INDIAN INSTITUTE OF INFORMATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

PROJECT PRESENTATION

on

Plant Leaf Classification and Information Retrieval Using Vision Transformers and Transfer Learning

Submitted by

Manish kumar

Roll No.:- 2101114

Dept:- CSE

Project Supervisor

Dr. Subhasish Dhal

Content

Background

- **Significance:** A reliable and efficient system for identifying plant species, especially in agriculture, forestry, and environmental conservation.
- **Manual Challenges:** Manual classification is time-consuming and requires expert knowledge.
- **Limitations of Existing Tools**: Existing tools may not be tailored for specific applications and often provide a broad range of results, which can be overwhelming and inaccurate.
- **Automation Benefits**: Advances in deep learning and computer vision provide faster, more reliable, and scalable plant classification.
- **Real-Time Impact**:Real-time classification, such as improving productivity for farmers and enabling rapid decision-making in critical situations like disease outbreaks in crops.

Literature review

1. Plant Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques[1]

- Their study focuses on automated plant species identification using plant leaf images, employing **Computer Vision** and **machine learning** techniques.
- The system involves four main steps: image acquisition, pre-processing, feature extraction, and classification using **Multiclass-Support Vector Machine (MSVM)**.
- The Swedish Leaf Dataset, containing **1,125 images** of **15 species**, was used, achieving a classification accuracy of **93.26%**.

[1] Kaur, Surleen & Kaur, Prabhpreet. (2019). Plant Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques. Journal of Multimedia Information System. 6. 49-60. 10.33851/JMIS.2019.6.2.49.

Cont’d

2. Deep Learning for Plant Species Classification[2]

- The paper presents an automated plant species classification system using a **Convolutional Neural Network (CNN)**.
- The model includes multiple layers, such as convolutional layers for extracting detailed features, pooling layers to reduce dimensionality, and fully connected layers for the final classification.
- The process involves **image pre-processing** , **feature extraction** (texture and color), and classification with a CNN implemented using **TensorFlow**.
- The model was trained on a dataset of **17 plant species**, achieving an accuracy of **94.26%**.

[2] Kiran S.Gawli Ashwini S. Gaikwad Mtech 2nd year Student Assistant professor Department Of Computer Science & Engg. Department Of Computer Science & Engg. Deogiri Institute Of Engg.& Management Deogiri Institute Of Engg.& Management Studies, Aurangabad. Studies, Aurangabad.

Cont’d

3. Alexey Dosovitskiy et al., 2021[3]

- This study explores the application of Vision Transformers (ViTs) for image classification tasks, traditionally dominated by convolutional neural networks (CNNs).
- They treat images as sequences of patches, similar to **NLP tokens**, avoiding CNN biases.
- ViTs can outperform CNNs like **ResNets** when pre-trained on large datasets (e.g., ImageNet-21k and JFT-300M).
- Achieved high accuracy: **88.55%** on **ImageNet** and **94.55%** on **CIFAR-100** with reduced computational effort.

[3] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. Journal of Images. viii, 4, 12

Shortcomings

- It uses a dataset of limited number of **plant species**, which restricts the model's ability to generalize effectively across a larger, more diverse range of plants.
- The dataset appears to be captured under controlled conditions, which might not accurately reflect the challenges of real-world environments, like different lighting and backgrounds.
- The research does not focus on real-time classification, which could limit the usability of the system in practical scenarios where immediate results are required.
- It lacks comparison to newer deep learning models, which could provide better performance and robustness.

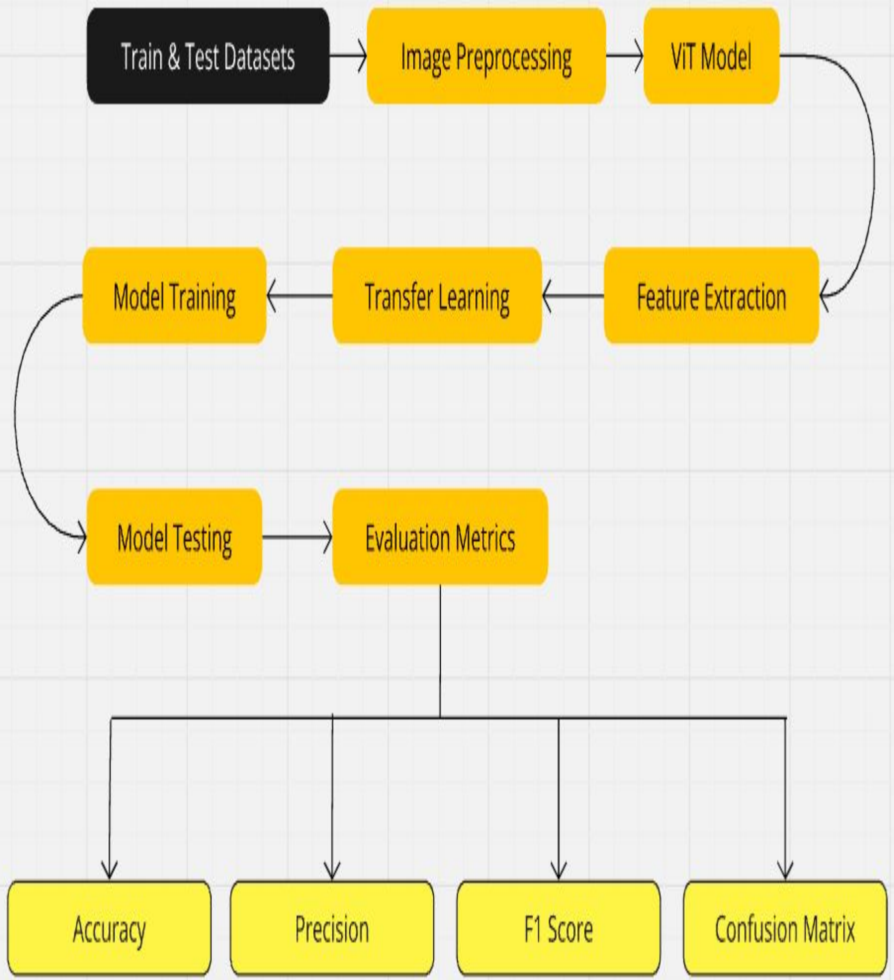
Problem Formulation

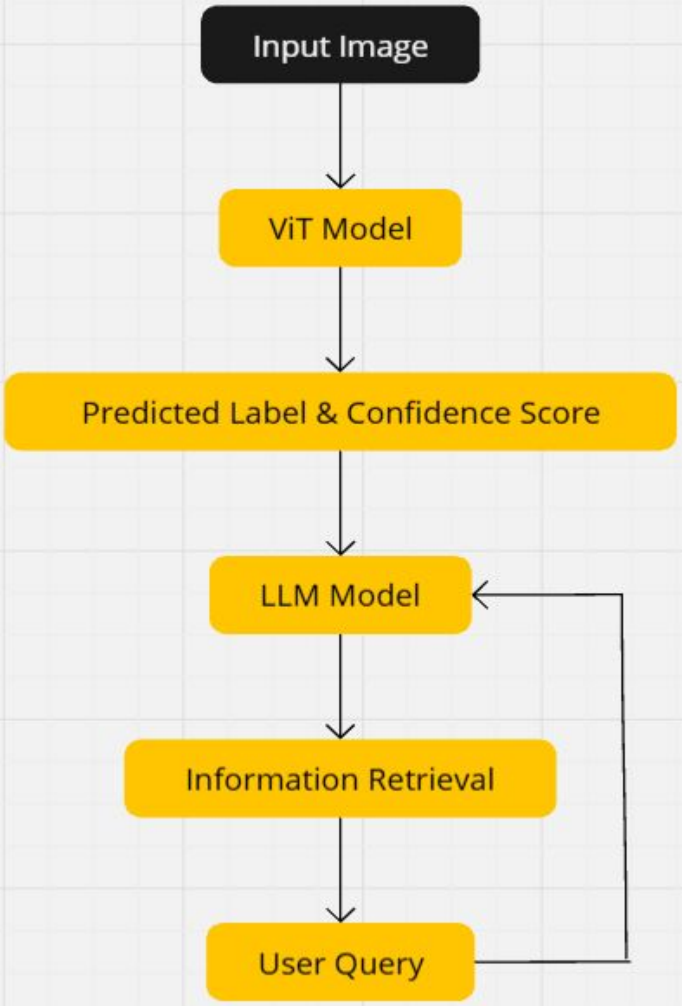
- Current plant leaf classification systems, including CNNs, struggle to capture **global features**, often overfit on limited datasets, and lack adaptability to varying scales, leading to reduced accuracy and practical usability.
- This project develops a real-time plant leaf classification system using **Vision Transformers (ViTs)[3]** and **transfer learning** for accurate species identification.
- ViTs use a **self-attention** mechanism to capture **global image features**, overcoming CNN limitations and improving adaptability across varying scales.
- By leveraging pretraining and transfer learning, ViTs reduce overfitting, allowing the model to achieve high performance with less data.
- An integrated chatbot powered by **Google Gemini Pro AI** enhances user experience by providing instant, detailed information on identified plant species.

Methodology

Building an advanced plant species classification system, integrating deep learning, transfer learning, and AI-driven information retrieval.

- **Model :** Fine-tuned Vision Transformer (ViT) for plant classification using transfer learning.
- **AI-Powered Chatbot:** Google Gemini Pro LLM provides detailed plant insights interactively.
- **Streamlit Interface:** Real-time plant classification via a user-friendly GUI.


Train & Test Datasets
Image Preprocessing
ViT Model
Feature Extraction
Transfer Learning
Model Training
Model Testing
Evaluation Metrics
Accuracy
Precision
F1 Score
Confusion Matrix


Input Image
ViT Model
Predicted Label & Confidence Score
LLM Model
Information Retrieval
User Query

Algorithm

- **Image Preprocessing:** Automate data handling with resizing, normalization, and augmentation (rotations, flips, shifts) for better generalization.
- **Feature Extraction:** Utilize Vision Transformers (ViT) pre-trained on ImageNet by dividing images into patches, removing top layers, and adding a custom classifier for plant species classification.
- **Transfer Learning:** Freeze initial layers of ViT, add a custom classifier (Linear layer (512 units) → ReLU activation → Dropout (0.5) → Output layer (matching number of classes)) and fine-tune final layers on the plant dataset for enhanced accuracy and faster convergence.
- **Model Training:** Employed Adam optimizer (learning rate 0.001), Cross-entropy loss, learning rate scheduler (0.1 reduction every 5 epochs), and early stopping (patience of 3 epochs) to optimize model performance and prevent overfitting.

Cont’d

- **Model Testing:** The model was evaluated on the test dataset using batch processing to calculate validation loss and accuracy, achieving a final accuracy score based on correct predictions against the total samples.
- **Generative AI Integration:** Utilizes Google Generative AI to provide detailed information on plant leaves, seamlessly integrated with a chatbot powered by Gemini Pro LLM for real-time interaction.
- **Streamlit GUI Integration:** Provides an interactive, web-based interface for users to upload plant leaf images.

TECHNOLOGY SUMMARIZATION

- **Framework:** PyTorch
- **Python Version:** 3.11.5
- **Key Libraries:** torch, torchvision, transformers, scikit-learn, matplotlib, seaborn
- **Generative AI Integration:** langchain-google-genai, langchain[openai,all], langchain-community, sentence-transformers
- **Hardware:** NVIDIA A100 GPUs with CUDA
- **GUI Development:** Streamlit

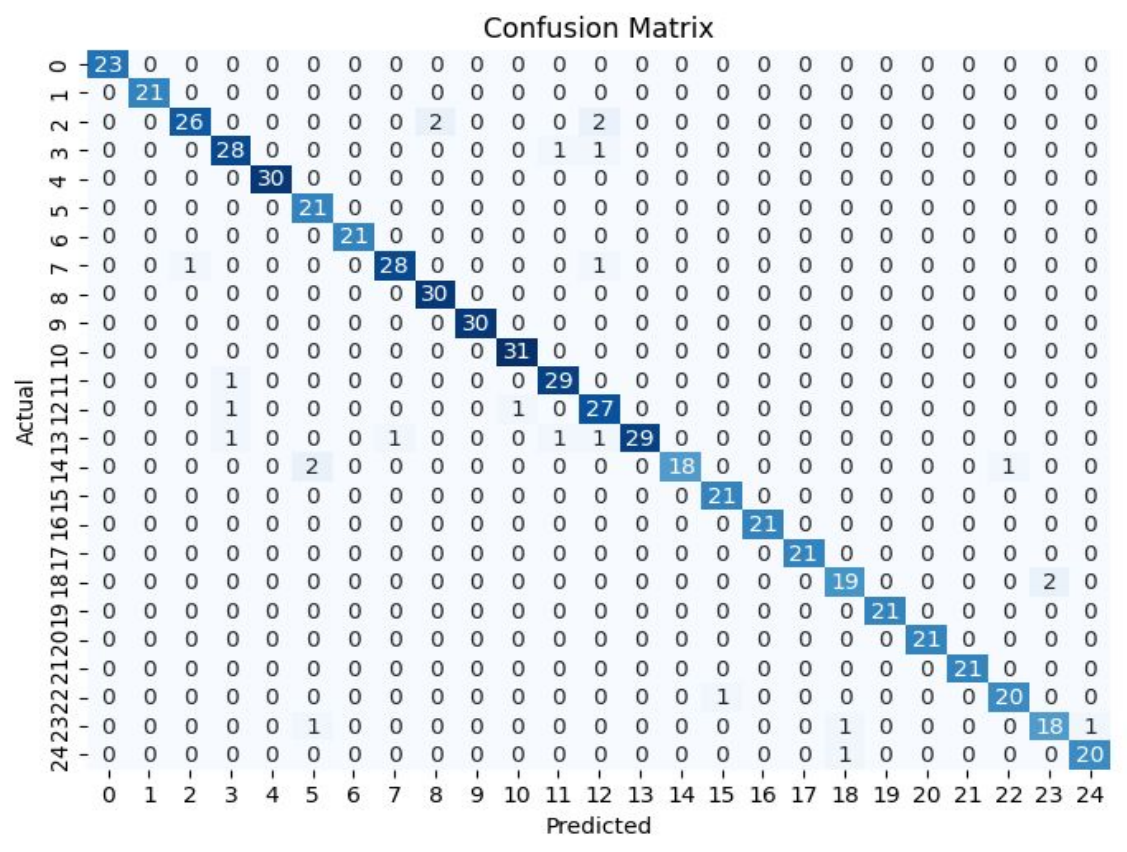
Results

Model Training Result

```
Epoch 1/20, Training Loss: 2.1874
Validation Loss: 0.9478, Accuracy: 0.8452
Epoch 2/20, Training Loss: 0.8483
Validation Loss: 0.4636, Accuracy: 0.8677
Epoch 3/20, Training Loss: 0.5456
Validation Loss: 0.3156, Accuracy: 0.9403
Epoch 4/20, Training Loss: 0.4128
Validation Loss: 0.2514, Accuracy: 0.9484
Epoch 5/20, Training Loss: 0.3164
Validation Loss: 0.2217, Accuracy: 0.9403
Epoch 6/20, Training Loss: 0.2795
Validation Loss: 0.1996, Accuracy: 0.9597
Epoch 7/20, Training Loss: 0.2570
Validation Loss: 0.1911, Accuracy: 0.9532
Epoch 8/20, Training Loss: 0.2506
Validation Loss: 0.1873, Accuracy: 0.9581
Epoch 9/20, Training Loss: 0.2501
Validation Loss: 0.1845, Accuracy: 0.9597
Epoch 10/20, Training Loss: 0.2447
Validation Loss: 0.1826, Accuracy: 0.9581
Epoch 11/20, Training Loss: 0.2372
Validation Loss: 0.1783, Accuracy: 0.9597
Epoch 12/20, Training Loss: 0.2412
Validation Loss: 0.1798, Accuracy: 0.9597
Epoch 13/20, Training Loss: 0.2375
Validation Loss: 0.1775, Accuracy: 0.9597
Epoch 14/20, Training Loss: 0.2353
Validation Loss: 0.1796, Accuracy: 0.9597
Epoch 15/20, Training Loss: 0.2330
Validation Loss: 0.1806, Accuracy: 0.9597
Epoch 16/20, Training Loss: 0.2372
Validation Loss: 0.1768, Accuracy: 0.9597
Epoch 17/20, Training Loss: 0.2330
Validation Loss: 0.1780, Accuracy: 0.9597
Epoch 18/20, Training Loss: 0.2379
Validation Loss: 0.1785, Accuracy: 0.9597
Epoch 19/20, Training Loss: 0.2311
Validation Loss: 0.1768, Accuracy: 0.9597
Early stopping triggered.
```

```
Precision: 0.96, Recall: 0.96, F1 Score: 0.96
```

Cont’d



Conclusion

The model trained on a dataset of **25 classes** with a total of **1,875 images**, achieved a high accuracy of **95.97%**, with precision, recall, and F1-score each at **0.96**, demonstrating the effectiveness of **Vision Transformers and transfer learning** in automating plant leaf classification. This significantly reduces manual effort in species identification.

Additionally, integrating a chatbot powered by Google Generative AI (Gemini Pro) enriches user interaction by delivering real-time, informative insights about identified plant species, enhancing both accessibility and engagement.

Future Works

- **Multilingual Support:** Implementing multiple language options in both the classification model and chatbot to broaden accessibility, especially in regions critical to plant conservation and research.
- **Disease Detection:** Expanding the model to identify plant diseases from leaf symptoms, enabling early intervention in agriculture and aiding conservation efforts.

References and Links :

[1] Kaur, Surleen & Kaur, Prabhpreet. (2019). Plant Species Identification based on Plant Leaf Using Computer Vision and Machine Learning Techniques. Journal of Multimedia Information System. 6. 49-60. 10.33851/JMIS.2019.6.2.49.

[2] Kiran S.Gawli Ashwini S. Gaikwad Mtech 2nd year Student Assistant professor Department Of Computer Science & Engg. Department Of Computer Science & Engg. Deogiri Institute Of Engg.& Management Deogiri Institute Of Engg.& Management Studies, Aurangabad. Studies, Aurangabad.

[3] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. Journal of Images. viii, 4, 12

THANK YOU