

Deep learning-based detection and classification of tumor and lymphocyte cells in breast cancer IHC

Elin Samuelsson¹, Jean-Eudes Le Douget¹, Jérémy Guillon¹, Benoit Schmauch¹, David Lin¹, Alexandre Filiot¹, Mina Farag¹, Corentin Wiscart¹, Hortense Deslandes¹, Zoé Vaquette¹, Katharina von Loga¹

¹ Waiv, formerly Owkin Dx, Paris, France.

Background

Tumor and lymphocyte cells carry critical prognostic and predictive information. Immunohistochemistry (IHC) is central to identifying relevant biomarkers, yet quantitative interpretation is time-consuming and variable. Current pipelines often interpret IHC signal globally, or focus exclusively on tumor cells, limiting lymphocyte-specific analysis.

Objectives

We present a deep learning framework for the combined detection and classification of tumor and lymphocyte cells in IHC whole-slide images (WSIs).

Data

- Two independent HER2 Breast IHC cohorts: one for model training, one for external validation.
- The training dataset consists of 56,606 cells (62% tumor) in 699 tiles from 49 WSIs. 191 (27%) tiles were further reviewed, with 2,174 cells annotated as lymphocytes.
- Our external independent validation dataset consists of 19,957 cells (59% tumor, 12% lymphocytes) in 245 tiles from 52 WSIs.

	Training		Validation
	Full cohort	Revised subset	
WSIs	49	43	52
Tiles	699	191	245
Cells	56,606	15,780	19,957
Cancer cells	35,273 (62%)	8,618 (55%)	11,774 (59%)
Lymphocytes	Unknown	2,186 (14%)	2,418 (12%)

Methods

Our model was trained using the CellViT [1] architecture with an internal IHC-specific foundation model as frozen backbone.

We trained two-class (tumor, rest) and three-class (tumor, lymphocyte, rest) models.

Soft cell-level class predictions were obtained with softmax over pixel class predictions, averaged over cell pixels.

Detection and segmentation performances are jointly measured with binary panoptic quality (bPQ).

Classification performance is measured with one-VS-rest AUC computed only on matched (IOU > 0.1) cells.

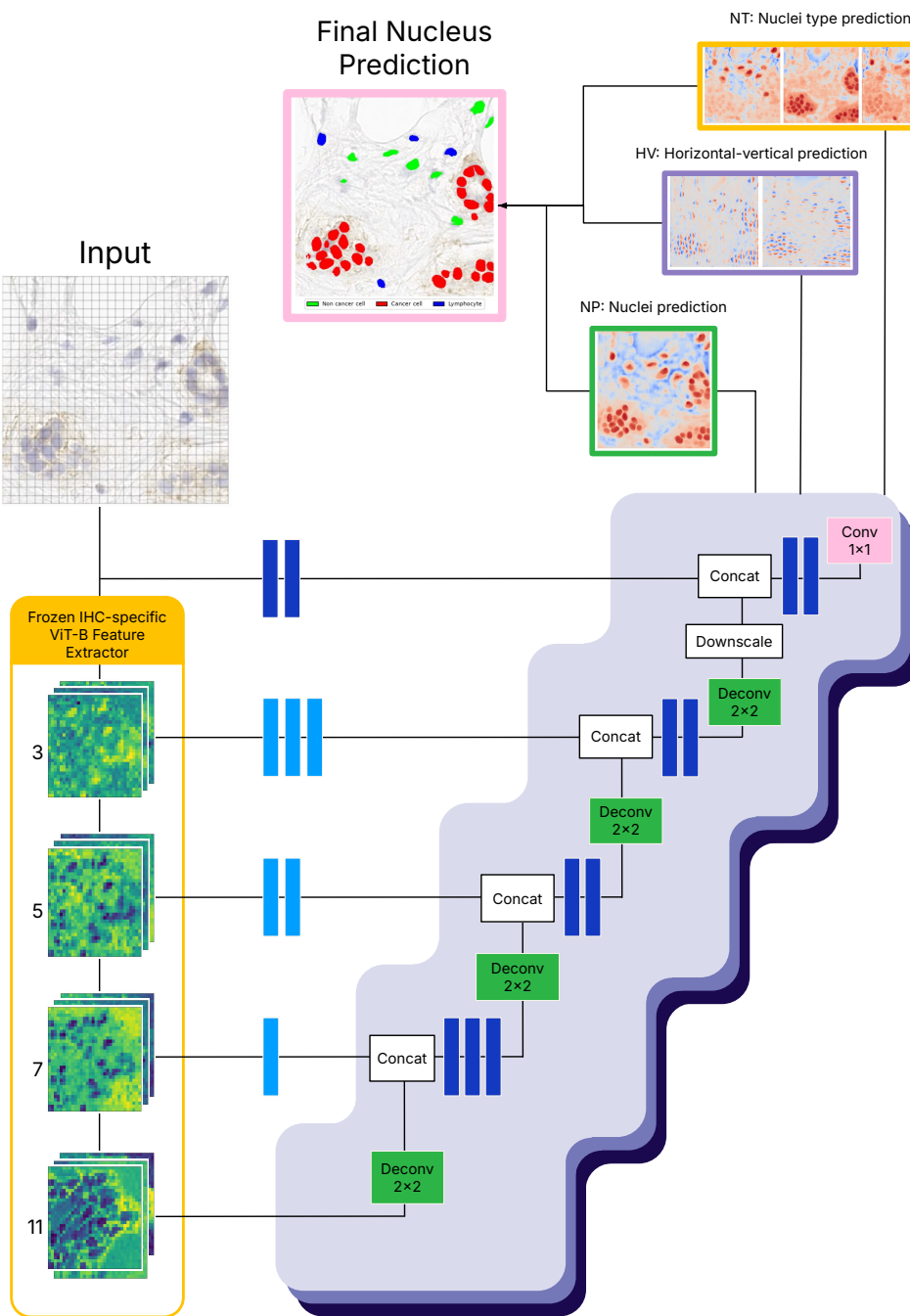


Figure 1: Overview of our CellViT derived model.

Results: 2-class vs 3-class

- The two-class and three-class models show consistent performances:
- 2-class model:** 87.9% matched cells with 70.5% bPQ and 93.1% tumor AUC
 - 3-class model:** 88.5% matched cells with 70.0% bPQ and 91.6% tumor AUC.

The 3-class model achieves **lymphocyte AUC of 92.2%**.

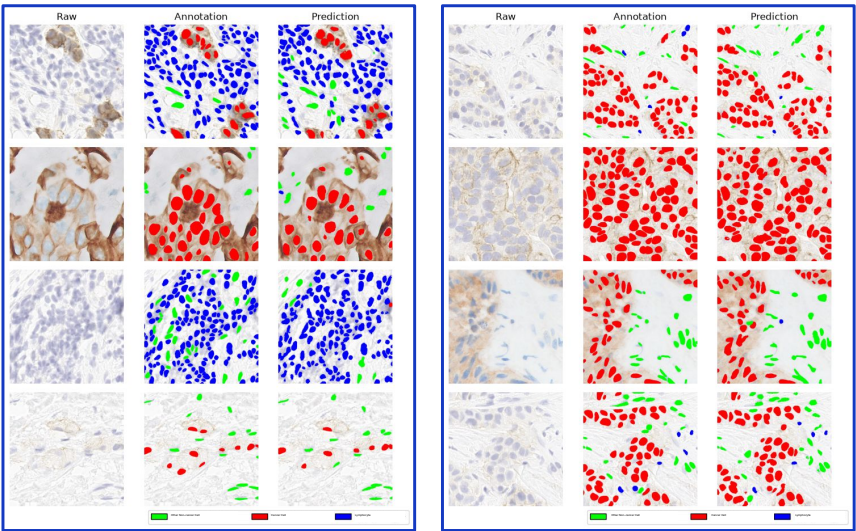


Figure 2: Examples of model results on the external validation dataset. Left column shows the raw tissue tile, center column shows the annotation as provided by an expert pathologist, and right column shows our 3-class model prediction. Tumor cells are represented in red, lymphocytes in blue, and other non-tumor cells in green.

Large Scale Example

When applied to whole-slide images, our lymphocyte-aware model allows for large-scale visualisation.

Our model thereby may enable precise spatial evaluation of lymphocyte-to-tumor relationships.

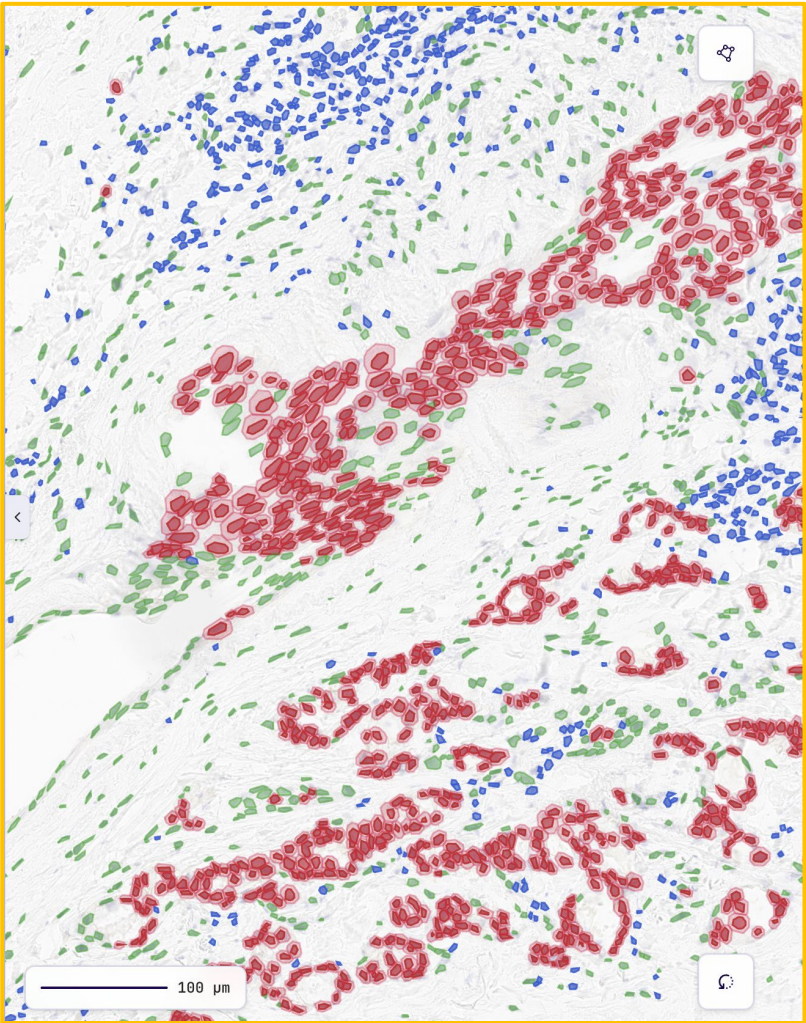


Figure 3: Example of large-scale inference. Tumor cells are represented in red, lymphocytes in blue, and other non-tumor cells in green. Tumor cells cytoplasms, computed through an additional post-processing step, are represented in light-red. around nuclei.

Conclusions

These results demonstrate the potential for **accurate analysis of tumor and lymphocytes populations in breast IHC WSIs**. Notably, impressive lymphocyte classification was achieved with few annotations and without compromising the accuracy of the two-class training. This approach facilitates reproducible cell phenotyping, supporting downstream biomarker analyses and objective spatial score quantification for clinical applications in breast cancer.

References

1. F. Hörst et al., "CellViT: Vision Transformers for precise cell segmentation and classification," *Medical Image Analysis*, May 2024.