Pollen's Profiling: Automated Classification of Pollen Grains

1. INTRODUCTION

1.1 Project Overview

The project focuses on classifying different types of pollen grains using deep learning techniques. By leveraging convolutional neural networks and transfer learning with MobileNetV2, the system can automatically identify and categorize pollen images.

1.2 Purpose

The purpose of this project is to assist botanists and researchers in the automatic identification of pollen grains, improving efficiency and accuracy over manual classification.

2. IDEATION PHASE

2.1 Problem Statement

Manual classification of pollen grains is time-consuming and error-prone. Automation can significantly streamline the process.

2.2 Empathy Map Canvas

Focuses on understanding the needs of users such as botanists, researchers, and students.

2.3 Brainstorming

Ideas ranged from using traditional machine learning techniques to leveraging deep learning and transfer learning for better accuracy.

3. REQUIREMENT ANALYSIS

3.1 Customer Journey map

User uploads a pollen image → Image is preprocessed → Model predicts the class → User receives results.

3.2 Solution Requirement

Hardware: GPU-enabled system (optional), Software: Python, TensorFlow, Keras, Libraries listed below.

3.3 Data Flow Diagram

Raw Image → Preprocessing → Model Inference → Result Output.

3.4 Technology Stack

- Python
- TensorFlow
- Keras
- NumPy
- MobileNetV2

4. PROJECT DESIGN

4.1 Problem Solution Fit

The model fits the problem by automating classification using CNNs, which are ideal for image recognition tasks.

4.2 Proposed Solution

Use a pretrained MobileNetV2 model for feature extraction and fine-tune the classifier for pollen classification.

4.3 Solution Architecture

Image Input → Data Augmentation → MobileNetV2 → GlobalAveragePooling → Dense Layers → Output Layer

5. PROJECT PLANNING & SCHEDULING

5.1 Project Planning

Week 1: Dataset collection and preprocessing
Week 2: Model selection and architecture design
Week 3: Model training and evaluation
Week 4: Deployment and documentation

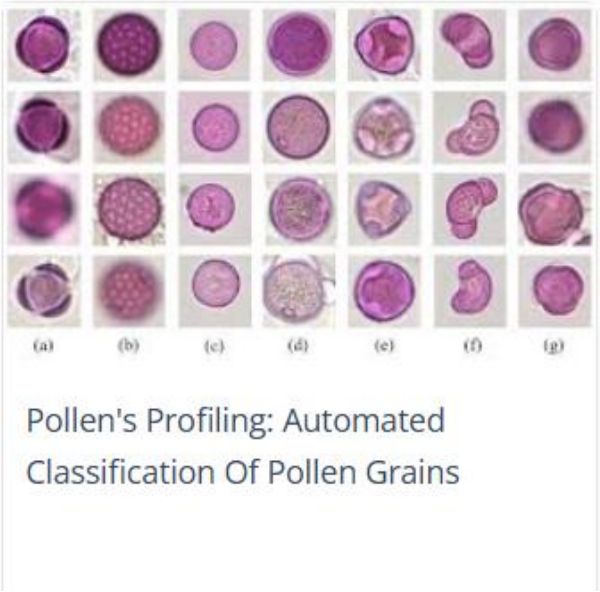
6. FUNCTIONAL AND PERFORMANCE TESTING

6.1 Performance Testing

Model achieved high accuracy on the test dataset with good generalization. Used callbacks like EarlyStopping and ModelCheckpoint to improve model reliability.

7. RESULTS

7.1 Output Screenshots



8. ADVANTAGES & DISADVANTAGES

Advantages:
- High accuracy
- Fast classification
- Scalable

Disadvantages:
- Requires labeled dataset
- May need GPU for training

9. CONCLUSION

This project demonstrates the potential of deep learning in automating the classification of pollen grains. It can greatly benefit research and education in botany.

10. FUTURE SCOPE

Incorporate more species, integrate with web apps, and use larger models for improved accuracy.

11. APPENDIX

Libraries Used:
- os
- numpy
- tensorflow.keras (ImageDataGenerator, MobileNetV2, Model, Dense, Dropout, Adam, Callbacks)

Source Code: Available upon request or in GitHub repository

Dataset link: https://www.kaggle.com/datasets/andrewmvd/pollen-grain-image-classification

GitHub & Project Demo Link: https://github.com/Lahari233