Fig. 5. Sample misclassified images from the multi-task EfficientNetB0 model, showing cases where minimal differences in texture, dryness, or color intensity confused closely related quality classes.

Despite this complexity, assigning a higher loss weight to the quality branch helped the model focus on these fine-grained cues. As a result, the multi-task network produced fewer quality-related mistakes and exhibited a cleaner diagonal structure in the quality confusion matrix, indicating more consistent predictions.

The improvement is mainly due to the model's ability to use contextual signals, such as how color and stalk position correlate with quality, to better resolve borderline cases. In summary, the multi-task framework demonstrates robust performance across all grading dimensions, while the remaining errors reflect intrinsic challenges in fine-grained tobacco leaf grading. Importantly, the absence of large off-diagonal error clusters indicates that the model does not exhibit systematic bias toward any specific class, further supporting the robustness of the proposed framework.

V. CONCLUSION AND FUTURE WORK

This study presented a Multi-Task Learning (MTL) framework to modernize the grading of Burley tobacco leaves. By modeling stalk group, quality, and color together, rather than treating them as separate predictions or combining them into a single flat label, the framework mirrors the way human graders evaluate leaves, but with the added benefits of consistency and speed. This shift from conventional single-task and multi-class models represents a meaningful engineering improvement in automated agricultural assessment. The preprocessing pipeline, particularly coin-based size normalization and orientation correction, also strengthened the consistency and reproducibility of feature extraction, supporting more stable model performance.

The results demonstrate the value of this approach. EfficientNetB0, used as the shared backbone, achieved 94.82%, outperforming both the multi-class baseline (89.94%) and the single-task approach (92.25%), while also reducing training time and delivering the fastest inference at 18 ms/frame. These improvements translate directly into practical advantages in real-world grading environments, where decisions must be both reliable and efficient.

Compared with existing studies, the proposed framework achieves substantial numerical gains across a more challenging grading space. Prior research addressed anywhere from 3 to 41 grades, yet often relied on small datasets (e.g., only 1588 and 1498 images for 41-grade evaluations) or large datasets covering only a few classes, such as 66,966 images for just six grades. Against this backdrop, the proposed model delivers higher accuracy while handling 35 expert-defined grades, demonstrating strong scalability and generalization.

Along with strong predictive performance, this study contributes a clear pathway for improving tobacco grading practices. The combination of a controlled imaging setup, a large expert-annotated dataset, and an MTL-based architecture provides a foundation that can be expanded and refined over time. With further development, such systems could help reduce disputes in buying stations, support fairer pricing for farmers, and enhance quality control for manufacturers.

In practical deployment scenarios, such as conveyor-belt grading systems, the operational design typically enforces single-leaf presentation per frame to ensure reliable inspection. Under this assumption, the current preprocessing pipeline based on the Largest Connected Component (LCC) segmentation remains effective and robust. However, in rare cases where partial overlap or occlusion may still occur due to handling or mechanical irregularities, adjacent leaves could be merged into a single region, potentially affecting segmentation quality and downstream classification. To enhance robustness in such situations, the proposed framework can be complemented with confidence-based rejection mechanisms, whereby samples producing low prediction confidence in one or more task branches (stalk group, quality, or color) are automatically flagged for manual review. This human-in-the-loop strategy allows ambiguous or non-ideal samples to be handled reliably without interrupting automated throughput, while preserving grading accuracy in high-volume industrial environments.

While the results are promising, several limitations should be noted. The model was trained on only 35 of the 58 official Burley grades, as some rare classes lacked sufficient samples, which limits full-grade generalization. In addition, all images were collected under controlled indoor conditions, which may not fully capture the variability encountered in real grading environments, such as changes in lighting, background, camera type, or acquisition setup. The system also relies solely on RGB images and does not incorporate other potentially informative cues, such as moisture content, chemical attributes, or multispectral information that human graders may implicitly consider during manual assessment.

Future work should address these limitations by expanding the dataset to include the complete 58-grade classification system, capturing images under diverse and uncontrolled field conditions, and exploring the integration of multimodal sensing, such as hyperspectral imaging or IoT-based leaf condition sensors. More advanced architectures, including Vision Transformers, hybrid CNN-ViT models, and self-supervised feature learning approaches, may further improve robustness to domain shifts and enhance interpretability.

Additionally, on-field testing and deployment within real-time IoT-enabled grading systems, such as conveyor-belt inspection units, would help transition the framework from controlled laboratory settings to practical use in tobacco-buying stations and supply chains. Future implementations should also consider security aspects, including secure data storage, access control, and protected deployment within closed or on-premise networks, to ensure data integrity and trustworthy operation in real-world grading environments, supporting scalable, sensor-driven automated grading. Reliability-aware deployment strategies, such as confidence-based decision thresholds and human-in-the-loop verification for borderline cases, will also be explored to further enhance dependable grading performance, supporting scalable, sensor-driven automated grading.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

CM led the entire research process, including conceptualizing the study, designing the methodology, developing the dataset, conducting all primary experiments, analyzing the results, and writing the full manuscript. CR and OA contributed by assisting with additional experiments, validating the findings, and providing critical feedback that strengthened the final analysis. All authors reviewed and approved the final version of the manuscript.

Acknowledgment

The authors wish to thank Universal Leaf Philippines Inc., De La Salle University, Don Mariano Marcos Memorial State University, the grading experts, and the dedicated tobacco farmers for their invaluable support and contributions to this study.

References

[1] L. Han, "Recognition of the part of growth of flue-cured tobacco leaves based on support vector machine," in *Proc. 7th World Congress on Intelligent Control and Automation*, 2008, pp. 3624–3627.

[2] D. S. Guru, P. B. Mallikarjuna, and S. Manjunath *et al.*, "Machine vision-based classification of tobacco leaves for automatic harvesting," *Intelligent Automation & Soft Computing*, vol. 18, no. 5, pp. 581–590, 2012.

[3] F. Zhang and X. Zhang, "Classification and quality evaluation of tobacco leaves based on image processing and fuzzy comprehensive evaluation," *Sensors*, vol. 11, no. 3, pp. 2369–2384, 2011.

[4] J. Liu, J. Shen, and Z. Shen *et al.*, "Grading tobacco leaves based on image processing and generalized regression neural network," in *Proc. Intelligent Conference on Intelligent Control, Automatic Detection and High-End Equipment*, 2012, pp. 89–93.

[5] S. K. Dasari, K. R. Chintada, and M. Patruni, "Flue-cured tobacco leaves classification: A generalized approach using deep convolutional neural networks," in *Cognitive Science and Artificial Intelligence*, Singapore: Springer Singapore, 2017, pp. 13–21.

[6] C. S. Marzan and C. R. Ruiz, "Automated tobacco grading using image processing techniques and a convolutional neural network," *International Journal of Machine Learning and Computing*, vol. 9, no. 6, pp. 807–813, 2019.

[7] L. Pei, J. Shen, and R. Liu, "Deep feature of image screened by improved clustering algorithm cascaded with genetic algorithm," in *Proc. 29th Chinese Control and Decision Conference,* 2017, pp. 452–455.

[8] G. Li, H. Zhen, and F. Jiao *et al.*, "Research on tobacco leaf grading algorithm based on transfer learning," in *Proc. International Conference on Artificial Intelligence and Computer Applications*, 2021, pp. 32–35.

[9] M. Lu, C. Wang, W. Wu *et al.*, "Intelligent grading of tobacco leaves using an improved bilinear convolutional neural network," *IEEE Access*, vol. 11, pp. 68153–68170, 2023.

[10] X. Xin, H. Gong, R. Hu *et al.*, "Intelligent large-scale flue-cured tobacco grading based on deep densely convolutional network," *Scientific Reports*, vol. 13, no. 1, 2023.

[11] Z. He, G. Chen, Y. Zhang *et al.*, "Pyramid feature fusion through shifted window self-attention for tobacco leaf classification," *Expert Systems with Applications*, vol. 230, 120601, 2023.

[12] H. Luo and C. Zhang, "Features representation for flue-cured tobacco grading based on transfer learning to hard sample," in *Proc. 14th IEEE International Conference on Signal Processing*, 2018, pp. 591–595.

[13] M. Lu, S. Jiang, C. Wang *et al.*, "Tobacco leaf grading based on deep convolutional neural networks and machine vision," *Journal of the ASABE*, vol. 65, no. 1, pp. 11–22, 2022.

[14] L. Mengyao, Z. Qiang, J. Shuwen *et al.*, "Flue-cured tobacco leaf grading method based on deep learning and multi-scale feature fusion," *Journal of Chinese Agricultural Mechanization*, vol. 43, no. 1, pp. 158–166, 2022.

[15] Q. Li, H. Lin, J. Hu *et al.*, "Automated tobacco leaf grading system based on deep learning," in *Proc. IEEE 3rd International Conference on Software Engineering and Artificial Intelligence*, 2023, pp. 40–46.

[16] W. Gu, C. Duan, X. Gai *et al.*, "Automatic grouping method of flue-cured tobacco based on MobileViT," *Chinese Tobacco Science*, vol. 45, no. 1, pp. 104–111, 2024. (in Chinese)

[17] C. Feng, S. Zhu, H. Huang *et al.*, "Research on high efficiency flue-cured tobacco grade identification model based on deep learning," *Journal of Southwest University Natural Science Edition*, vol. 47, no. 1, pp. 213–225, 2025.

[18] Y. Chen, C. Lian, X. Chen *et al.*, "Fine-grained visual classification of flue-cured tobacco leaf grades based on vein information," in *Proc. 10th International Conference on Systems and Informatics*, 2024, pp. 1–6.

[19] D. Wang, Y Li, C. Wang *et al.*, "Tobacco leaf grading algorithm based on expert ViT model," in *Proc. 2025 Asia-Europe Conference on Cybersecurity, Internet of Things and Soft Computing*, 2025, pp. 234–238.

[20] L. Huabo, L. He, and L. Tao, "The barrel theory-based decision-making algorithm and its application," in *Proc. 2009 International Conference on Computational Intelligence and Natural Computing,* 2009, pp. 11–14.

[21] X. Ma, J. Shen, R. Liu *et al.*, "Choice of tobacco leaf features based on selected probability of particle swarm algorithm," in *Proc. 2016 Chinese Control and Decision Conference*, 2016, pp. 3041–3044.

[22] S. Yang, C. Dong, F. Wang *et al.*, "Fisher's tobacco leaf grading method based on image multi-features," in *Proc. 2022 International Conference on Artificial Intelligence and Computer Information Technology*, 2022, pp. 1–6.

[23] D. Wang, G. Li, H. Zhen *et al.*, "Intelligent classification of tobacco leaves based on residual network," in *Proc. International Conference on Artificial Intelligence and Computer Applications*, 2022, pp. 156–159.

[24] C. Zhang, L. Yun, and Z. Chen, "Flue-cured tobacco grading method based on a convolutional neural network," in *Proc. CCF National Conference on Computer Applications*, 2023, pp. 54–66.

[25] R. Surendran, C. P. Sree, and K. Sumathy, "Texture- and symptom-based coffee leaf disease classification using CNN and ResNet algorithms," in *Proc. 4th International Conference on Advancement in Electronics & Communication Engineering*, 2024, pp. 1287–1292.

[26] P. Sree, R. Surendran, and R. Selvanarayanan, "Leveraging CNN and AlexNet algoritms for improved coffee leaf disease identification and detection," in *Proc. 2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering,* 2024, pp. 1–6.

[27] T. Tamilvizhi, R. Surendran, K. Anbazhagan *et al.,* "Quantum-behaved particle swarm optimization-based deep transfer learning model for sugarcane leaf disease detection and classification," *Mathematical Problems in Engineering,* vol. 1, 3452413, 2022.

[28] D. Chen, Y. Zhang, Z. He *et al.*, "Feature-reinforced dual-encoder aggregation network for flue-cured tobacco grading," *Computers and Electronics in Agriculture*, vol. 210, 107887, 2023.

[29] Q. He, X. Zhang, J. Hu *et al.*, "CMENet: A cros-modal enhancement network for tobacco leaf grading," *IEEE Access*, vol. 11, pp. 109201–109212, Oct. 2023.

[30] M. Xu, J. Gao, Z. Zhang *et al.*, "Multi-channel and multi-scale separable dilated convolutional neural network with attention mechanism for flue-cured tobacco classification," *Neural Computing and Applications*, vol. 35, no. 21, pp. 15511–15529, 2023.

[31] T. Hu, K. Hou, Y. Su *et al.*, "Research on synchronous recognition method of multiple tobacco leaf grades based on MDCNet," in *Proc.*

5th International Conference on Frontiers Technology of Information and Computer, 2023, pp. 483–490.

[32] S. Wang, R. Zhuang, W. Lu *et al.*, "Intelligent classification of flue-cured tobacco based on improved network model of six-channel ResNet," *Journal of Biotech Research*, vol. 20, pp. 98–110, 2025.

[33] H. Benrong, F. Zhaofeng, W. Fei *et al.*, "Intelligent grading of flue-cured tobacco based on VGG16-DenseNet integrated model," *Chinese Tobacco Science*, vol. 45, no. 3, pp. 102–112, 2024.

[34] A. Bensaoud and J. Kalita, "Deep multi-task learning for malware image classification," *Journal of Information Security and Applications*, vol. 64, 103057, 2022.

[35] Y. J. Wu, C. M. Tsai, and F. Shih, "Improving leaf classification rate via background removal and ROI extraction," *Journal of Image and Graphics,* vol. 4, no. 2, pp. 93–98, 2016.

[36] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed., Switzerland: Springer, 2022.