Position Papers of the 20th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 69–73
DOI: 10.15439/2025F2775
ISSN 2300-5963 ACSIS, Vol. 44

Integrating Real-ESRGAN with CNN Models for UAV Image Based Plant Disease Detection

Sravya Malladi
Independent Researcher
San Jose, California, USA
sravya.malladi@gmail.com

Pranav Kulkarni
Stanford University
Palo Alto, California, USA
pranavsk@stanford.edu

***Abstract*—The integration of deep learning models with UAV captured images for plant disease detection has been explored in many papers and has the potential to revolutionize commercial precision agriculture, by allowing for early and efficient detection and classification of crop disease stages. In order to address the limitations posed by low-resolution aerial imaging, this paper proposes the additional integration of an Enhanced Super Resolution Generative Adversarial Network (ESRGAN) with a Convolutional Neural Network model for field monitoring through UAV captured imagery. UAVs are a cost effective method of monitoring large swaths of agricultural land; however, it is difficult to capture images of a high enough quality and clarity to be adequately analyzed by a CNN. The images typically lack the necessary resolution for accurate classification, especially for diseases with smaller, less noticable symptoms. The Real-ESRGAN model is employed to generate a dataset of high-resolution images, from low-resolution inputs, allowing the disease detection CNN to more accurately and effectively identify and classify disease stages in Armillaria afflicted cherry trees. This solution offers a solution to the problem posed by traditional UAV based approaches that enhances classification accuracy even in suboptimal conditions. Through this integrated approach, the model was able to reach an increased validation accuracy, as well as significantly decreased loss values due to the ESRGAN enhanced imagery allowing for clearer detection of early stage Armillaria symptoms. This integrated system provides a practical scalable solution for commercial agriculture, allowing for more comprehensive and efficient crop disease monitoring. Future research can be explored to optimize the architecture of this model and expand its applicability to other crops and environmental conditions, allowing more efficient precision agriculture and paving the way for more sustainable farming practices.**

***Index Terms*—Crop Monitoring, Enhanced Super Resolution GAN, Deep Learning, UAV imagery, Precision Agriculture**

I. Introduction

THE advancement of deep learning has impacted commercial agriculture significantly, particularly in the classification and recognition of plant diseases. Traditional crop inspection methods are prone to human errors such as psychological and cognitive biases [1]. Furthermore, the vastness of agricultural land and the scarcity of trained plant pathologists make manual monitoring impractical [2]. Deep learning, specifically Convolutional Neural Networks (CNNs), offers a promising alternative by automating disease detection and classification tasks with high accuracy [3]. In order to be properly analyzed by a CNN, UAV captured images must have a high enough resolution and clarity so that the symptoms of a plant disease can be seen. This problem is amplified when said plant disease has very small symptoms, or when the farm area is very large. The proposed solution to this limitation is targeted sampling of a field, where images are acquired from a small section of the field's area. While this is a solution would provide farmers more information than the traditional methods of scouting for diseases, the difficulties faced by a model when analyzing UAV images image still stand. The Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) model detailed in the section below presents a method for crop disease classification on low resolution images. The model is used to generate high resolution images from low resolution crop images. This is called Image Super Resolution (SR). Though there are SR methods besides ESRGAN, the paper shows that the ERAGAN model, and its successor the Real-ESRGAN model, generate higher visual quality images than other methods used [4][5].

II. Literature Review

A. Literature Review

In commercial agriculture, identifying disease severity is crucial for making timely and effective decisions to reduce financial losses and fight plant infections[6]. Machine Learning models that classify different stages of a plant disease, are therefore, most helpful. For example, the regression model proposed in Detection and Characterization of Stressed Sweet Cherry Tissues Using Machine Learning identifies different stages of Amarilma, a devastating cherry tree disease that causes annually 8 million dollars in losses in the United States alone[7]. Commercial farmers use aerial and satellite imagery to monitor their crop fields. According to The application of small unmanned aerial systems for precision agriculture: a review, UAV captured aerial imagery is a cost effective solution that can be used for crop disease detection, reducing the need for in person monitoring [8][9]. The use of UAVs paired with detection technologies, is a transformative practice that will greatly facilitate the practicality of precision agriculture[9]. These technologies enhance crop monitoring, optimize resource use, and minimize the environmental footprint of farming by reducing the application of fertilizers and pesticides. However, despite their success, these techniques face challenges when applied in real-world agricultural conditions. In order to be properly analyzed by a model, UAV captured images must have a high enough resolution and clarity so that the symptoms of a plant disease can be seen. An image of this quality is hard to capture with a UAV which is subject to wind, lighting conditions, and other environmental challenges [9][11]. As detailed in Millimeter-Level Plant Disease Detection From Aerial Photographs via Deep Learning and Crowdsourced Data, this problem is amplified when said plant disease has very small symptoms, or when the farm area is very large [12]. To address low-resolution imagery, several approaches can be considered. One example is hyperspectral imagery, which captures information across dozens or hundreds of narrow spectral bands. This can reveal subtle physiological and biochemical changes in plants that are not visible in standard RGB images, improving disease detection even at lower spatial resolutions. However, specialized hyperspectral cameras are expensive, and the data they produce is extremely large and complex, requiring extensive preprocessing, calibration, and storage [13]. These requirements make hyperspectral imaging difficult to implement efficiently for routine agricultural applications. Another alternative is SRCNN (Super-Resolution Convolutional Neural Network). SRCNN is of the earliest deep learning based super resolution models that employs a three-layer CNN architecture to upscale images. However, although this architecture is straightforward and computationally efficient, it has limited capacity to capture complex textures or fine details, which are critical for early detection of diseases with small and difficult to see symptoms [14]. Another proposed solution for this problem is to use GAN models for data augmentation [2][15][16]. In contrast, GAN-based super-resolution methods offer a highly practical alternative. GANs can be used to enhance the resolution of standard RGB images captured by UAVs, allowing models to detect fine disease symptoms without the need for specialized sensors, requiring no extra hardware beyond a conventional camera. Generated images and lesions are used for data augmentation, where the model is trained on an expanded dataset. GAN-based models are capable of reconstructing realistic textures and subtle visual cues. Therefore augmenting data with GAN super resolution models can greatly improve the effectiveness of plant disease classification models[15][16][17] The best GAN model for this purpose seems to be the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) model. Super-resolution models are designed to reconstruct high-resolution (HR) images based on low-resolution (LR) inputs. When given a low-resolution image $X \in \mathbb{R}^{h' \times w' \times c}$, the model generates a corresponding high-resolution image $Y \in \mathbb{R}^{h \times w \times c}$, where $h > h'$ and $w > w'$[18]. The ESRGAN model improves upon the earlier SRGAN framework by emphasizing perceptual realism. Its generator network employs deep Residual-in-Residual Dense Blocks (RRDBs) to extract important features and recreate fine details in the images. The model also uses a Relativistic average GAN (RaGAN) discriminator, which compares the generated images to real images. A crop disease detection and classification model integrated with ESRGAN has been shown to be better at detecting crop disease than the models that use other Super Resolution methods, with a higher classification accuracy due to greater visual quality [4][19]. The REAL-ESRGAN model further expands upon the ESRGAN model, generating images of an even greater visual quality. The generator network of this model is trained using a combination of content, perceptual, and adversarial losses [5]. Content loss measures pixel-wise similarity between real-world and generated images:

$$L_{\text{content}} = \frac{1}{N} \sum_{i=1}^{N} \|G(X_i) - Y_i\|_2^2$$

Perceptual loss encourages high-level similarity using feature maps from a pre-trained network, measuring how real the images appear in terms of patterns, textures, and shapes:

$$L_{\text{perceptual}} = \frac{1}{N} \sum_{i=1}^{N} \|\phi_j(G(X_i)) - \phi_j(Y_i)\|_1$$

Adversarial loss guides the generator to produce images that are difficult for the discriminator to distinguish from real HR images:

$$\begin{aligned} L_{\text{GAN}}(G, D) = \mathbb{E}_Y \Big[\log(D(Y) - \mathbb{E}_X[D(G(X))]) \Big] \\ + \mathbb{E}_X \Big[\log(1 - (D(G(X)) - \mathbb{E}_Y[D(Y)])) \Big] \end{aligned}$$

The overall generator loss is a weighted combination of these components:

$$L_G = L_{\text{content}} + \lambda \cdot L_{\text{perceptual}} + \eta \cdot L_{\text{GAN}}$$

where λ and η balance the contributions of perceptual and adversarial losses.

The discriminator outputs a probability map indicating the likelihood that each pixel belongs to a real image. The discriminator loss is calculated using a weighted sum over all pixels. Spectral normalization is added in Real-ESRGAN to regularize the discriminator weights:

$$\hat{W} = \frac{W}{\sigma(W)}$$

Through these enhancements, the model is better able to handle real-world image degradation, making it more reliable for practical image restoration applications and an effective approach for improving the quality of aerial images [5][19][20][21].

III. Methodology

The dataset used in this study is the Cherry Tree Disease Detection dataset from the article above, Detection and Characterization of Stressed Sweet Cherry Tissues Using Machine Learning, which contains both hyperspectral and standard JPG images of cherry trees at different stages of Armillaria infection. The stages represented are healthy, stage 1, and stage 2. To prepare the dataset for analysis, the images that were originally organized by the day of data collection were consolidated into broader categories corresponding to each disease stage. This reorganization facilitated the removal of

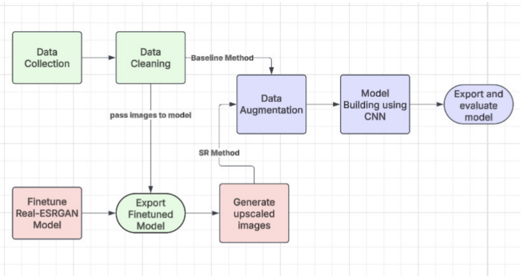

Fig. 1. Methodology

irrelevant or redundant images and ensured that the dataset was structured consistently for model training and evaluation.

The first stage of the project employed a Convolutional Neural Network (CNN) in TensorFlow to classify images of cherry trees into the three categories described above. Data augmentation techniques, such as resizing and rescaling, were applied to increase the size and diversity of the training dataset, therefore increasing the model's robustness. After data cleaning and augmentation, the images were divided into training and validation sets, which were subsequently used for model development and evaluation.

The model architecture consisted of 64 convolutional layers, including pooling and fully connected layers. The structure used allowed the model to learn the features and patterns in the images associated with each different stage of the disease.

In the second stage of the project, the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) was integrated and used to generate a new higher resolution version of the original cherry tree dataset. The ESRGAN model was fine-tuned on a super resolution dataset composed of paired high-resolution and low-resolution images, thereby enhancing image resolution and clarity.

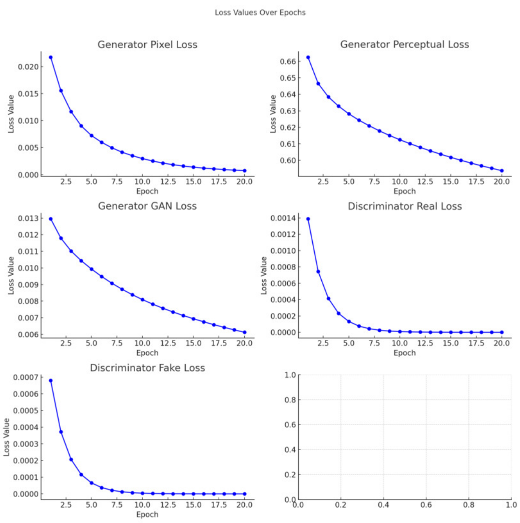

Fig. 2. Graph of various losses over epochs

The fine-tuning process began with downloading the publicly available Real-ESRGAN model from its GitHub repository. The options file was then modified so that the training and validation sections used the custom dataset of paired high-resolution and low-resolution images. Once the training script was executed, the dataset was iteratively adjusted to improve super resolution performance. After several refinements, this process yielded an optimized balance between image clarity and model runtime.

Fig. 3. Image up-scaling

By running the finetuned ESRGAN model on the original Cherry Tree Dataset, high-resolution versions of images from the dataset were generated. The generated images were used as input for classification, allowing for a direct comparison of performance and accuracy between the baseline CNN and the ESRGAN augmented CNN.

A. Results

Fig. 4. Baseline model

The integrated system achieved a 94 percent validation accuracy in classifying cherry tree disease stages, compared to 83 percent for the original CNN model. Images enhanced by ESRGAN consistently produced higher accuracy and greater confidence values during classification. Additionally, loss values decreased significantly when compared directly with the

Fig. 5. New model

original CNN trained on the standard dataset. This is due to the fact that the enhanced model could detect subtle symptoms of Armillaria that were difficult to discern in lower-resolution images. Therefore, the combined use of the two models allowed the system to classify plant disease stages more effectively, even from lower-quality images, compared to using the original CNN alone.

B. Discussion and Applications

The increased validation accuracy achieved by integrating the two models demonstrates its effectiveness in discerning between healthy cherry trees and those in various stages of Armillaria. These results show the importance of high resolution imaging in plant disease detection, validating the performance of the integrated system.

This accuracy level shows that the ESRGAN enhanced images provide greater visual clarity for the CNN to make proper classifications, compared to the original images. The results show that the CNN will be able to make reliable classifications, even when the original UAV captured images lack resolution,because of the integration of the ESRGAN Model, suggesting that the system has successfully mitigated the challenges posed by the difficulty of getting high resolution crop images from UAVs.

The results reflect that the upscaled images provided by ESRGAN significantly improve the model's ability to detect Armillaria symptoms, which is crucial for timely intervention in commercial cherry tree farming. The high accuracy and confidence levels reflected by this system means that commercial farms could rely on this to detect and classify Armillaria disease stages with few errors, increasing the efficiency of crop inspection, and allowing for better monitoring of plant disease. This in turn allows for early stage detection, allowing for better disease management, minimizing yield loss.

This method solves the problems from previous papers that struggled with the limits posed by the resolution of UAV captured images. By introducing a super resolution model, however, these become much less challenging. Unlike previous models that required time consuming targeted sampling to be practical for commercial farming, this system allows for more efficient analysis of broader areas, with a stable accuracy.

So, the results prove that the system is both effective and practical for commercial farming in the real world. By enhancing image clarity, ESRGAN allows the CNN to easily identify lesions that would be too hard for models that take in lower resolution images, proving that super resolution techniques can help improve ML models in agriculture.

C. Conclusion

The research presents a large advancement in crop disease detection through the integration of ESRGAN model and CNN model to classify UAV captured images. By successful enhancement and accurate classification, the model shows near perfect accuracy in identifying the stages of Armillaria in cherry trees. The integrated hybrid approach addresses the challenge posed by low quality images captured by UAVs, providing a proper solution for real world applications in commercial agriculture.

Areas for future research would include expanding the dataset to include a wider range of diseases. Additionally, working on detection for other, more widely grown crop types would both increase the model's robustness, but also its real world applicability. Exploration of different enhancement techniques along with ESRGAN could cause improvements in classification accuracy. It would be prudent to investigate the model's performance in various environmental conditions, like harsh weather, or unclear lighting, to see if it would still perform as well, providing insights into practicality.

Another avenue for future research includes optimization of the model architecture. Exploring different configurations and techniques, like using larger pretrained models that would have to be finetuned (YOLO) could enhance the system's performance [22]. This model would not only classify different images, but also identify the specific locations of each individual lesion. This would help the practicality of the system proposed in this paper immensely, because it would allow images to be taken over a broader region, and would allow for more precise disease identification. This was not used in this paper due to lack of available data.

Using UAV data for predictive modeling has strong applications in the future as well. In conclusion, the research establishes a foundation for leveraging Super Resolution image augmentation techniques for Agricultural disease classification, paving the way for solutions that enhance productivity and sustainability in farming.

References

[1] C. H. Bock, G. H. Poole, P. E. Parker, and T. R. Gottwald, "Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging," Crit. Rev. Plant Sci., vol. 29, no. 2, pp. 59–107, 2010. [Online]. Available: https://doi.org/10.1080/07352681003617285

[2] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," Biosyst. Eng., vol. 172, pp. 84–91, 2018. [Online]. Available: https://doi.org/10.1016/j.biosystemseng.2018.05.013

[3] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Comput. Electron. Agric., vol. 145, pp. 311–318, 2018. [Online]. Available: https://doi.org/10.1016/j.compag.2018.01.009

[4] X. Wang et al., "ESRGAN: Enhanced super-resolution generative adversarial networks," in Proc. Eur. Conf. Comput. Vis. Workshops, 2018, pp. 63–79. [Online]. Available: https://doi.org/10.1007/978-3-030-11021-5_5

[5] X. Wang, L. Xie, and C. Dong, "Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data," in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops, 2021, pp. 1905–1914. [Online]. Available: https://doi.org/10.1109/ICCVW54120.2021.00217

[6] X. Zeng and Y. Ma, "GANs-based data augmentation for citrus disease severity detection using deep learning," Agronomy, vol. 10, no. 12, p. 1939, 2020. [Online]. Available: https://www.mdpi.com/2073-4395/10/12/1939

[7] C. Chaschatzis et al., "Detection and characterization of stressed sweet cherry tissues using machine learning," Remote Sens., vol. 12, no. 3, p. 531, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/3/531

[8] S. Zhang and J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: A review," Precision Agric., vol. 13, pp. 693–712, 2012. [Online]. Available: https://link.springer.com/article/10.1007/s11119-012-9274-5

[9] H. Zhu et al., "Intelligent agriculture: Deep learning in UAV-based remote sensing imagery for crop diseases and pests detection," Front. Plant Sci., vol. 15, 2025. [Online]. Available: https://doi.org/10.3389/fpls.2024.1435016

[10] A. D. Boursianis et al., "Internet of Things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: A comprehensive review," Internet Things, vol. 18, p. 100187, 2022. [Online]. Available: https://doi.org/10.1016/j.iot.2020.100187

[11] J. Agrawal and M. Y. Arafat, "Transforming farming: A review of AI-powered UAV technologies in precision agriculture," Drones, vol. 8, no. 11, p. 664, 2025. [Online]. Available: https://doi.org/10.3390/drones8110664

[12] O. Bongomin et al., "UAV image acquisition and processing for high-throughput phenotyping in agricultural research and breeding programs," Plant Phenome J., vol. 7, no. 1, e20096, 2025. [Online]. Available: https://doi.org/10.1002/ppj2.20096

[13] A. Purushothaman and R. Pannerselvam, "Hyperspectral imaging and its applications: A review," Front. Bioeng. Biotechnol., vol. 12, pp. 123456, Jun. 2024. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11253060/

[14] S. Ramesh and S. Srinivas, "RDA-CNN: Enhanced Super Resolution Method for Rice Plant Disease Classification," Comput. Syst. Sci. Eng., vol. 42, no. 1, pp. 273–287, 2022. [Online]. Available: https://www.techscience.com/csse/v42n1/45755

[15] S. A. Wahabzada et al., "Millimeter-level plant disease detection from aerial photographs via deep learning and crowdsourced data," Front. Plant Sci., vol. 9, p. 1453, 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpls.2018.01453/full

[16] K. Sathya and M. Rajalakshmi, "RDA-CNN: Enhanced Super Resolution Method for Rice Plant Disease Classification," Comput. Syst. Sci. Eng., vol. 42, no. 1, pp. 33–47, Jul. 2022. [Online]. Available: https://www.techscience.com/csse/v42n1/45755/html

[17] A. ul Haq and S. Kaur, "Super resolution image based plant disease detection and classification using deep learning techniques," Propuls. Tech. J., vol. 45, no. 1, pp. 1020–1022, 2024. [Online]. Available: https://www.propulsiontechjournal.com/index.php/journal/article/view/4108

[18] L. Bi and G. Hu, "Improving image-based plant disease classification with generative adversarial network under limited training set," Front. Plant Sci., vol. 11, 2020. [Online]. Available: https://doi.org/10.3389/fpls.2020.583438

[19] J. Wen, Y. Shi, X. Zhou, and Y. Xue, "Crop disease classification on inadequate low-resolution target images," Sensors, vol. 20, no. 16, p. 4601, 2020. [Online]. Available: https://doi.org/10.3390/s20164601

[20] Ş. B. Çetin, "Real-ESRGAN: A deep learning approach for general image restoration and its application to aerial images," Advanced Remote Sensing, vol. 3, no. 2, pp. 90–99, 2023. [Online]. Available: https://publish.mersin.edu.tr/index.php/arsej/article/view/1072

[21] F. Rezapoor Nikroo et al., "A comparative analysis of SRGAN models," arXiv preprint, arXiv:2307.09456, 2023. [Online]. Available: https://arxiv.org/abs/2307.09456

[22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 779–788. [Online]. Available: https://doi.org/10.1109/CVPR.2016.91