9] Nayak, S. K., & Panda, C. S. (2021). Dynamic task scheduling using nature inspired algorithms. J. Math. Comput. Sci., 11(1), 893-913.

[10] NAYAK, S. K., & PANDA, C. S. (2021). DYNAMIC TASK SCHEDULING USING NATURE INSPIRED ALGORITHMS.

[11] Sasmita Kumari Nayak, Chandra Sekhar Panda, Dynamic task scheduling using nature inspired algorithms, J. Math. Comput. Sci., 11 (2021), 893-913 http://hdl.handle.net/10603/464925

[12] Nayak, S. K. Multiprocessor Scheduling Using Nature Inspired Optimization. Thesis. Sambalpur University, Odisha, 2021. Web. http://hdl.handle.net/10603/464925

[13] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Multiprocessor Scheduling using Krill Herd Algorithm (KHA). International Journal of Computer Sciences and Engineering, 6(6), 7-17.

[14] Nayak, S. K., Panda, C. S., & Padhy, S. K. (2018). Multiprocessor Scheduling using Krill Herd Algorithm (KHA).

[15] S.K. Nayak, C.S. Panda, S.K. Padhy, "Multiprocessor Scheduling using Krill Herd Algorithm (KHA)," International Journal of Computer Sciences and Engineering, Vol.6, Issue.6, pp.7-17, 2018. CrossRef-DOI: https://doi.org/10.26438/ijcse/v6i6.717

[16] Nayak, S. K., & Panda, C. S. (2019). Multiple Processor Scheduling with Optimum Execution Time and Processor Utilization Based on the SOSA. International Journal of Recent Technology and Engineering, ISSN: 2277-3878. Volume-8, Issue-2, July, 5463-5471.

[17] Nayak, S. K., & Panda, C. S. Multiple Processor Scheduling with Optimum Execution Time and Processor Utilization Based on the SOSA.

[18] Barik, Swati & Nayak, Sasmita & Panda, Stitiprajna & Tripathy, Aeisuriya & Mohapatra,Gourav. (2020). Human Face Recognition using LBPH. International Journal of RecentTechnology and Engineering. 8. 3208-3212. 10.35940/ijrte.F8117.038620

[19] Swain, S., Nayak, S. K., & Barik, S. S. (2020). A review on plant leaf diseases detection and classification based on machine learning models. Mukt shabd, 9(6), 5195-5205.

[20] S Sucharita, S Nayak, S Panda et al., "Human Face Recognition using LBPH", International Journal of Recent Technology and Engineering, vol. 8, no. 6, pp. 3208-3212, 2020.

[21] Stitiprajna Panda*, Swati Sucharita Barik, Sasmita Kumari Nayak, Aeisuriya Tripathy, and Gourav Mohapatra. 2020. Human Face Recognition using LBPH. International Journal of Recent Technology and Engineering (IJRTE) 8, 6 (2020), 3208-3212. DOI:https://doi.org/10.35940/ijrte.f8117.038620

[22] S. Panda, S. Sucharita Barik and S. Kumari Nayak, "Human Face Recognition using LBPH", International Journal of Recent Technology and Enginee, vol. 8, no. 6, pp. 3208-3212, 2020.

[23] Sucharita S , Nayak S , Panda S , Human Face Recognition using LBPH. International Journal of Recent Technology and Engineering, 2020, 8(6):3208-3212.

[24] Jena, T. R., Barik, S. S., & Nayak, S. K. (2020). Electricity consumption & prediction using machine learning models. Acta Tech. Corviniensis-Bull. Eng, 9, 2804-2818.

[25] Nayak, S. K., Barik, S. S., & Beura, M. (2020). Analysis of Infectious Hepatitis Disease with High Accuracy Using Machine Learning Techniques. TEST Engineering & Management, 83, 83.

[26] Nayak, S. K. (2020). Analysis and high accuracy prediction of coconut crop yield production based on principle component analysis with machine learning models. International Journal of Modern Agriculture, 9(4), 359-369.

[27] Sasmita Kumari Nayak, Swati Sucharita Barik, Mamata Beura," Weather Forecasts Based on Rainfall Prediction Using Machine Learning Methodologies," Adalya Journal 9 (6), Page No: 72 – 80, ISSN NO: 1301-2746.

[28] Mishra, S. P., Siddique, M., Beura, M., & Nayak, S. K. (2021). Analysis of Indian Food Based on Machine learning Classification Models. Journal of Scientific Research and Reports, 27(7), 1-7.

[29] Nayak, S. K. (2023). Classification of cyclones using machine learning techniques. World Journal of Advanced Research and Reviews, 20(2), 433-440.

[30] Sasmita Kumari Nayak. (2020). CONSTRICTING THE BIVARIATE ANALYSISOFHIGHEST CROP YIELD PRODUCTION BASED ON DIFFERENT ZONES OF INDIA. International Journal of Modern Agriculture, 9(4), 216 - 226. Retrieved from https://modern-journals.com/index.php/ijma/article/view/205

[31] Sasmita Kumari Nayak, Mohammed Siddique. (2020). EFFECT OF STOCK INDEX PARAMETERS ON FORECASTING THE HIGH STOCK VALUE OF VISA STEEL USING DEEP LEARNING NEURAL NETWORK MODEL. International Journal of Modern Agriculture, 9(4), 227 - 236. Retrieved from https://modern-journals.com/index.php/ijma/article/view/207

[32] Sowmya Jagadeesan, B. B. (2022). A Perishable Food Monitoring Model Based on IoT and Deep Learning to Improve Food Hygiene and Safety Management. IJFANS International Journal of Food and Nutritional Sciences, 11(8), 1164-1178.

[33] Jagadeesan, S., Barman, B., Agarwal, R. K., Srivastava, Y., Singh, B., Nayak, S. K., & Venu, N. A Perishable Food Monitoring Model Based On Iot And Deep Learning To Improve Food Hygiene And Safety Management. interventions, 8, 9.

[34] Rajesh, J., Ashraf, M. S., Kaur, L., Rout, S., Nayak, S. K., Kaur, G., & Saikanth, D. R. K. APPLICATION OF FUZZY LOGIC IN SMART AGRICULTURE TO RECOGNISE TOMATO FRUIT RIPENESS. IJFANS International Journal of Food and Nutritional Sciences, 11(1), 2360-2367.

[35] N. Dash, S. K. Nayak and J. Majumdhar, "Detection of Cut Transition in Videos Using Optical Flow and Clustering," 2021 Asian Conference on Innovation in Technology (ASIANCON), PUNE, India, 2021, pp. 1-7, doi: 10.1109/ASIANCON51346.2021.9544553.

[36] Majumdhar, J., & Nayak, S. K. (2021, August). A Novel Method on Summarization of Video Using Local Ternary Pattern and Local Phase Quantization. In 2021 2nd International Conference on Range Technology (ICORT) (pp. 1-6). IEEE.

[37] Chakravarthy, A., Panda, B. S., & Nayak, S. K. (2023). Review and Comparison for Alzheimer's Disease Detection with Machine Learning Techniques. International Neurourology Journal, 27(4), 403-409.

[38] Nayak, S. K. (2023). Nature inspired algorithms in dynamic task scheduling: A review. World Journal of Advanced Research and Reviews, 20(2), 829–833.

[39] Nayak, S. K. (2023). Exploring and forecasting of solar radiation with machine learning methods. World Journal of Advanced Research and Reviews, 20(2), 824–828.

[40] Bhumika S.Prajapati, Vipul K.Dabhi Harshadkumar, B.Prajapati, “A Survey on Detection and Classification of Cotton Leaf Diseases”, International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016.

[41] M.Malathi, K.Aruli , S.Mohamed Nizar, A.Sagaya Selvaraj, “A Survey on Plant Leaf Disease Detection Using Image Processing Techniques”,International Research Journal of Engineering and Technology (IRJET),Volume: 02 Issue: 09, Dec 2015.