Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Hybrid GNN- PDP Model for Leaf Disease Detection

Meenalochini. M[1], Dr. P. Amudha[2*]
[1]*Research scholar, Department Of CSE, School Of Engineering, Avinashilingam Institute For Home Science And Higher Education For Women, Coimbatore, India*
meenalochinicse@gmail.com
[2*]*Professor, Department Of CSE, School Of Engineering, Avinashilingam Institute For Home Science And Higher Education For Women, Coimbatore, India*
amudha_cse@avinuty.ac.in

ARTICLE INFO

Received: 29 Dec 2024
Revised: 15 Feb 2025
Accepted: 24 Feb 2025

ABSTRACT

Introduction: As an important vegetable crop, cauliflower (Brassica oleracea var. botrytis) is afflicted by many diseases that significantly reduce yield and quality. Although advancements have taken place in agricultural technologies, a disease detection system based on deep learning yet remains to be developed for the specific purpose of cauliflower.

Objectives: The specific aim of this research is to develop an automated expert system that integrates the Internet of Things feature for cauliflower disease early detection with the use of deep learning. Such a system could facilitate farmers in identifying infections by employing mobile or portable-based devices to relay the results.

Methods: The mobile and IoT-enabled devices have aided in the collection of 750 images of cauliflower plants. The Cat Swarm Optimization (CSO) technique was utilized to segment the affected areas visible in the given images. Feature extraction methods were explicitly engaged to acquire the statistical and co-occurrence features of the images. The study was limited to four diseases—in particular, bacterial soft rot, white rust, black rot, and downy mildew. The performance of the proposed model, termed GNN-PDP (Graph Neural Network-based Plant Disease Prediction), was observed against CNN, DNN, Random Forest, Decision Tree, LDA, and PCA classifiers.

Results: The GNN-PDP model achieved almost 89% accurate output, outperforming other models in disease classification.

Conclusion: An effective solution for real-time detection of cauliflower diseases, allowing early intervention and promoting sustainable agriculture, is offered by the proposed GNN-based system.

Keywords: Graph Neural Network, Plant Disease Prediction, Deep Learning, Internet of Things.

INTRODUCTION

Cauliflower has become highly popular and one of the most broadly cultivated crops in agricultural fields due to its nutritious and economically important aspects. Cauliflower is rich in fiber, B vitamins, and cancer-fighting phytonutrients, which impart a productive benefit to human health. Like many other crops, [it] is susceptible to a lot of diseases which immensely affect yield, quality, and profits from the economically valued crop. Early detection and treatment are of utmost importance in maintaining these crops for sustainable production. Traditional manual inspection methods are generally time-consuming and inaccurate, often requiring expert knowledge which may not be available to farmers, especially in rural areas. Given all these challenges, there is an increasing demand for advancements in technology that will help farmers in rapid and accurate diagnosis of plant disease for reducing crop losses and enhancing productivity.

This proposal hence presents an IoT-enabled deep learning framework for cauliflower disease detection specifically. This system consists of images taken by mobile phones or other portable IoT devices from cauliflower plants and

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

inputted for analysis. The system detects the four possible diseases and gives an instant diagnosis report - bacterial soft rot, white rust, black rot, and downy mildew.

A Graph Neural Network (GNN) based approach has been adopted for processing and analyzing the images which will distinguish healthy from infected regions. More sophisticated image processing techniques will extract relevant statistical and co-occurrence features from the segmented regions. These features will be conducted as input to the disease classification model which will enable the exact disease kind to be identified and inform timely decision-making for farmers.

BACKGROUND STUDY OF PLANT DISEASE DETECTION AND PREDICTION MODELS

Deep learning techniques have emerged as key players for plant disease detection in diverse crops, as revealed by recent studies. In the methods developed, autoencoders integrated with CNNs are used for diagnosis through leaf features [1], while CNN-based systems find utility in the early detection of rice crops [2] or, for instance, the assessment of nitrogen deficiency [3]. In the disease prediction of pearl millet, a deep-learning-and-IoT-integration-based architecture "Custom-Net" has been developed [4]. Spatial RNNs have been used for crop disease forecasting [5]. Autonomously working systems are usually classified according to disease-specific features [6], while some approaches integrate raw features with CNNs into a more fine detection framework [7]. Probabilistic learning models such as naïve Bayes are used to point out the misclassification done by the classifier [8], and the integration of cluster-based fuzzy C-means with an improved LSTM paradigm has outperformed the traditional models on this task [9].

Fuzzy inference was also developed to assist farmers who speak Urdu to identify disease [10], while ANN methods have sometimes achieved higher efficiency for cauliflower disease identification [11]. Deep transfer learning with Agri-ImageNet has also benefitted in real-life plant disease identification. Frame variance modeling, adaptive deep learning techniques [12], and CFD-based temperature analysis in controlled environments like plant factories [13] are among future research directions in this domain. Furthermore, studies have pointed out the significance of weight optimization in LSTM models and the inception of IoT-based water quality monitoring systems. Therefore, efficient integration of deep learning goals, GNNs, and IoT will keep on enhancing the accuracy and efficiency of such plant disease detection systems.

PROPOSED MODEL

Figure 1 portrays the overall architecture of the proposed GNN-PDP (Graph Neural Network and Plant Disease Prediction) model for detecting diseases in cauliflower plants. The process starts with collecting 760 images, 660 taken locally under different light and field conditions, and 100 sourced from online. The obtained images undergo various preprocessing methods for quality enhancement and consistency. The preprocessing methods included contrast enhancement, histogram equalization for balancing intensity distribution, resizing all images to a common size of 300×300 pixels, and converting from RGB to Lxy color space for better perceptual uniformity.

Following preprocessing, K-means clustering (with k = 3) is used to segment the areas of interest in the images, helping to distinguish between possibly diseased and healthy tissues. The detection of affected regions is made much more precise after the segmentation phase with the implementation of Cat Swarm Optimization (CSO). The CSO is fundamental in offering even better details of which regions related to the disease are best worked on and shrinks the image space to only the most relevant features by imitating the cat behavior searching for optimum position.

Once segmentation is completed, texture-based features are constructed using Local Binary Patterns (LBP), and measure the micro-texture of the leaf surface and the diseased area. Together with statistical features (mean, variance, and skewness), and co-occurrence-based features from the Gray-Level Co-occurrence Matrix (GLCM) that assess the spatial relationship of pixel intensities, serving as chief measures for diseased patterns. Hence, these defined features constitute a solid representation for every image served in the classification.

Now the bipartite Graph Neural Network (GNN) comes into play, excelling in modeling relationships between features and diseases. The Graph Convolutional Layer of the network captures the underlying structural patterns and updates the nodes' representations based on adjacency relationships using learned weight matrices and activation functions that maintain local and global contextual information. The rich feature embeddings from the GNN allow

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

the model to discern small visual differences between diseases, thereby improving its accuracy in classification. This is represented as:

$$H^{(l+1)} = \sigma\left(W^{(l)}H^{(l)}A\right) \tag{1}$$

where W stands for weight matrix, A refers adjacency matrix, and σ represents activation function.

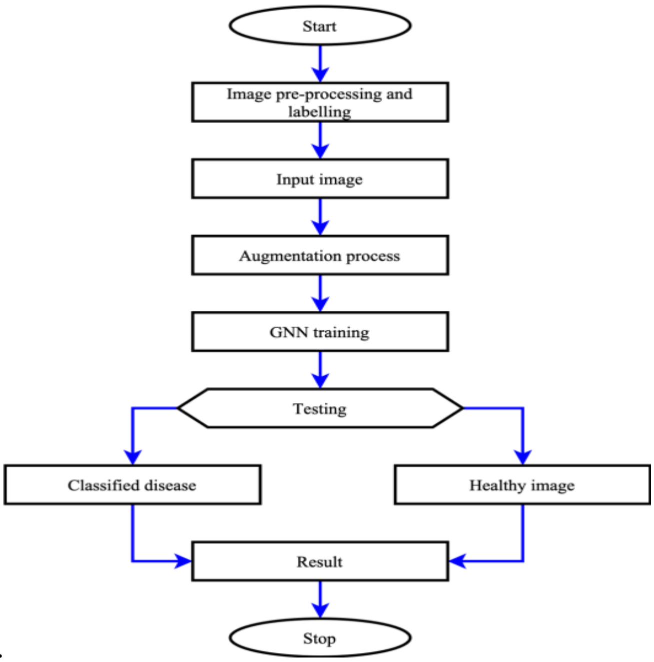


Figure 1: GNN-based framework

In the second step of the GNN-PDP framework, data is aggregated from nearby nodes in the graph structure. Here, the nodes represent either a feature or an image segment, while edges define the relationships or similarities between them. Aggregation is one of the key operations in Graph Neural Networks (GNNs), whereby the information coming from nearby nodes of a node is gathered and combined to enhance the node representation. This allows the model to not only learn from local features but also get a bigger picture from where it stands in the structural context regarding the disease-related patterns in the image:

$$H^{(l+1)} = \sigma\left(W_1H^{(l)} + W_2H^{(l)}A^2\right) \tag{2}$$

The next step is pooling, which is an important aspect of the GNN-PDP architecture to rank, filter, and condense the number of graph nodes while keeping the most meaningful and discriminative features. Pooling in the graph neural networks, much like in CNNs, is a pooling operation suitably modified to suit applications with non-Euclidean data. In this step, the network samples on a subset of nodes by their importance, based on some scoring function or learned weights, so that the inclusion in the final classification is only by the pertinent features.

This ranking mechanism ensures the elimination of redundancy while suppressing noise coming from insignificant nodes and retains the necessary topological structure of the input graph. At the same time, it makes the graph more compact, thus improving computational efficiency without any compromise on performance.

The filtered tensor, parameter Z, represents the enriched and compressed feature representation of the input graph. This tensor then gets conveyed through a set of convolutional layers, resulting in yet more refined features through the learned filters that capture high-level disease characteristics. These convolutional layers are then followed by a set of fully connected (dense) layers that act as interpreters of the extracted features and thus make the final decision in classification:

$$Z = ReLU\left(WZ^{(l)} + b\right) \tag{3}$$

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

where b indicates the bias term. Then softmax classifies the disease labels:

$$P(y_i|Z) = \frac{e^{Z_i}}{\sum j e^{Z_j}} \tag{4}$$

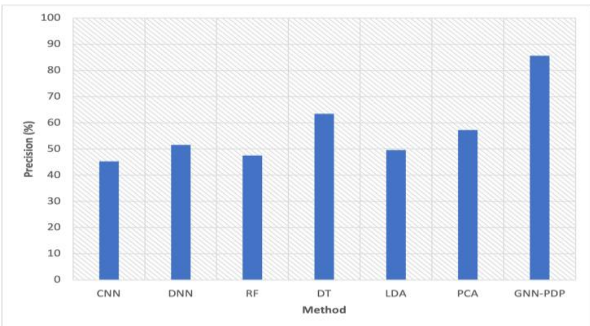
Inside the GNN-PDP framework, dropout layers are deployed to prevent overfitting by randomly dropping certain neurons and their connections during training, forcing a robust learning of features and a generalization. This maintains that the model does not become too dependent on specific features and performs rather well when exposed to previously unseen data. Post dropout, the model employs cross-entropy loss function to improve the accuracy in classification by measuring the difference between predicted probabilities and actual class labels. This loss function is ideal for multi-class problems: cauliflower diseases such as black rot, white rust, and others. The loss gradient is backpropagated for the updating of the network weights for improved prediction accuracy. Finally, it sends the predicted disease labels to an IoT-based interface such as a mobile application so that farmers can get real-time feedback for taking actions toward disease and crop health monitoring:

$$L = -\sum_I y_i \log P(y_i|Z) \tag{5}$$

The predictions are then sent to the IoT-based mobile app for real-time monitoring of plant health for farmers and agricultural professionals.

COMPUTATIONAL ANALYSIS AND COMPARATIVE ASSESSMENT

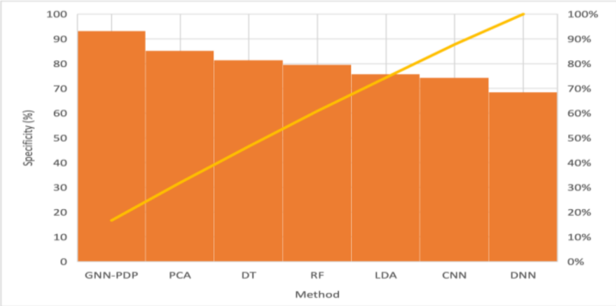
The GNN-PDP model has been implemented in a setup relying on Ubuntu 18.04, equipped with Intel Core i7 9820X processor, having 128GB RAM, and an NVIDIA GeForce GPU, using PyTorch and CUDA 9.0 for deep learning. Batch sizes were 16 and 8 for learning and testing, respectively, with training capped at 50 iterations and at a learning rate of 0.0001. Figure 2 shows the overall performance metrics to evaluate this. Subject this model to PCA, LDA, DT, RF, DNN, and CNN as references (Fig. (a)), and it gives the most accurate results of all, courtesy of its inclusion of GNN and CSO. Such efficiency was further proved in specificity and sensitivity analyses (Fig.(b), Fig. (c)). The comparison between True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), and False Negative Rate (FNR) (Fig. (d), Fig. (e)) improved disease prediction .The classifying accuracy analysis (Fig.(f)) shows and confirms that GNN-PDP is better than all classical models of classification due to its real-time monitoring and much more enhanced feature extraction through CSO when integrated with IoT and preservation of accurate classification and detection of diseases affecting cauliflower.Indeed, this integrated system enhances diagnostic precision, as well as gives the timing and detail at which farmers need to act. In fine, it greatly promotes sustainable farming and reduces losses after crop damage. Scalable thus makes it adaptable even to many crops monitoring or smart agriculture applications.



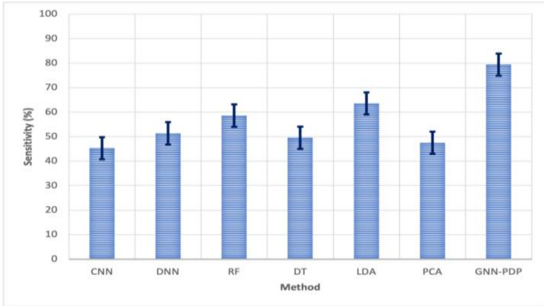
(a): Accuracy

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

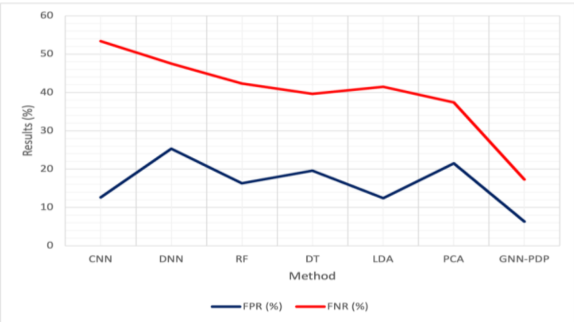
Research Article



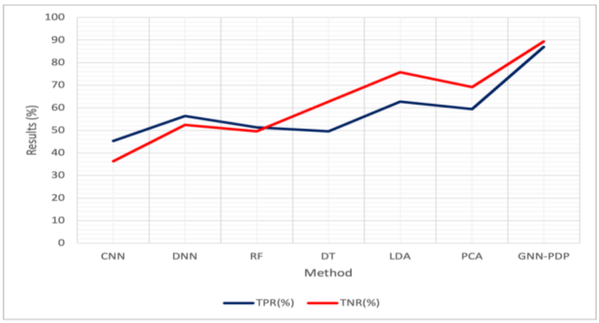
(b): Specificity



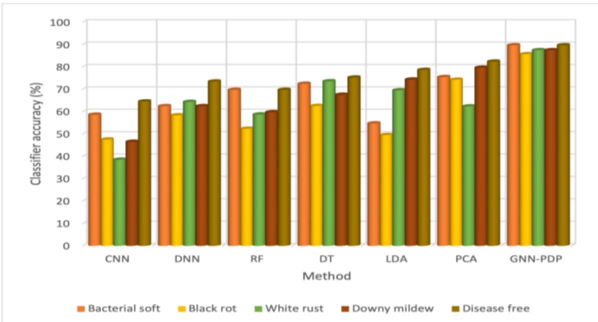
(c): Sensitivity



(d): FPR and FNR



(e): TPR and TNR



(f): Classification accuracy

Figure 2: Performance Evaluation

CONCLUSION

This study presents a GNN-PDP model that combines IoT technologies with deep learning methods. The research established an intelligent system utilizing deep learning to identify diseases in cauliflower, a significant and extensively produced vegetable. Experimental study employed a total of 750 photographs. K-means classification technique was utilized to delineate disease-affected areas in the acquired images. Two separate feature sets were utilized to diagnose cauliflower illnesses. Image processing methodologies were employed to derive features from photos. Following feature extraction, the components were classified into six distinct categories, with GNN classification attaining highest performance, exhibiting an average accuracy of 89%. Applying DL algorithms to an extensive image dataset may improve the efficacy of cauliflower illness identification; nevertheless, additional research is required. The primary research objective is to create a sophisticated intelligent system that can autonomously identify diseases in diverse crops.

Acknowledgements

The editors and reviewers are commended for their contributions by the author. This study did not receive any dedicated funding from public, private, or nonprofit organizations.

Contribution of Authors

The work was equally contributed to by all authors.

Competing Interests

The authors declare that they have no conflicts of interest.

REFRENCES

[1] Huang, Q., Yamada, M., Tian, Y., Singh, D., & Chang, Y. (2023). GraphLIME: Local interpretable model explanations for graph neural networks. *IEEE Transactions on Knowledge and Data Engineering*, *35*(7), 6968–6972. doi:10.1109/tkde.2022.3187455

[2] Khamparia, A., Saini, G., Gupta, D., Khanna, A., Tiwari, S., & de Albuquerque, V. H. C. (2020). Seasonal crops disease prediction and classification using deep convolutional encoder network. *Circuits, Systems, and Signal Processing*, *39*(2), 818–836. doi:10.1007/s00034-019- 01041-0

[3] Sharma, R., Das, S., Gourisaria, M. K., Rautaray, S. S., & Pandey, M. (2020). A model for prediction of paddy crop disease using CNN. In *Advances in Intelligent Systems and Computing* (pp. 533–543). Singapore: Springer Singapore.

[4] Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Nitrogen deficiency prediction of rice crop based on convolutional neural network. *Journal of Ambient Intelligence and Humanized Computing*, *11*(11), 5703–5711. doi:10.1007/s12652-020-01938-8

[5] Kundu, N., Rani, G., Dhaka, V. S., Gupta, K., Nayak, S. C., Verma, S., … Woźniak, M. (2021). IoT and interpretable machine learning based framework for disease prediction in pearl millet. *Sensors (Basel, Switzerland)*, *21*(16), 5386. doi:10.3390/s21165386

[6] Xu, W., Wang, Q., & Chen, R. (2018). Spatio-temporal prediction of crop disease severity for agricultural emergency management based on recurrent neural networks. *GeoInformatica*, *22*(2), 363–381. doi:10.1007/s10707-017-0314-1

[7] Das, S., & Sengupta, S. (2020). Feature extraction and disease prediction from paddy crops using data mining techniques. In *Computational Intelligence in Pattern Recognition* (pp. 155–163). Singapore: Springer Singapore. doi:10.1007/978-981-15-2449-3_13

[8] Picon, A., Seitz, M., Alvarez-Gila, A., Mohnke, P., Ortiz-Barredo, A., & Echazarra, J. (2019). Crop conditional Convolutional Neural Networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions. *Computers and Electronics in Agriculture*, *167*(105093), 105093. doi:10.1016/j.compag.2019.105093

[9] Hernández, S., & López, J. L. (2020). Uncertainty quantification for plant disease detection using Bayesian deep learning. *Applied Soft Computing*, *96*(106597), 106597. doi:10.1016/j.asoc.2020.106597

[10] Farooqui, Nafees Akhter, Mishra, A. K., & Mehra, R. (2022). Concatenated deep features with modified LSTM for enhanced crop disease classification. *International Journal of Intelligent Robotics and Applications*. doi:10.1007/s41315-022-00258-8

[11] Toseef, M., & Khan, M. J. (2018). An intelligent mobile application for diagnosis of crop diseases in Pakistan using fuzzy inference system. *Computers and Electronics in Agriculture*, *153*, 1–11. doi:10.1016/j.compag.2018.07.034

[12] Ratna, C. L., & Y Srinivas. (2023). A hybrid of NHPP and Generalized Gaussian Mixture Model: A combinatorial approach for background elimination. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *34*(1), 1–14. doi:10.37934/araset.34.1.114

[13] Sabudin, S., Zulkarnaen, M. E., Mohammed, A. N., & Mohd Faizal Bin Mohideen Batcha. (2022). Numerical investigation of temperature distribution in a container-type plant factory. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *28*(2), 90–101. doi:10.37934/araset.28.2.90101