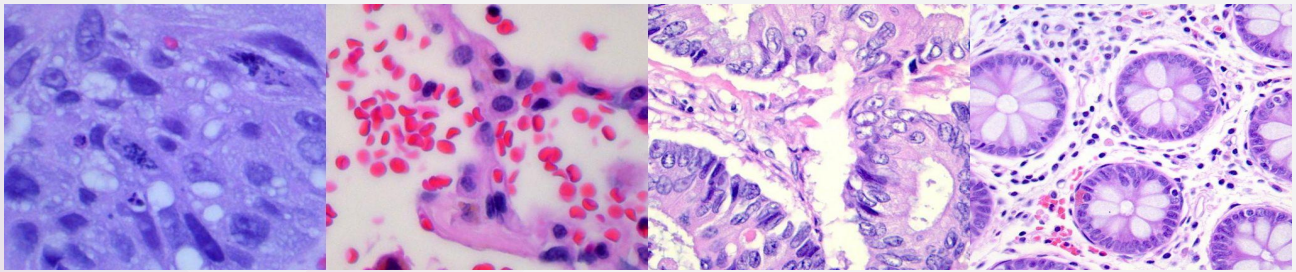
Universität Stuttgart

Dual Approach: DINO and CLIP for Lung and Colon Cancer Image Analysis

Presenters:

Hananeh Shirzadnia
Remziye Maral Demirsecen
Viraat Saaran

Winter Term 2024/2025



Motivation

- Improving early and accurate lung and colon cancer detection
- Handling the complexity of histopathological images
- Comparing the results of **DINO** and **CLIP** to determine which model yields more precise results.

Contents

01 Related Work

02 Key Contributions

03 Methodology: DINO

04 Performance Insights: DINO

05 Methodology: CLIP

06 Performance Insights: CLIP

07 Summary

Related Work

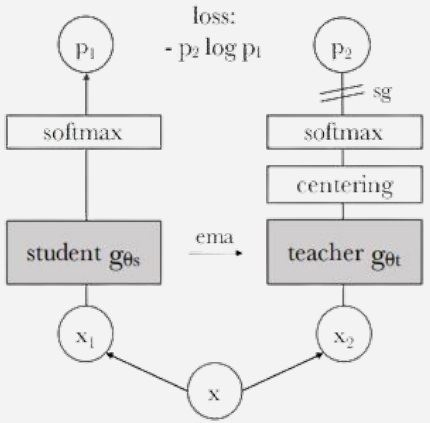
- **Self-supervised learning:** Analyzing large and unlabeled histopathological image datasets
- **Vision Transformers (ViTs):** Modeling local and global features in tissue images
- **TransPath**: Pre-training vision transformers for robust histopathological classification
- **Virchow Foundation Model**: Pan-cancer detection with large-scale pre-training

Key Contributions

Objectives

- Applying **DINO** (self-supervised learning) for feature extraction in histopathological images of lung and colon cancer

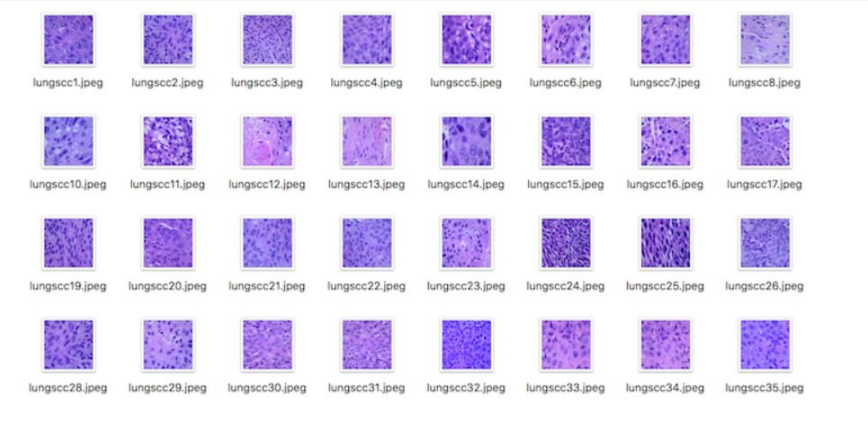
- Utilizing **CLIP** to understand the semantic relationship between histopathology images and text pairs



Methodology

LC25000 Dataset

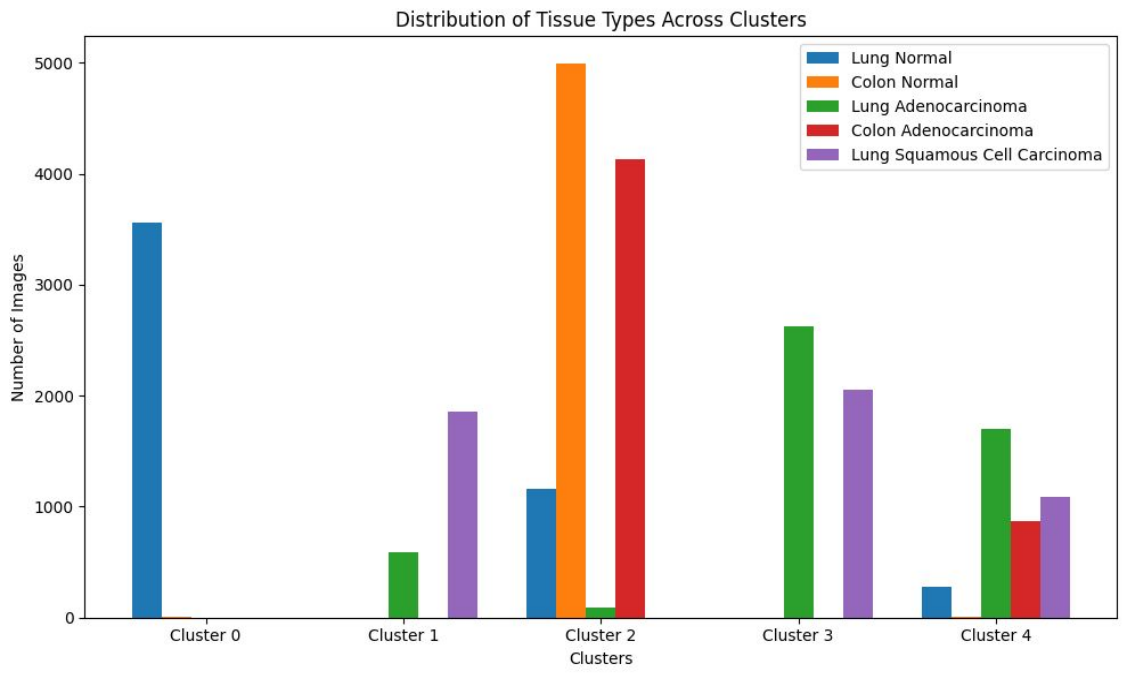
- **Images**: 25,000 total, equally distributed across 5 classes
- **Colon classes**: adenocarcinoma, normal
- **Lung classes**: adenocarcinoma, squamous cell carcinoma, normal



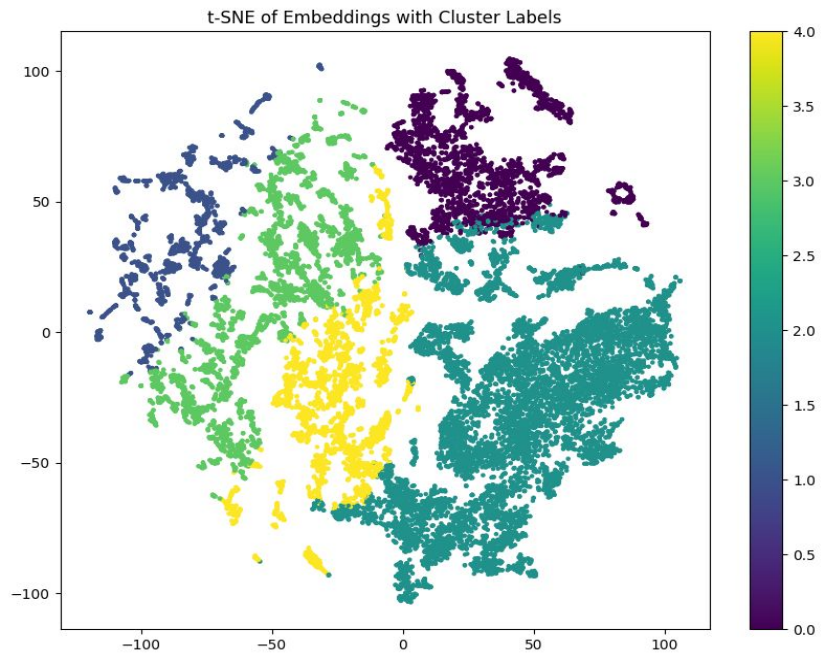
Methodology: DINO

- Combining transformers and CNNs for better performance
- Grouping similar features effectively
- Learning from unlabeled image data

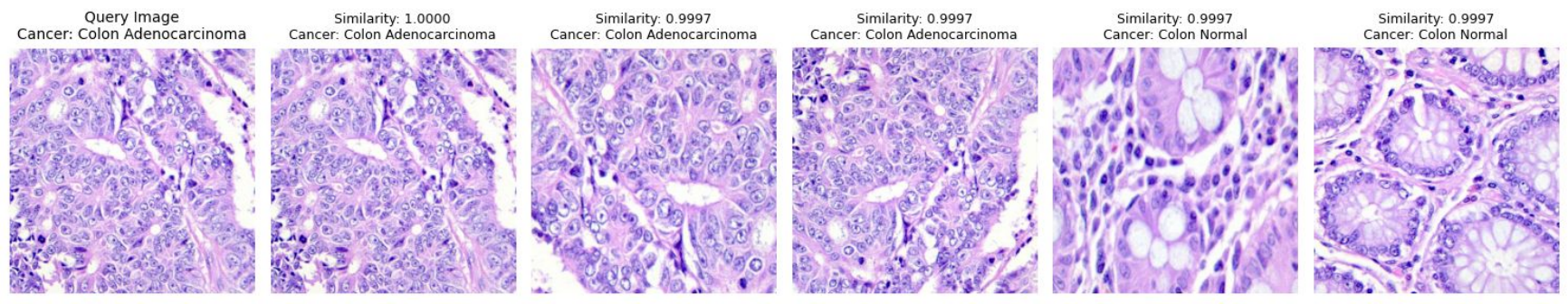
Performance Insights: DINO



Performance Insights: DINO



Performance Insights: DINO

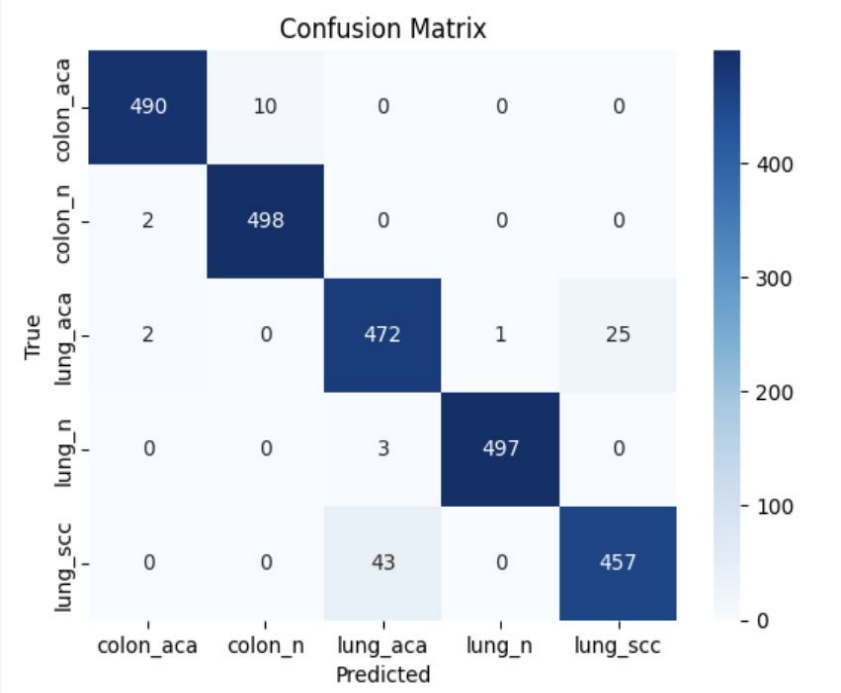


Performance Insights: DINO

Methodology: CLIP

- Learning from image-text pairs
- Applying transformers for text and Vision Transformers (ViT) for images
- Aligning embeddings in a shared space through contrastive learning

Performance Insights: CLIP



Performance Insights: CLIP

Classification Analysis

	precision	recall	f1-score	support
colon_aca	0.99	0.98	0.99	500
colon_n	0.98	1.00	0.99	500
lung_aca	0.91	0.94	0.93	500
lung_n	1.00	0.99	1.00	500
lung_scc	0.95	0.91	0.93	500
accuracy			0.97	2500
macro avg	0.97	0.97	0.97	2500
weighted avg	0.97	0.97	0.97	2500

Summary

- Pretraining DINO and CLIP models on 25,000 histopathology images from lung and colon tissues
- Applying clustering and t-SNE for visual insights
- Achieving 92% classification accuracy using Random Forest
- Attaining 96% classification accuracy with the CLIP model

Thank you!

Any Questions?

Resources

Dataset:

- Lung and Colon Cancer Histopathological Images

Image:

- DINO Model

Papers:

- Kumar, R. et al., "Detection and Classification of Cancer from Microscopic Biopsy Images," *BioMed Research International*, 2015.
- Deininger, L. et al., "A Comparative Study Between Vision Transformers and CNNs in Digital Pathology," *F. Hoffmann-La Roche AG*, 2023.
- Wang, X. et al., "TransPath: Transformer-based Self-supervised Learning for Histopathological Image Classification," *CVPR*, 2021.
- Vorontsov, E. et al., "A Foundation Model for Clinical-Grade Computational Pathology and Rare Cancers Detection," *arXiv preprint arXiv:2309.07778*, 2023.
- Borkowski, A. A. et al., "Lung and Colon Cancer Histopathological Image Dataset (LC25000)," *arXiv preprint arXiv:1912.12142*, 2019.