The 10th Conference on Mathematics in Engineering: Theory and Applications
Novi Sad, December 12-14th 2025

MATHEMATICAL MODELS IN VINE DISEASE DETECTION

Levente Tamás, [1] **Tibor Lukić** [2] **and Stefan Gubo** [3]

https://doi.org/10.24867/META.2025.12 Review article

Abstract. This paper presents a short description of an automated proximal aerial image-based vine disease detection approach, developed specifically for the region of the Carpathian Basin. Particular focus is given to the description of the applied Convolutional Neural Network architecture. Several experimental results are shown comparing results obtained from data collected in different geographical regions.

AMS Mathematics Subject Classification (2020): 68U10, 42B10

Key words and phrases: Vine Disease Detection, Convolution, Drone Imaging, Convolutional Neural Networks.

1. Introduction

Precision viticulture has evolved into a multidisciplinary research area where autonomous robots, multisensory perception, and deep learning are central to real-time plant health monitoring. Efficient vineyard management requires early detection of foliar diseases, accurate canopy segmentation, and robust cross-regional generalization. The reviewed works [1, 2, 3] collectively demonstrate how UAV-based remote sensing and ground-level proximal imaging contribute to building adaptable vision systems for viticulture.

A recurring challenge is the spatial and temporal variability of vineyards: grape cultivars, canopy morphology, disease expression, and illumination conditions vary across regions. The three studies address these challenges through multi-domain datasets, convolutional neural network (CNN) architectures, and domain transfer techniques.

1.1. Geographically Distributed Vine Disease Detection

The work in [1] examines disease detection in vineyards of the Carpathian Basin using low-altitude UAV platforms equipped with RGB cameras. The study emphasises the need for regional transfer learning because visual disease symptoms differ across cultivars and climatic zones. The authors constructed a large dataset exceeding 10,000 annotated images and employed YOLOv7 with COCO pretrained weights to detect diseased leaf regions.

[1]Automation Department, Technical University of Cluj-Napoca, Romania, e-mail: Levente.Tamas@aut.utcluj.ro

[2]Department of Fundamental Sciences, Faculty of Technical Sciences, University of Novi Sad, Serbia, e-mail: tibor@uns.ac.rs

[3]Faculty of Economics and Informatics, J. Selye University, Komárno, Slovakia, e-mail: gubos@ujs.sk

1.2. Segmentation of Diseased Leaf Regions

The authors of the paper [2] study segmentation methods for detecting the affected areas on individual vine leaves. The authors compare five techniques: Otsu thresholding, SegNet, Mask R-CNN, Feature Pyramid Networks, and MobileNetV3. Their dataset includes laboratory images (PlantVillage) and in-field leaf crops collected from multiple locations.

1.3. Proximal Canopy Segmentation Using FPN

The work [3] presents an FPN-based approach for proximal canopy segmentation using drone and ground-robot imagery. Unlike leaf-level segmentation, canopy segmentation must account for heterogeneous foliage, occlusions, and mixed backgrounds.

2. Mathematical background and CNN architectures of the methods

The automatic detection of vine leaf diseases in aerial close-range images relies heavily on mathematical models rooted in modern computer vision and deep learning. At the core of these models lie convolutional operators, which act as learnable filters capable of extracting hierarchical image features. A convolutional operator computes a weighted sum between a small kernel and localized image regions, enabling the detection of fundamental structures such as edges, color transitions, and texture variations. When stacked in multiple layers, these operators provide increasingly abstract representations that map low-level pixel patterns to meaningful biological indicators of leaf abnormalities. This mechanism is particularly important in vineyards, where disease symptoms manifest as subtle chromatic or structural changes on leaf surfaces.

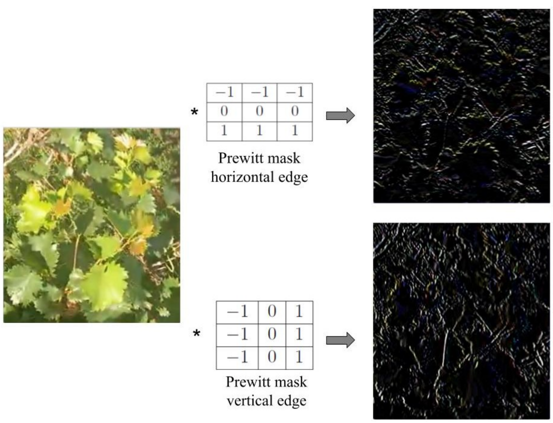

Figure 1: Extracting horizontal and vertical edges by applying two convolutional filters.

Building upon convolutional operators, Convolutional Neural Networks (CNNs) form the backbone of the applied detection models. CNNs exploit the shift-invariance and compositional nature of convolutions to achieve robust feature extraction across diverse grape varieties and geographical regions. Figure 1 shows an example of extracting horizontally and vertically oriented image features from an image belonging to the training set. Early layers typically capture local gradients, while deeper layers learn disease-specific morphology, such as lesion boundaries or necrotic spot textures. Pooling operators, often max-pooling, complement convolutions by introducing spatial downsampling, thereby reducing computational load while preserving the most informative activations. The general pipeline of the applied CNN model is presented in Figure 2. Nonlinearities such as ReLU further enhance the representational power of CNNs by enabling the modeling of complex, non-linear disease patterns.

In this work, CNN-based architectures are employed through the YOLOv7 detection framework, which integrates convolutional backbones with real-time object localization heads. The backbone network uses a sequence of convolutional blocks to build a multiscale representation capable of detecting both large and small disease spots. The detection head predicts bounding boxes and associated class probabilities using feature maps produced by these convolutional layers. YOLO-type networks rely on anchor-based or anchor-free mechanisms, but in either case, convolutional operators translate learned spatial patterns into actionable predictions. A crucial mathematical element is the IoU-based loss function, which quantifies the overlap between predicted and ground-truth regions and guides the optimization process during training. The model constantly updates convolutional parameters by minimizing this loss using stochastic gradient descent.

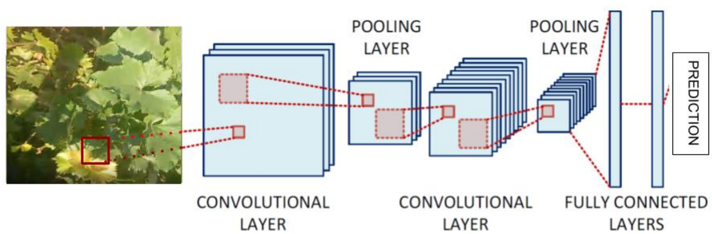

Figure 2: General pipeline of the applied Convolutional Neural Network model. The prediction (or output) contains a list of bounding box coordinates corresponding to the diseased leaf patches.

The extended pipeline also incorporates data augmentation, which can be viewed as applying mathematical transformations such as rotations, scalings, or color perturbations to expand the training distribution. These transformations improve the generalization of the convolutional filters, especially when the dataset contains images from geographically diverse wine regions. Transfer learning plays an additional role: convolutional backbones pretrained on large image datasets provide generic filters that are later fine-tuned to the specific spectral and structural characteristics of grape leaves. Because many disease features share cross-regional similarities, convolutional operators trained in one region can adapt effectively to datasets acquired elsewhere. This adaptability demonstrates the fundamental strength of CNNs as universal feature extractors.

Overall, the mathematical foundation of convolutional filtering, nonlinear activations, pooling, and gradient-based optimization [4, 5] enables reliable modeling of disease signatures in complex vineyard environments. The integration of these elements in CNNs such as YOLOv7 ensures that the system captures both fine-grained lesion patterns and broader contextual cues from aerial images. As a result, convolution-based architectures provide a scalable and region-transferable framework for automatic vine disease detection in precision agriculture.

3. Experimental results

Aerial data collection was conducted in four viticultural districts of the Carpathian Basin (Transylvania wine region in Romania, Eger wine region in Hungary, Východoslovenská wine region in Slovakia, and Fruška Gora wine region in Croatia) utilizing remotely operated quadcopters. All flights were conducted under mild weather circumstances (light wind and no precipitation) using manual piloting at low altitudes (maximum 6 meters above ground level).

For recording the dataset in a vineyard at Cluj-Napoca (Transylvania wine region), Szücsi (Eger wine region) and Ilok (Fruška Gora wine region), we used a small and light quadcopter, DJI Mini 3 with 12 MP integrated camera. At a vine parcel near Hrhov (Východoslovenská wine region), a DJI Phantom 3 Professional drone was employed. A DJI P4 Multispectral drone was also used at Cluj-Napoca. In all four locations, the video recordings and images were captured in the maximum resolution possible.

We compared various models trained in different geographical regions with our base model from Transylvania, which contains more than 10,000 annotated images and over 30,000 labels of manually annotated and validated grape leaves exhibiting abnormalities. This large-scale dataset enables geographically distributed analyses and comparisons.

We investigated the cross-domain relationship between the base Transylvanian region and the data collected in Szücsi, Hrhov, and Ilok regions. The results in terms of the Mean Average Precision (mAP) are shown in Figure 3. The mAP metrics [6] measures how well an object-detection model identifies and localizes objects by averaging the precision values computed at different recall levels. It summarizes the model's overall detection performance by combining both classification accuracy and bounding-box quality into a single score, form 0 to 1. Zero means the model detects nothing correctly, while one means perfect detection and localization. The mAP50 metric refers to the mean average precision at an intersection-over-union (IoU) threshold of 0.5, which measures how well the model predicts bounding boxes with at least 50% overlap with ground truth.

The presented results show that the best performance (highest precision) is achieved for the Cluj-Napoca region, which is a consequence of the larger available dataset used for the training process. The lower performance observed for the Szücsi region can be attributed to the fact that the grape leaves there differ significantly in texture, shape, and color from those in the other regions.

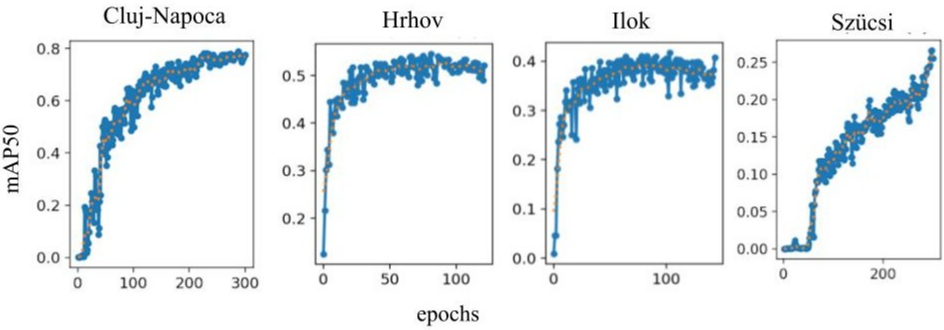

Figure 3: Mean Average Precision (mAP) metrics based on data collected in the four considered different geographical regions.

4. Conclusion

This paper presents research focused on examining disease detection in vineyards across the Carpathian Basin using low-altitude UAV platforms equipped with RGB cameras. Transfer learning plays a specific role, as convolutional backbones pretrained on large image datasets (Cluj-Napoca region) provide generic filters that are later fine-tuned to the specific spectral and structural characteristics of grape leaves in the other analysed regions (Hrhov, Ilok and Szücsi). Since many disease features share cross-regional similarities, we conclude that convolutional operators trained in one region can adapt effectively to datasets acquired in other regions. CNNs provide an excellent platform for this transfer-learning approach. As a possible future research direction, we foresee extending the field of application of the developed CNN-based algorithm, particularly to problems related to disease detection in various agricultural settings, especially for crops with characteristic leaves such as tobacco, cabbage or cotton.

Acknowledgement

The authors acknowledge the support received from the Hungarian Academy of Sciences through the Domus Hungarica project.

References

[1] L. Tamás, S. Gubo, and T. Lukić, "Vine Diseases Detection Trials in the Carpathian Region with Proximity Aerial Images," *Proceedings of the SAMI 2024*, 2024.

[2] S. Molnár and L. Tamás, “Segmentation methods evaluation on grapevine leaf diseases,” *Proceedings of the FedCSIS 2023*, 2023.

[3] S. Molnár, B. Keresztes, and L. Tam ́as, “Feature pyramid network based proximal vine canopy segmentation,” *Proceedings of the IFAC 2023*, 2023.

[4] T. Lukić and A. Lukity, “A Spectral Projected Gradient Optimization for Binary Tomography,” *in Computational Intelligence in Engineering, vol. 313 of SCI, pp. 263–272, Springer-Verlag*, 2010.

[5] T. Lukić and P. Balázs, “Moment preserving tomographic image reconstruction model,” *Image and Vision Computing, vol. 146, p. 105036*, 2024.

[6] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. da Silva, “A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit,” *Electronics, vol. 10, no. 3*, 2021.