**Paper Title:** A deep learning based approach for automated plant disease classifcation using vision transformer

**Paper Link:** <https://www.nature.com/articles/s41598-022-15163-0>

1. **Summary**

The research paper introduces a novel deep learning approach for automated plant disease classification using Vision Transformer (ViT). The objective is to aid farmers in early disease diagnosis to ensure healthier crop yields. The proposed method is compared with traditional Convolutional Neural Networks (CNN) and a combined CNN-ViT model for disease classification. The study reveals that attention blocks enhance accuracy and slow down predictions. However, combining attention blocks with CNN blocks proves effective in maintaining speed. The paper emphasizes the importance of early disease diagnosis in agriculture and underscores the role of AI in providing accurate diagnoses to reduce harmful effects. The ViT structure, inspired by human focus mechanisms, is highlighted for its efficacy in image classification. Challenges associated with large datasets required for optimal parameter values in pre-designed architectures like ResNet and ViT are discussed. The paper also notes that transfer learning, while achieving high accuracy in similar domains, faces difficulties in diverse domains. It outlines future directions, including exploring object localization networks, multi-label classification, and further dataset collection.

* 1. **Motivation**

The research aims to supply visual information to empower farmers in implementing preventive measures against plant diseases, underscoring the critical role of early disease diagnosis in ensuring the production of healthy agricultural products. It emphasizes the potential of artificial intelligence to assist farmers in accurately identifying and diagnosing plant diseases. The study introduces the Vision Transformer (ViT) structure to enhance plant disease classification and proposes a lightweight deep learning approach specifically designed for real-time automated plant disease classification.

.

* 1. **Contribution**

The research significantly contributes to automated plant disease classification by proposing a lightweight deep-learning approach. The study encompasses implementing and evaluating Vision Transformer (ViT), both independently and in conjunction with classical convolutional neural networks (CNN), across diverse datasets. Notably, attention blocks enhance accuracy but at the expense of prediction speed, a limitation addressed by combining them with CNN blocks. The research underscores the pivotal role of early plant disease diagnosis for ensuring the production of healthy agricultural products. Additionally, it delves into the application of pre-designed architectures like ResNet and ViT for computer vision tasks, discusses the challenges of transfer learning on datasets with varying domains, and explores hand-designed architectures as potential solutions for optimal performance. The identified need for object localization networks and multi-label classification for automatic disease detection reflects a forward-looking perspective. The study introduces the matrix-based convolutional neural network (M-bCNN) structure, specifically tailored for wheat leaf disease detection. It conducts a comparative performance analysis of AlexNet and SqueezeNet for real-time applications. Overall, these contributions collectively advance the field of automated plant disease classification, offering valuable insights and potential solutions for enhancing accuracy and efficiency in agricultural practices.

* 1. **Methodology**

The methodology of the research centers around equipping farmers with visual information to facilitate proactive measures against plant diseases and ensuring the early detection of these diseases for the production of healthy agricultural products. To achieve these goals, the study employs artificial intelligence as a powerful tool to accurately diagnose plant diseases. The introduction of the Vision Transformer (ViT) structure is a key component of the methodology, aimed at enhancing the classification of plant diseases through a focus on specific image areas. Additionally, the research proposes a lightweight deep learning approach tailored for real-time automated plant disease classification, emphasizing efficiency in processing and predicting disease instances. This methodological framework combines advanced technologies, such as ViT and lightweight deep learning, to empower farmers with timely and accurate information for effective disease prevention and management in agriculture.

* 1. **Conclusion**

In conclusion, the research findings demonstrate that the Vision Transformer (ViT) outperforms traditional Convolutional Neural Networks (CNN) and hybrid models regarding classification accuracy. Notably, ViT exhibits superior accuracy with a reduced number of parameters compared to existing literature. However, the analysis indicates that the attention blocks within the ViT architecture were slower in prediction across all devices when compared to convolutional blocks. A strategic enhancement to address this limitation involved combining attention blocks with convolutional blocks, improving prediction speed and accuracy. Looking ahead, the study highlights the necessity for incorporating object localization networks and embracing multi-label classification methodologies further to advance the capabilities of automatic disease detection systems. These conclusions collectively contribute valuable insights to plant disease classification, emphasizing the trade-offs and potential enhancements associated with attention mechanisms in Vision Transformer architectures.

.

1. **Limitations**
   1. **First Limitation**

The study reveals that attention blocks play a pivotal role in improving the accuracy of the model for plant disease classification. However, a notable trade-off is observed, as the integration of attention blocks is found to introduce a slowdown in prediction speed. While attention mechanisms contribute to the model's ability to focus on relevant features and enhance accuracy, the increased computational complexity delays the prediction process. This trade-off underscores the need for careful consideration of the computational efficiency of attention mechanisms, especially in real-time applications where timely predictions are crucial. Balancing accuracy and computational efficiency remains a key challenge in optimizing deep learning models for plant disease diagnosis.

* 1. **Second Limitation**

Moreover, the research sheds light on the computational burden of pre-designed architectures. These architectures, known for their robustness and effectiveness in various tasks, are characterized by heavy models with extensive computational requirements. The computational overhead poses challenges, particularly in resource-constrained environments or applications with strict latency requirements. As such, the study underscores the importance of exploring lightweight architectures or optimization techniques to balance model complexity and computational efficiency. This consideration becomes particularly relevant in deploying automated plant disease classification systems in real-world scenarios where computational resources may be limited.

1. **Synthesis**

Researchers are continuously finding new ways to implement deep learning and machine learning in early detection of plant disease. All these techniques use pattern recognition, which showed promising results; the proposed method architecture in the paper can be extended to handle more complex plant disease image classifications as breakthroughs keep occurring in computer vision. A thorough investigation must be done on multi-modal data by combining various plant leave images.