**Report on Vision Transformer (ViT) for Mango Leaf Disease Classification (MangoLeafBD Dataset)**

**Displays the ViT model’s predictions on mango leaf images.**

***Examples:***

* **Mango leaf with "Anthracnose" → Predicted: Anthracnose | True: Anthracnose**
* **Mango leaf with "Bacterial Canker" → Predicted: Healthy | True: Bacterial Canker**

A collage of a leaf

AI-generated content may be incorrect.

**2. Attention Visualization**

Highlights which regions the model used to make its decision.

*Examples:*

* Discoloration spots
* Irregular veins or edge damage

A chart of a graph

AI-generated content may be incorrect.

**Interpretation of Attention Maps**

The attention maps from the ViT model show it tends to focus on:

* **Discolored areas**, especially dark patches or fungal-like growths.
* **Edges of the leaf**, where bacterial or fungal infection symptoms often first appear.
* **Lesion textures**, such as raised or sunken areas that indicate disease presence.

Unlike CNNs, which learn primarily from spatial proximity, ViTs consider **global context**, which helps them focus on **non-contiguous but related regions**.

**How does attention differ from CNN feature maps?**

| **Feature** | **CNN** | **Vision Transformer (ViT)** |
| --- | --- | --- |
| Architecture | Convolutional layers | Self-attention layers |
| Focus | Local (small receptive fields) | Global (entire image patches) |
| Learning Style | Bottom-up (local to global) | Simultaneous patch-level relationships |
| Feature Visualization | Feature maps | Attention heatmaps |

CNNs extract features progressively and are excellent at textures and edges. ViTs, however, capture both **local and global patterns** at once, which is very useful for disease detection where **context across the leaf** matters.

**Challenges in Classifying Mango Leaf Diseases:**

1. **Visual Similarity** – Diseases like Bacterial Canker and Anthracnose can have similar symptoms.
2. **Inconsistent Lighting** – Variations in image brightness and contrast affect generalization.
3. **Background Noise** – Leaves photographed in varied environments (soil, sky, other plants).
4. **Disease Staging** – Symptoms can differ across stages (early vs. late Anthracnose).
5. **Class Imbalance** – Some diseases are underrepresented, affecting learning balance.