بسم الله الرحمن الرحیم

# CLIPath

## Clip for multimodal model based on WSI image and report

Mohammad Hosein Movasaghinia, Reza Abbasi, Hasan Alikhani, Hosein Jafarinia,
Mahdi Shadroy, Ali Salmani

Natural Language Processing course
Dr. Asgari

**Sharif University of Technology**
**March 2023**

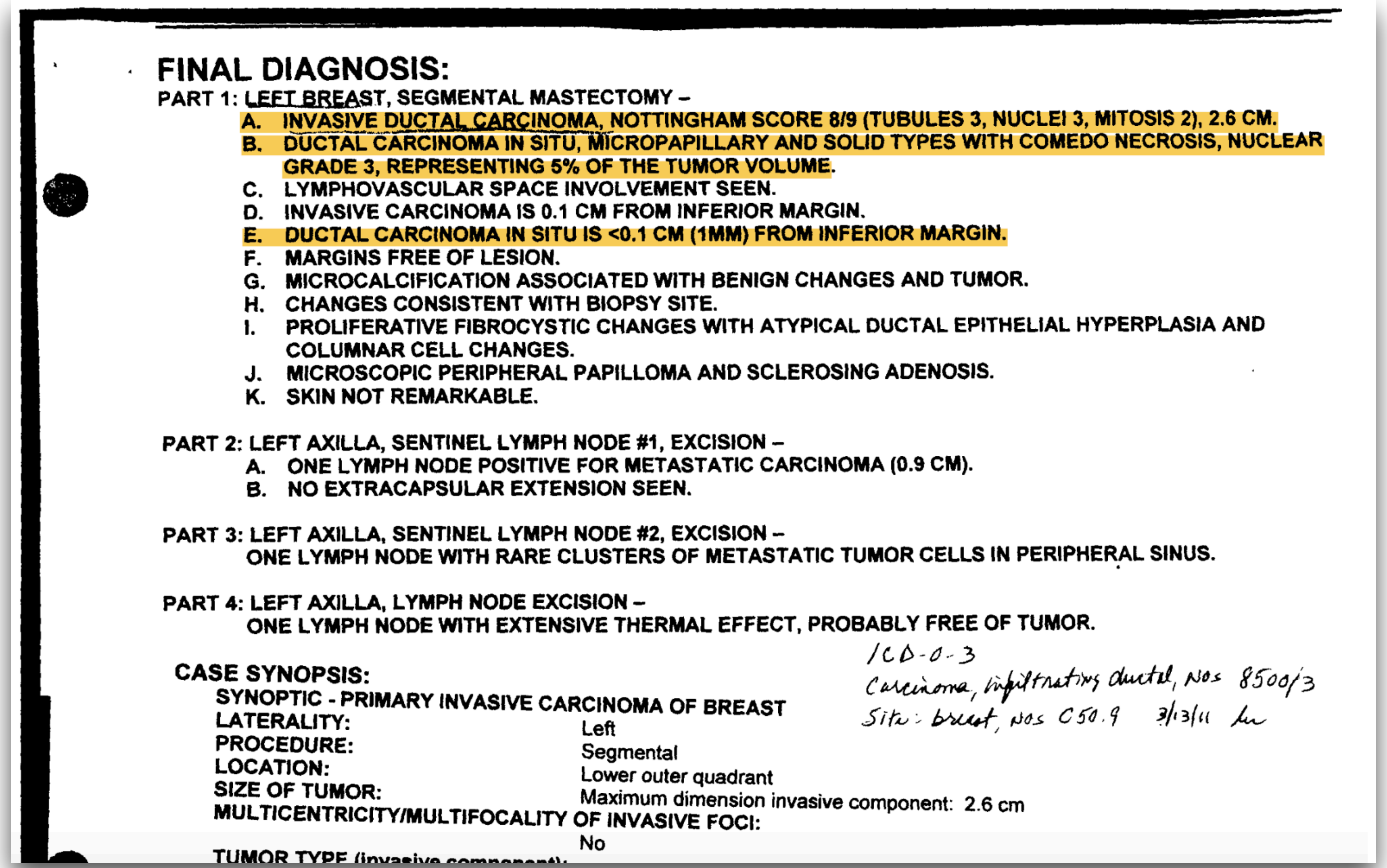
# Introduction

# WSI (Whole Slide Image)


120K×40K

TCGA dataset [1]

# Reports


FINAL DIAGNOSIS:

PART 1: LEFT BREAST, SEGMENTAL MASTECTOMY –

A. INVASIVE DUCTAL CARCINOMA, NOTTINGHAM SCORE 8/9 (TUBULES 3, NUCLEI 3, MITOSIS 2), 2.6 CM.
B. DUCTAL CARCINOMA IN SITU, MICROPAPILLARY AND SOLID TYPES WITH COMEDO NECROSIS, NUCLEAR GRADE 3, REPRESENTING 5% OF THE TUMOR VOLUME.
C. LYMPHOVASCULAR SPACE INVOLVEMENT SEEN.
D. INVASIVE CARCINOMA IS 0.1 CM FROM INFERIOR MARGIN.
E. DUCTAL CARCINOMA IN SITU IS <0.1 CM (1MM) FROM INFERIOR MARGIN.
F. MARGINS FREE OF LESION.
G. MICROCALCIFICATION ASSOCIATED WITH BENIGN CHANGES AND TUMOR.
H. CHANGES CONSISTENT WITH BIOPSY SITE.
I. PROLIFERATIVE FIBROCYSTIC CHANGES WITH ATYPICAL DUCTAL EPITHELIAL HYPERPLASIA AND COLUMNAR CELL CHANGES.
J. MICROSCOPIC PERIPHERAL PAPILLOMA AND SCLEROSING ADENOSIS.
K. SKIN NOT REMARKABLE.

PART 2: LEFT AXILLA, SENTINEL LYMPH NODE #1, EXCISION –

A. ONE LYMPH NODE POSITIVE FOR METASTATIC CARCINOMA (0.9 CM).
B. NO EXTRACAPSULAR EXTENSION SEEN.

PART 3: LEFT AXILLA, SENTINEL LYMPH NODE #2, EXCISION –

ONE LYMPH NODE WITH RARE CLUSTERS OF METASTATIC TUMOR CELLS IN PERIPHERAL SINUS.

PART 4: LEFT AXILLA, LYMPH NODE EXCISION –

ONE LYMPH NODE WITH EXTENSIVE THERMAL EFFECT, PROBABLY FREE OF TUMOR.

ICD-O-3
Carcinoma, infiltrating ductal, NOS 8500/3
Site: breast, NOS C50.9 3/13/11

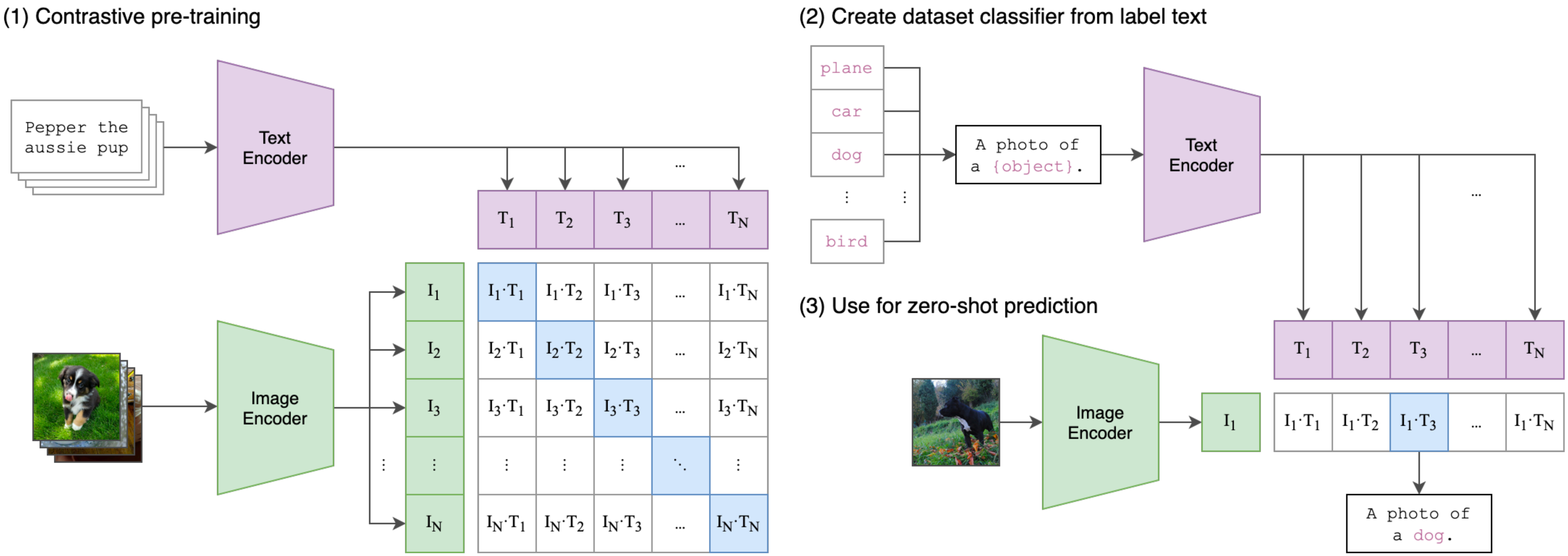
CASE SYNOPSIS:

SYNOPTIC - PRIMARY INVASIVE CARCINOMA OF BREAST
LATERALITY: Left
PROCEDURE: Segmental
LOCATION: Lower outer quadrant
SIZE OF TUMOR: Maximum dimension invasive component: 2.6 cm
MULTICENTRICITY/MULTIFOCALITY OF INVASIVE FOCI: No
TUMOR TYPE (invasive component):

INVASIVE DUCTAL CARCINOMA, DUCTAL CARCINOMA IN SITU, MICROPAPILLARY AND SOLID TYPES WITH COMEDO NECROSIS. DUCTAL CARCINOMA IN SITU IS NEAR INFERIOR MARGIN.

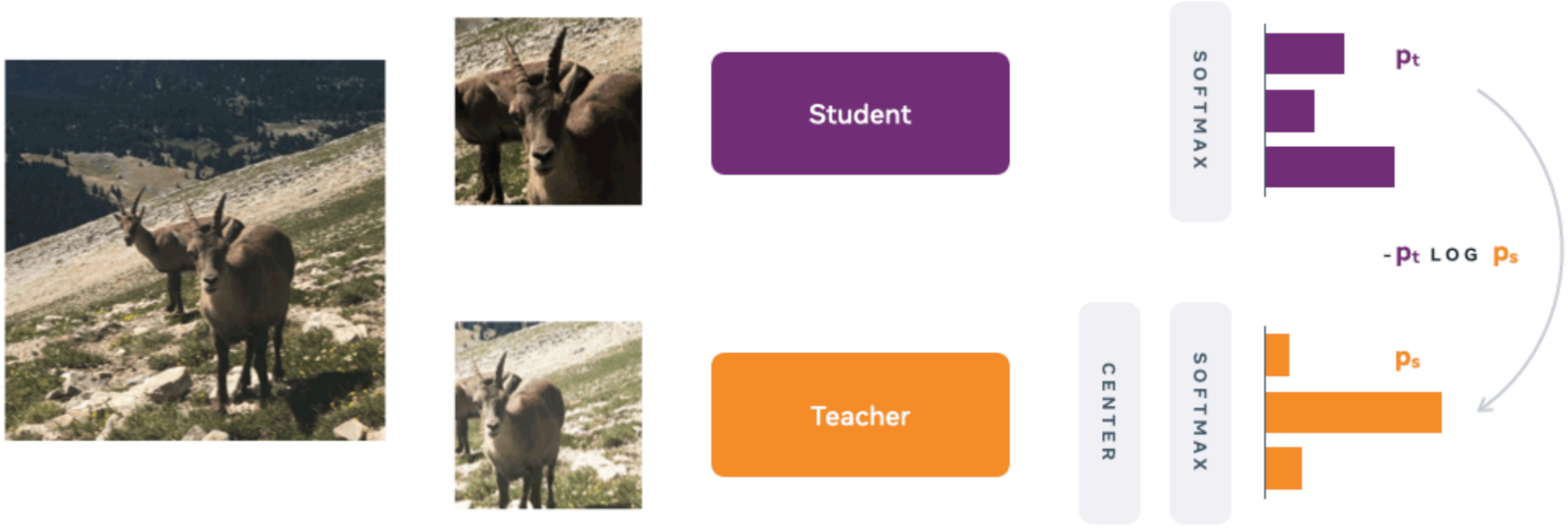
TCGA dataset [1]

# CLIP



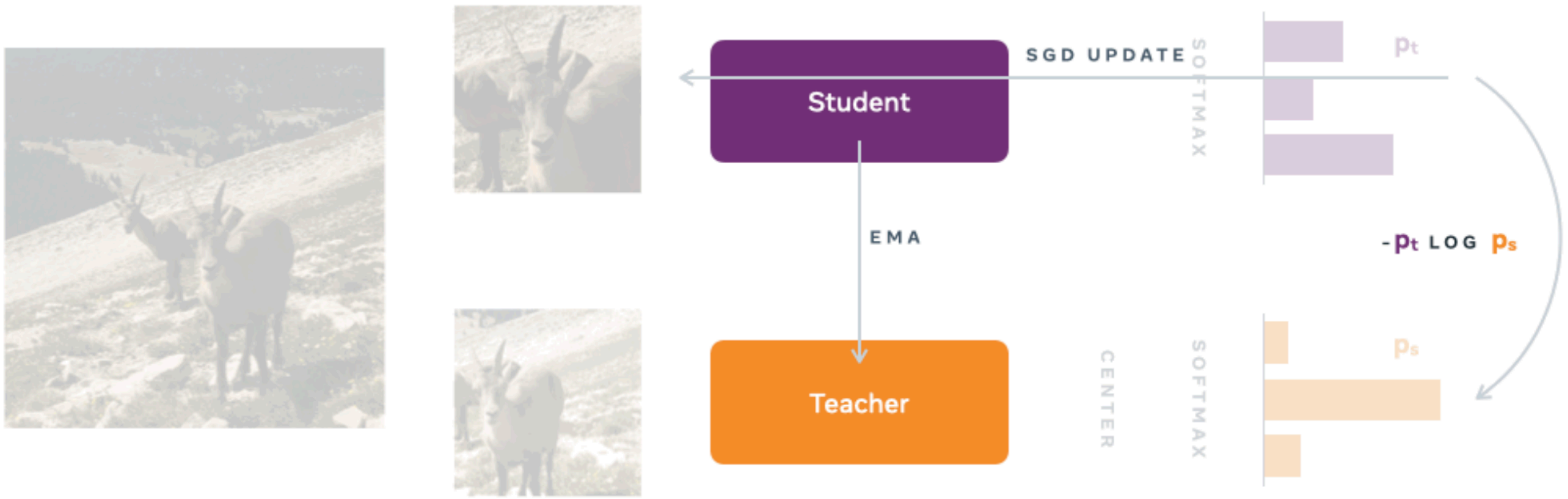

Learning transferable visual models from natural language supervision (CLIP) [2]

# DINO




Emerging properties in self-supervised vision transformers (DINO) [3]

# DINO




Emerging properties in self-supervised vision transformers (DINO) [3]

# Pathology BERT

## PathologyBERT - Pre-trained Vs. A New Transformer Language Model for Pathology Domain

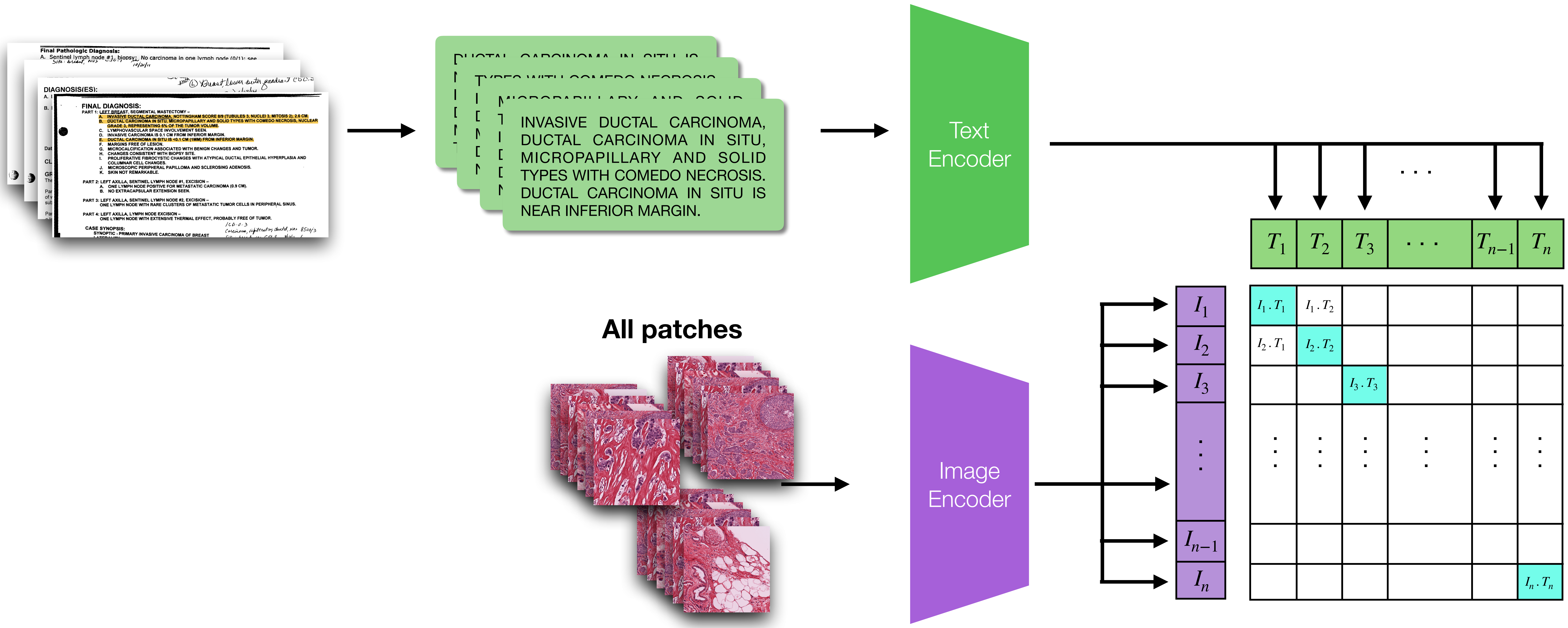
**Thiago Santos, MS[1], Amara Tariq, Ph.D.[2], Susmita Das, MS[3], Kavyasree Vayalpati, MS[4], Geoffrey H. Smith, MD[5], Hari Trivedi, MD[6], Imon Banerjee, Ph.D[2,4]**
**[1]Emory University, Department of Computer Science, Atlanta, Georgia, USA; [2]Mayo Clinic, Phoenix, Arizona, USA;[3]Indian Institute of Technology (IIT), Centre of Excellence in Artificial Intelligence, Kharagpur, West Bengal, India;[4]Arizona State University, School of Computing and Augmented Intelligence, Tempe, Arizona, USA;[5]Emory University, Department of Pathology, Atlanta, Georgia, USA;[6]Emory University, Department of Radiology, Atlanta, Georgia, USA;**

**Abstract** *Pathology text mining is a challenging task given the reporting variability and constant new findings in cancer sub-type definitions. However, successful text mining of a large pathology database can play a critical role to advance 'big data' cancer research like similarity-based treatment selection, case identification, prognostication, surveillance, clinical trial screening, risk stratification, and many others. While there is a growing interest in developing language models for more specific clinical domains, no pathology-specific language space exist to support the rapid data-mining development in pathology space. In literature, a few approaches fine-tuned general transformer models on specialized corpora while maintaining the original tokenizer, but in fields requiring specialized terminology, these models often fail to perform adequately. We propose PathologyBERT - a pre-trained masked language model which was trained on 347,173 histopathology specimen reports and publicly released in the Huggingface[1]repository. Our comprehensive experiments demonstrate that pre-training of transformer model on pathology corpora yields performance improvements on Natural Language Understanding (NLU) and Breast Cancer Diagnose Classification when compared to nonspecific language models.*

A bert model generates diagnostically relevant semantic embeddings from pathology synopses with active learning (PathologyBERT) [4]

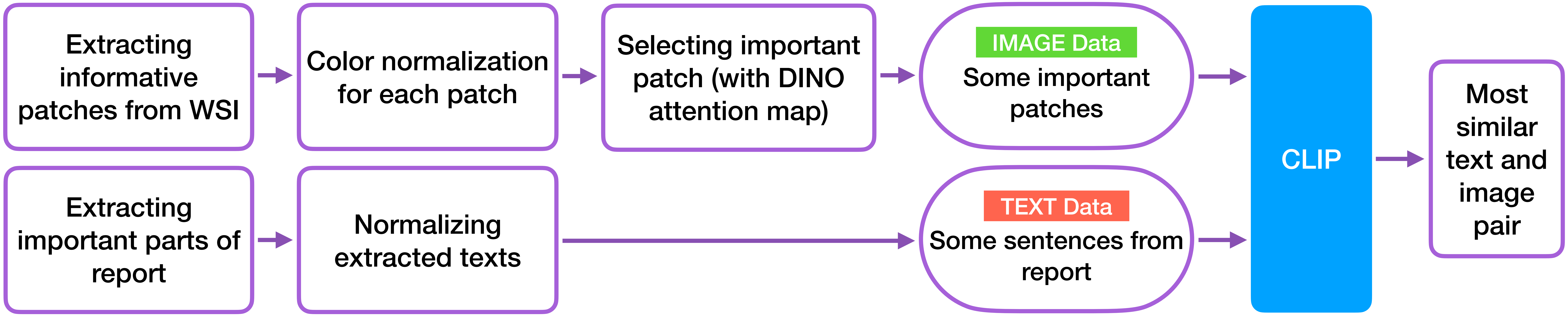
# Baseline

# Baseline model

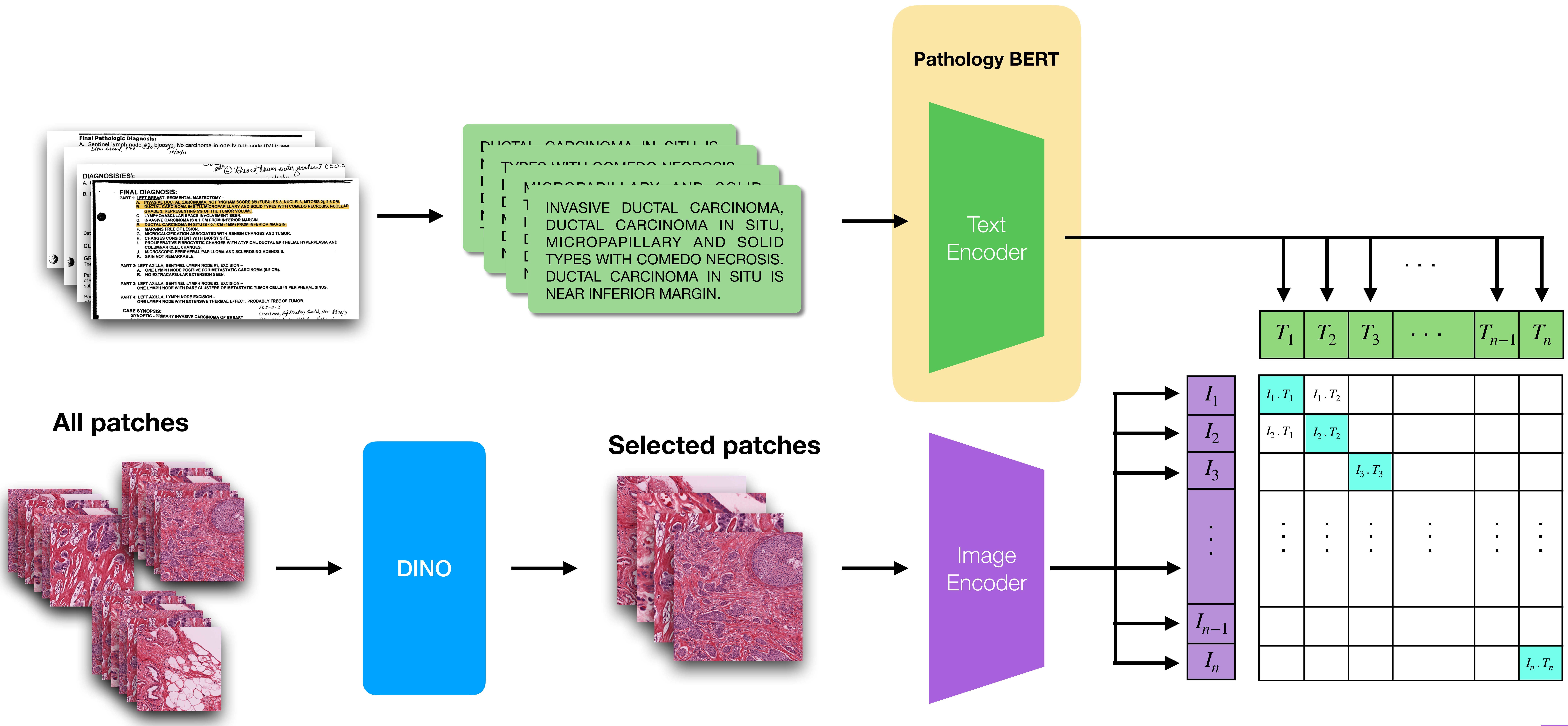


# Proposed Model

# Pipeline


Extracting informative patches from WSI
Color normalization for each patch
Selecting important patch (with DINO attention map)
IMAGE Data
Some important patches
Extracting important parts of report
Normalizing extracted texts
TEXT Data
Some sentences from report
CLIP
Most similar text and image pair

# Proposed model



# Results

# Discussion

# Discussion

- We create new dataset (selected part of report and informative patch of WSI)
- Our proposed model get better Accuracy and lower Loss within all reports
- The limitation of computing resources was one of the challenges of this project, which did not allow testing better and more methods.
- There is still a lot of work to be done and it is possible to provide better solutions to improve the model.

# Future works

# Future works

- More effective way for first stage of patch selection
- Using other method for main model (e.g. MIL, Ava. Pool, …)
- Using creative method to select important part of report automated (e.g. MIL)
- Using more powerful computational units (e.g. GPU with more memory)

# References

# References


[1] National Cancer Institute 2023, The Cancer Genome Atlas Program (TCGA), accessed 9 March 2023, https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga

[2] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In *International conference on machine learning* (pp. 8748-8763). PMLR.

[3] Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 9650-9660).

[4] Mu, Y., Tizhoosh, H. R., Tayebi, R. M., Ross, C., Sur, M., Leber, B., & Campbell, C. J. (2021). A bert model generates diagnostically relevant semantic embeddings from pathology synopses with active learning. *Communications Medicine*, *1*(1), 11.

Thank For
Your
Attention :)