**Title:** Scratch Vision Transformer Model for Diagnosis Grape Leaf Disease.

**Author name:** Sajib Bin Mamun, Md. Taimur Ahad, Md. Monzur Morshed, Nafiull Hossain, Yousuf Rayhan Emon.

**Abstract:** Recently, vision transformer has gained significant attention in the field of precision agriculture for its ability to transfer knowledge from pre-trained deep models for downstream tasks, particularly with limited datasets. However, very small studies shed light on the capabilities of grape disease detection. Grape is an important fruit worldwide, and early diagnosis and detection of grape diseases are crucial for ensuring plant health and preventing yield and quality reductions in the grape-growing industry. To fill the gap, this study developed a vision transformer model for the diagnosis of grape leaf disease. To provide a comprehensive understanding of ViT, in this study, the experiments were conducted with different image sizes and patch sizes of grape leaf images. Moreover, to extend the capabilities, both augmented and non-augmented datasets were used in the experiment configurations on two types of datasets, and the proposed model can provide quite similar output for both data (augmented and without augmented) in classifying grape disease, with approximately 98–99% validation and testing accuracy. This achievement highlights the potential of integrating advanced deep learning tools into the nation's agricultural practices. As worldwide fruit demands are increasing, this study provides a foundation for how machine learning techniques can be implemented to increase fruit production by identifying diseases.

**Keyword:** Vision Transformers, Precision, F1 Score, Loss function, Training Time.

**DOI:** https://doi.org/10.1007/978-981-97-1923-5_8