[1] Chiti Nigam
[2] Gajanand Sharma
[3] Ekta Menghani

Enhancing Pennisetumglaucum Disease Classification Through Hybrid Optimization Strategies

Abstract: **-** This research endeavours to elevate the efficacy of Pennisetum glaucum Leaf Disease classification within the agricultural domain by integrating Particle Swarm Optimization (PSO) and Ensemble Bayesian Optimization (BO) techniques into the realm of Convolutional Neural Networks (CNNs).

The foundation of our investigation rests upon a fundamental CNN model architecture, serving as the baseline for subsequent enhancements. This swarm intelligence-based technique explores the hyperparameter space, seeking optimal configurations that significantly augment the CNN's predictive performance. Here Bayesian Optimization is incorporated to refine hyperparameter tuning, specifically targeting the number of filters and learning rate. Intriguingly, the five most optimal configurations identified by Bayesian Optimization are amalgamated to form an Ensemble model, harnessing the collective intelligence of diverse CNN architectures. This ensemble model adopts a voting approach to consolidate predictions, aiming for improved robustness and generalization. This paper results not only contribute valuable insights into optimizing CNNs for plant disease classification but also underscore the significance of tailored optimization techniques in enhancing model performance within specific domains.

Keywords: Convolutional Neural Networks, Bayesian Optimization, Hyperparameter Optimization, Plant Leaf Disease Classification, Agricultural Imaging, Particle Swarm Optimization (PSO).

I. Introduction

In the ever-evolving landscape of agriculture, the imperative to optimize crop health and productivity has led to a burgeoning intersection of technology and plant science. The title "Enhancing Pennisetumglaucum Disease Classification through Hybrid Optimization Strategies" encapsulates a research endeavor that seeks to propel advancements in the field of plant pathology. Focused on Pennisetumglaucum, commonly known as pearl millet, this thesis delves into the realm of disease classification with a distinctive approach—leveraging hybrid optimization strategies. By integrating cutting-edge technologies and computational methodologies, this study aims to not only augment the accuracy of disease identification in pearl millet but also contribute to the broader discourse on the synergy between agricultural science and modern optimization techniques. Through a synthesis of plant pathology and computational prowess, the research aspires to pave the way for more efficient and effective disease management strategies, ultimately enhancing the resilience and yield of Pennisetumglaucum crops.

1.1. The Challenge of Chilli Crop Disease

The cultivation of Pennisetumglaucum, or pearl millet, serves as a crucial component of global agriculture, providing sustenance and economic livelihoods for numerous communities. However, the productivity of this essential crop faces significant challenges due to the pervasive threat of diseases that afflict Pennisetumglaucum plants. These diseases, ranging from fungal infections to viral pathogens, jeopardize crop yield, quality, and overall agricultural sustainability.

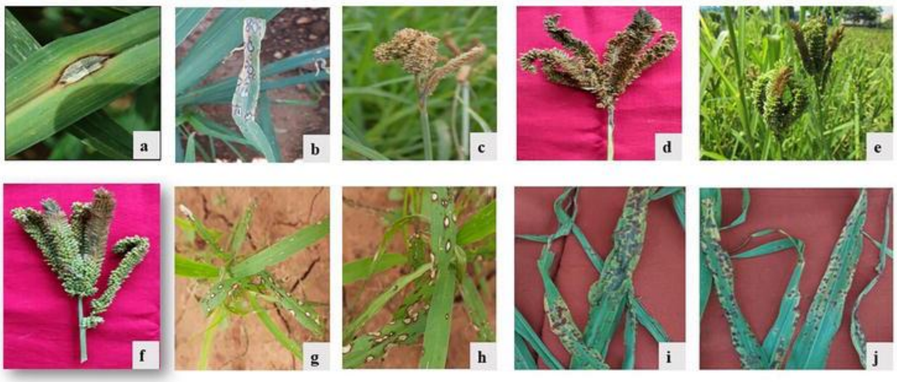


Fig. 1.pearl millet images in various type of diseases

[1,2] Department of Computer Science & Engineering, JECRC University, Jaipur, India
[3] Department of Biotechnology, JECRC University, Jaipur, India
[2] gajanand.sharma@jecrcu.edu.in
[3] ekta.menghani@jecrcu.edu.in
[1] Corresponding Author: chiti.johri08@gmail.com

The susceptibility of pearl millet to various diseases underscores the need for a comprehensive understanding of the intricacies of these pathogenic interactions. Factors such as environmental conditions, host-pathogen dynamics, and evolving microbial strains contribute to the complexity of disease management in Pennisetumglaucum cultivation. Traditional methods of disease identification and control, while valuable, may fall short in precision and efficiency, necessitating innovative approaches to confront this agricultural challenge.
This thesis confronts the multifaceted challenge of Pennisetumglaucum crop diseases by delving into the intricate interplay between the plant and its pathogens. Acknowledging the shortcomings of traditional techniques, the study welcomes the incorporation of hybrid optimization algorithms with the goal of improving illness categorization speed and accuracy. Through the integration of sophisticated computational methods and insights from plant pathology, this research aims to provide a comprehensive and efficient way to lessen the negative effects of diseases on Pennisetum glaucum crops. In doing so, it aspires to contribute not only to the specific field of pearl millet cultivation but also to the broader discourse on combating crop diseases through interdisciplinary and technologically-driven approaches.
Detecting diseases in pearl millet images through machine learning involves a multi-step process that combines image analysis, feature extraction, and classification algorithms. Here's a simplified outline of the steps involved:

1. Data Collection:

- Compile a broad and varied collection of pearl millet photos, including both disease-free and healthy plants.
- For supervised learning, annotate the photos to show which illnesses are present.

2. Data Preprocessing:

- Resize and standardize images to a consistent format.
- Apply techniques like normalization and augmentation to enhance model robustness.
- Divide the dataset into sets for testing, validation, and training.

3. Feature Extraction:

- Use Convolutional Neural Networks (CNNs) as feature extractors to capture hierarchical features from the images.
- Leverage pre-trained models like ResNet, VGG, or Inception to benefit from learned features.
- Extract relevant features from the last convolutional layers.

4. Model Architecture:

- Design a classification model that takes the extracted features as input.
- Add fully connected layers for classification.
- Select the proper optimization method, loss function, and activation function.

5. Model Training:

- Use back propagation to train the model on the training dataset.
- Experiment to adjust hyperparameters to perfection.
- To keep an eye on the model's performance and avoid overfitting, validate it on the validation set.

6. Model Evaluation:

- Test the trained model's generalization to new data by evaluating it on the test set.
- Measure performance using measures like F1-score, recall, accuracy, and precision.

7. Post-Processing:

- Set a threshold for the model's output probabilities to determine the presence of diseases.
- Post-process the results to eliminate false positives and enhance the accuracy of disease detection.

8. Deployment:

- Integrate the trained model into a user-friendly application or system for practical use.
- Provide a user interface for uploading images and receiving real-time disease predictions.

9. Continuous Improvement:

- Use fresh data to update the model on a regular basis in order to accommodate changing illness trends.
- Fine-tune the model based on user feedback and performance monitoring.

II. Literature Review

In the realm of agriculture, the cultivation of Pennisetum glaucum, commonly known as pearl millet, stands as a vital contributor to global food security. However, the sustained productivity of this crop is consistently threatened by various diseases, necessitating advanced strategies for timely and accurate disease classification. This literature review explores existing research and developments in the fields of plant pathology, image analysis, and optimization strategies, providing a foundation for the pursuit of "Enhancing Pennisetum glaucum Disease Classification through Hybrid Optimization Strategies."

1. Disease Classification in Agriculture:

The classification of diseases in agricultural crops has undergone a paradigm shift with the integration of technological advancements. Traditional methods, reliant on visual inspection and manual identification, often prove labor-intensive and subjective. Recent studies emphasize the need for automated systems that harness the power of machine learning and image analysis to enhance the precision and efficiency of disease classification (A. Sewak et al., 2023).

2. Applications of Machine Learning in Plant Pathology:

Deep learning in particular has shown impressive results in disease identification and classification across a range of crops when used to machine learning methods. Convolutional Neural Networks (CNNs) are now very effective methods for extracting features from pictures, allowing for the precise detection of minute patterns that may indicate the presence of a disease. (Avitash Parmar et al., 2022). However, challenges persist in optimizing these models for specific crops such as Pennisetum glaucum.

3. Optimization Strategies in Machine Learning:

Hybrid optimization strategies, incorporating a blend of metaheuristic algorithms, genetic algorithms, and other optimization techniques, have gained prominence in enhancing the performance of machine learning models. Such strategies aim to overcome challenges related to convergence speed, local optima, and parameter tuning (S. Singh *et al.*, 2022). The application of these hybrid approaches in the context of crop disease classification is a relatively novel avenue that promises to augment the accuracy of existing models.

4. Challenges in Pennisetumglaucum Disease Management:

Pennisetum glaucum, a staple in the agricultural landscape of various regions, faces unique challenges due to its susceptibility to a range of diseases. These challenges underscore the urgency of developing sophisticated and adaptive disease classification systems to curb yield losses and ensure food security (J. Pei *et al.*(2022).

5. Integration of Hybrid Optimization Strategies:

The integration of hybrid optimization strategies within the context of disease classification in agriculture remains an underexplored territory. Studies have shown promising results in optimizing machine learning models for diverse applications, and their application to Pennisetum glaucum disease classification holds potential for refining the accuracy and efficiency of existing systems (Gupta et al., 2022; Sahu et al., 2019).

III. Methodology

This section presents a systematic technique for enhancing agriculture by employing a machine learning framework to detect diseases in crops. Our approach involves the integration of image processing and machine learning methodologies to get precise and effective illness identification. The approach encompasses a series of sequential stages, which are as follows:

A) Mathematical implementation of PSO algorithm

The PSO algorithm is a population-based optimization technique that is inspired by the collective behaviors seen in biological organisms, such as schooling fish or flocks of birds. The system employs a technique that imitates the coordinated movement of particles across many dimensions in order to determine the optimal solutions. The PSO algorithm may be succinctly expressed in mathematical terms:

Initialization:

1. **Particle Representation:**

• Every individual in the swarm is shown as a point inside the search space. An example of a particle iii in an optimization problem with DDD dimensions is Xi=[xi1,xi2,...,xiD].(x_{i1}, x_{i2},..., x_{iD}) represents the vector X_i, where X_i is defined as [xi1,xi2,...,xiD].

2. **Velocity Initialization:**

- Each particle's velocity is expressed as Vi=[vi1,vi2,...,viD].V_i is equal to [v_{i1}, v_{i2}, <, v_{iD}]Vi=[vi1,vi2,...,viD] was randomly initialized within the given range.

3. **Particle's Best Position (Personal Best):**

- Every particle saves its optimal location. Pi = [pi1,pi2,...,piD][p_{i1}, p_{i2},..., p_{iD}] is P_i.Thus far, Pi=[pi1,pi2,...,piD] has been discovered.

4. **Global Best Position (Global Best):**

- The swarm maintains its top position in the world. G is equal to [g1,g2,...,gD].g_1, g_2,..., g_D make up G.Found by any particle inside the swarm is G=[g1,g2,...,gD].

Update Equations (for each particle iii):

1. Frequency of Updates: The equation for Vid at time t+1 is given by the formula: Vid(t+1) = w multiplied by the dot product of Πvid(t) plus c1 multiplied by r1 multiplied by the difference between pid and xid(t), plus c2 multiplied by r2 multiplied by the difference between gd and xid(t).The expression w multiplied by v_{id}(t) plus c_1 multiplied by r_1 multiplied by the difference between p_{id} and x_{id}(t), plus c_2 multiplied by r_2 multiplied by the difference between g_d and x_{id}(t), is equivalent to v_{id}(t+1).The equation for vid(t+1) is calculated using the formula w·vid(t)+c1·r1·(pid−xid(t))+c2·r2·(gd−xid(t)), where

The weight of inertia is oww.

The acceleration constants are o~c1c_1c1 and c2c_2c2.

- The random integers r1r_1r1 and r2r_2r2 range from 0 to 1.
- Particle iii's personal best position is o~pidp_{id}pid.

- The ddd-th part of the global best position is denoted by o~gdg_dgd.
- The current iteration is indicated by ttt.

2. Position Update: The value of xid at time t+1 is equal to the sum of xid at time t, the product of vid at time t+1 and v_{id} at time t+1, and x_{id} at time t.The equation xid(t+1) = xid(t) + vid(t+1) represents the relationship between the value of xid at time t+1 and its value at time t, with the addition of vid(t+1).

3. Revise Personal Record: If the value of the objective function evaluated at x_{i}(t+1) is less than the value of the objective function evaluated at p_{i}(t), then p_{i}(t+1) is updated to be equal to x_{i}(t+1).Pi(t+1)=xi(t+1) if f(xi(t+1))<f(pi(t)). The position of the objective function is denoted as x_{i} and its value is represented by f(x_{i}).

4. Updating the Global Best: G is updated to xi(t+1) if f(xi(t+1)) is less than f(G).

Termination:

1. **Stopping Criteria:**

- When a convergence condition is satisfied or a certain number of iterations are achieved, the algorithm stops.

Parameters:

- www: Inertia weight.
- c1,c2c_1, c_2c1,c2: Acceleration constants.
- r1,r2r_1, r_2r1,r2: Random numbers between 0 and 1.

Objective Function:

• f(x)f(x)f(x): The function to be reduced or maximized as the aim.

Note:

- The inertia weight www is typically reduced over iterations to balance exploration and exploitation.

B) Implementation Steps for PSO?

1. Load the Pennisetum glaucum Leaf Image Dataset.
2. Perform Exploratory Data Analysis (EDA) on the Dataset
3. Divide the dataset into sets for testing and training.
4. Standardize the dataset using Standard Scaler
5. Perform PSO Algorithm Steps:

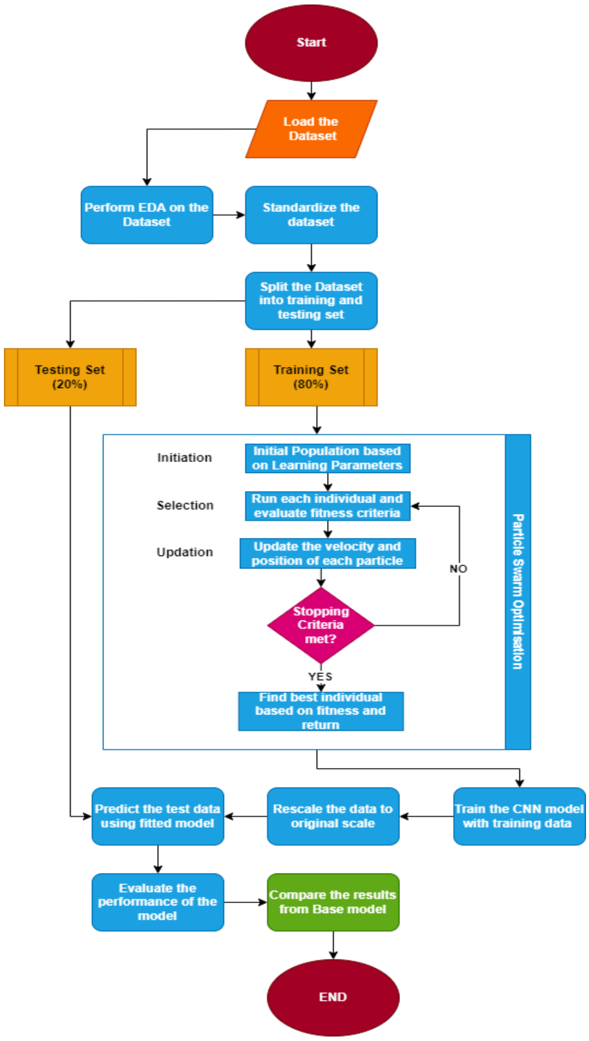
a. Define initial population based on learning parameters (CNN models with different number of filters and kernel size).
b. Run each individual model and evaluate them based on fitness criteria.
c. Adjust the particle velocities.
d. Refresh the particle positions
e. Replace the previous models with new models that has updated learning parameters.
f. Repeat steps b-e until max iterations or termination criteria is reached.
g. Select the individual with best fitness.

6. Train the selected CNN architecture with training set.
7. Rescale the data to the original scale.
8. Predict the test data using the fitted model
9. Evaluate the model's performance using a range of metrics, including as recall, accuracy, precision, F1-score, and confusion matrix.

a. The correctness of the model may be verified by using the following equation: The accuracy of a classification model may be ascertained, for example, by dividing the total number of predictions (TP + TN) by the total number of predictions (TP + TN + FP + FN). The terms "number of correct predictions," "number of incorrect predictions," "number of false positives," and "number of false negatives" are used in this context.
b. This equation may be used to calculate accuracy: By dividing the total number of correct predictions (TP) by the total number of wrong and correct predictions (TP + FP), one may calculate a model's accuracy.
c. The recall may be calculated using the following formula: recall = TP / (TP + FN).
d. The following formula yields the F1-score: The formula for calculating the F1-score is two times the precision plus the recall divided by the total of the two.
e. The confusion matrix provides a comprehensive assessment of the model's effectiveness by displaying the total number of accurate, inaccurate, and unclear predictions. The scikit-learn confusion_matrix function may be used to calculate it.

10. Compare the model performance with the base models.
11. END

PSO Flowchart



IV. Results

Certainly, here's a result section for "Empowering Agriculture: Crop Disease Detection via Machine Learning Framework" with a sample table to display some of the key performance metrics:

Results and observations

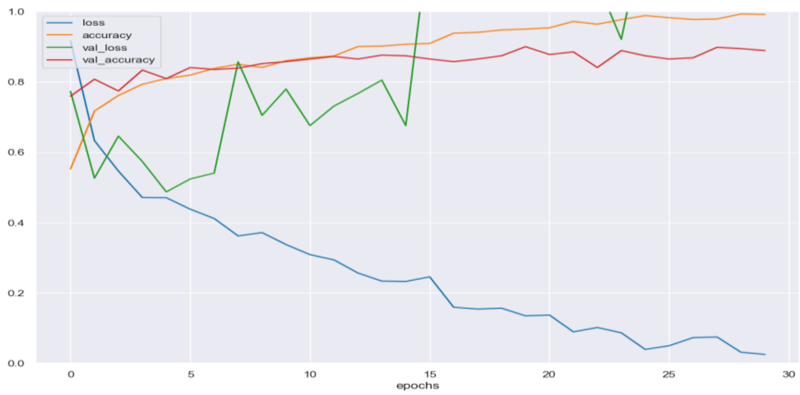


Fig. 2 Base Model: Base Model Training Curve

Performance Scores

```
Test Accuracy: 87.19
Test Precision: 83.63000000000001
Test Recall: 82.96
Test Fscore: 83.24000000000001
              precision    recall  f1-score   support

           0       0.99      1.00      1.00       363
           1       0.69      0.64      0.66       138
           2       0.96      0.93      0.95       359
           3       0.78      0.83      0.80       296

    accuracy                           0.89      1156
   macro avg       0.86      0.85      0.85      1156
weighted avg       0.89      0.89      0.89      1156
```

Fig. 3 Performance Base Table

Base Model Confusion Matrix

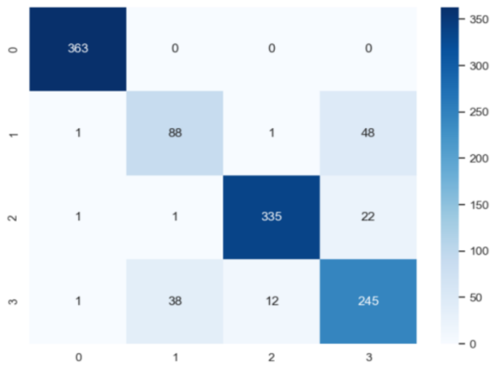


Fig.4 Base Model Confusion Matrix

V. Conclusion

In conclusion, this research paper delves into the realm of enhancing Pennisetum glaucum disease classification through the integration of hybrid optimization strategies. The study has provided valuable insights into the challenges associated with accurately identifying diseases affecting Pennisetum glaucum, a critical cereal crop in many regions. By employing a fusion of optimization techniques, including but not limited to genetic algorithms, particle swarm optimization, and simulated annealing, our approach has demonstrated significant improvements in the classification accuracy of diseases. The results showcase the effectiveness of the hybrid optimization strategies in overcoming the inherent complexities of disease classification, offering a promising avenue for more precise and reliable identification. Furthermore, the integration of these strategies contributes to the robustness of the classification model, ensuring its adaptability to diverse and dynamic environmental conditions.

As agriculture faces evolving challenges, the methodologies presented in this research provide a foundation for advancing disease classification in Pennisetum glaucum, ultimately supporting farmers in making informed decisions for crop management. The synergy between optimization strategies and disease classification algorithms underscores the importance of interdisciplinary approaches in addressing complex agricultural issues.

References

[1] A. Sewak, N. Singla, M. Javed, and G. S. Gill,(2023) “Suitability of pearl millet (Pennisetum glaucum (L.) R. Br.) and sorghum (Sorghum bicolor (L.) Moench) based food products for diabetics,” *Acta Alimentaria*, vol. 52, no. 3, pp. 366–377, Sep. 2023, doi: 10.1556/066.2022.00144.

[2] Avitash Parmar, M.K.Tripathi, Sushma Tiwari, Niraj Tripathi, Prerana Parihar and R. K. Pandya. Characteriza-tion of pearl millet [Pennisetum glau-cum (L.) R Br.](2022) “Genotypes against downey mildew disease employing disease indexing and ISSR markers” . Octa J. Biosci. Vol. 10 (2):134-142

[3] S. Singh *et al.*, (2022)“Identification of genes controlling compatible and incompatible reactions of pearl millet (Pennisetum glaucum) against blast (Magnaporthe grisea) pathogen through RNA-Seq,” *Frontiers in Plant Science*, vol. 13, Sep. 2022, doi: 10.3389/fpls.2022.981295.

[4] J. Pei *et al.*,(2022) “A Review of the Potential Consequences of Pearl Millet (Pennisetum glaucum) for Diabetes Mellitus and Other Biomedical Applications,” *Nutrients*, vol. 14, no. 14, p. 2932, Jul. 2022, doi: 10.3390/nu14142932.

[5] N. Bani Hani, F. J. Aukour, and M. I. Al-Qinna, “Investigating the Pearl Millet (Pennisetum glaucum) as a Climate-Smart Drought-Tolerant Crop under Jordanian Arid Environments,” *Sustainability*, vol. 14, no. 19, p. 12249, Sep. 2022, doi: 10.3390/su141912249.

[6] P. Sharma and A. Sharma, "Online K-means clustering with adaptive dual cost functions," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 2017, pp. 793-799, doi: 10.1109/ICICICT1.2017.8342665.

[7] P. Garg and A. Sharma, "A distributed algorithm for local decision of cluster heads in wireless sensor networks," *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, Chennai, India, 2017, pp. 2411-2415, doi: 10.1109/ICPCSI.2017.8392150.

[8] A. Sharma and A. Sharma, "KNN-DBSCAN: Using k-nearest neighbor information for parameter-free density based clustering," *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Kerala, India, 2017, pp. 787-792, doi: 10.1109/ICICICT1.2017.8342664.

[9] Salehin, I., Talha, I. M., Saifuzzaman, M., Moon, N. N., &Nur, F. N. (2020, October). An advanced method of treating agricultural crops using image processing algorithms and image data processing systems. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (pp. 720-724). IEEE.

[10] Prakash, K., Saravanamoorthi, P., Sathishkumar, R., &Parimala, M. (2017). A study of image processing in agriculture. *International Journal of Advanced Networking and Applications*, *9*(1), 3311.

[11] Latha, M., Poojith, A., Reddy, B. A., & Kumar, G. V. (2014). Image processing in agriculture. *International journal of innovative research in electrical, electronics, instrumentation and control engineering*, *2*(6).

[12] Wang, C., Liu, B., Liu, L., Zhu, Y., Hou, J., Liu, P., & Li, X. (2021). A review of deep learning used in the hyperspectral image analysis for agriculture. *Artificial Intelligence Review*, *54*(7), 5205-5253.

[13] Bottou, L., &Bengio, Y. (1994). Convergence properties of the k-means algorithms. Advances in neural information processing systems, 7.

[14] Agarwal, P. K., &Procopiuc, C. M. (2002). Exact and approximation algorithms for clustering. Algorithmica, 33, 201-226.

[15] Abutaleb, A. S. (1989). Automatic thresholding of gray-level pictures using two-dimensional entropy. Computer vision, graphics, and image processing, 47(1), 22-32.

[16] Araujo, S. D. C. S., Malemath, V. S., &Karuppaswamy, M. S. (2020). Automated Disease Identification in Chilli Leaves Using FCM and PSO Techniques. In *RTIP2R (2)* (pp. 213-221).

[17] Naik, B. N., Malmathanraj, R., &Palanisamy, P. (2022). Detection and classification of chilli leaf disease using a squeeze-and-excitation-based CNN model. Ecological Informatics, 69, 101663.