IEEE INTERNATIONAL SYMPOSIUM

Consumer Technology

Paper ID: #39 (1571158148)

Indonesia, Bali

Lightweight Dual-Backbone Framework for Tea Leaf Disease Detection

Presented by,

Mr. Seyong Jin, Department of Artificial Intelligence, CVPR Lab., Sejong University

Co-Authors: G.H Lee, W. Choi, J. W. Kim, J. M. Oh, H. J. Moon

SEJONG UNIVERSITY

CVPR LAB

COMPUTER VISION & PATTERN RECOGNITION LAB

PRESENTATION OUTLINE

- Introduction
- Related Work
- Methodology
- Implementation & Data
- Results and Discussion
- Conclusion

References

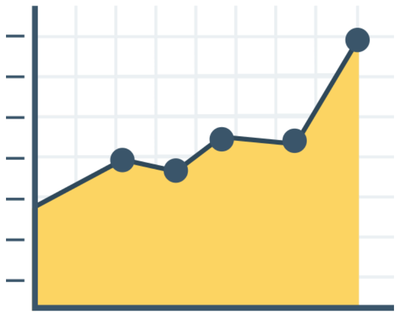
INTRODUCTION

Research Background

Global production significance
Tea is one of the most widely consumed beverages globally, with annual production exceeding 6 million tons. This industry is a cornerstone of rural development, poverty alleviation, and food security.

Economic impact analysis
The tea market has a global economic volume of over $50 billion annually and is considered a high value-added crop with consistently growing demand.

Vulnerability due to climate change
Increasing temperatures and humidity levels are accelerating the spread of tea leaf diseases, raising the need for advanced diagnostic methods and sustainable agricultural practices.



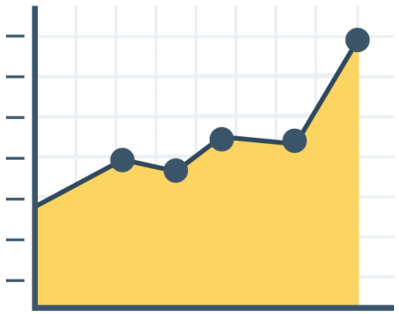
INTRODUCTION

Study Contributions

Lightweight architecture
TinyGFNet combines MobileNetV3-Small and EfficientFormer-L1 backbones to strike a balance between computational efficiency and expressive capability, tailored for edge deployment.

Novel Gated Fusion module
The study introduces a gated mechanism that dynamically adjusts the integration of local and global features, optimizing disease diagnosis accuracy.

Revolutionary implications
TinyGFNet can be utilized on edge devices like drones and smartphones, presenting scalable and eco-friendly solutions for real-world smart farming systems.


Related Work

Performance Limitations of CNN and ViT

Local feature limitations of CNN

CNNs excel at capturing local patterns but struggle to contextualize complex global features, which hampers performance in diverse agricultural applications.

Computational inefficiencies of ViT

The Vision Transformers demand substantial computational resources and extensive datasets, making them less suitable for lightweight and real-time systems on edge devices.

Balancing accuracy and efficiency

Traditional tea leaf disease classification approaches have difficulties maintaining high predictive performance while minimizing computational costs when relying solely on CNN or ViT structures.

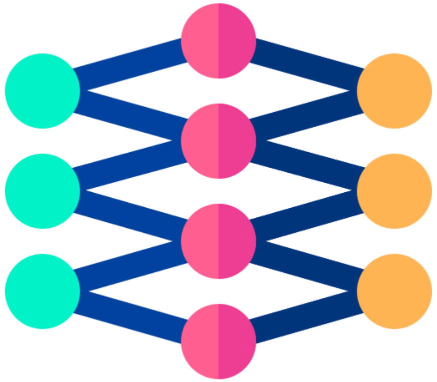
Related Work

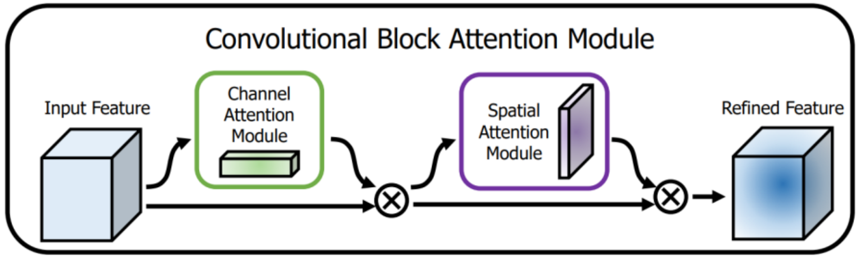
Innovations in Disease Detection

LSTM-CNN hybrid approaches: Innovative models combine spatial patterns from CNN with temporal sequence data via LSTM, achieving high accuracy but lacking real-time inference suitability for edge applications.

CBAM and RFFB structures: Methods employing Convolutional Block Attention Modules (CBAM) and Residual Feature Fusion Blocks (RFFB) enhance feature extraction and computational efficiency, though they may struggle with complex lesion patterns.

Integration of YOLO-based models: The augmented YOLO architectures incorporate lightweight and highly specific modules for improved environmental adaptability but reveal limitations in precision for complex disease detection tasks.

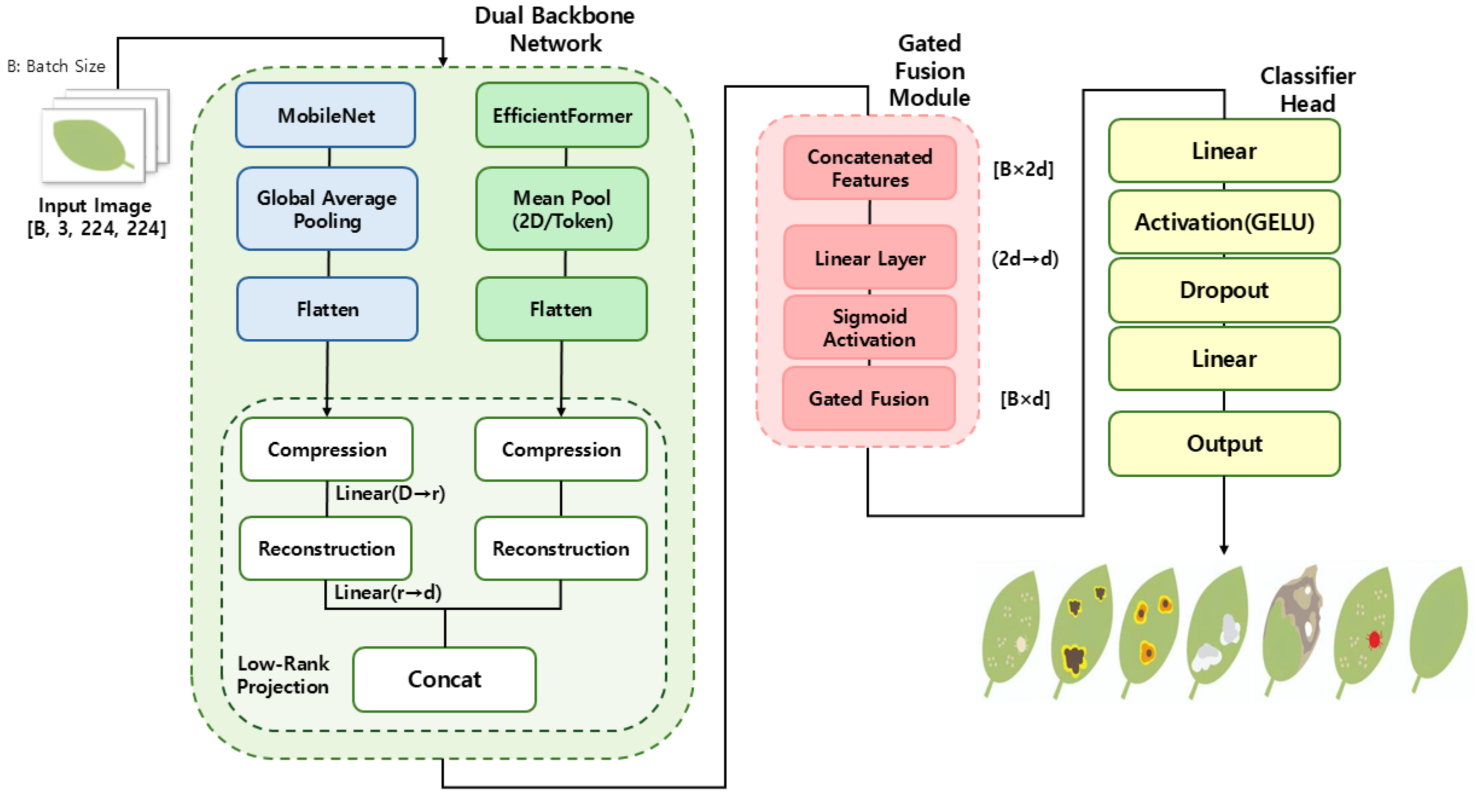




Methodology

TinyGFNet Architecture

- **Combines MobileNetV3-Small and EfficientFormer-L1 as complementary backbones**
- **Lightweight and designed for edge devices**
- **Hybrid feature extraction**



Methodology

Gated Fusion Module

Adaptive weighting mechanism

The Gated Fusion module dynamically adjusts the importance of local and global features via a gate vector, enhancing the integration of heterogeneous information.

$$F_{\text{fused}} = G \odot F_{\text{m}} + (1 - G) \odot F_{\text{e}}$$

Sigmoid-based gating operations

A trainable sigmoid function is used to selectively emphasize meaningful features and suppress irrelevant ones, optimizing the fusion process.

$$G = \sigma(W_g[F_{\text{m}}; F_{\text{e}}])$$

Low-rank projection for dimensional alignment

Before fusion, feature vectors are aligned to a unified dimension using a two-step low-rank projection, ensuring computational efficiency and proper feature alignment.

$$F_{\text{m}} = W_{\text{m}}^{(2)}(W_{\text{m}}^{(1)} F_{\text{mobile}})$$
$$F_{\text{e}} = W_{\text{e}}^{(2)}(W_{\text{e}}^{(1)} F_{\text{eformer}})$$

Methodology

Classification Head

Two-layer classifier structure
The final classification layer comprises intermediate and output layers, utilizing techniques like dropout and GELU activation to mitigate overfitting and improve stability.

Seven-class prediction
The head is specifically optimized for the classification of seven tea leaf disease categories, including healthy and diseased classes.

Softmax for probability distribution
Outputs are processed using a softmax function, providing class probabilities for accurate and interpretable predictions across all categories.

$$\hat{y} = \mathrm{Softmax}(W_{\mathrm{cls}} \cdot \mathrm{Dropout}(\mathrm{GELU}(W_{\mathrm{mid}} \cdot F_{\mathrm{fused}})))$$

Implementation and Data

Dataset Overview

TeaLeafBD dataset composition

The TeaLeafBD dataset consists of seven specific disease classes, including Tea Algal Leaf Spot, Brown Blight, Gray Blight, Helopeltis, Red Spider, Green Mirid Bug, and Healthy leaves. It ensures a balanced distribution of approximately 1,000 images per class to prevent bias.

Image acquisition conditions

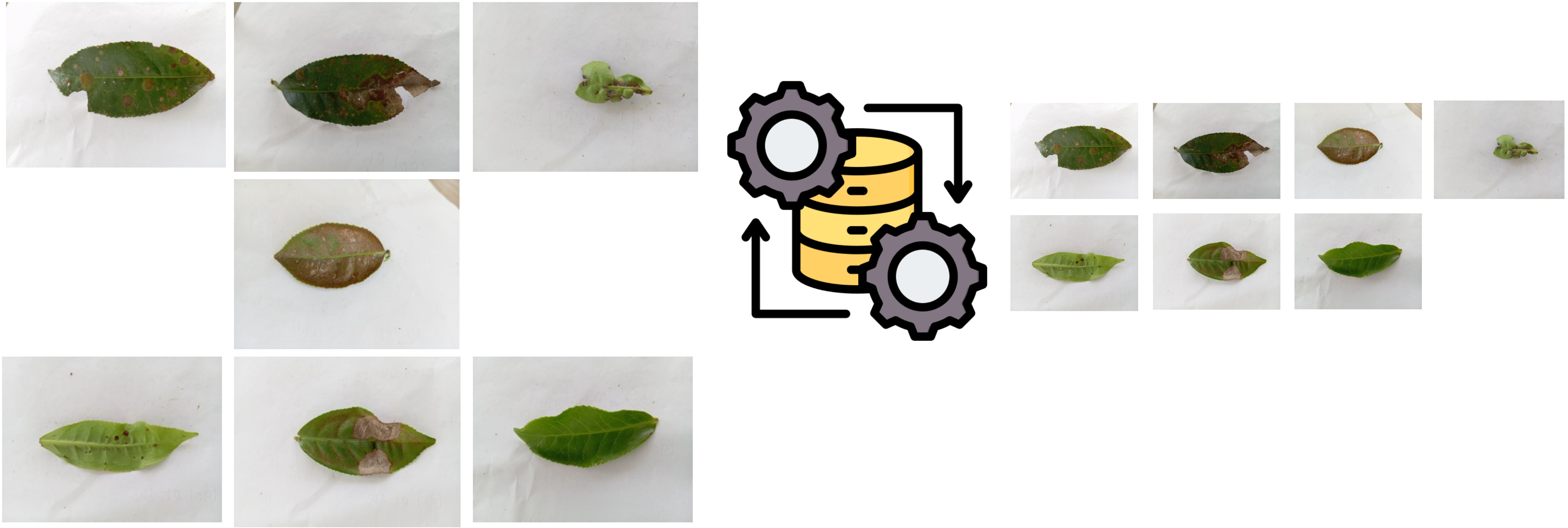
The dataset was collected under varied conditions like different lighting, background environments, and resolutions, providing robust representativeness for training models effectively.

Implementation and Data

Dataset Overview

Image preprocessing approach

Images were resized to 224x224 pixels uniformly, enabling compatibility with deep learning frameworks while maintaining key disease-specific visual features.



Implementation and Data

Training Procedures

Hyperparameter selection

The model was optimized using the Adam algorithm with a learning rate of 0.001, a batch size of 64, and a dropout rate of 0.3 to prevent model overfitting during training cycles.

Early stopping and adaptive learning decay

Early stopping ensured efficient training by terminating processes after 25 epochs of validation stagnation, while weight decay helped stabilize the optimization when training loss plateaued.

Model checkpoints and reproducibility

Best-performing models were automatically saved during training, preserving reproducibility and enabling further performance analysis or fine-tuning in subsequent research.

Results and Discussion

Performance Comparison

Benchmarking results: TinyGFNet demonstrated superior F1-score (0.95) compared to CNN (0.92), ViT (0.95), MobileNet (0.80), and EfficientFormer-L1 (0.93). This consistent performance underscores its effectiveness for tea leaf disease classification.

Computational efficiency: TinyGFNet achieved high accuracy using only 12.88M parameters and 1.34 GFLOPs, making it significantly lighter than ViT while delivering comparable performance.

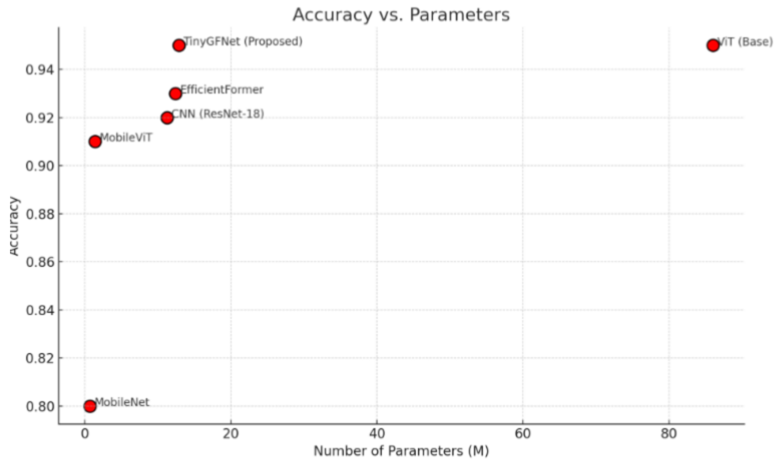
Feature integration advantage: By leveraging the complementary strengths of MobileNet and EfficientFormer, TinyGFNet balances local feature extraction and global context understanding, bolstering classification reliability.

TABLE I
COMPARISON OF CLASSIFICATION PERFORMANCE ACROSS DIFFERENT MODELS

Model	Precision	Recall	F1-score
CNN [1]	0.91	0.93	0.92
ViT [2]	0.95	0.95	0.95
MobileNet [3]	0.82	0.81	0.80
EfficientFormer [4]	0.93	0.93	0.93
MobileViT [5]	0.92	0.90	0.91
TinyGFNet(Proposed)	**0.95**	**0.95**	**0.95**

TABLE II
MODEL COMPLEXITY, COMPUTATIONAL COST, AND CLASSIFICATION ACCURACY

Model	Params(M)	FLOPs(G)	Accuracy
CNN [1]	11.24	1.82	0.92
ViT [2]	86.00	12.02	0.95
MobileNet [3]	**0.70**	**0.024**	0.80
EfficientFormer [4]	12.42	1.31	0.93
MobileViT [5]	1.40	0.254	0.91
TinyGFNet(Proposed)	12.88	1.34	**0.95**



Results and Discussion

Confusion Matrix Analysis

Class-wise accuracy insights

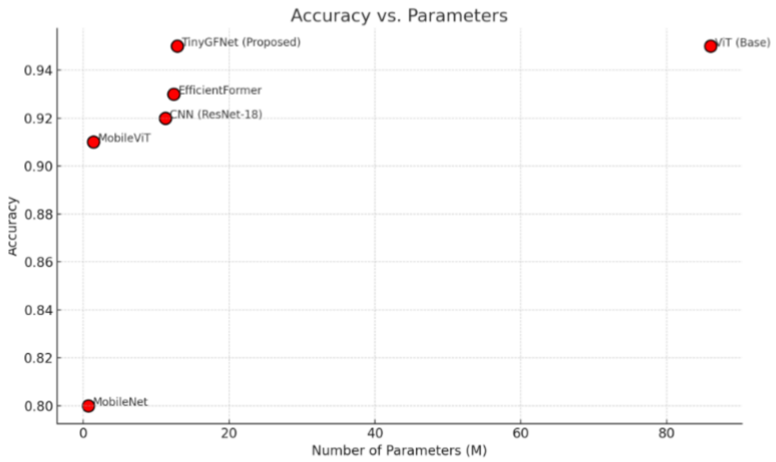
TinyGFNet exhibited robust classification across seven tea leaf disease categories, with minimal bias and consistent performance.

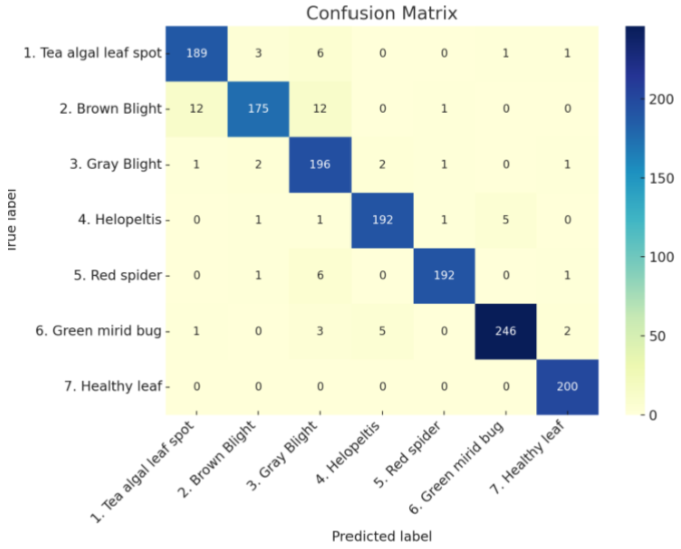
Misclassification patterns

Minor errors were observed between closely related diseases, such as Brown Blight and Gray Blight, due to their visual similarities in lesion texture and color.

Improvement areas

Enhanced feature discriminative modules or fine-grained visual representations could further reduce errors in visually similar disease categories, improving the system's precision further.





Results and Discussion

Contribution to Agriculture

Reduction in nitrogen usage

TinyGFNet facilitates early and accurate disease detection, enabling optimized nitrogen fertilizer application and mitigating associated environmental risks.

Sustainability impact

By minimizing pollution stemming from inefficient fertilizer use, TinyGFNet supports sustainable agricultural ecosystems.

Integration with smart farming

Its lightweight architecture makes it viable for real-time deployment on platforms such as drones and smartphones, promoting eco-friendly farming practices and reducing resource inefficiencies.

Results and Discussion

Future Directions

Integrated ecosystem modeling

Research could delve into combining disease detection with predictive ecological models to foresee long-term agricultural sustainability impacts.

Robust dataset expansion

New datasets with broader environmental variations and disease types would enhance TinyGFNet's generalization capabilities across diversified crops and conditions.

Results and Discussion

Future Directions

Investigating real-time reasoning

Future research could focus on achieving faster on-device inference performance to enable real-time disease detection, ensuring proactive measures in dynamic agricultural environments.

Exploring edge scalability

Emphasis should be placed on optimizing TinyGFNet for scalability across diverse edge devices, including low-power platforms, UAVs, and smartphones for smart farming applications.

Developing metaverse-based frameworks

Building immersive agricultural education systems within the metaverse could enhance disease management training, enabling farmers to visualize and understand disease patterns virtually.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for imagerecognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp.770–778, 2016.
[2] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T.Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, and N.Houlsby, “An image is worth 16×16 words: Transformers for imagerecognition at scale,” arXiv preprint arXiv:2010.11929, 2020.
[3] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang,Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, “Searching forMobileNetV3,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp.1314–1324, 2019.
[4] Y. Li et al., “Efficientformer: Vision transformers at mobilenet speed,”Advances in Neural Information Processing Systems, vol. 35, pp.12934–12949, 2022.
[5] S. Mehta and M. Rastegari, “MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer,” arXiv preprintarXiv:2110.02178, 2021.
[6] R. Geirhos, P. Rubisch, C. Michaelis, M. Bethge, F. A. Wichmann,and W. Brendel, “ImageNet-trained CNNs are biased towards texture;increasing shape bias improves accuracy and robustness,” in Proc. Int.Conf. Learn. Represent. (ICLR), 2019.
[7] Y. Pan, “Leaf disease detection and classification with a CNN-LSTMhybrid model,” Applied and Computational Engineering, vol. 104, pp.7–20, 2024.
[8] J. Lin, X. Chen, R. Pan, T. Cao, J. Cai, Y. Chen, X. Peng, T. Cernava,and X. Zhang, “GrapeNet: A lightweight convolutional neural networkmodel for identification of grape leaf diseases,” Agriculture, vol. 12, no.6, p. 887, 2022.
[9] A. Balasundaram, P. Sundaresan, A. Bhavsar, M. Mattu, M. S. Kavitha,and A. Shaik, “Tea leaf disease detection using segment anything modeland deep convolutional neural networks,” Results in Engineering, vol.25, p. 103784, 2025.
[10] Y. Miao, W. Meng, and X. Zhou, “SerpensGate-YOLOv8: An enhancedYOLOv8 model for accurate plant disease detection,” Frontiers in PlantScience, vol. 15, p. 1514832, 2025.

[11] X. Huang, P. Lakshmanan, W. Zhang, X. Wang, B. Liu, K. Ni, J. Ruan,X. Shi, X. Chen, and F. Zhang, “Large loss of reactive nitrogen andthe associated environmental damages from tea production in China,”Agriculture, Ecosystems & Environment, vol. 377, p. 109252, 2025.
[12] M. El Sakka, M. Ivanovici, L. Chaari, and J. Mothe, “A review of CNNapplications in smart agriculture using multimodal data,” Sensors, vol.25, no. 2, p. 472, 2025.
[13] B. Zhan, M. Li, W. Luo, P. Li, X. Li, and H. Zhang, “Study on the teapest classification model using a convolutional and embedded iterativeregion of interest encoding transformer,” Biology (Basel), vol. 12, no. 7,p. 1017, Jul. 2023, doi: 10.3390/biology12071017.
[14] A. C. Bhowmik, M. T. Ahad, and Y. R. Emon, “Machine learning-basedJamun leaf disease detection: A comprehensive review,” arXiv preprintarXiv:2311.15741, 2023.
[15] J. Wang, P. Xu, H. Lin, S. Guo, Z. Han, and J. Zou, “Gaseous reactivenitrogen losses from orchards, vegetables and tea plantations,” Frontiersof Agricultural Science and Engineering, vol. 10, no. 2, pp. 155–166,2023.
[16] G. Pruntseva, B. Danylyshyn, N. Popadynets, O. Kopylyuk, and I.Kotsan, “Substaminale development in agricultural investments and the food security system,” in Proc. E3S Web of Conferences, vol. 567, p.01020, 2024.
[17] O. M. Yetunde et al., “Comparative analysis of tea and non-tea con-sumers in South West, Nigeria,” Int. J. Tea Sci., vol. 16, no. 01, pp.1–4, 2022.
[18] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in Proc.IEEE Conf. Comput. Vis. Pattern Recognit., pp. 7132-7141, 2018.
[19] R. K. Srivastava, K. Greff, and J. Schmidhuber, “Highway networks,”arXiv preprint arXiv:1505.00387, 2015.
[20] M. Wang, B. Liu, and H. Foroosh, “Factorized convolutional neuralnetworks,” in Proc. IEEE Int. Conf. Comput. Vis. Workshops, pp. 545-553, 2017.

Q&A

THANK YOU FOR YOUR ATTENTION

E-mail davjin@sju.ac.kr

Lab. cvpr.sejong.ac.kr