WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

AN ENHANCED IMAGE CLASSIFICATION MODEL BASED ON GRAPH CLASSIFICATION AND SUPERPIXEL-DERIVED CNN FEATURES FOR AGRICULTURAL DATASETS: A REVIEW

Dr. Gaurav Srivastava[1], Bindoo Lal[2]

[1] Guide Name
[2] Author Name
Author / Corresponding Author: Bindoo Lal
BANSAL INSTITUTE OF ENGINEERING & TECHNOLOGY APJ ABDUL KALAM TECHNICAL UNIVERSITY (AKTU)

ABSTRACT :

Agricultural image classification supports plant disease recognition, fruit maturity assessment, crop monitoring, and agricultural product grading. Convolutional Neural Networks (CNNs) learn hierarchical visual features from images, but conventional CNNs primarily process data on regular pixel grids and do not explicitly represent arbitrary relationships among image regions. Superpixel segmentation can transform an image into a compact set of coherent regions, while graph neural networks (GNNs) learn from relationships between these regions. This review examines the integration of superpixel segmentation, CNN-derived node features, and graph-level classification for agricultural image analysis. It discusses graph construction, feature representation, graph convolution, pooling, dataset design, evaluation metrics, robustness, computational cost, and reproducibility. A published 2025 [1] study has already demonstrated this general combination on four agricultural datasets, making it important for subsequent reviews to critically assess the wider evidence rather than present the same architecture as novel. Key unresolved issues include sensitivity to segmentation granularity, possible background bias, weak external validation, data leakage risks, and insufficient end-to-end efficiency reporting. Future research should prioritize adaptive multiscale graphs, field-level validation, uncertainty estimation, interpretable predictions, and reproducible experimental protocols. This manuscript is a critical review draft; a systematic database search and documented study-screening process must be completed before it is described as a systematic review.

Keywords: agricultural image classification; superpixels; convolutional neural networks; graph neural networks; graph convolutional networks; precision agriculture; plant disease detection; graph classification.

1. Introduction

Agriculture increasingly uses computer vision and artificial intelligence to support crop monitoring, disease diagnosis, fruit grading, and resource management. Manual inspection remains valuable but can be time-consuming and difficult to scale. Automated image classification offers a way to process images consistently, although its practical reliability depends on data quality, environmental conditions, and evaluation design.

CNNs such as VGG, ResNet, and EfficientNet learn visual representations of color, texture, shape, and spatial patterns. They have been widely applied to plant disease and fruit classification. However, conventional convolution operates on regular image grids and does not directly encode arbitrary relations among meaningful image regions. This can matter when a class depends on the distribution of lesions, the arrangement of healthy and damaged tissue, or spatial relationships among surface features.

Graph-based learning offers a complementary representation. In a graph, nodes represent image regions and edges represent relationships between regions. Superpixel segmentation groups neighboring pixels into approximately homogeneous regions, reducing the number of units to process. CNN-derived descriptors can be assigned to these regions, and a graph neural network can aggregate information across connected regions to make an image-level prediction.

A 2025 article by Nguyen et al. presented a CNN-feature and graph-convolution framework using superpixel-based representations for agricultural datasets. Its existence means that a new paper should not claim the basic combination of CNN features, superpixels, and graph classification as an original contribution. The value of a review instead lies in comparing related approaches, examining the strength of the evidence, identifying limitations, and defining useful research questions.

1.1 Review objectives

- To examine CNN feature extraction, superpixel segmentation, and graph-based image classification in agricultural applications.

- To compare graph construction strategies, node features, message passing, and graph-level pooling.
- To assess dataset design, evaluation metrics, external validity, robustness, and computational efficiency.
- To identify reproducibility gaps and propose a practical evaluation framework for future studies.

1.2 Review questions

- RQ1: How does superpixel segmentation affect the representation of agricultural images for graph learning?
- RQ2: What benefits and limitations arise from combining CNN features with graph classification?
- RQ3: How do node features, edge construction, and graph pooling affect classification?
- RQ4: How should robustness, generalization, efficiency, and reproducibility be assessed?
- RQ5: Which future directions are most promising for deployment in real agricultural environments?

2. Technical Background

2.1 CNN-based feature extraction

CNNs learn hierarchical representations through convolution, nonlinear activation, and downsampling. Early layers tend to capture low-level patterns, while deeper layers represent more abstract features. In agricultural images, these features may encode lesion texture, leaf morphology, fruit color, and surface defects. Transfer learning can be useful when labeled datasets are small. However, learned representations can also encode background, illumination, or camera-specific shortcuts; high performance on a controlled dataset therefore does not establish field robustness.

2.2 Superpixel segmentation

Superpixel algorithms group neighboring pixels according to appearance and spatial proximity. SLIC (Simple Linear Iterative Clustering) is a common method because it balances regional coherence and computational efficiency. Let an image segmentation be $S = \{S_1, S_2, \ldots, S_k\}$, where each S_i is a superpixel. The number of superpixels and the compactness parameter affect whether small lesions are preserved or merged with healthy tissue. Too few regions can erase diagnostic detail; too many increase graph size and may create redundant nodes.

2.3 Graph representation

An image can be represented as G = (V, E, X), where V is the set of nodes, E is the set of edges, and X contains node features. Each superpixel can become a node with CNN features, color statistics, texture descriptors, region size, centroid coordinates, or shape information. A Region Adjacency Graph (RAG) connects superpixels that share a boundary. Spatial k-nearest-neighbor graphs connect regions based on position, while feature-based graphs connect regions based on visual similarity. These constructions encode different assumptions and should be compared empirically.

2.4 Graph convolution and graph-level prediction

A standard graph convolutional update can be written as:

$$\mathbf{H}^{(l+1)} = \sigma(\tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2} \mathbf{H}^{(l)} \mathbf{W}^{(l)})$$

Here $\tilde{A}$ is the adjacency matrix with self-connections, $\tilde{D}$ is its degree matrix, $H^{(l)}$ is the node-feature matrix at layer l, $W^{(l)}$ is a trainable weight matrix, and σ is a nonlinear activation. Node embeddings can be combined using mean, sum, or attention pooling to obtain an image-level representation. Repeated message passing can cause over-smoothing, in which node representations become too similar. Graph depth and connectivity therefore require careful evaluation.

3. Review Methodology and Reporting Requirements

This draft uses a thematic critical-review structure. It does not claim to be an exhaustive systematic review because a complete database search, duplicate removal, screening log, and verified included-study count have not yet been documented. To support publication as a systematic review, the authors should search Scopus, Web of Science, IEEE Xplore, ScienceDirect, SpringerLink, and Google Scholar, subject to access.

3.1 Candidate search strategy

("agricultural image classification" OR "plant disease detection" OR "fruit ripeness classification") AND ("superpixel" OR "region adjacency graph") AND ("graph neural network" OR "graph convolutional network" OR "CNN-GCN")

3.2 Eligibility criteria to apply

Include peer-reviewed studies that investigate agricultural image classification or directly relevant methods for CNN-based features, superpixel representations, graph construction, or graph learning. Record the agricultural task, dataset, split strategy, model, metrics, code availability, and external validation. Exclude duplicate records, unrelated applications, and papers that lack sufficient methodological detail; document each exclusion category.

3.3 Transparency

Report the search date, exact database queries, records retrieved, duplicates removed, title/abstract screening, full-text screening, reasons for exclusion, and final included studies. Use the PRISMA 2020 checklist and flow diagram if the work is conducted and reported as a systematic review. Do not invent screening counts or imply that a search was performed when it was not.

4. Comparative Analysis of Approaches

4.1 Conventional CNN classifiers

CNNs provide strong learned visual features and mature training pipelines. Their weaknesses include possible reliance on spurious background cues and performance degradation under shifts in lighting, crop variety, camera, or acquisition environment. Strong CNN baselines are necessary to determine whether graph modeling adds measurable value.

4.2 Superpixel-based graph classification

Superpixels provide a compact region-level representation, while graph edges describe local or nonlocal relationships. This can encode spatial context beyond a flat feature vector. However, segmentation errors may remove important lesions, and noisy edges can propagate irrelevant information. Graph classification should therefore be treated as a hypothesis to test, not an automatic improvement over CNNs.

4.3 Hybrid CNN–GCN models

Hybrid models use CNNs to extract appearance features and graph networks to aggregate contextual information among superpixels. Nguyen et al. (2025) report a framework evaluated on Tomato Leaf Disease, Dragon Fruit, Tomato Ripeness, and Dragon Fruit and Leaf datasets. Their reported accuracies are 96.57%, 99.63%, 95.19%, and 90.00%, respectively. These are results from that particular study and protocol; they should not be treated as independently reproduced or generalized to all agricultural environments.

4.4 Related graph and geometric representations

Work on geometric superpixel representations and multiscale wavelet superpixels provides useful ideas for graph construction and region representation. Such work can inform agricultural model design, but results from general image benchmarks should not be presented as direct evidence of agricultural effectiveness without agricultural evaluation.

Component	**Potential benefit**	**Key limitation**
CNN feature extraction	Rich visual descriptors	Can encode background and acquisition shortcuts
SLIC superpixels	Compact region representation	Sensitive to granularity and compactness
Region adjacency graph	Preserves local spatial adjacency	May miss nonlocal relationships
Spatial/feature k-NN	Adds spatial or appearance-based links	Can introduce noisy edges
GCN message passing	Aggregates regional context	Over-smoothing and noisy propagation
Graph pooling	Produces an image-level representation	May suppress small, important lesions
Multiscale graphs	Captures patterns at different scales	More computational and methodological complexity

5. Datasets and Evaluation

5.1 Plant disease datasets

PlantVillage is widely used for plant disease classification. Mohanty et al. (2016) demonstrated the potential of deep learning for image-based plant disease identification. Controlled datasets are useful for benchmarking, but field images can contain complex backgrounds, changing illumination, occlusion, and overlapping leaves. Controlled-test accuracy alone is insufficient evidence of field readiness.

5.2 Fruit maturity and quality

Fruit ripeness classification may depend on color, texture, shape, and surface defects. Dataset labels and maturity stages should be defined consistently. Where possible, near-duplicate images of the same fruit, plant, or capture session should not be split across training and test sets.

5.3 Metrics

Accuracy should be supplemented by balanced accuracy, macro-F1, per-class precision and recall, and a confusion matrix. ROC-AUC or precision–recall AUC may be appropriate depending on class balance and the task. Report training time, preprocessing time, inference latency, peak memory, and model size. For disease detection, missed diseased cases may be especially costly, so recall and the intended use case should guide metric selection.

5.4 Leakage and external validity

Images from the same plant, farm, camera session, or acquisition environment can be highly correlated. Random image-level splitting may inflate performance when related samples occur in both training and test partitions. Group-based splits and external evaluation on a separate dataset or field environment provide stronger evidence of generalization.

6. Research Gaps

- **Segmentation sensitivity:** Fixed superpixel settings may fail to preserve small lesions or may create unnecessary graph nodes. Evaluate performance across segmentation settings and consider adaptive or multiscale segmentation.
- **Graph construction is under-isolated:** Node features, edge design, and pooling should be tested through controlled ablations rather than changed simultaneously.
- **Field generalization:** Evaluate cross-farm, cross-season, cross-camera, and cross-dataset performance where suitable data are available.
- **End-to-end computational cost:** Measure segmentation, feature extraction, graph construction, and inference separately; a smaller graph does not guarantee a faster pipeline.
- **Interpretability and uncertainty:** Validate whether highlighted regions correspond to meaningful symptoms; assess calibration and uncertainty rather than relying on visually plausible explanations.
- **Reproducibility:** Publish code, data split details where permitted, preprocessing, hyperparameters, random seeds, and confidence intervals.

7. Future Research Directions

- **Adaptive multiscale superpixel graphs:** Compare fine and coarse region graphs to preserve small lesions while representing larger structural patterns.
- **Feature-aware graph connectivity:** Combine spatial distance, appearance similarity, and learned semantic relations, while testing whether extra edges add noise.
- **Hybrid CNN–GNN–attention models:** Use attention or transformer components only when improvements justify the extra complexity.
- **Explainable classification:** Evaluate node importance and region attribution against expert-identified symptoms or validated annotations.
- **Robustness under distribution shift:** Test changes in illumination, blur, compression, occlusion, background, camera, season, and farm.
- **Edge deployment:** Study graph sparsification, compact embeddings, quantization, and lightweight CNN backbones using measured latency and memory.

8. Recommended Experimental Framework

1. Baseline comparison: conventional CNN, pretrained CNN with a linear classifier, and superpixel CNN–GCN.
2. Segmentation ablation: compare multiple superpixel counts and compactness settings.
3. Connectivity ablation: compare RAG, spatial k-NN, and feature-aware graphs.
4. Feature ablation: compare CNN-only features with CNN plus color, texture, and shape descriptors.
5. Pooling ablation: compare mean pooling, sum pooling, and attention-based readout.
6. Robustness evaluation: test controlled corruptions and, where possible, an external field dataset.
7. Efficiency evaluation: report preprocessing time, graph construction time, inference latency, peak memory, and model size.
8. Statistical validation: use multiple seeds, confidence intervals, and paired comparisons where appropriate.

Keep the test set separate from hyperparameter selection. Fit any data-dependent preprocessing only on training data. Do not report numerical improvements until experiments have actually been performed.

9. Limitations of This Review Draft

This manuscript is a critical-review draft rather than a completed systematic review. The candidate search strategy and thematic synthesis require a documented, current literature search. The cited literature should be checked against publisher or DOI records, and the comparative analysis should be expanded using a verified study-extraction table. The draft does not claim new experiments, a meta-analysis, or independent reproduction of the reported agricultural accuracy values.

10. Conclusion

Superpixel-based graph learning provides a structured way to combine local visual descriptors with relationships among agricultural image regions. CNNs offer discriminative appearance features, superpixels reduce the image to coherent regions, and graph neural networks aggregate information among connected regions. This combination is promising for plant disease recognition and fruit maturity classification, but its effectiveness depends on segmentation, node features, connectivity, pooling, dataset quality, and evaluation design.

Future research should move beyond accuracy-only comparisons and prioritize reproducible experiments, strong CNN baselines, external field validation, robustness under environmental variation, explainability, and end-to-end efficiency. These elements are necessary to establish whether graph-based representations provide a reliable practical advantage in precision agriculture.

REFERENCES

1. Nguyen, T. P. T., Nguyen, T. T., Nguyen, H. Q., Nguyen, T. D., Nguyen, C. K., & Cu, N. G. (2025). An Enhanced Image Classification Model Based on Graph Classification and Superpixel-Derived CNN Features for Agricultural Datasets. Computers, Materials & Continua, 85(3), 4899–4920. https://doi.org/10.32604/cmc.2025.067707
2. Cosma, R. A., Knobel, L., van der Linden, P. A., Knigge, D. M., & Bekkers, E. J. (2023). Geometric Superpixel Representations for Efficient Image Classification with Graph Neural Networks. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 109–118. https://doi.org/10.1109/ICCVW60793.2023.00018
3. Vasudevan, V., Bassenne, M., Islam, M. T., & Xing, L. (2023). Image classification using graph neural network and multiscale wavelet superpixels. Pattern Recognition Letters, 166, 89–96. https://doi.org/10.1016/j.patrec.2023.01.003
4. Kipf, T. N., & Welling, M. (2017). Semi-Supervised Classification with Graph Convolutional Networks. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1609.02907
5. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. Frontiers in Plant Science, 7, 1419. https://doi.org/10.3389/fpls.2016.01419
6. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(11), 2274–2282. https://doi.org/10.1109/TPAMI.2012.120
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778. https://doi.org/10.1109/CVPR.2016.90
8. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71