AN IDENTIFICATION OF PLANT LEAF DISEASE DETECTION USING HYBRID ANN AND KNN

Abstract

Plant diseases are a regular cause of low yields and lessen the income for farmers. Detecting plant diseases is an important task in the agribusiness sector because plant diseases are common. To detect diseases in the leaves, it is necessary to constantly monitor the plants. However, manual detecting of diseases consumes a more period and effort. Hence, it is better to have an automated system. This analysis describes identification of plant leaf disease detection using hybrid Artificial Neural Networks (ANN) and Convolutional Neural Network (CNN). The characteristics are detected for the non-infected and infected regions of the leaf. The complete data set containing of 300 images and classified for training and testing. These images are executed with the described technique as well as it is classified as infected or non-infected. Hybridized model achieves 98% of Accuracy, 97% of Precision and 96% of Recall. Observational output shows that plant infections will be classified exactly.

Keywords: Plant diseases, Image analysis, feature extraction, Artificial Neural Networks (ANN) and Convolutional Neural Network (CNN).

Authors

Sai Srinivas Vellela
Assistant Professor
Department of CSE-Data Science
Chalapathi Institute of Technology
Guntur, Andrapradesh, India
sais1916@gmail.com

Dr K Kiran Kumar
Professor
Department of CSE
Chalapathi Institute of Technology
Guntur, Andrapradesh, India
kommineni.kiran11@gmail.com

Dr .M Venkateswara Rao
Professor
Department of CSE
NRI Institute of Technology
Pothavarappadu
Vijayawada, Andrapradesh, India
muvva.venky@gmail.com

Venkateswara Reddy B
Assistant Professor
Department of CSE - Cyber Security
Chalapathi Institute of Technology
Guntur, Andrapradesh, India
bvreddymtech@gmail.com

Khader Basha Sk
Assistant Professor
Department of CSE-Data Science
Chalapathi Institute of Technology
Guntur, Andrapradesh, India
khaderbasha576@gmail.com

Roja D
Assistant Professor
Department of CSE-Data Science
Chalapathi Institute of Technology
Guntur Andrapradesh, India

I. INTRODUCTION

Agriculture is the major employment in India. Agriculture is significant to economy for every country. Advances in the agricultural sector are mainly aimed at meeting the demands of the growing population [1]. At current situation agriculture sector requires an upgrade for survival. Crops are infected by bacterial and fungal disorders. Due to this, there is a major loss in production of the farmers. To get optimum production, the crop must be healthy.

In a country like India, whose economy is heavily dependent on agriculture, technology plays an important role for humans when it comes to food production. Much of the world's populations are dependent on their huge economy. Also economic development performs a significant part in growth of any country and it's GDP (Gross Domestic Product). The effect of this economy is completely dependent on agriculture. Hence, various aspects of farming affect the quality and quantity of grains and vegetables [2]. These grains and vegetables are susceptible to various infections because of various climatic conditions in various locations. Hence, the output of growers in any country faces high damages due to these diseases.

Diagnosing diseases with our eyes always remains a complicated process. To do that, it is essentially constant to monitor the farm [3]. This is a mechanical process. It is also expensive when the land size is more. Because of difficulty of the agricultural specialists are unable to identify the infection simply and detect an answer for the issue. An automated model for detecting plant diseases can be significant benefit for farmers. This model acts as a device to notify farmers at the exact time and make important precautions. Different infections that affecting the plants, which causes damage to plant parts like leaves, fruits, seeds etc. These infections are specific for various parts of the herb. Leaves are the significant region of the herb. If a plant's leaf is infected, it disrupts the herbs life cycle. Infections which are regularly affect leaves are bacterial, fungal disease etc. Therefore, early diagnosis of Infection is very important.

Various approaches have been developed to improve plant disease classification results, including modified versions of well-known DL (Deep Learning) models, different training methods, data augmentation methods, cascaded versions of two successful DL architectures, etc. For example, the popular GoogLeNet model has been enhanced to get better test accuracy for detecting maize leaf blight in a short period of duration because of a smaller number of aspects [4]. Correspondingly, a modified CNN architecture inspired by the Alexnet model is proposed, which has a smaller number of filters and nodes in the convolutional layers, which clearly reduces the total parameters compared to the original model and successfully detects the disease in tea leaves [5].

Machine learning, particularly CNN or its variant, has appeared as a new system in image classification and object prediction in recent times [6]. Spectral information is significant in categorization work in Hyperspectral Imaging (HSI). Anyhow, some bands will risk executing work, with fewer bands containing more significant data than others. Some of the traditional machine learning algorithms used to extract needful data from spectral bands are principal component analysis, independent component discriminant analysis, and linear discriminant analysis [7]. There are many methods to predict various kinds of diseases in plants in its early period. The aim of this paper is to detect the plant leaf disease

implementation using two various ML models namely, CNN and ANN. Conventional methods of plant disease detection include naked eye observation techniques and are not effective for large crops. Disease detection in plants using digital image processing and machine learning is efficient, consuming short duration and accurate. This technique saves time, effort, labor and use of pesticides. This approach will become a little contribution for agriculture fields.

The remaining part is organized as follows: Section II expands the concepts of Related Work. The described plant leaf disease detection system is described in Section III. Section IV analyzed the observational outputs and describes about performed accuracy. In Section V, the paper is concluded.

II. LITERATURE SURVEY

Saradhambal. G, Dhivya.R, Latha.S, R.Rajesh, et al. [8] proposed recent techniques which are utilized for disease prediction. They described Otsu threshold algorithm as well as K-means algorithm for image segmentation. Color co-occurrence techniques as well as leaf color extraction by utilizing H as well as B components for feature extraction are explained. And compared ANN and *Back Propagation Neural Network* (*BPNN*)classifiers to classify the diseases.

Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, et. al. [9] proposed Convolutional Neural Network based innovative rice disorder identification technique with image preprocessing to detect and identify the diseases using 500 original images from rice empirical field. In addition, the described technique showed high accuracy because of 10-fold cross-validation scheme, higher feasibility and efficiency, faster convergence rate and greater detection capability than conventional techniques of ML.

Amara, J., Bouaziz, B., Algergawy, A., et. al. [10] made utilization of CNN named LeNet for classification of disorder on banana leaf. For the image processing part, image resizing and leaf images were converted to gray scale to change the images to a standard scale for later processing. In categorization system, fully connected layers utilize neurons which are connected to every neuron for the followed layer, thus neurons has entire connections to all feature maps produced by the pervious layer. The described framework is capable to provide better outputs in terms of accuracy 98%.

Sengupta S., Das A.K., et al. [11] made a categorization system for detecting disorders in rice. They utilized incremental classification in a similar way that performs on the principle that knowledge obtained from information will be utilized with currently obtainable information to generate an upgraded classification. The classifiers were made following a rule based model called Particle Swarm Optimization (PSO). In this design, optimized classification rules are created for dynamically increasing information utilizing the concept of association rule mining and Particle Swarm Optimization algorithm. Characteristics like pattern and quality have been used to characterize rice disease; this characteristic includes spot pattern, region, circumference, etc. The accuracy of the model is 85%.

PrabhjeetKaur, Sanjay Singla, Sukhdeep Singh, et al. [12] described a technique for detecting disorder that will perform in the cucumber plants. Segmentation of non-infected as

well as affected regions is accomplished using constant shape detection model. The characteristics like color, shape, and texture will be obtained. The characteristics are provided to SVM (Support Vector Machine) that executes the important categorization. They concluded that an output executed from the Support Vector Machine is better when compared that obtains from neural networks.
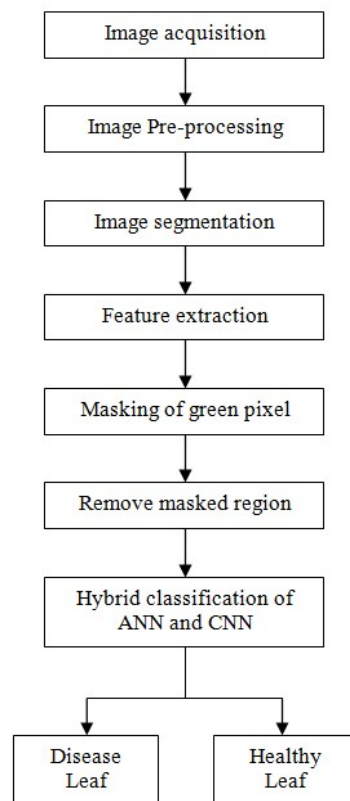
M. Jhuria, A. Kumar and R. Borse, et. al. [13] describes the neural networks models to predict as well as observe the disorder for fruits herbs from plantation to harvesting. The number of three features vectors, which are obtained includes color, morphology and texture. The morphology characteristic obtains 90% of accurate outputs that compared with two different vectors.

Revathi, P., Hemalatha, M. et. al. [14] described a novel model for prediction of visual disorder of the plant. Digital pictures of herbs are generated and pre-processed. The obtained techniques, like edge detection, color space and textural elements, are executed. The characteristics which are obtained are forwarded to classifiers. This analysis targets the affected area for detection in cotton leaf by utilizing image processing.

Camargo, A., Smith, et. al. [15] analyzed an Algorithm for Perpetuating Visual Traits of Plant Diseases by Color Image Analysis. The described model starts at changing Gray scale image of affected leaf for color changing like I3a, I3b and H. Initially, the changed picture is segmented by using intensity distribution in histogram. This method is particularly utilized to aim a particular picture which has high intensity distribution. If the segmentation method is over, the part executed is then performed to remove pixel parts, which are out of the focused area. The outputs achieved the basis and were effectual in detecting plant pathogens.

III. HYBRID PLANT LEAF DISEASE DETECTION

The framework of identification of plant leaf disease detection using hybrid Artificial Neural Networks (ANN) and Convolutional Neural Network (CNN) is represented in below Fig. 1.

```
┌─────────────────────┐
│  Image acquisition  │
└─────────────────────┘
           │
┌─────────────────────┐
│ Image Pre-processing│
└─────────────────────┘
           │
┌─────────────────────┐
│ Image segmentation  │
└─────────────────────┘
           │
┌─────────────────────┐
│ Feature extraction  │
└─────────────────────┘
           │
┌─────────────────────┐
│Masking of green pixel│
└─────────────────────┘
           │
┌─────────────────────┐
│ Remove masked region│
└─────────────────────┘
           │
┌─────────────────────┐
│Hybrid classification│
│   of ANN and CNN    │
└─────────────────────┘
        │       │
┌──────────┐ ┌──────────┐
│ Disease  │ │ Healthy  │
│  Leaf    │ │  Leaf    │
└──────────┘ └──────────┘
```

Figure.1: Framework of Plant Leaf Disease Detection

The leaves of the plant are susceptible to many diseases. This may be because of humidity and natural situations. Common disorders include viral, bacterial and fungal disease. This leads to a transform in color patterns. The changes are hard to distinguish due to same patterns. Hence the early prediction of these disorders can prevent damage. A ML method for plant disoder classification is described in this analysis.

The images were collected from the rural area of Panpoli village in Shenkottaitaluk, Tirunelveli district, Tamil Nadu. A number of 300 leaf sample leaves were taken from normal as well as affected parts from field. Pictures are captured by utilizing the Redmi Note 5 camera with high resolution and then cropped to 256x256pixels. The image is in RGB format.

A preprocessing method is executed to separate any noise that enhances the image features. Pictures are pre-processed using contrast improvement. It improves image contrast by mapping the input intensity to a recent value. At the time of image pre-processing, Red Green Blue pictures are transformed to hue, saturation, and value pictures, because performing with HSV (hue, saturation, and value) is easier to distinguish colors.

Image segmentation is third step in presented method. Image segmentation is a method executed to classify a digital image into many segments that aid in simple detection. The segmented images are clustered into different sectors using K-Means classifier. The choices of K are highly significant in the K-means clustering method. Selection of three values by utilizing trial and error method, third value of K was determined.

This step is critical for image classification. Alternative of selecting the complete picture, they obtained features from the affected area only. A Gray Co-occurrence Matrix (GLCM) is utilized to obtain characteristics. It is utilized to detect the spatial variation of Frey level in a picture. Few features which can be obtained from a picture using GLCM are Contrast, Correlation, Energy, and Homogeneity. Remaining characteristics like Mean, Standard Deviation, Entropy, Root Mean Square (RMS), Variance, Smoothness, Kurtosis, and Skewness are extracted using Matlab commands.

At this stage mostly green pixels are masked. Here, they calculate threshold value utilized for these pixels. Then highly green pixels are masked as follows: If pixel intensity is lower than a precomputed threshold value, Red, Green and Blue(RGB) components of this pixel assigned value is zero. In affected groups, boundaries separate masked cells.

Hybridized classification model is used in this approach and used classifications are ANN and CNN.

An ANN is a computational system depended on pattern and purposes of biological neural networks. It obtains knowledge by instructing. Therefore, it should be trained to enhance the accuracy of classification. Artificial Neuron Network contains of three layers that are internally connected. The initial layer forward the information to the second layer and it forward the described information to the third layer. In the described model, three hidden layers are selected with 'N' numbers of input (N=1, 2,….8). Hence, result layers produce categorization output as usual and a diseased leaf picture. If estimated value is higher than 0.5, the healthy leaf determined otherwise disease affected.

Convolutional neural network is variation among the ANN and is highly utilized for categorization, image processing, segmentation tasks, etc. Convolution means sliding a filter over the picture to know few significant characteristics of the input image. They realized that a picture is matrix filled with few numerical values. The invisible layer of a CNN consists of convolutional layers, pooling layers, fully connected layers, and a normalization layer, where result of invisible layers is fed to an activation operation. A twisted two-dimensional layer consists of a flat layer and a dense layer connecting between them. For this model ReLU or Rectified Linear Units are used in the form of enable function. In the framework, softmax is described, based on activation high likelihood for forecasting. The equation for the SoftMax function is given below:

$$P(x) = \frac{e^{x^T W^l}}{\sum_{k=1}^{k} e^{x^T W^l}} \dots \dots (1)$$

Here, W signifies x and W's internal product. The ANN and CNN classifiers classify the collected image which is normal or infected.

IV. RESULT ANALYSIS

In this section, categorization outputs of tomato and potato leaf pictures are described by performing a set of analysis on a dataset that includes two classification as uninfected and infected leaf pictures to represent the execution of the observational outputs. A total of 300 leaf samples were taken from healthy as well as affected parts of field. Pictures are captured using the Redmi Note 5 camera with high resolution and cropped to 256x256pixels. 70% of

leaf pictures were utilized for the training dataset and the rest of the 30% were used for the testing dataset.

Accuracy, Precision and Recall are different parameters used in this analysis for performance evaluation.

- **Accuracy:** Accuracy means the ratio of exactly estimates for complete detections. Accuracy can be mentioned as capability to accurately detect result of a situation.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \dots (2)$$

- **Recall:** True positive rate, sensitivity, or recall defined as measure which explains the proportion of true instances that exactly have diabetes with the true positive instances.

$$Recall = \frac{TP}{TP + FN} \dots (3)$$

- **Precision:** Positive detected value or precision is number of exact positive scores divided by number of true scores detected by the categorization algorithm represented in below equation (4)

$$Precision = \frac{TP}{TP + FP} \dots (4)$$

Where,
- True Negative (TN) – A percentage of healthy leaves which are exactly grouped.
- True Positive (TP) – A percentage of diseased leaves which are exactly grouped.
- False Positive (FP) - The percentage of healthy leaves which are not exactly grouped like diseased leaves.
- False Negative (FN) - The percentage of diseased leaves which are not exactly grouped as healthy leaves.

The performance of leaf disease detection systems based on different classifiers such as hybridized model (ANN+CNN), Support Vector Machine (SVM) and K- Nearest Neighbor (KNN) are compared in below Table 1.

**Table 1: Comparative Performance Analysis of Different Classifiers**

Parameters	SVM	KNN	ANN+CNN
Accuracy	87	89	98
Precision	86	87	97
Recall	88	86	96

Graphical representation of Recall and Precision values of the different classifiers are shown in below Fig. 2.
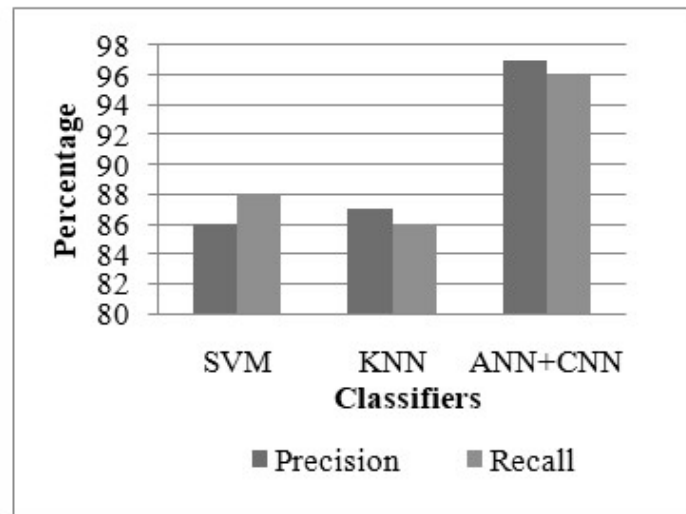
Figure 2: Performance Comparison

Graphical representation of Accuracy of hybridized model (ANN+CNN), SVM and KNN is represented in below Fig. 3.
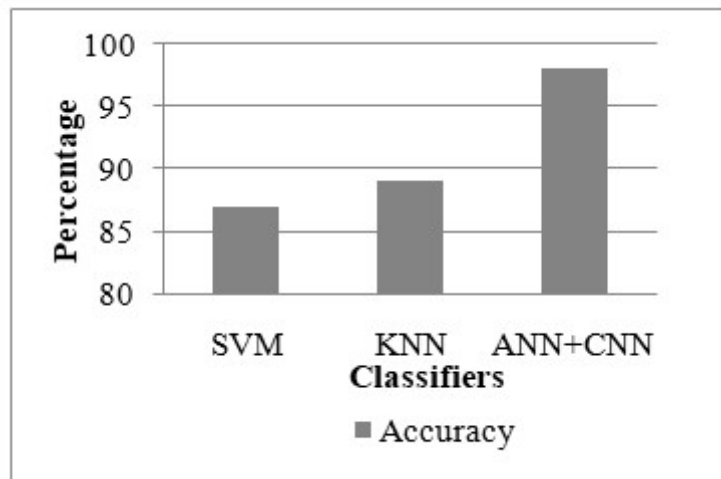
Figure 3: Accuracy Analysis of Different Classifiers

Therefore from results it is clear that, hybridized model (ANN+CNN) has high efficiency than the other machine learning models as CNN and Multi layer perceptron (MLP) in terms of Accuracy, Precision and Recall. Hybridized model achieves 98% of Accuracy, 97% of Precision and 96% of Recall.

V. CONCLUSION

In this analysis, identification of plant leaf disease detection using hybrid Artificial Neural Networks (ANN) and Convolutional Neural Network (CNN) is described. Farmers are facing issues due to not being able to identify disorders in plants manually. Also, specialist instruction is difficult for farmers. It would be beneficial to have an automated mechanism to detect plant diseases. The image was captured for healthy as well as diseased region using Redmi Note 5 camera with high resolution. The segmented images are clustered into different sectors using K-Means classifier. Gray Co-occurrence Matrix (GLCM) is generated to obtain characteristics. Accuracy, Precision and Recall are different parameters used in this analysis

for performance evaluation. Comparative performance analysis is made among three classification models as described hybridized model (ANN+CNN), Support Vector Machine (SVM) and K- Nearest Neighbor (KNN). Hybridized model achieves 98% of Accuracy, 97% of Precision and 96% of Recall. The described hybrid model classifies the leaves very efficiently than other classifications. At present, they classify the diseases as infectious. However, for future work, they plan to categorize every disease by its name.

REFERENCES

[1]   R. Kavitha, M. Kavitha, R. Srinivasan, "Crop Recommendation in Precision Agriculture using Supervised Learning Algorithms", 2022 3rd International Conference for Emerging Technology (INCET), Year: 2022

[2]   Zihan Jiang, YuboGuo, Kaiyang Jiang, Mingrui Hu, Zimin Zhu, "Optimization of Intelligent Plant Cultivation Robot System in Object Detection", IEEE Sensors Journal, Volume: 21, Issue: 17, Year: 2021

[3]   Vellela, S. S., & Krishna, A. M. (2020). On Board Artificial Intelligence With Service Aggregation for Edge Computing in Industrial Applications. Journal of Critical Reviews, 7(07).

[4]   Lili Li, Shujuan Zhang, Bin Wang, "Plant Disease Detection and Classification by Deep Learning—A Review", IEEE Access, Volume: 9, Year: 2021

[5]   DraškoRadovanović, Slobodan Đukanović, "Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms", 2020 24th International Conference on Information Technology (IT), Year: 2020

[6]   Fatma Marzougui, Mohamed Elleuch, Monji Kherallah, "A Deep CNN Approach for Plant Disease Detection", 2020 21st International Arab Conference on Information Technology (ACIT), Year: 2020

[7]   Vellela, S. S., &Balamanigandan, R. (2022, December). Design of Hybrid Authentication Protocol for High Secure Applications in Cloud Environments. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 408-414). IEEE.

[8]   Madhuri, A., Jyothi, V. E., Praveen, S. P., Sindhura, S., Srinivas, V. S., & Kumar, D. L. S. (2022). A New Multi-Level Semi-Supervised Learning Approach for Network Intrusion Detection System Based on the 'GOA'. Journal of Interconnection Networks, 2143047.

[9]   Vellela, S. S., Reddy, B. V., Chaitanya, K. K., &Rao, M. V. (2023, January). An Integrated Approach to Improve E-Healthcare System using Dynamic Cloud Computing Platform. In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 776-782). IEEE.

[10] S Phani Praveen, RajeswariNakka, AnuradhaChokka, VenkataNagarajuThatha, SaiSrinivasVellela and UddagiriSirisha, "A Novel Classification Approach for Grape Leaf Disease Detection Based on Different Attention Deep Learning Techniques" International Journal of Advanced Computer Science and Applications(IJACSA), 14(6), 2023. http://dx.doi.org/10.14569/IJACSA.2023.01406128

[11] Praveen, S. P., Sarala, P., Kumar, T. K. M., Manuri, S. G., Srinivas, V. S., &Swapna, D. (2022, November). An Adaptive Load Balancing Technique for Multi SDN Controllers. In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 1403-1409). IEEE.

[12] Vellela, S. S., BashaSk, K., &Yakubreddy, K. (2023). Cloud-hosted concept-hierarchy flex-based infringement checking system. International Advanced Research Journal in Science, Engineering and Technology, 10(3). Vellela, S. S., &Balamanigandan, R. (2023).

[13] Sk, K. B., Roja, D., Priya, S. S., Dalavi, L., Vellela, S. S., & Reddy, V. (2023, March). Coronary Heart Disease Prediction and Classification using Hybrid Machine Learning Algorithms. In 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA) (pp. 1-7). IEEE.

[14] VenkateswaraRao, M., Vellela, S., Reddy, V., Vullam, N., Sk, K. B., &Roja, D. (2023, March). Credit Investigation and Comprehensive Risk Management System based Big Data Analytics in Commercial Banking. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 2387-2391). IEEE.

[15] Vellela, S.S., Balamanigandan, R. Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. Multimed Tools Appl (2023). https://doi.org/10.1007/s11042-023- 15926-5

[16] Vullam, N., Vellela, S. S., Reddy, V., Rao, M. V., SK, K. B., &Roja, D. (2023, May). Multi-Agent Personalized Recommendation System in E-Commerce based on User. In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 1194-1199). IEEE.

[17] Rao, K. N., Gandhi, B. R., Rao, M. V., Javvadi, S., Vellela, S. S., & Basha, S. K. (2023, June). Prediction and Classification of Alzheimer's Disease using Machine Learning Techniques in 3D MR Images. In 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) (pp. 85-90). IEEE.

[18] Rao, D. M. V., Vellela, S. S., Sk, K. B., & Dalavai, L. (2023). Stematic Review on Software Application Under-distributed Denial of Service Attacks for Group Website. Dogo Rangsang Research Journal, UGC Care Group I Journal, 13.

[19] Venkateswara Reddy, B., &KhaderBashaSk, R. D. Qos-Aware Video Streaming Based Admission Control And Scheduling For Video Transcoding In Cloud Computing. In International Conference on Automation, Computing and Renewable Systems (ICACRS 2022).

[20] Madhuri, A., Praveen, S. P., Kumar, D. L. S., Sindhura, S., &Vellela, S. S. (2021). Challenges and issues of data analytics in emerging scenarios for big data, cloud and image mining. Annals of the Romanian Society for Cell Biology, 412-423.

[21] Vellela, S. S., Balamanigandan, R., & Praveen, S. P. (2022). Strategic Survey on Security and Privacy Methods of Cloud Computing Environment. Journal of Next Generation Technology (ISSN: 2583- 021X), 2(1).

[22] Venkateswara Reddy, B., Vellela, S. S., Sk, K. B., Roja, D., Yakubreddy, K., & Rao, M. V. Conceptual Hierarchies for Efficient Query Results Navigation. International Journal of All Research Education and Scientific Methods (IJARESM), ISSN, 2455-6211.

[23] Sk, K. B., Vellela, S. S., Yakubreddy, K., & Rao, M. V. (2023). Novel and Secure Protocol for Trusted Wireless Ad-hoc Network Creation. Khader Basha Sk, Venkateswara Reddy B, Sai Srinivas Vellela, Kancharakunt Yakub Reddy, M Venkateswara Rao, Novel and Secure Protocol for Trusted Wireless Ad-hoc Network Creation, 10(3).

[24] Sk, K. B., & Vellela, S. S. (2019). Diamond Search by Using Block Matching Algorithm. DIAMOND SEARCH BY USING BLOCK MATCHING ALGORITHM", International Journal of Emerging Technologies and Innovative Research (www. jetir. org), ISSN, 2349-5162.

[25] Rao, D. M. V., Vellela, S. S., Sk, K. B., &Dalavai, L. (2023). Stematic Review on Software Application Under-distributed Denial of Service Attacks for Group Website. DogoRangsang Research Journal, UGC Care Group I Journal, 13.

[26] S. S. Priya, S. Srinivas Vellela, V. R. B, S. Javvadi, K. B. Sk and R. D, "Design And Implementation of An Integrated IOT Blockchain Framework for Drone Communication," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205659.

[27] N. Vullam, K. Yakubreddy, S. S. Vellela, K. Basha Sk, V. R. B and S. Santhi Priya, "Prediction And Analysis Using A Hybrid Model For Stock Market," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-5, doi: 10.1109/CONIT59222.2023.10205638.