August 22, 2024

Aly Khan Nuruddin

Expression-Based Molecular Subtyping & Classification of Bladder Cancer TCGA Data

Artificial Intelligence in Medicine (AIM) Lab

A Consensus Molecular Classification of Muscle-Invasive Bladder Cancer (MIBC)

Aurelie Kamoun, Aurelien de Reynies, Yves Allory, Gottfrid Sjodahl, A. Gordon Robertson, Roland Seiler, Katherine A. Hoadley, Clarice S. Groeneveld , Hikmat Al-Ahmadie, Woonyoung Choi, Mauro A.A. Castro, Jacqueline Fontugne, Pontus Eriksson, Qianxing Mo, Jordan Kardos, Alexandre Zlotta, Arndt Hartmann, Colin P. Dinney, Joaquim Bellmunt, Thomas Powles, Nuria Malats, Keith S. Chan, William Y. Kim, David J. McConkey, Peter C. Black, Lars Dyrskjøt, Mattias Hoglund, Seth P. Lerner, Francisco X. Real, Francois Radvanyi

European Association of Urology (EAU).

Overview

Bladder Cancer:

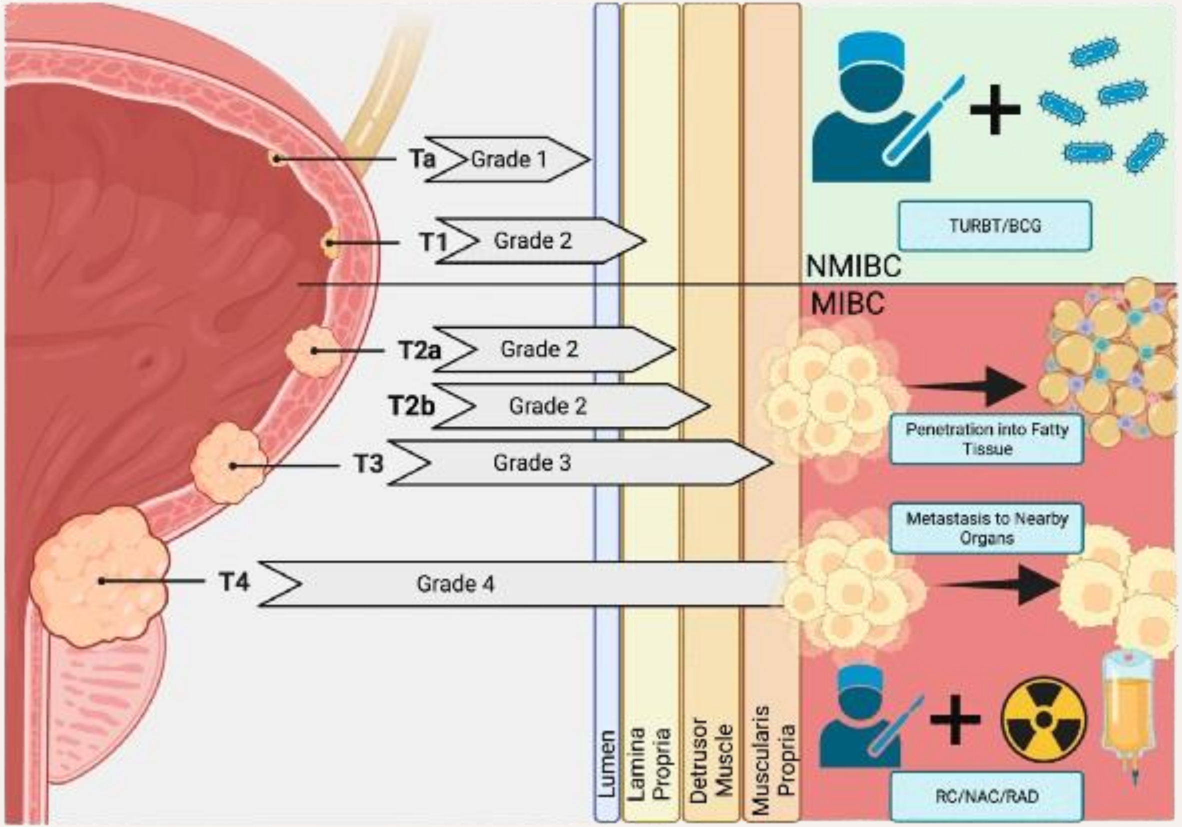
Urothelial Carcinoma Originating in Tissues of the Urinary System

Non-Muscle Invasive Bladder Cancer

Superficial Cancer Confined to Mucosa Layer

Muscle-Invasive Bladder Cancer

Aggressive Cancer Penetrating Muscle Layer



https://www.mdpi.com/2227-9059/11/2/539

Introduction

Existing Classification Systems:

Diversity of Molecular Subtypes Impedes Clinical Application

⇓

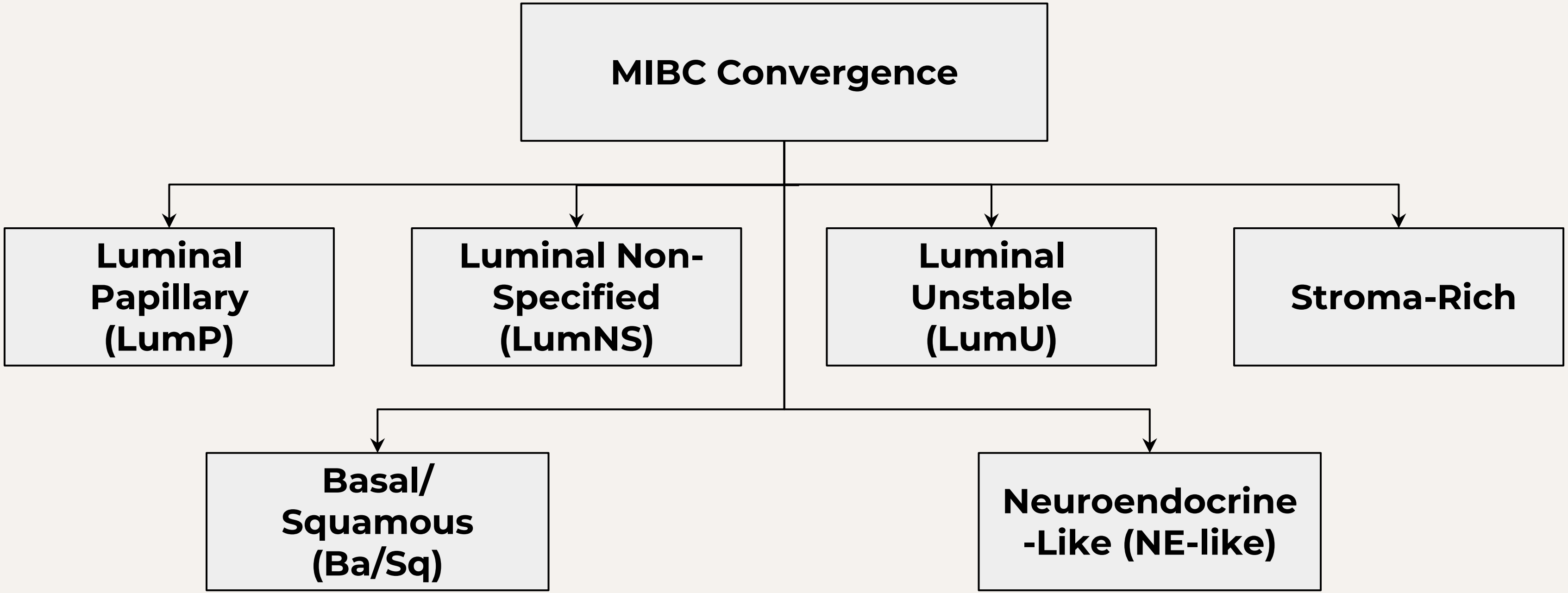
Lack of Naming Convention + Data Variability

Research Objective:

“Achieve International Consensus of MIBC Molecular Subtypes that Reconciles Published Classification Schemes”

Consensus Classification

Implementation: Nearest-Centroid Transcriptomic Classifier in R



Analytical Workflow

Patch Extraction

⇩

Feature Extraction

⇩

Multiple Instance Learning (MIL)

⇩

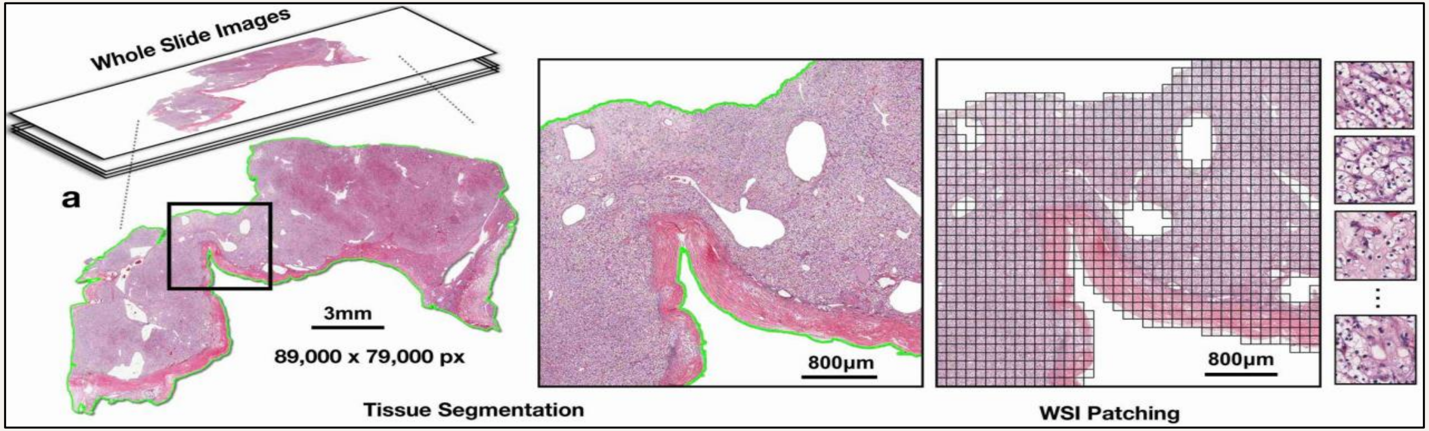
Information Visualization

Patch Extraction

Purpose: Facilitate Localized Deep Learning

Input: Masks Pre-Selected from HistoQC

Output: 1,500 Patches Per Whole Slide Image



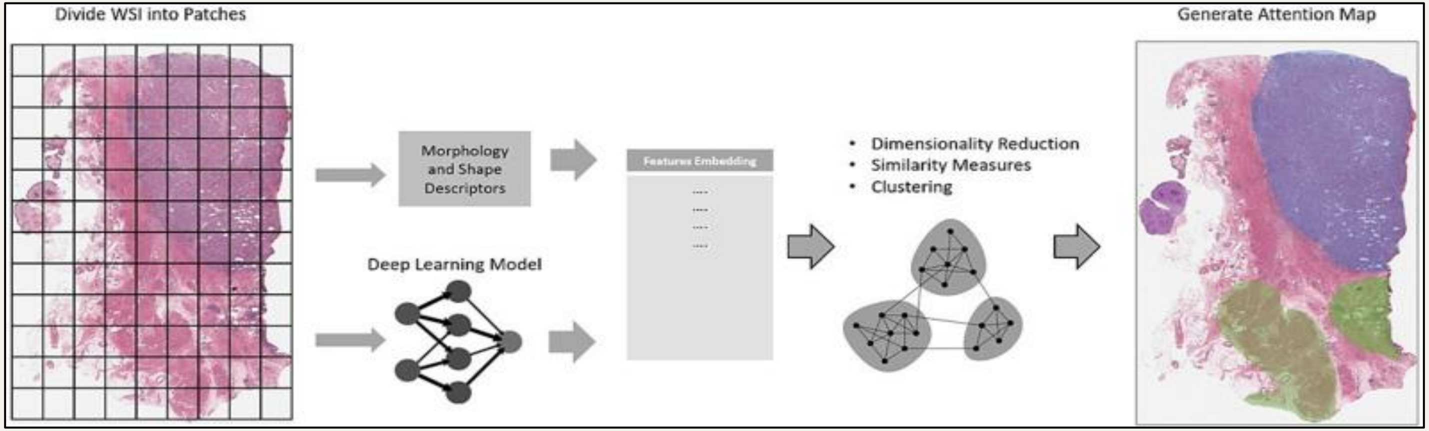
https://pixelscientia.com/articles/from-patches-to-slides/

Feature Extraction

Purpose: Meaningfully Represent Tissue Characteristics

Input: 1,500 Patches Per Whole Slide Image

Output: 458 Diagnostic + 463 Frozen .h5 Files



https://www.sciencedirect.com/science/article/pii/S215335392300149

Feature Encoders

CTransPath • Pre-Trained Vision Transformer for Unsupervised Contrastive Learning	**Lunit Dino** • Pre-Trained Vision Transformer for Self-Supervised Learning
Phikon • Pre-Trained Vision Transformer for Self-Supervised Learning	**PLIP** • Pre-Trained Vision Transformer for Pathology Image Retrieval
UNI • Pre-Trained Vision Transformer for Self-Supervised Learning	**vit** • Pre-Trained Vision Transformer for Self-Supervised Learning

Multiple Instance Learning

Purpose: Analyse Extracted Features at Bag-Level

Input: 458 Diagnostic + 463 Frozen .h5 Files for 20x Mag

Output: 90 Checkpoint + Output .pt Files per Model

Models:

1. DeepMIL
2. TransMIL
3. ClamSB

Data Folds:

1. Fold #1
2. Fold #2
3. Fold #3

Relevant Metrics:

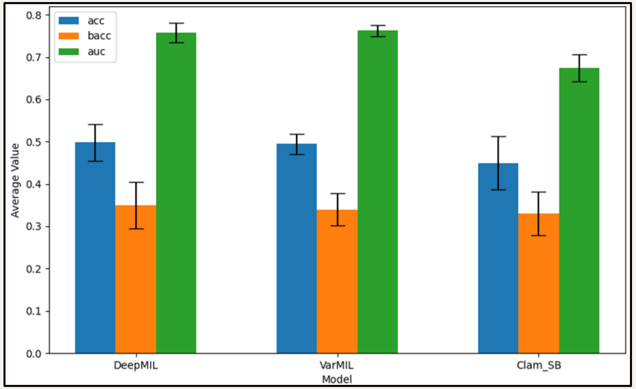
1. Accuracy
2. Bal. Accuracy
3. AUC

Others:

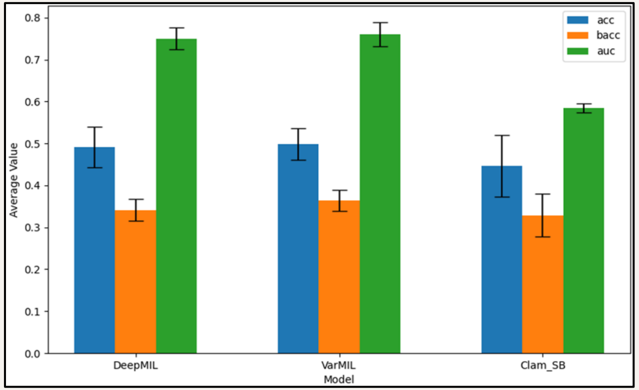
1. Time
2. Recall
3. Precision
4. F1-Score

DIAGNOSTIC

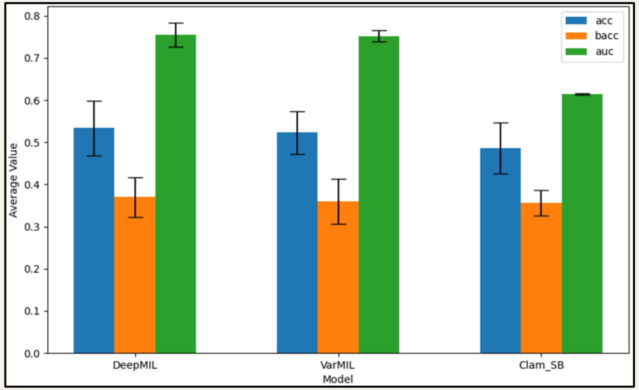
CTransPath



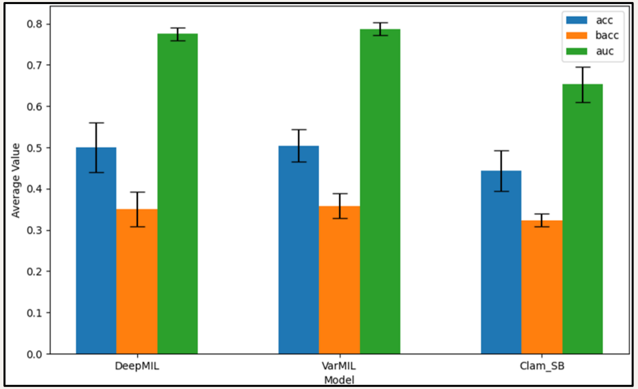
Phikon



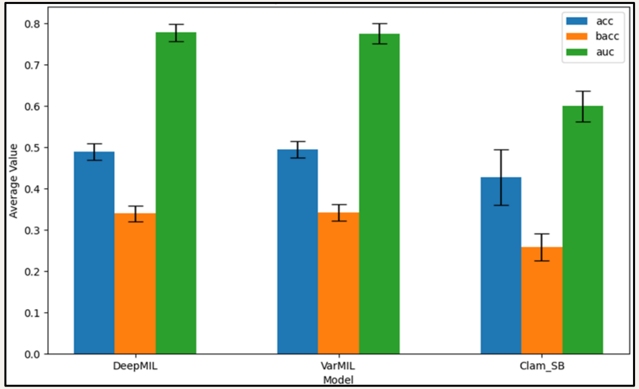
UNI



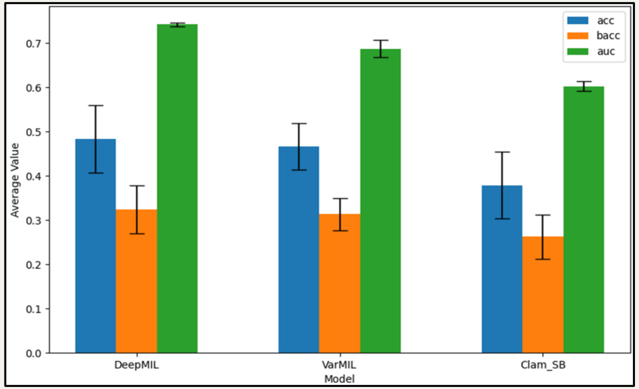
Lunit-Dino



PLIP

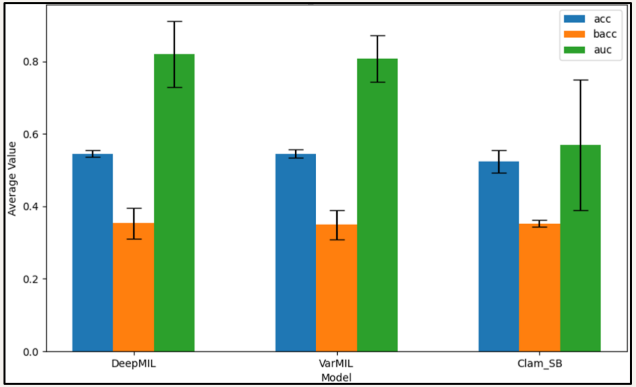


vit

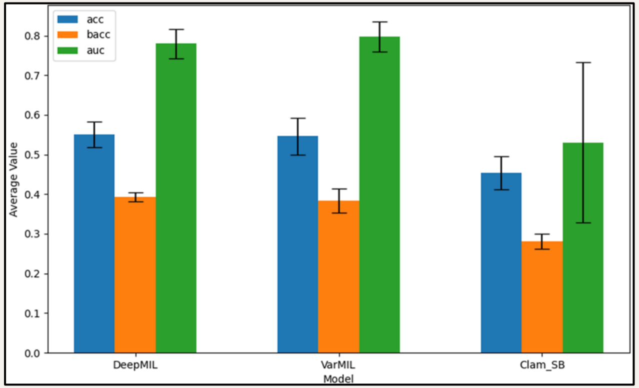


FROZEN

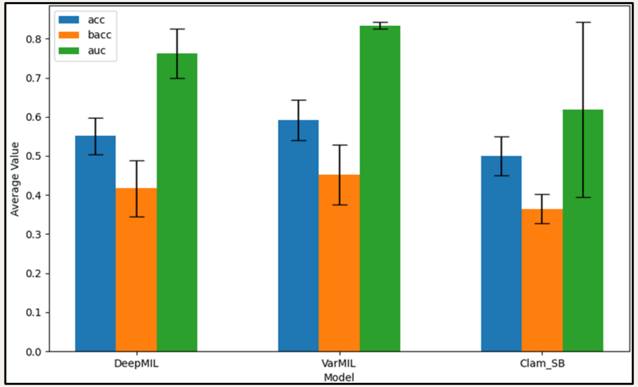
CTransPat



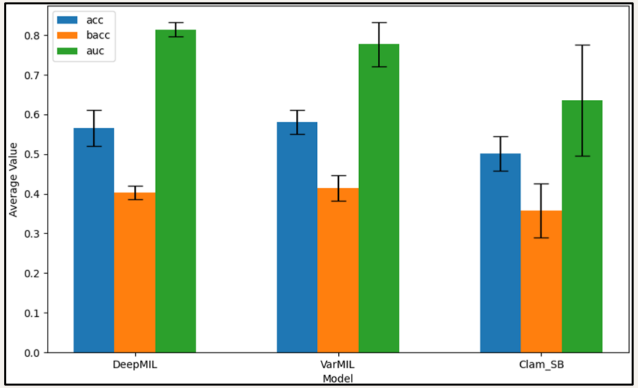
Phikon



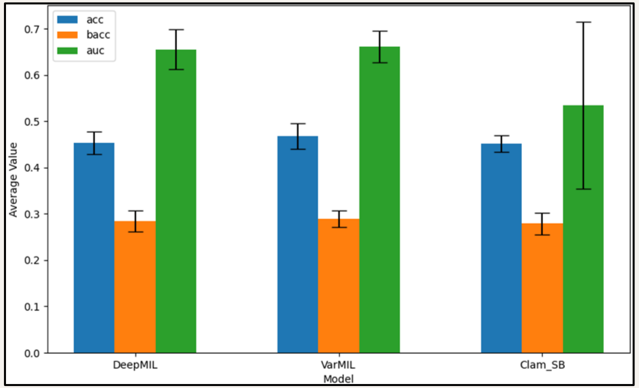
UNI



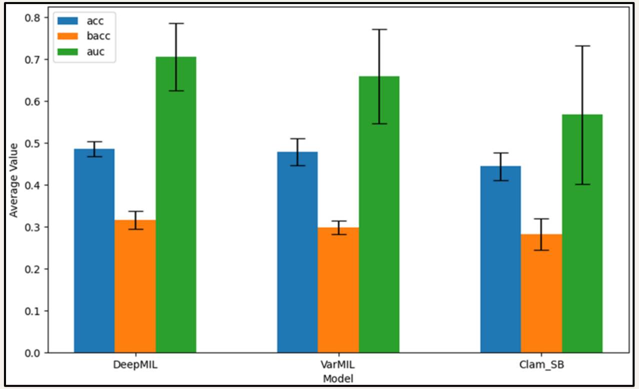
Lunit-Dino



PLIP



vit



Summary of Findings

Observations:

- CTransPath, Phikon and UNI were the Highest Performing Encoders
- DeepMIL and VarMIL Consistently Provided the Highest Metrics
- AUC Varied the Most Between Models while BACC Varied the Least

Future Work:

- Implement Hyperparameter Tuning (Learning Rate, Weight Decay etc.)
- Experiment with Multi-Headed Attention Mechanisms or Hybrid Variants
- Use Cross-Validation to Prevent Over-Fitting and Increase Robustness

Information Visualization

Purpose: Determine Relationships Between Subtypes

Input: 458 Diagnostic + 463 Frozen .h5 Files for 20x Mag., Manifest File (patient_id, slide_id, subtype, slide_path)

Output: UMAP Plots

Slide-Level

Mean Pooling of 1,500 Patches + Filtering of Features with No Subtypes

Eg: (1,500, 768) → (1, 768)

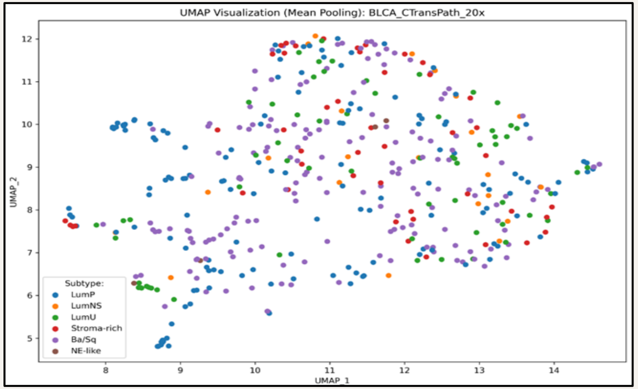
Patch-Level

Random Selection of 150 Patches + Filtering of Features with No Subtypes

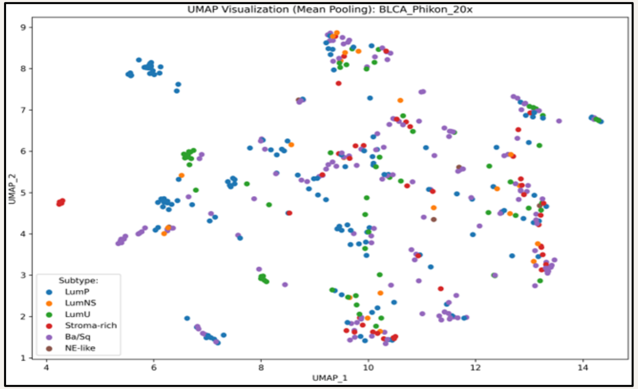
Eg: (1,500, 768) → (150, 768)

DIAGNOSTIC (Slide-Level)

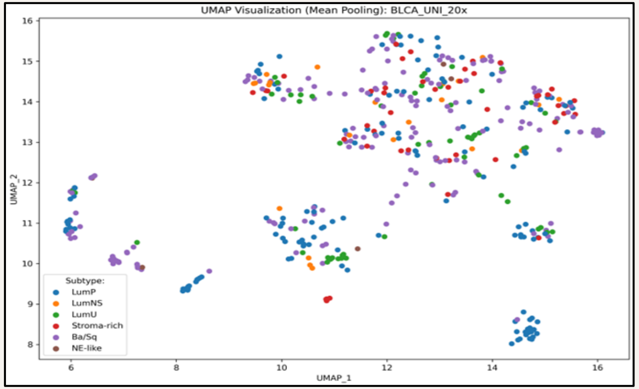
CTransPath



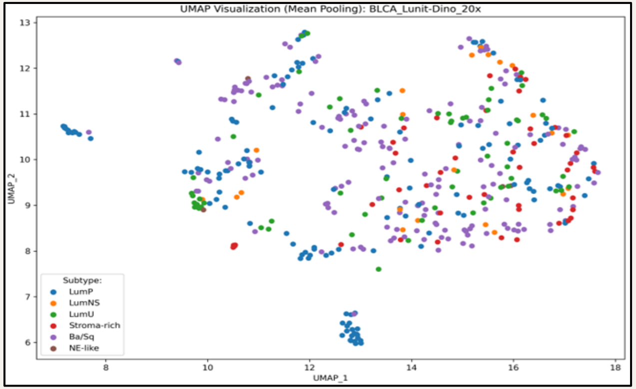
Phikon



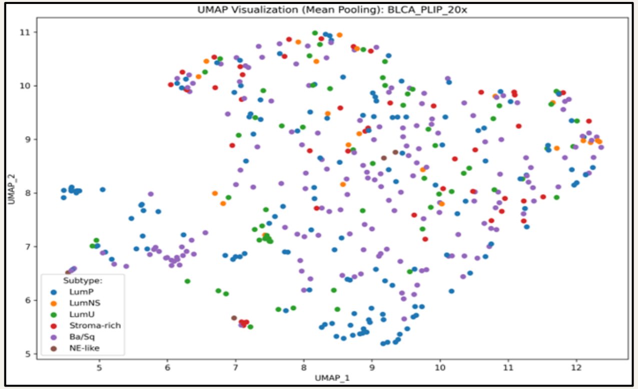
UNI



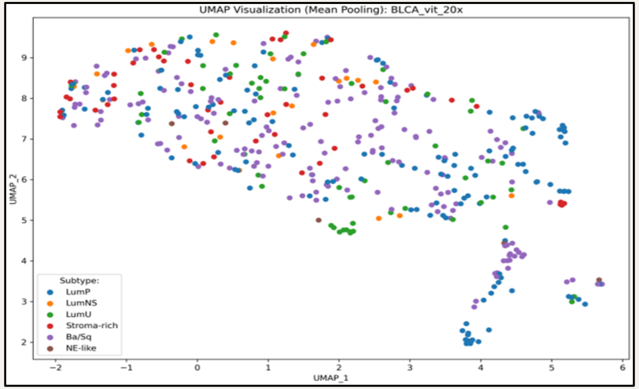
Lunit-Dino



PLIP

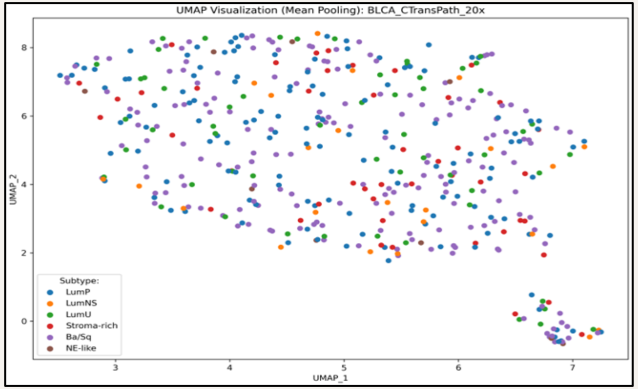


vit

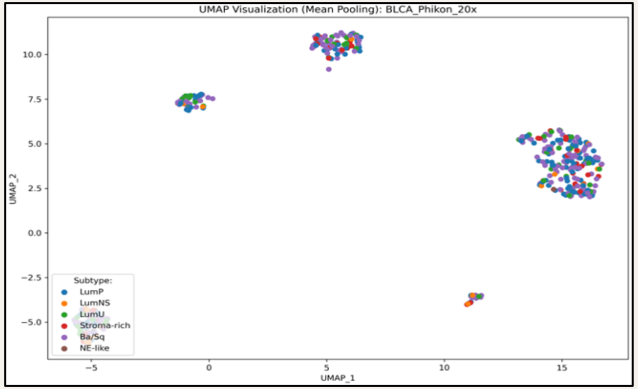


Frozen (Slide-Level)

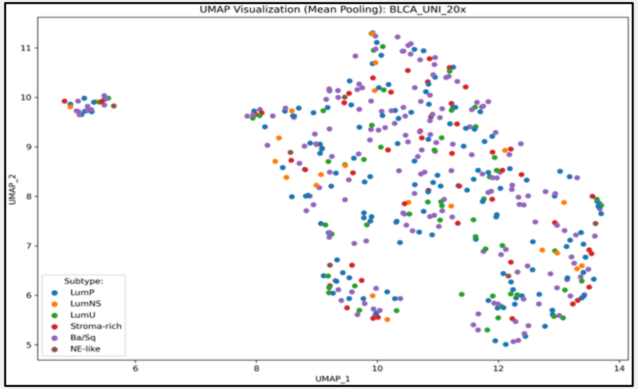
CTransPath



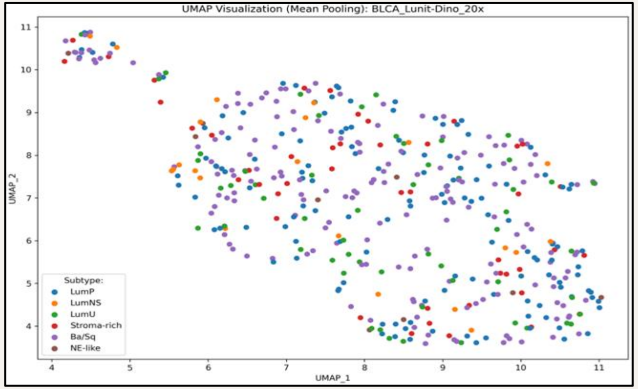
Phikon



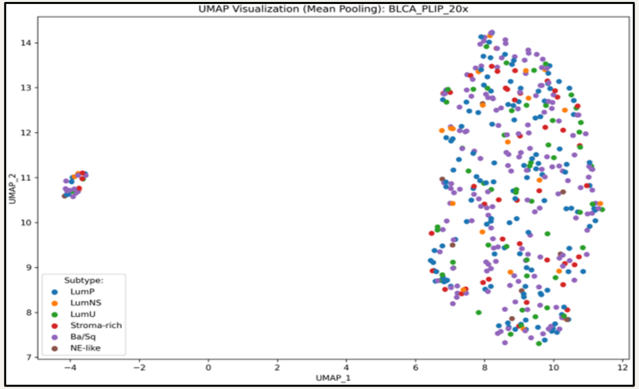
UNI



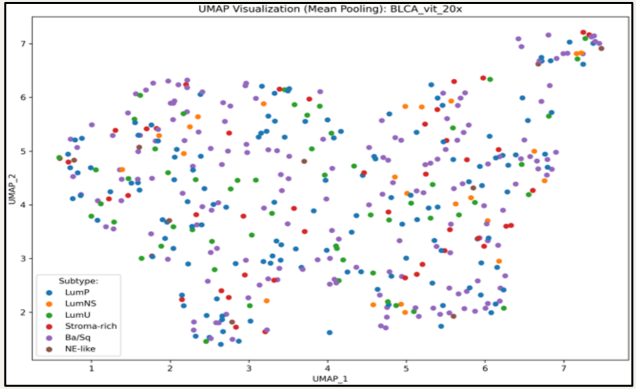
Lunit-Dino



PLIP

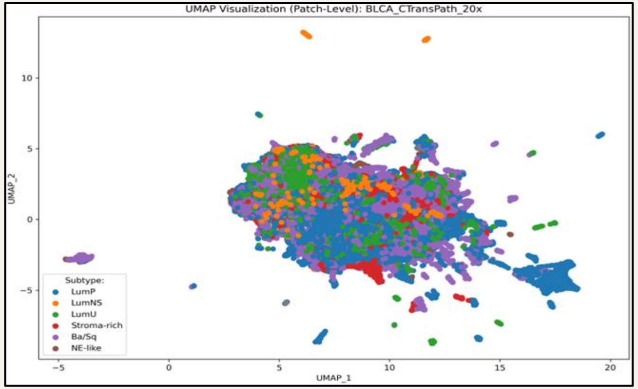


vit

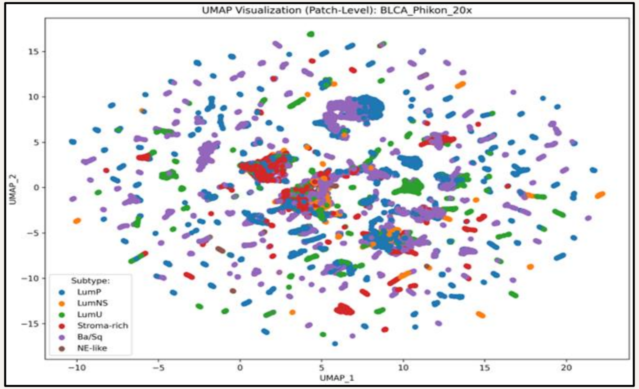


DIAGNOSTIC (Patch-Level)

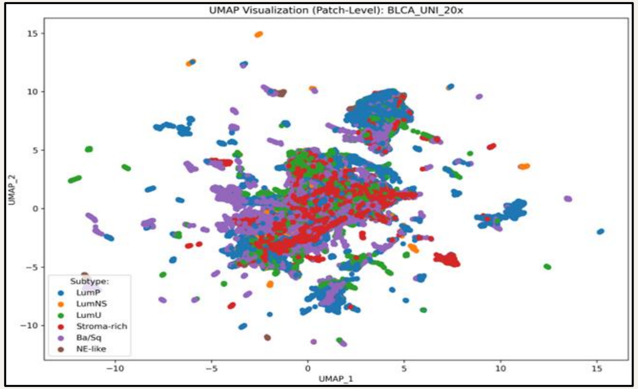
CTransPath



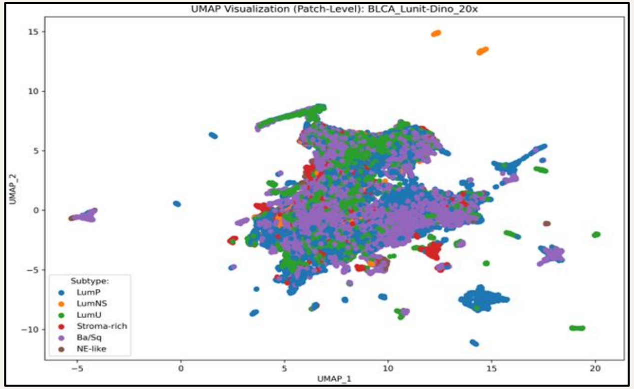
Phikon



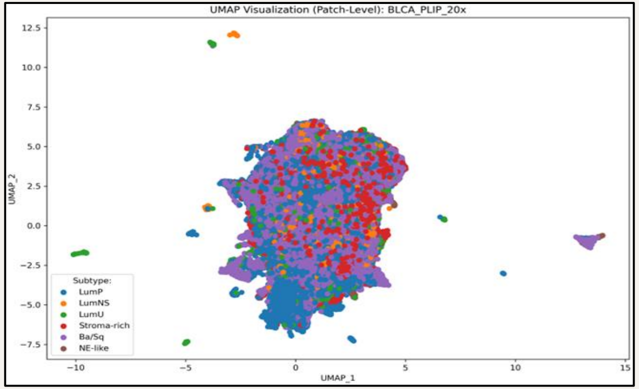
UNI



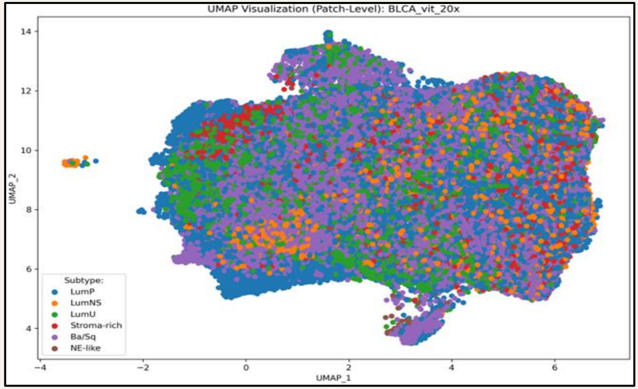
Lunit-Dino



PLIP

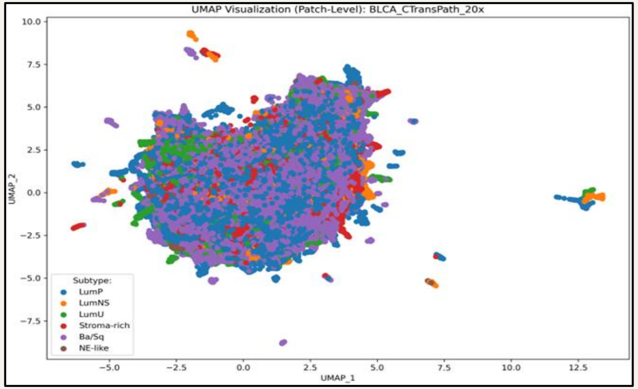


vit

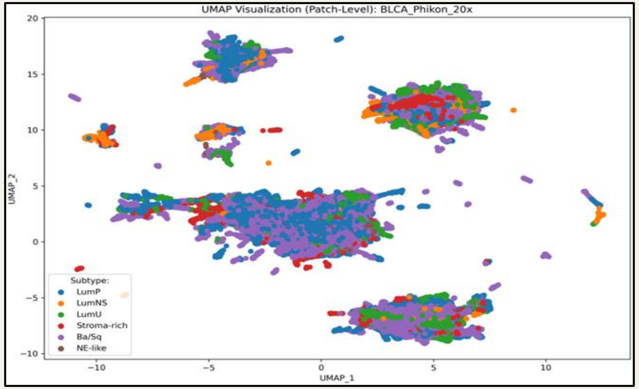


FROZEN (Patch-Level)

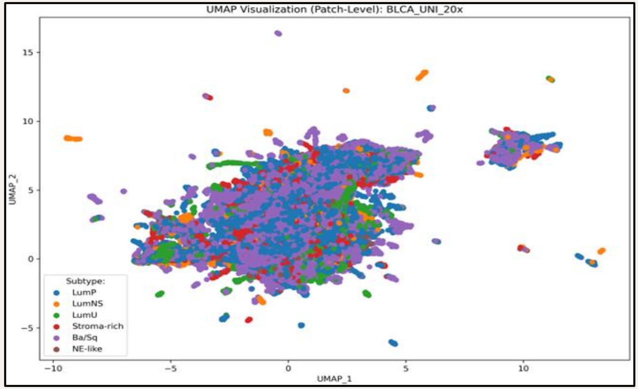
CTransPath



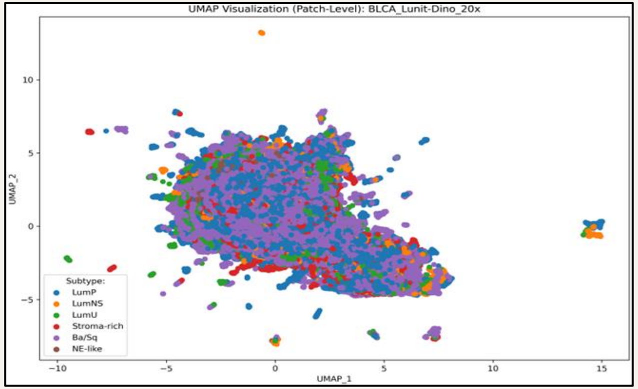
Phikon



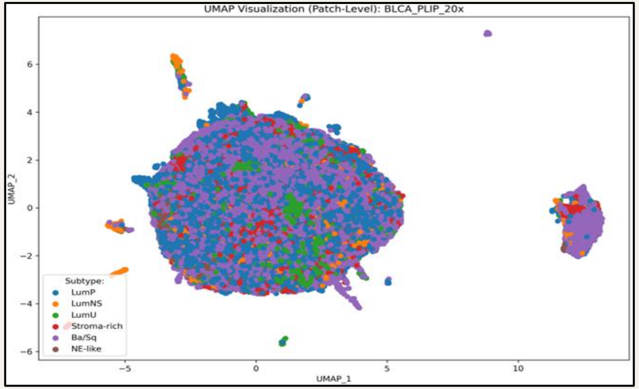
UNI



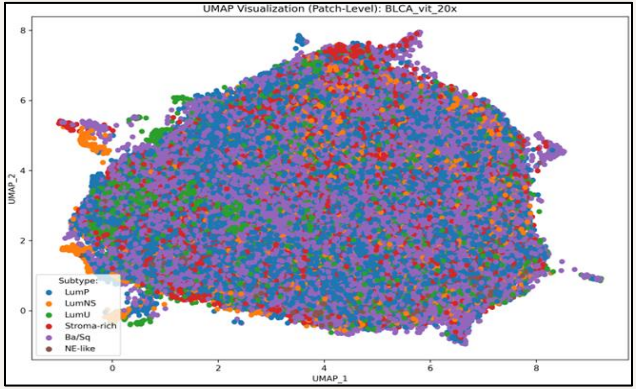
Lunit-Dino



PLIP



vit



Summary of Findings

Observations:

- Patch-Level UMAPs Provided Higher Granularity than Slide-Level UMAPs
- Patch-Level UMAPs Capture Data Diversity Better than Slide-Level UMAPs
- Patch-Level UMAPs Retain Local Patterns Better than Slide-Level UMAPs

Future Work:

- Stratified Sampling Based on Patches of Interest + Extract More Samples
- Implement Weighted Aggression or Hierarchical Pooling at Slide-Level
- Perform Dimensionality Reduction (PCA, t-SNE) Prior to UMAP Plotting

Acknowledgements

Questions?