POL-ID: AI-Based Honey Authentication

Introduction

Honey authentication is critical for consumer trust and trade. Manual identification of pollen under a microscope is time-consuming, labour-intensive and error-prone. Our goal is to automate pollen detection and classification to support apiculture and food traceability.

Objectives

We aim to develop a region-specific deep learning pipeline capable of detecting pollen grains in microscopy images, classifying South African pollen taxa, and clustering novel or unlabeled types. This integrated approach will be tailored to South African pollen.

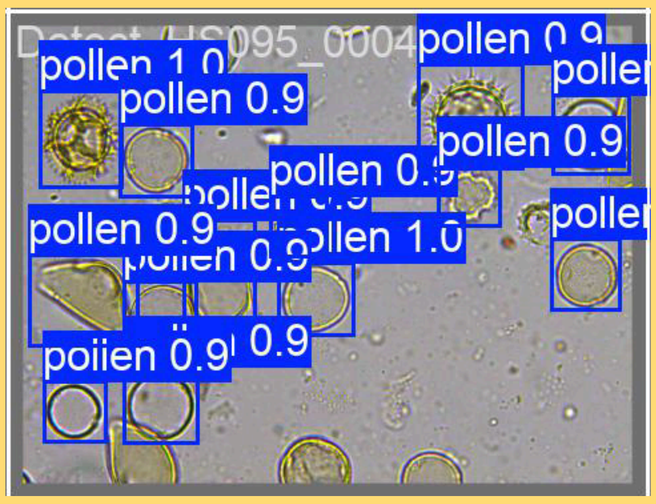
Methodology

We use **YOLOv11** to detect pollen grains and generate bounding boxes in microscopy images. Classification is carried out with three complementary models: a **Convolutional Neural Network (CNN), a Vision Transformer (ViT), and a hybrid CNN–Transformer model**. To identify pollen types not present in training, **HDBSCAN** clustering is applied to feature embeddings, enabling the discovery of novel groups. Performance is evaluated using **mAP, precision, recall, F1-score, and silhouette score**, ensuring both accurate classification and meaningful clustering results.

Results

Detection

98.8% mAP@50



98.4% Recall

92.8% mAP@50-95

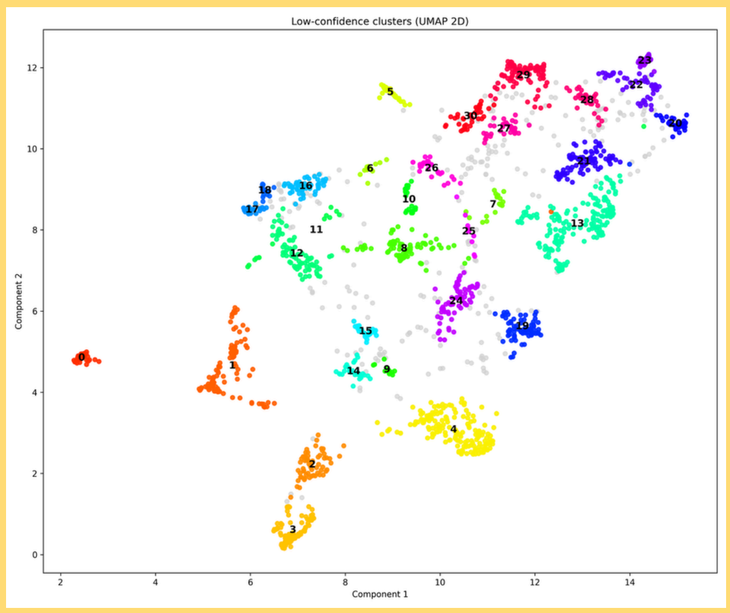
Classification

CNN Classifier
94.43% test accuracy
0.916 F1-score

Hybrid Classifier
96.5% test accuracy
0.960 F1-score

VIT Classifier
92.64% test accuracy
0.902 F1-score

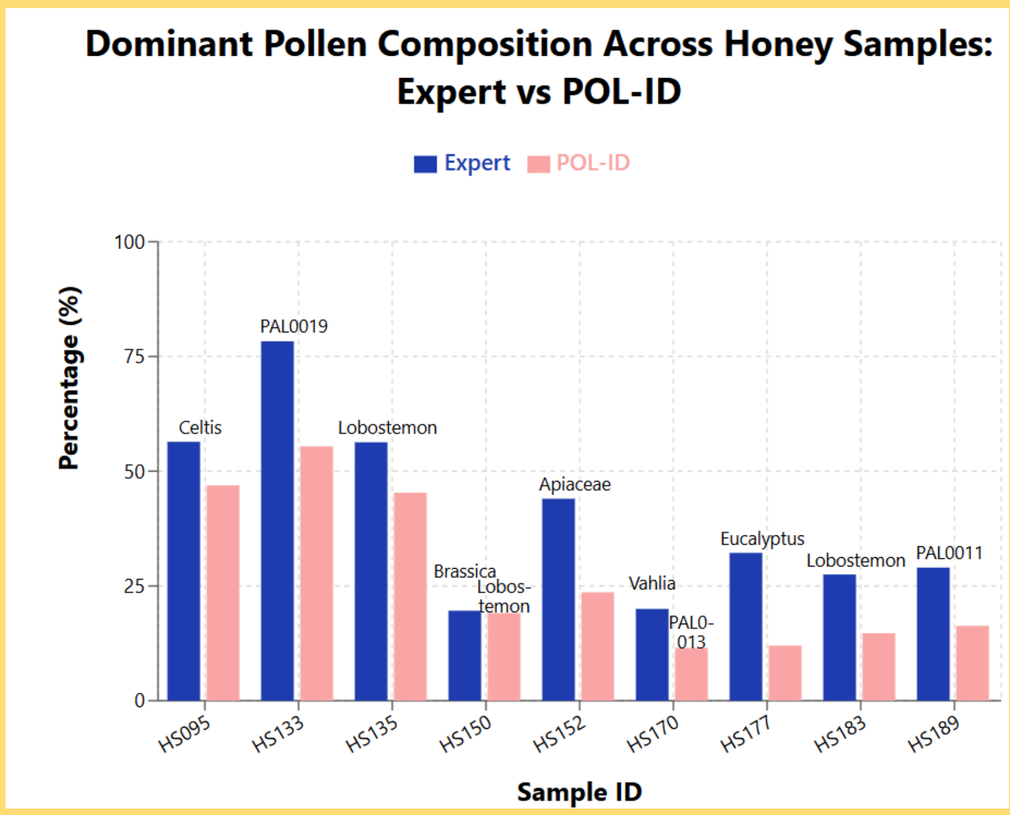
Clustering



Noise fraction = 12.8%

Pipeline

- Best performing pipeline was the CNN pipeline
- All 9 samples correctly classified according to ICBB standards



- Many clusters were dominated by a specific pollen class
- Class imbalances need to be reduced + increased dataset sizes

Authors
Maryam Mather: MTHMAR046@myuct.ac.za
Lillian Mtumanje: MTMLIL003@myuct.ac.za
Yash Ramklass: RMKYAS002@myuct.ac.za

Supervisor: Patrick Marais
Co-supervisors: Cesarina Edmonds-Smith and Janais Delport

SCHOOL OF IT