Members: **Gowtham Nimmal Haribabu, Gaaurav Lath**

Mask-Free Nanoscale Feature Detection Using Self-Supervised Learning

1. Introduction

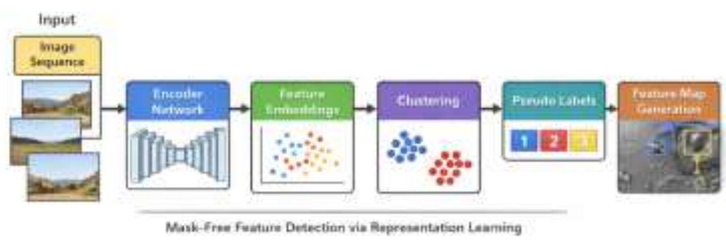
High-resolution microscopy images contain rich nanoscale information critical for materials characterization, including defects, grain boundaries, and nanoparticles. Manual analysis of such data is slow and does not scale. This project proposes a **mask-free, deep learning–based feature detection pipeline** that automatically identifies salient nanoscale regions without requiring pixel-level annotations or supervised model training.

The approach leverages self-supervised Vision Transformer embeddings to enable scalable and annotation-efficient analysis.

2. Pipeline Overview

The end-to-end workflow consists of:

1. Preprocessing and HDF5 dataset construction from raw TIFF microscopy images
2. Feature extraction using a frozen DINOv2 Vision Transformer (ViT-B/14)
3. Mask-free top-k patch pooling to emphasize informative regions
4. Embedding analysis, visualization, and unsupervised clustering



3. Data Preparation

Microscopy images are processed as follows:

- Conversion to single-channel grayscale
- Percentile-based intensity normalization (1–99% or 2–98%)
- Storage in HDF5 format with labels extracted from filename prefixes

Both full-image and fixed-size patch (64×64) datasets are supported, along with metadata such as file paths and label names.

GitHub Link: https://github.com/nhgowtham/hackathon_gg.git

4. Feature Representation with DINOv2

A frozen **DINOv2 ViT-B/14** model is used as a universal feature extractor. Grayscale images are converted to RGB by channel replication, resized to 224×224, and normalized using ImageNet statistics. The model outputs patch-level embeddings (768D) that capture local structural information.

5. Mask-Free Top-k Patch Pooling

To avoid explicit ROI masks, a saliency-driven pooling strategy is applied:

- Patch tokens are ranked by L2 norm
- The top-k most informative patches are selected
- Selected embeddings are averaged into a single image-level vector

A sweep over k ∈ {8, 16, 32, 64} shows that k = 32 provides stable and discriminative representations.

6. Analysis and Unsupervised Discovery

The resulting embeddings are evaluated using:

- UMAP (or PCA fallback) for 2D visualization
- A separation proxy based on between- and within-class variance
- Bulk-centric anomaly scoring using cosine distance from the bulk centroid
- Unsupervised clustering via K-means and agglomerative clustering

These analyses demonstrate that the embeddings support meaningful separation and discovery of nanoscale features.

7. Conclusion

This work presents a compact, mask-free pipeline for nanoscale feature detection in microscopy images. By combining robust preprocessing, self-supervised ViT embeddings, top-k patch pooling, and unsupervised analysis, the approach enables scalable and annotation-efficient materials characterization. The method provides a strong foundation for future extensions such as domain-specific fine-tuning or hybrid localization techniques.