SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

ATTENTION BASED MULTIPLE INSTANCE LEARNING FOR GENERALIZING ON THE UBC-OCEAN

Ovarian cancer subtyping and outlier detection

Karthik G Shanmugasundaram
(PA2212044010007)
Dr. S. Ganesh Kumar
(Professor/Department of Data Science and Business Systems)

Saturday, May 10th, 2025 (Final Review)

Abstract

Epithelial ovarian cancer (EOC) is one of the deadliest cancers faced by women across the world. Early and accurate diagnosis of EOC is crucial for timely intervention. Accurate **subtype classification** driven by Artificial Intelligence (AI) will address subjective variability of assessments and is key to how treatment is managed. To improve applicability as required by diverse data, rare subtypes that may be present in various populations need to be flagged as outliers for further investigation. **Whole slide images** (WSIs) and **Tissue Microarray (TMA)** images from the **UBC-OCEAN** dataset form the most extensive collection of images from centres across the world for ovarian cancer subtyping. TMA images are more varied in size for the unseen public and private test sets in the Kaggle challenge which introduced this dataset. WSI images are large (600MB+), and **diseased tissues may be sparsely distributed**, requiring careful feature extraction and use. **Multiple instance learning approaches (MIL)**, a type of **weakly supervised learning model**, are key to effectively utilizing WSI for generalizability. Neural approaches have been favoured more than classical MIL because of trainability and interpretability. **ABMIL (Attention mechanism-based MIL)** has shown great performance for this problem. In this work, results from an architecture with a **double vision transformer backbone with two patch sizes, and ABMIL** is presented and optimized to generalize on the unique challenge of the UBC-OCEAN dataset.

Problem identification and description

Motivation & Problem Statement

EOC is one of the deadliest cancers in women

Late diagnosis is typically fatal

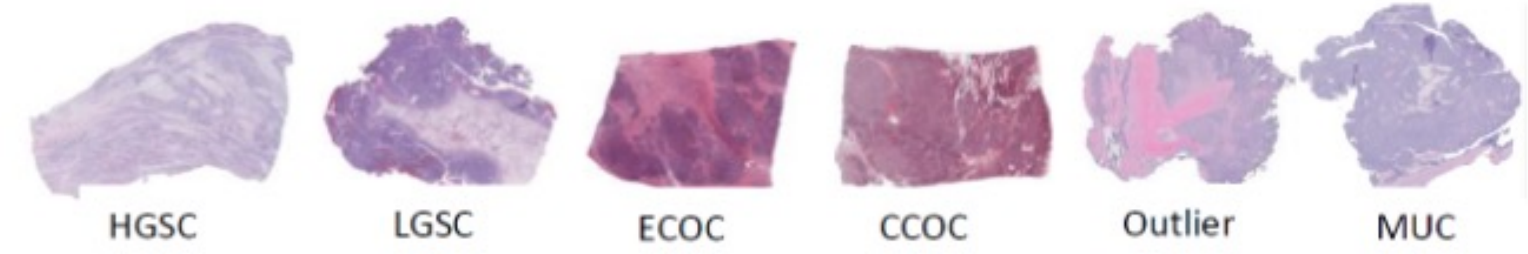
Subtype classification and outlier detection

Subjective differences are an issue

AI can

- enhance accuracy
- deal with inter-observer agreement issue
- widen applicability

HGSC(High Grade Serous Carcinoma), LGSC (Low Grade), ECOC(Endometrioid), CCOC(Clear Cell), MUC (Mucinous)



UBC-OCEAN challenge

- Ovarian cancer
  - Whole slide images (513 WSIs) & Tissue Microarray (25 TMAs)
- To build an architecture that will generalize to unseen test sets (1915 images, many of them TMAs)
- To solve the problem of how the sparsity of the disease tissue in huge images affects classification.

Objectives

Improve generalization on UBC-OCEAN dataset

To match the top scores under available computational constraints

Literature Review (1/3)

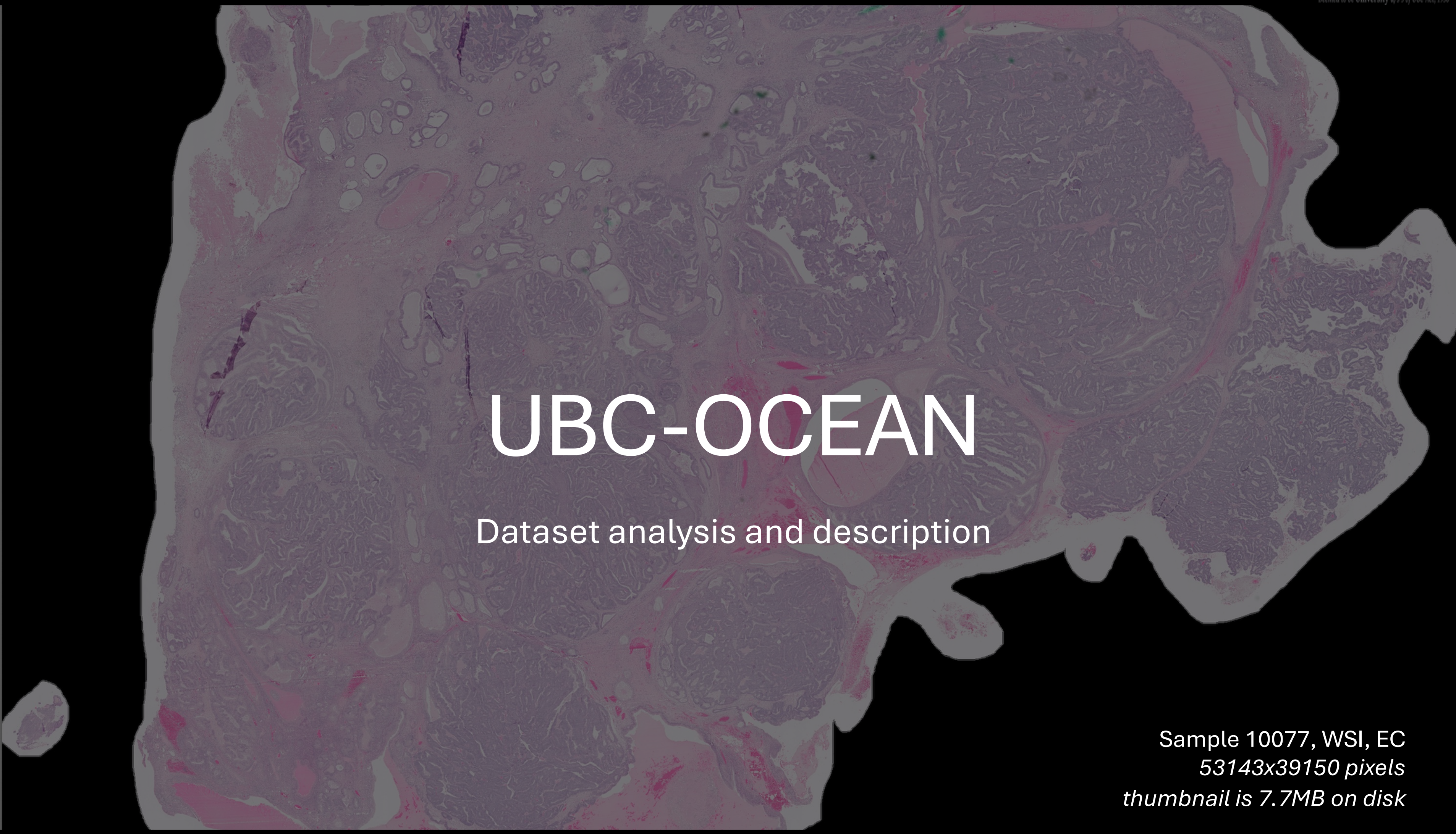
Title of the Paper	Journal	Dataset Used	Contributions	Other Observations
1. **Machine Learning-Driven Histotype Diagnosis of Ovarian Carcinoma: Insights from the OCEAN AI Challenge**	***medRxiv****2024* (UBC AIM Lab)	UBC-OCEAN	WSI/TMA dataset curation, description, outlining the challenge	Detailed class distribution of public and private test sets is available providing pointers for design
2. Deep Learning-Based Histotype Diagnosis of Ovarian Carcinoma Whole-Slide Pathology Images	***Modern Pathology*** *2022* IF: 7.1 CiteScore: 11.2	OVCARE + External dataset (became some part of UBC-OCEAN)	Putting the dataset in context and explicating key techniques including MIL	1 stage-TL with VGG19 and tiling outperformed MIL. Potential for AI to augment diagnostician noted.
3. Benchmarking weakly-supervised deep learning pipelines for whole slide classification in computational pathology	***Medical image analysis*** 79 (2022) IF: 10.7 CiteScore: 22.1	WSI data of different types of cancer from 2980 patients	Systematic review of weakly-supervised and MIL techniques on whole slide images	Other WSL methods are also considered and observed to be competitive.
4. A Comprehensive Evaluation of Histopathology Foundation Models for Ovarian Cancer Subtype Classification	***arXiv preprint*** *2024* Leeds University Hospital	UBC-OCEAN Transcanadian study	Performance comparison of foundation models	SOTA/application-level performance drastically improved by foundation models
5. Scaling Self-Supervised Learning for Histopathology with Masked Image Modeling	***medRxiv*** *2023* Owkin UK Ltd.	UBC-OCEAN	Winning entry in Kaggle competition inspiring overall architecture and methodology.	TMA detector, outlier detection method, 60% of the train TMAs used for validation

Literature Review (2/3)

Title of the Paper	Journal	Dataset Used	Contributions	Other Observations
6. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images	***Nature Medicine* 2019** IF: 58.7	CAMELYON16 (*breast cancer metastasis detection)*	MIL based deep learning using only reported diagnoses and avoiding pixel-wise annotations	AUC above .98 for all cancer types Comparison of slide aggregation approaches
7. Attention-based Deep Multiple Instance Learning	**35th International Conference on Machine Learning (ICML 2018)**	MNIST, Breast Cancer & Colon Cancer cellular H&E images	Introduces Attention-based MIL pooling	Classical MIL is ”predefined and non-trainable”
8. Classification of Epithelial Ovarian Carcinoma Whole-Slide Pathology Images Using Deep Transfer Learning	**Medical Imaging with Deep Learning (MIDL) 2020**	305 Ovarian cancer WSI from Vancouver GH	Two stage deep transfer learning algorithm and progressive resizing	Downsamples a patch at different rates, modifies VGG19 accordingly, propagates weights.
9. Data-efficient and weakly supervised computational pathology on whole-slide images	**Nature Biomedical Engineering** IF: 27.7	Assortment of public and private WSIs from BW Hospital (Harvard)	Introduces CLAM (Clustering-constrained-attention multiple-instance learning)	Addressing data inefficiency in standard weakly supervised learning
10. Multiple instance learning for histopathological breast cancer image classification	**Expert Systems With Applications** IF: 7.5	BreakHis breast cancer dataset	Comparison of state of the art MIL methods	Review of the MIL paradigm

Literature Review (3/3)

Title of the Paper	Journal	Dataset Used	Contributions	Other Observations
11. Kaggle 1st place solution report	***Owkin Inc writeup***	UBC-OCEAN	Winner, unique methodology	Architectural choices such as matter detection, data splitting etc.
12. Benchmarking self-supervised learning on diverse pathology datasets	**Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023**	BACH CRC Pcam MHIST CoNSeP	Domain specific large scaled pretraining. Compares different SSL methods (	33 million patches with necessary augmentations, largest-scale study known
13. What makes transfer learning work for medical images: Feature reuse & other factors	**Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023**	IMAGENET (Source) Targets: APTOS2019, CBIS-DDSM, ISIC2019, CHEXPERT, PATCHCAMELYON	Investigates domain specific pre-training and its impact on transfer learning for different datasets	Benefits from transfer learning increase with 'smaller distance between source and target', ViT/CNN comparisons
14. Emerging properties in self-supervised vision transformers	**ICCV 2021**	ImageNet	Introduces DINO. Synergy between ViT and DINO	Self-supervised ViT features explicitly contain object boundaries
15. An image is worth 16x16 words: Transformers for image recognition at scale	**ICLR 2020**	ImageNet, CIFAR-100, VTAB	Introduces ViT – extending Transformers to images	ViTs become the new state of art displacing CNNs



UBC-OCEAN

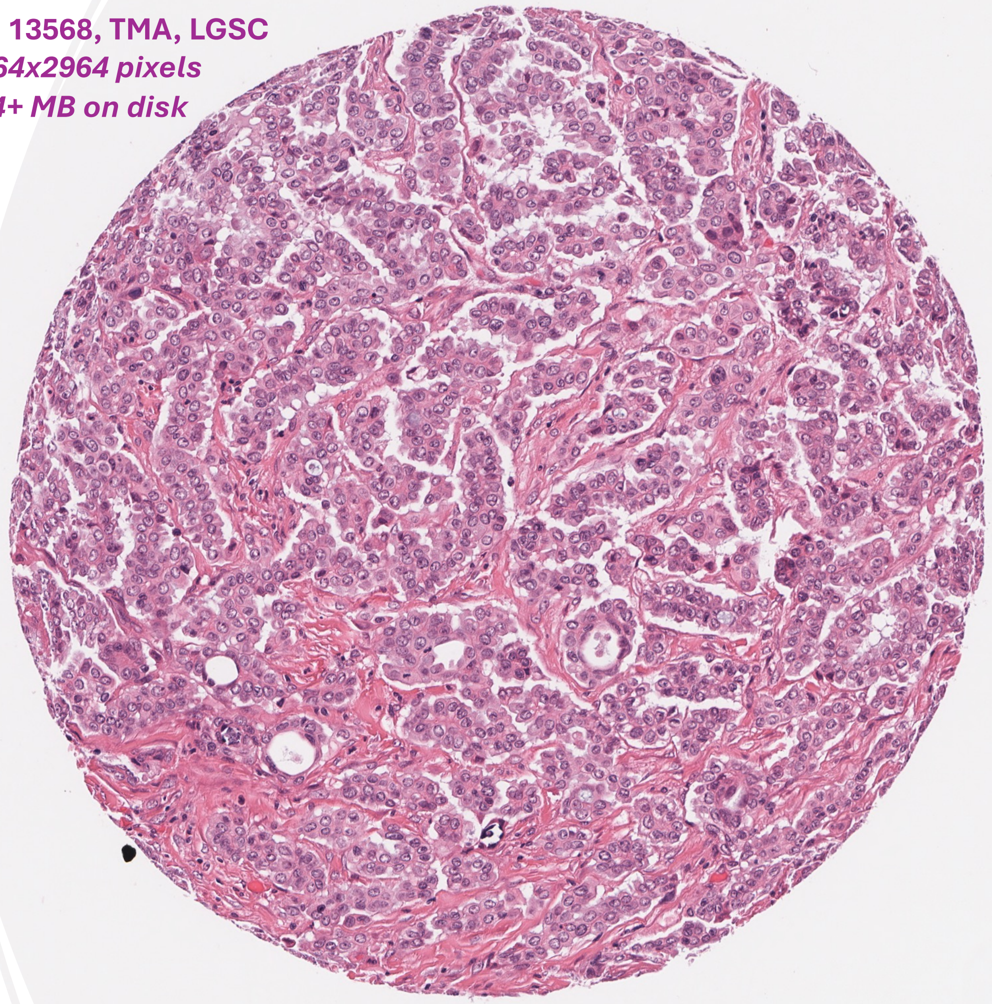
Dataset analysis and description

Sample 10077, WSI, EC
53143x39150 pixels
thumbnail is 7.7MB on disk

WSIs/TMAs

- 513 Whole slide images (WSIs) and 25 Tissue Microarray (TMAs) images of Ovarian cancer are available for training
  - WSIs are huge at 20x magnification
    - (avg. 30000x30000pixels)
    - thumbnails are available in train and test
  - TMAs are smaller (~3000x3000 pixels), at 40x magnification, whole images are available

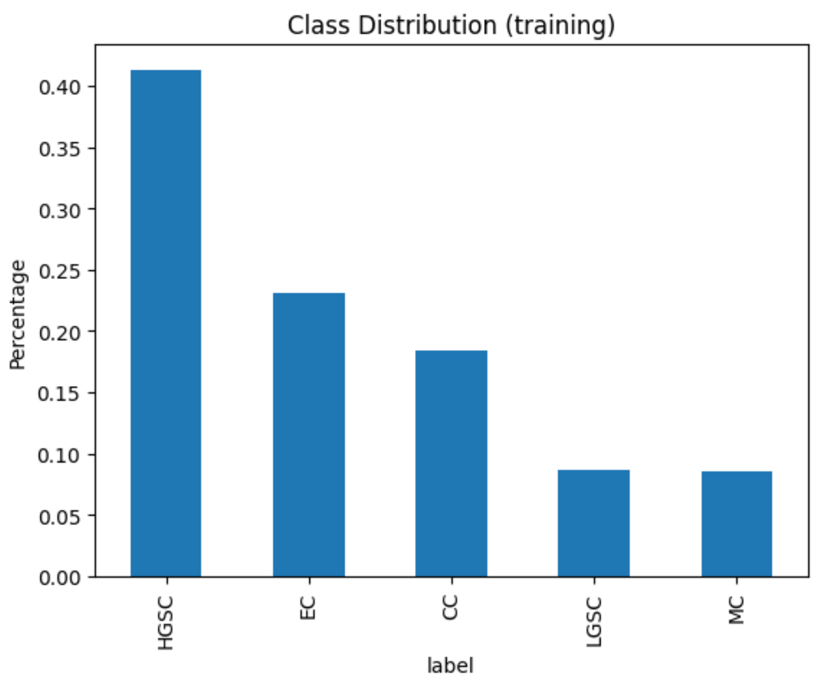
Sample 13568, TMA, LGSC
2964x2964 pixels
14+ MB on disk



Dataset overview

- 2438 images (WSI + TMA)
- 20+ global centers
- 5% TMAs in train set
- 80% TMAs in private test set
- High variance in image size and source

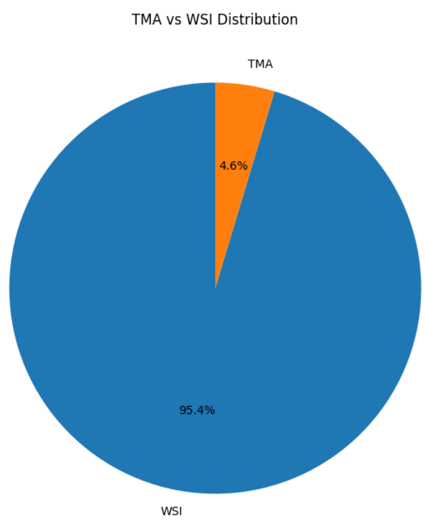
Data distribution



		Train	Public Test	Private Test
TMA	CC	**5**	**64**	**175**
	EC	5	41	210
	HGSC	5	51	492
	LGSC	5	36	212
	MC	5	20	56
	Other	0	31	44
WSI	CC	94	23	61
	EC	119	35	70
	HGSC	217	68	71
	LGSC	42	10	39
	MC	41	9	35
	Other	0	49	23
		538	437	1488

TMAs are underrepresented in train!

- Possibilities:
  - Splitting strategy reserving significant proportion of TMAs for validation
    - 70/30 or 80/20 for WSI
    - 40/60 for TMA [11]
  - Augment TMAs by interpolating portions of WSI
    - Borrowed augmentation data



Dataset	Train	Public	Private
%of TMAs	5%	56%	80%
WSI	513	194	299
TMA	25	243	1189
Total	538	437	1478

Requirements (functional and non-functional)

Training:

30 hours per week of Nvidia GPU P100 time provided by Kaggle.

Training notebooks can run for 12 hours.

Inference:

2x GPU T4 provided by Kaggle (100 submissions)

Inference notebooks can run for 9 hours.

Optimization within these constraints is what is required next.

Existing and proposed methodology

Existing solutions (pros and cons)

Classical MIL

- Simple to implement
- Not trainable
  - does not adapt to data
  - not good at generalizing
- Not interpretable

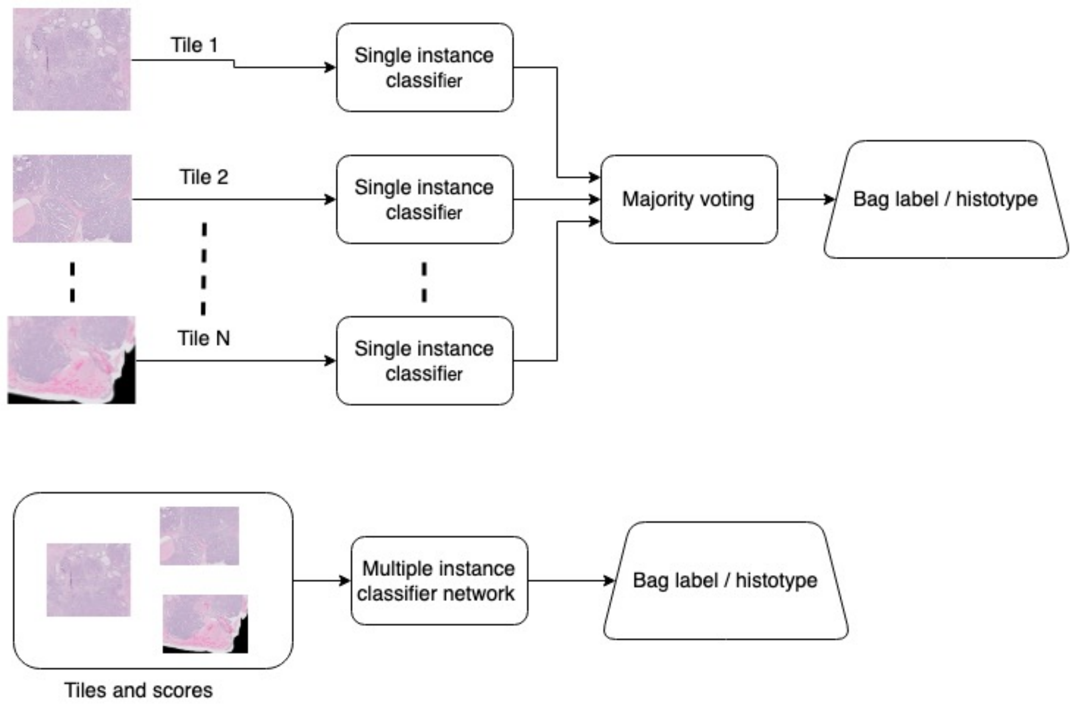
Attention based MIL

- Trainable
- Interpretable (produces a score for each tile of slide)
- Needs careful design and optimization
- Training and inference computational load

Proposed methodology and module description

Weak Supervision & MIL

MIL vs Single instance classification

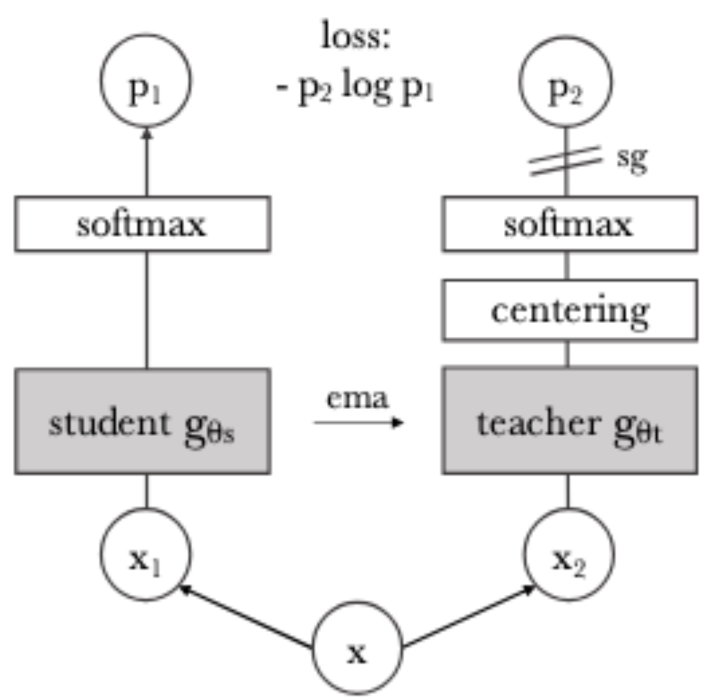


- Labels at slide level only
- WSI = Bag of tiles
- MIL: tile-level learning via attention

Transfer Learning & Vision Transformers

- Used Lunit DINO ViT (patch sizes 8 & 16)
- Lunit DINO
  - Pretrained via SSL on histopathology
- Effective in limited-label scenarios

SSL ViT training



Baseline Approach

InceptionV3 with 512x512 tiles

No attention, lower accuracy

Struggled with TMAs in test set

System architecture and design

Proposed Architecture

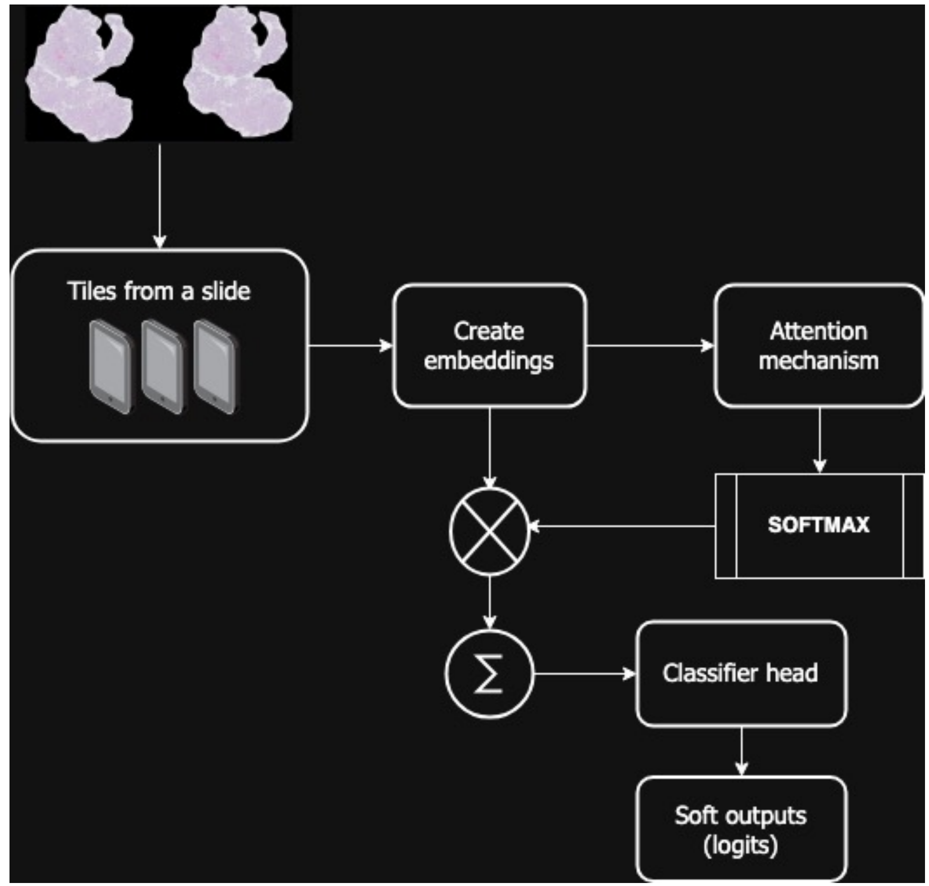
Dual ViT-DINO extractor (patch 8 & 16)

ABMIL attention for tile relevance

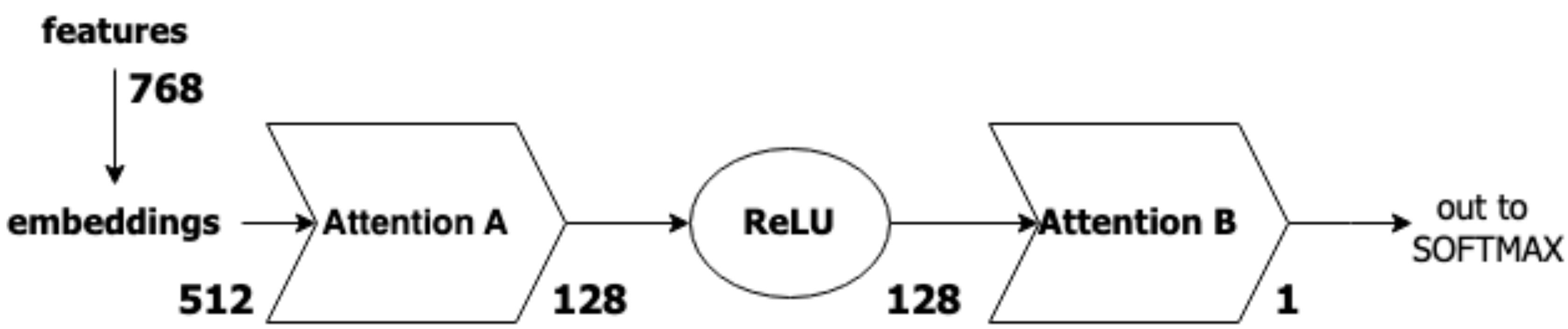
400 tiles per image, 224x224 resolution

ABMIL implementation

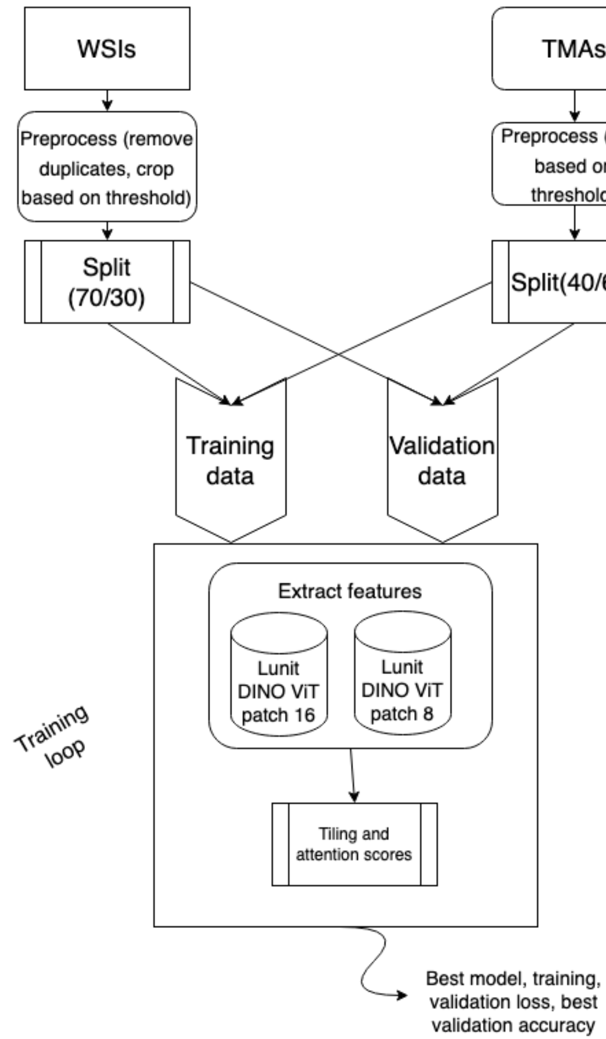
- **Innovation:** From scratch implementation and tuning of attention mechanism
- Customized deep learning workflow to fit within resource constraints
- Incorporating best practices from existing solutions



Optimized attention mechanism implemented for usecase



Detailed training methodology



Inference methodology:

Best results were obtained when preprocessing was avoided

Feature extractors

- Baseline – InceptionV3
- Vision Transformer backbones have been shown to be better at many computer vision tasks compared to CNNs [3]
- Pre-trained domain specific weights
  - Reduce training time
- Double ViT architecture (Lunit DINO)
  - pre-trained vision transformer weights
  - on 33 million histology patches from various pathology datasets
  - Self-supervised learning (SSL) method known as DINO (self-DIstillation with NO labels). [14] Facebook AI research

Implementation Details

- PyTorch + timm, pyvips for fast loading
- Trained on Kaggle P100
- 60% TMA validation
- Eval metric: Balanced Accuracy
- Key: Data loading and processing parallelization
- Implemented in PyTorch and uses timm for pre-trained models and loading.

Performance measures and evaluation

Balanced accuracy

average recall across all classes

Required by competition

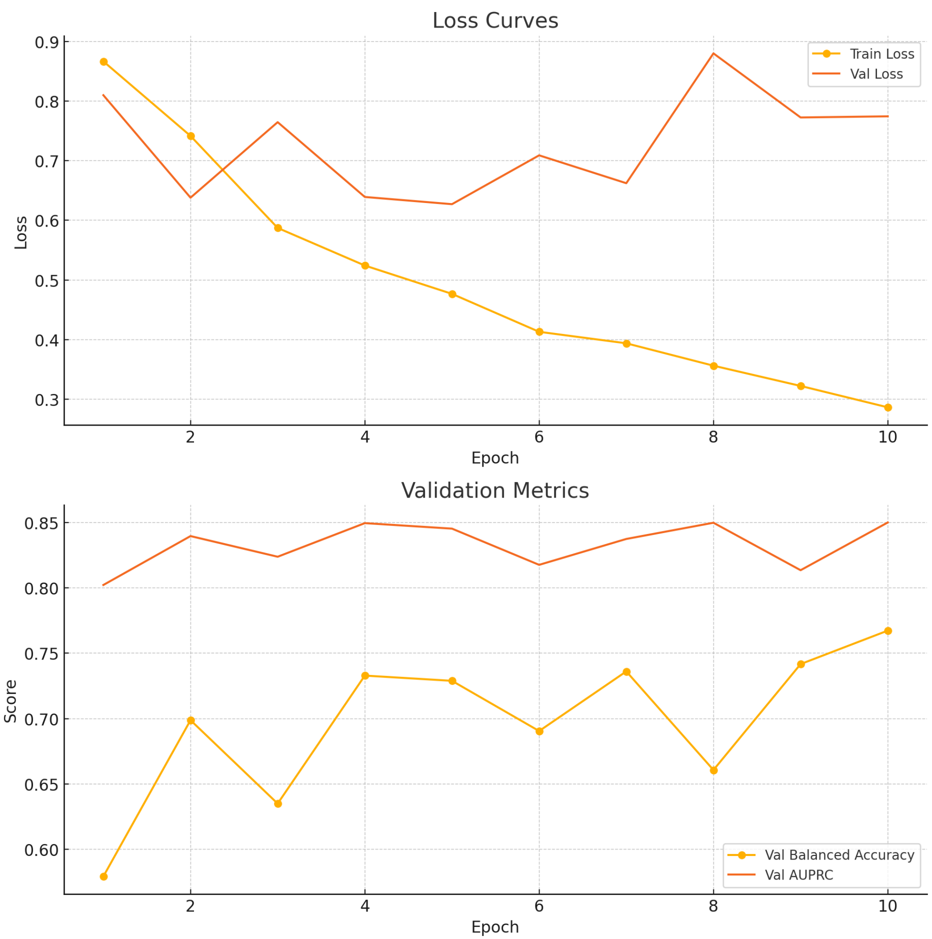
AUPRC (being imbalanced dataset)

Cross entropy loss for training

Focal loss was also used as an experiment

Results and analysis

Training Curves



Training/Validation loss & AUPRC curves

- Loss and AUPRC curves across epochs
- May appear to be overfitting (loss curves) but criterion for is val balanced accuracy
- Val bal accuracy guided stopping
- Val bal accuracy showed increasing trend until 11 epochs.

Baseline & Best results

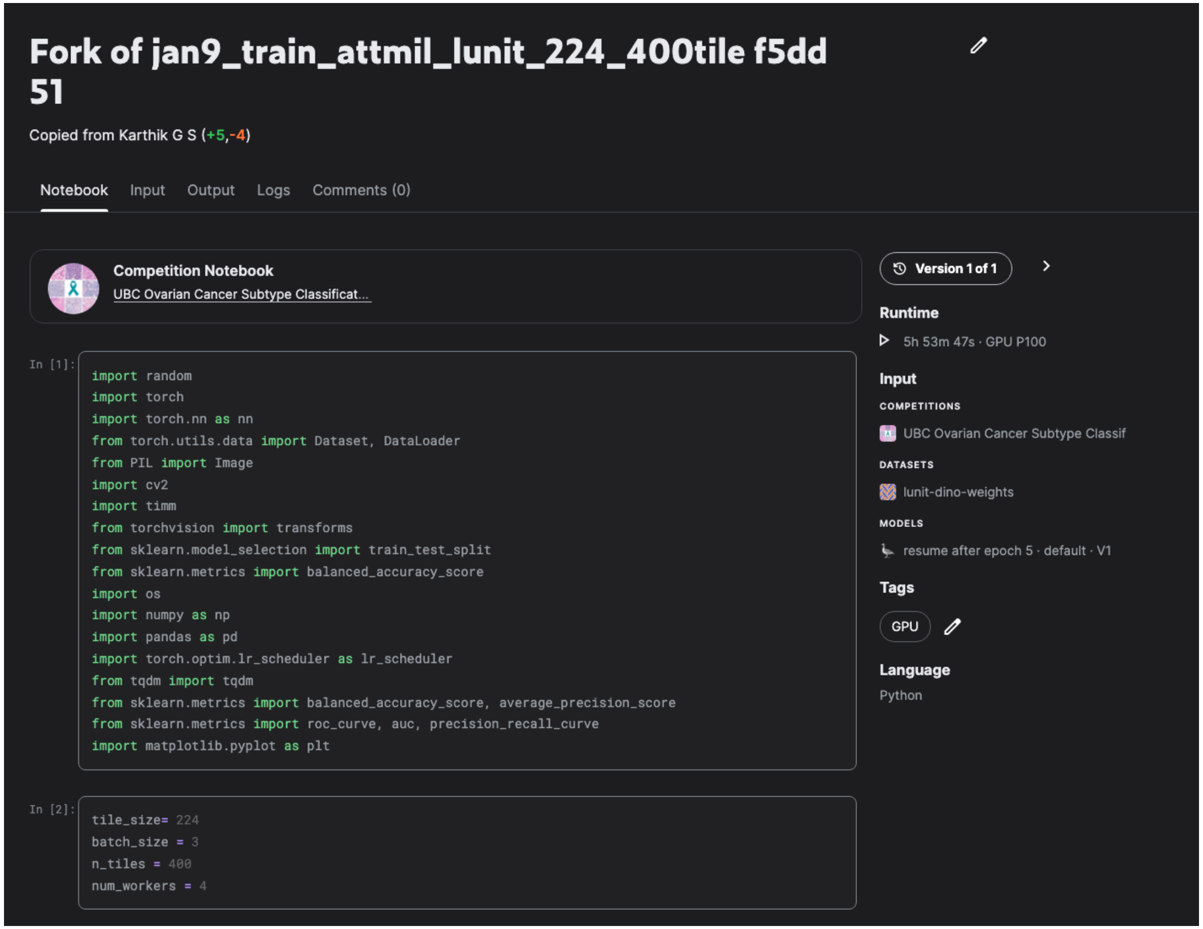
- Baseline: Public 0.27 / Private 0.24
- ABMIL ViT (400 tiles): Public 0.48 / Private 0.42
- Private test score matched top 10% leaderboard.
- Validation accuracy improved to 0.77

Key results

Solution		Public test score	Private test score	Best validation balanced accuracy (1-fold)
Baseline (InceptionV3) 512		0.27	0.24	0.54
Attention based MIL (InceptionV3) with 384, 20 tiles		0.36	0.33	0.72
Attention based MIL (Double ViT, Lunit DINO)	224x224, 100 tiles random	0.42	0.35	0.778
Attention based MIL (Double ViT, Lunit DINO)	224x224, 400 tiles, optimized, structured and complete	0.48	0.42	0.77

Project demo

Best train model notebook

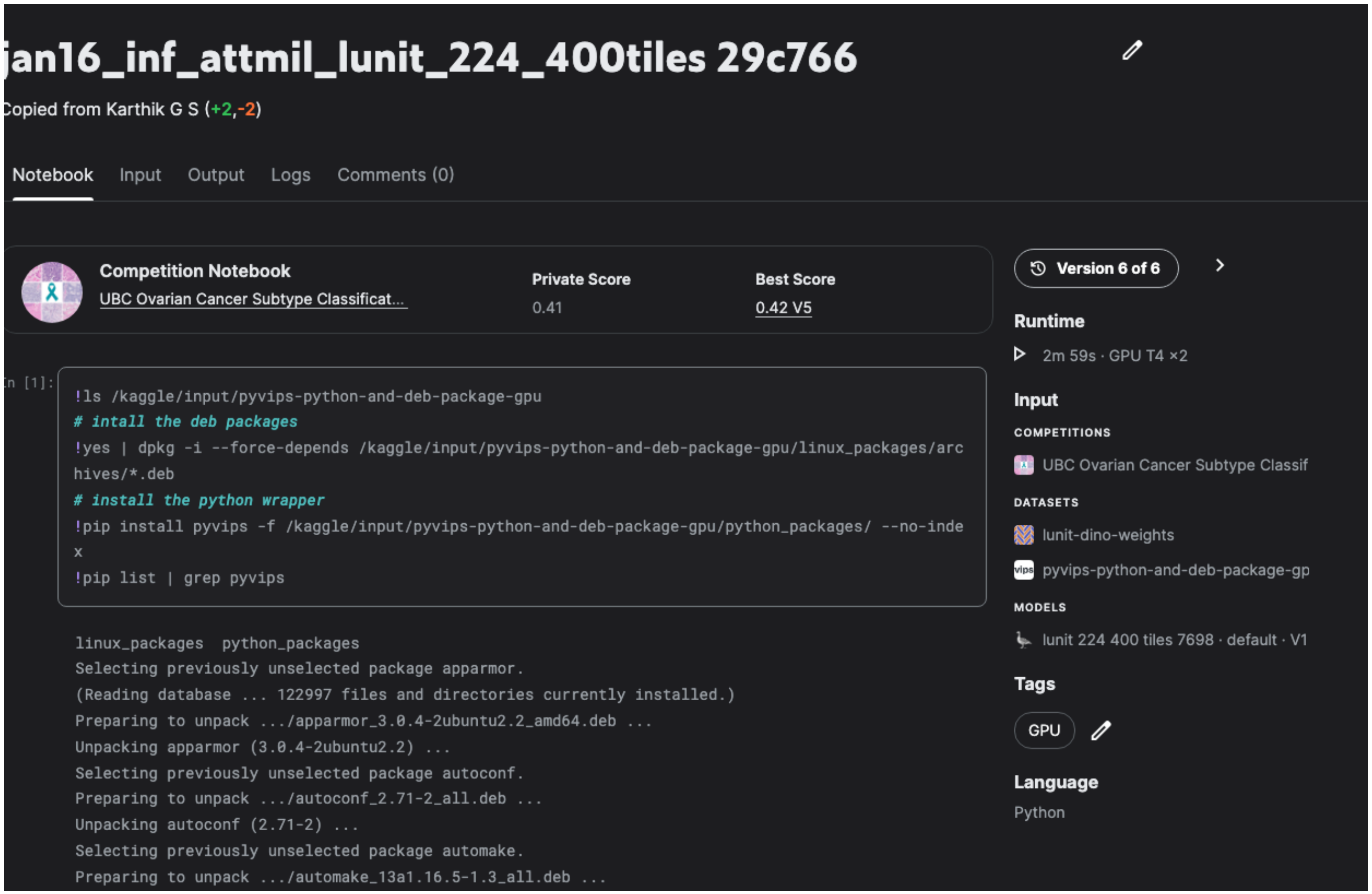


https://www.kaggle.com/code/skarthikguruswamy/fork-of-jan9-train-attmil-lunit-224-400tile-f5dd51

Best train model notebook results

```
Epoch 1 - Train Loss: 0.4766, Val Loss: 0.6272, Val Balanced Acc: 0.7289, Val AUPRC: 0.8453
  * Best model saved.

Epoch 2/6
Epoch 2 - Train Loss: 0.4129, Val Loss: 0.7092, Val Balanced Acc: 0.6905, Val AUPRC: 0.8176

Epoch 3/6
Epoch 3 - Train Loss: 0.3936, Val Loss: 0.6624, Val Balanced Acc: 0.7361, Val AUPRC: 0.8374
  * Best model saved.

Epoch 4/6
Epoch 4 - Train Loss: 0.3560, Val Loss: 0.8804, Val Balanced Acc: 0.6607, Val AUPRC: 0.8498
  * Best model saved.

Epoch 5/6
Epoch 5 - Train Loss: 0.3220, Val Loss: 0.7728, Val Balanced Acc: 0.7417, Val AUPRC: 0.8135
  * Best model saved.

Epoch 6/6
Epoch 6 - Train Loss: 0.2861, Val Loss: 0.7747, Val Balanced Acc: 0.7673, Val AUPRC: 0.8500
  * Best model saved.
```

Best inference result



- https://www.kaggle.com/code/skarthikguruswamy/jan16-inf-attmil-lunit-224-400tiles-29c766

Submissions sorted by private test set score

UBC Ovarian Cancer Subtype Classification and Outlier Detection (UBC-OCEAN)

Late Submission ···

Overview Data Code Models Discussion Leaderboard Rules Team **Submissions**

evaluated submission with the best Private Score is used for your final score.

■ Submissions evaluated for final score

All Successful Selected Errors

Private Score ▾

Submission and Description	Private Score	Public Score	Selected
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 2 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 2	0.42	0.38	☐
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 5 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 5	0.42	0.48	☐
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 3 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 3	0.41	0.47	☐
jan4_inf_attmil_lunit_224_400tiles - Version 2 Succeeded (after deadline) · Karthik G S · 4mo ago · Notebook jan4_inf_attmil_lunit_224_400tiles | Version 2	0.41	0.45	☐
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 6 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 6	0.41	0.47	☐
jan4_inf_attmil_lunit_224_400tiles - Version 10 Succeeded (after deadline) · Karthik G S · 4mo ago · Notebook jan4_inf_attmil_lunit_224_400tiles | Version 10	0.41	0.45	☐
jan4_inf_attmil_lunit_224_400tiles - Version 11 Succeeded (after deadline) · Karthik G S · 4mo ago · Notebook jan4_inf_attmil_lunit_224_400tiles | Version 11	0.41	0.45	☐
jan14_inf_attmil_lunit_224_400tiles_focal - Version 1 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan14_inf_attmil_lunit_224_400tiles_focal | Version 1	0.40	0.42	☐
Fork of jan4_inf_attmil_lunit_224_400tiles - Version 3 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook Fork of jan4_inf_attmil_lunit_224_400tiles | Version 3	0.39	0.42	☐

Submission sorted by public test score

UBC Ovarian Cancer Subtype Classification and Outlier Detection (UBC-OCEAN) Late Submission ···

Overview Data Code Models Discussion Leaderboard Rules Team **Submissions**

All Successful Selected Errors Public Score

Submission and Description	Private Score	Public Score	Selected
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 5 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 5	0.42	0.48	☐
march16_model5_inf_attmil_lunit_224_400tiles - Version 1 Succeeded (after deadline) · Karthik G S · 1mo ago · Notebook march16_model5_inf_attmil_lunit_224_400tiles | Version 1	0.38	0.47	☐
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 6 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 6	0.41	0.47	☐
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 3 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 3	0.41	0.47	☐
Fork of jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 4 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook Fork of jan16_inf_attmil_lunit_224_400tiles 29c766 | Versio...	0.37	0.46	☐
jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 4 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook jan16_inf_attmil_lunit_224_400tiles 29c766 | Version 4	0.39	0.46	☐
jan4_inf_attmil_lunit_224_400tiles - Version 11 Succeeded (after deadline) · Karthik G S · 4mo ago · Notebook jan4_inf_attmil_lunit_224_400tiles | Version 11	0.41	0.45	☐
jan4_inf_attmil_lunit_224_400tiles - Version 10 Succeeded (after deadline) · Karthik G S · 4mo ago · Notebook jan4_inf_attmil_lunit_224_400tiles | Version 10	0.41	0.45	☐
jan4_inf_attmil_lunit_224_400tiles - Version 2 Succeeded (after deadline) · Karthik G S · 4mo ago · Notebook jan4_inf_attmil_lunit_224_400tiles | Version 2	0.41	0.45	☐
Fork of jan16_inf_attmil_lunit_224_400tiles 29c766 - Version 5 Succeeded (after deadline) · Karthik G S · 3mo ago · Notebook Fork of jan16_inf_attmil_lunit_224_400tiles 29c766 | Versio...	0.38	0.45	☐

Paper publication details

ICMBNT 2025 acceptance

ICMBNT-2025 Acceptance Notification- (Paper ID: ICMBNT_002) External Inbox ×

ICMBNT2025
to me, ganeshk1

Thu, Mar 20, 11:57 AM

Dear Authors,

Congratulations!! **(Paper ID: ICMBNT_002)**
Your submitted article entitled: **"Optimizing an attention-based MIL architecture for generalizing on the UBC-OCEAN"** has been accepted based on the positive recommendations given by the ICMBNT technical committee.

Comments:
1. Send paper in word format
2. Citation should be in order of reference like [1], [2], [3]

Participants are requested to register for the (Conference with Fee of 3500 INR) either through Demand Draft (DD) or online payment on or

INTERNATIONAL CONFERENCE ON MULTIDISCIPLINARY BREAKTHROUGHS AND NEXTGEN TECHNOLOGIES (ICMBNT – 2025)
Hybrid Conference (Inperson +Virtual)
VENUE : SRM HOTEL - MARAIMALAI NAGAR - CHENNAI - INDIA
Organized by
SOCIETY FOR CYBER INTELLIGENT SYSTEMS
Puducherry – India

Date of Conference: 26.4.2025 & 27.4.2025

ELSEVIER Scopus Web of Science Journal Indexing

Conference Theme

ICMBNT 2025 focuses on the integration of knowledge, methods and approaches from different fields or disciplines like engineering, Medicine, Management, social Science, Humanities to address complex problems and to develop holistic solutions for global issues.
On the other hand, NextGen Technologies centers around fully automated machines, future Technological advancements and developments that are poised to reshape industries, societies, and economies. It Includes advanced robotics, AI, IoT, Cybersecurity, quantum computing, 3-D printing, 5G wireless networks, virtual reality augmented reality and blockchain.
This theme sets the stage for discussions around the disruptive technologies that will shape the future, fostering innovation and collaboration across disciplines

Broad Areas

STEM (Science ,Technology,Engineering,Mathematics)
HEALTH AND LIFE SCIENCE
FINANCE, BUSINESS,
MANGEMENT ECONOMIC AND ACCOUNTING
ARTS, HUMANITIES AND SOCIAL SCIENCES
EDUCATION, TEACHING LEARNING AND ASSESSMENT
SPORTS AND PHYSIOTHERAPHY
NEXT -GEN TECHNOLOGIES

Important Dates

Manuscript Submission Deadline	: 15.3.2025
Acceptance	: 25.3.2025
Registration Date	: 05.4.2025
Date of Conference	: 26.4.2025 & 27.4.2025

+91-9445690101
icmbnt2025@gmail.com
www.icmbnt2025.societycis.org

REGISTRATION

CONFERENCE FEE (Excluding Publication Fee)

Research scholars/Students	: 2750 INR
Academicians	: 3500 INR
Industry and Others	: 4000 INR
Foreign Authors	: 100 USD

BANK A/C DETAILS :

MELANGE PUBLICATIONS
Bank A/C No : 736805000791
IFSC Code : ICIC0007368
SWIFT Code : ICICINBBCTS
Bank Name : ICICI BANK Branch
VILLIANUR, PUDUCHERRY

Key Note Speaker

Dr. Sam Goundar
Senior Lecturer in Information Technology
RMIT University, Hanoi, Vietnam

Madeshwaran Sekkarapatti Ramasamy
Research Associate at University of Cambridge, UK

ICMBNT2025 proceedings

PROCEEDINGS

ISSN : 2456-1983

International Conference on Multidisciplinary Breakthroughs and NextGen Technologies

(ICMBNT - 2025)

26 April 2025

Organized by

SOCIETY FOR CYBER INTELLIGENT SYSTEMS

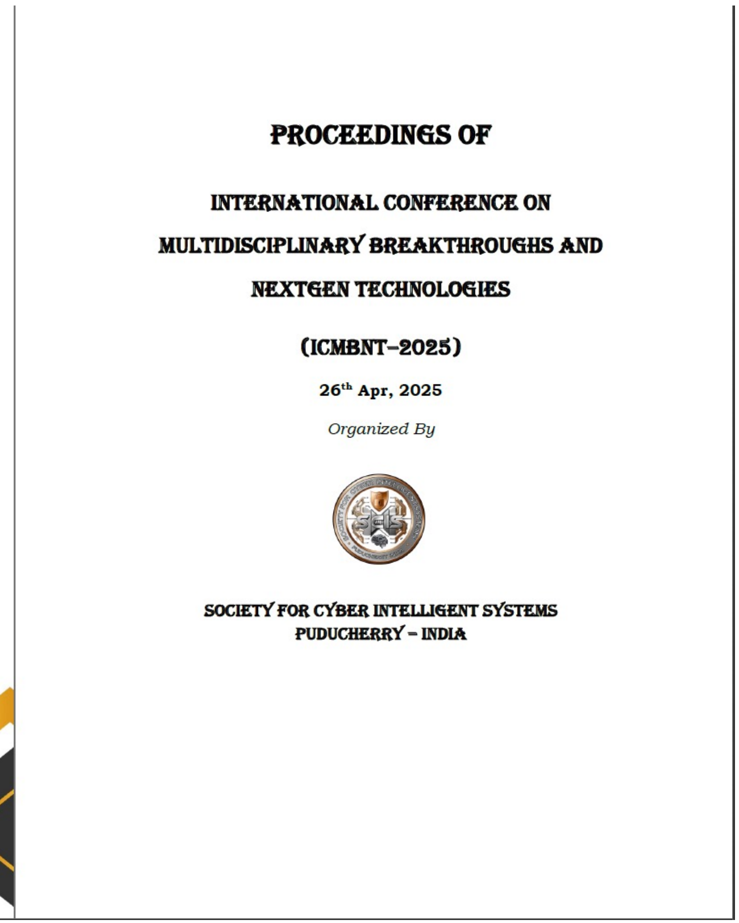
Reg.No. 78 of 2025

Puducherry - india

www.societycis.org

Melange Publications

www.melangepublications.com



PROCEEDINGS OF

INTERNATIONAL CONFERENCE ON MULTIDISCIPLINARY BREAKTHROUGHS AND NEXTGEN TECHNOLOGIES

(ICMBNT–2025)

26th Apr, 2025

Organized By

SOCIETY FOR CYBER INTELLIGENT SYSTEMS

PUDUCHERRY – INDIA

ICMBNT proceedings (paper)

BACKGROUND

Until October 2024, there was a publication embargo on the UBC-OCEAN challenge which is now lifted. The present literature review studies the methods adopted for cancer subtyping of whole slide images, especially MIL approaches.

Ovarian cancer is one of the deadliest cancers faced by women all over the world. Asian countries face a significant burden of this disease and early diagnosis and intervention is crucial. Demand for qualified histopathologists is also high and artificial intelligence could be a valuable assistant in this regard helping expand availability and to address the problem of inter-observer variability. Ovarian cancer typically affects the skin of the ovaries and epithelial ovarian cancer presents itself as a heterogeneous disease with five common types: high-grade serous carcinoma (HGSC), clear cell ovarian carcinoma (CCOC), endometrioid carcinoma (ENOC), low- grade serous carcinoma (LGSC) and mucinous carcinoma (MUC). HGSC is the most common accounting for 70% of the cases. [1] [4] Accurate classification is important as treatment management is dependent on the histotype with each type exhibiting distinctive characteristics.

A major benchmarking study that addresses the challenge of subtyping, tumour localization and grading using whole slide images "implemented and systematically compared six different methods" that include the use of variety of feature extractors (ResNets, EfficientNet, Vision Transformers) as well as varieties of weakly supervised (WSL) and multiple instance learning (MIL) methods such as classical MIL, attention based MIL and CLAM (Clustering-constrained attention MIL). [6] They also evaluated these methods on six different WSI datasets covering different cancerous tissues (renal, colorectal, gastric, bladder).

Another study [2] focused on the development of deep learning-based MIL systems that used only the labels and showed high performance on an extensive variety of cancer datasets without requiring pixel-wise annotations. This showed the potential for applying deep learning to drive decision- making in clinical settings. The effectiveness of MIL techniques for histopathological image classification on breast

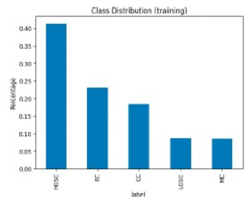


Figure 2. Imbalance in ovarian cancer subtypes
(UBC-OCEAN training)

cancer images has been demonstrated by [7]. In [8], different transfer learning approaches for classifying epithelial ovarian cancer WSI using a VGG19 backbone were studied.

Attention based MIL which produces a scalar attention score for each tile with two fully connected layers was introduced by [9].

CLAM, introduced by [10] has been empirically shown to outperform classical MIL by them. Among the five datasets considered by [6], their results show that ABMIL outperformed CLAM on five datasets as

measured by the area under the receiver operating curve (AUROC). Both ABMIL and CLAM were shown to outperform classical MIL.

Transfer learning has now been accepted as a crucial part of tasks on medical images and [11] have studied how CNNs and ViTs can be used effectively for domain transfer of weights.

Choice of pre-trained domain specific weights over standard IMAGENET weights has been shown to be critical to achieve the best performance.

2. Methodology and Experiments

Methodological considerations

Generalization performance is the main consideration for this dataset. As already noted, the training dataset has 5% TMAs while the private test set has up to 80% TMAs. A variety of approaches were adopted by the top entries in the competition. One approach was to use 60% of the TMAs for validation as adopted by [12]. The choice of appropriate feature extractors and the use of transfer learning plays an important role in generalization performance as well as in reducing training time which is a relevant constraint for this work. [13]

Feature extractors: Choice of feature extractor decides the details of the architecture. Some of the common choices available that are state of the art in many use cases are: ResNets, Inception networks, ViTs, EfficientNets.

Evaluation: Evaluation of the multi-class classification was done using balanced accuracy score as required by the competition. [3] Balanced accuracy is computed as the average recall across all classes and is necessary to accurately assess the performance of imbalanced classification.

Baseline

For baseline, an InceptionV3 network was chosen. The idea was to choose something different from ViTs used in our target architecture for the sake of comparison.

Our initial choice for baseline and target architecture was efficient nets as these are designed to perform well in resource constrained environments. The difficulty with these experiments was that the way they scaled up did not fit well with the GPU/memory constraints especially when our ABMIL algorithm which requires heavy parallelization was applied.

Eventually, InceptionV3 was used as it could also fit a minimum number of tiles required to demonstrate ABMIL. These experimental choices were necessary to study the performance improvements of key modules of the target architecture.

Architecture

Fig. 3 shows the overall training architecture.

Preprocessing:

A significant proportion of the WSIs in the training set had repetitions of the same image and these were cropped out. Both WSIs and TMAs were also cropped to fit to improve training efficiency.

Splitting:

60% of the TMAs were reserved for validation in line with the need to generalize on TMAs expected in the public and private test sets. WSIs were split at 70/30 which is a more conventional split.

Feature extraction:

Pre-trained vision transformer weights from Lunit DINO with image sizes of 224x224 and using two patch sizes (8 and 16) were used to extract features from each image and were concatenated to produce the input to the ABMIL classifier. [5] The DINO training methodology was introduced by [14] and is noted to be a natural fit with vision transformers.

Certificate of participation

INTERNATIONAL CONFERENCE ON MULTIDISCIPLINARY BREAKTHROUGHS AND NEXTGEN TECHNOLOGIES (ICMBNT - 2025)

CERTIFICATE

Of PARTICIPATION

This is to certify that Mr/Ms. Karthik G Shanmugasundaram *of* SRM Institute of Science and Technology, Chennai, India *has presented the paper titled* Optimizing an Attention-based MIL Architecture for Generalizing on the UBC-OCEAN *in the International Conference on Multidisciplinary Breakthroughs and NextGen Technologies (ICMBNT - 2025) held on 26 April 2025, Hybrid Conference (In-person + Virtual) SRM HOTEL - Maraimalai Nagar - CHENNAI – India.*

Organized By

SOCIETY FOR CYBER INTELLIGENT SYSTEMS

Reg. No. 78 of 2025

Puducherry - India

PRESIDENT

SOCIETY FOR CYBER INTELLIGENT SYSTEMS

Future scope

- Foundation models tuned to domain
- Attention heatmaps for interpretability
- Explore ensembling and augmentation
- Studying and improving outlier detection

Conclusion

- Top 10% private leaderboard
- Efficient ViT + ABMIL pipeline
- Trained within Kaggle constraints
- Insights and observations
  - ABMIL + ViTs = strong generalization
  - Cropping helped training, not inference
  - Full tiling improved test scores

References (1/4)

[1] H. Farahani et al., ‘Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images’, Modern Pathology, vol. 35, no. 12, pp. 1983–1990, 2022.

[2] G. Campanella et al., ‘Clinical-grade computational pathology using weakly supervised deep learning on whole slide images’, Nature Medicine, vol. 25, no. 8, pp. 1301–1309, 2019.

[3] UBC (Kaggle Profile). UBC-OCEAN research competition homepage. Retrieved Mar 12, 2025 - https://www.kaggle.com/competitions/UBC-OCEAN

[4] M. Asadi-Aghbolaghi et al., ‘Machine Learning-driven Histotype Diagnosis of Ovarian Carcinoma: Insights from the OCEAN AI Challenge’, medRxiv, 2024.

[5] Kang, M., et al, ‘Benchmarking self-supervised learning on diverse pathology datasets’, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 3344–3354.

[6] N. G. Laleh et al., ‘Benchmarking weakly-supervised deep learning pipelines for whole slide classification in computational pathology’, Medical Image Analysis, vol. 79, p. 102474, 2022.

References (2/4)

[7] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, 'Multiple instance learning for histopathological breast cancer image classification', Expert Systems with Applications, vol. 117, pp. 103–111, 2019.

[8] Y. Wang et al., 'Classification of epithelial ovarian carcinoma whole- slide pathology images using deep transfer learning', arXiv preprint arXiv:2005.10957, 2020.

[9] M. Ilse, J. Tomczak, and M. Welling, 'Attention-based deep multiple instance learning', in International Conference on Machine Learning, 2018, pp. 2127–2136.

[10] M. Y. Lu, D. F. K. Williamson, T. Y. Chen, R. J. Chen, M. Barbieri, and F. Mahmood, 'Data-efficient and weakly supervised computational pathology on whole-slide images'. Nature Biomedical Engineering, vol. 5, pp. 555–570, 2021.

[11] Christos Matsoukas, Johan Fredin Haslum, Moein Sorkhei, Magnus Söderberg, and Kevin Smith. 'What makes transfer learning work for medical images: Feature reuse & other factors', in CVPR, pp. 9225–9234, 2022.

[12] Owkin. Kaggle 1st place solution report. Retrieved Jan 31, 2025 from https://www.kaggle.com/competitions/UBC-OCEAN/discussion/466455

References (3/4)

[13] A. Filiot et al., ‘Scaling Self-Supervised Learning for Histopathology with Masked Image Modeling’, medRxiv, 2023.

[14] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. ‘Emerging properties in self-supervised vision transformers.’ in ICCV, pp. 9650–9660, 2021.

[15] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. ‘An image is worth 16x16 words: Transformers for image recognition at scale.’ in ICLR, 2020.

[16] Zznznb (Kaggle Profile). Kaggle 2nd place solution report. Retrieved Mar 6, 2025 from https://www.kaggle.com/competitions/UBC-OCEAN/discussion/465410

[17] J. Breen, K. Allen, K. Zucker, L. Godson, N. M. Orsi, and N. Ravikumar, ‘Histopathology Foundation Models Enable Accurate Ovarian Cancer Subtype Classification’, arXiv preprint arXiv:2405.09990, 2024.

References (4/4)

[18] M1dsolo (Kaggle Profile). Kaggle 7th place solution report. Retrieved Apr 30, 2025 from https://www.kaggle.com/competitions/UBC-OCEAN/discussion/465697

[19] Gunes Evitan (Kaggle Profile). Kaggle 10th place solution report. Retrieved Apr 30, 2025 from https://www.kaggle.com/competitions/UBC-OCEAN/discussion/465455

[20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.

[21] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. 36th Int. Conf. Mach. Learn. (ICML), Long Beach, CA, USA, Jun. 2019, pp. 6105–6114.

Thank You