

## Introduction

### Context & Goal:

This challenge focused on epithelial lesions of the uterine cervix and featured a unique collection of thousands expert-labeled WSIs collected from medical centers across France.

It gave access to a sizable dataset (3.3 TB) of extremely high-resolution images. Given the scale of the dataset, handling the data efficiently is a critical problem to solve in the process of developing an accurate approach to diagnosis.

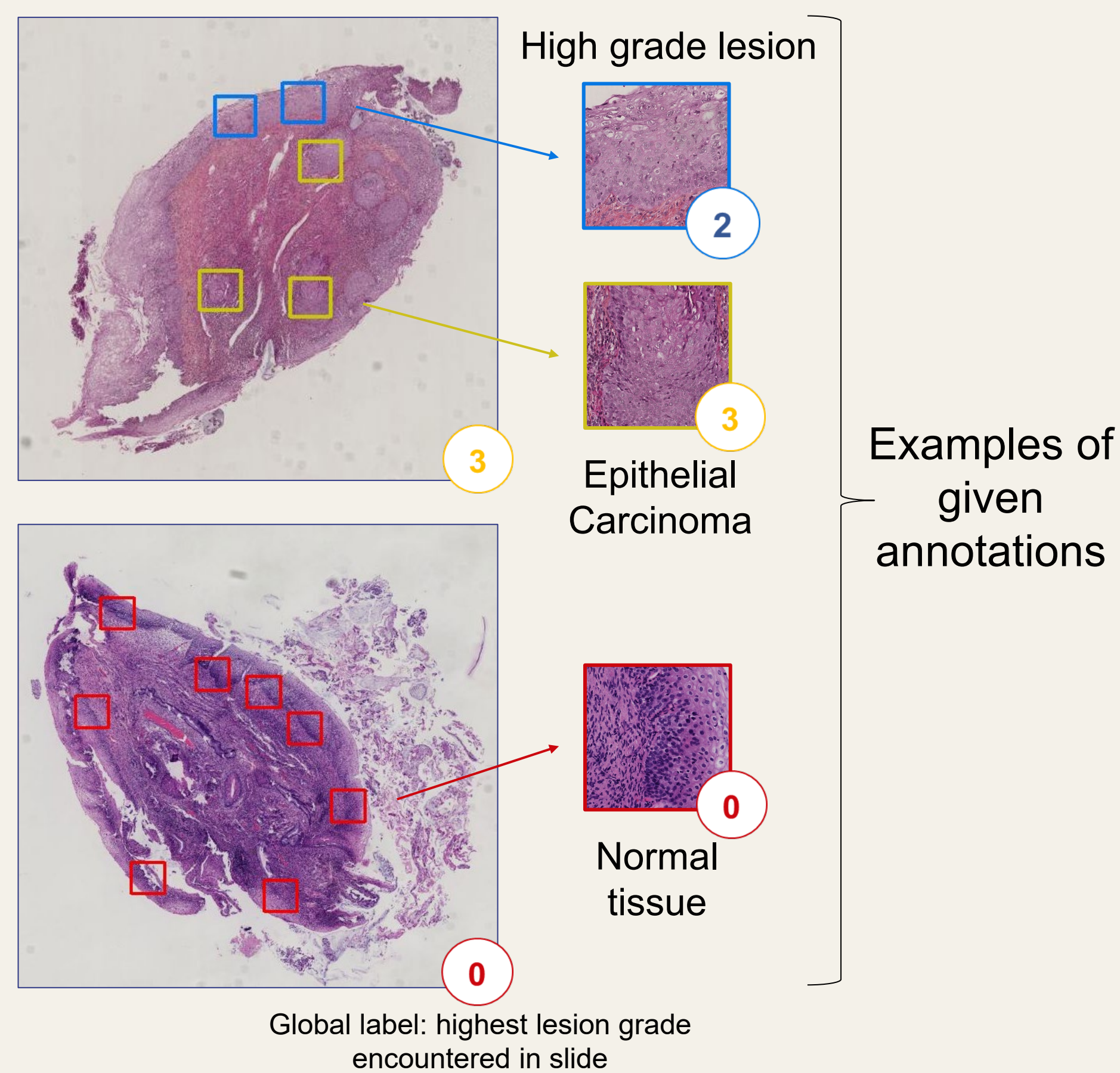
The objective is to **classify** each image according to the most severe category of epithelial lesion present in the sample.

The classes are as follows:

- 0 Normal
- 1 Low grade
- 2 High grade
- 3 Invasive cancer

### Available training data:

- 1015 slides with a global label
- 5926 local square annotations within the labeled slides
- 2048 unlabeled slides (not used in our model)



### Evaluation:

#### Testing data:

- 1500 hidden, labeled slides used for algorithm evaluation on a remote server

#### Scoring:

The score for each prediction equals 1 minus the error from the following **error table** (some errors are more penalizing than others):

		Predicted			
		Normal	Low gr.	High gr.	Invasive
Actual	Normal	0	0.1	0.7	1
	Low gr.	0.1	0	0.3	0.7
	High gr.	0.7	0.3	0	0.3
	Invasive	1	0.7	0.3	0

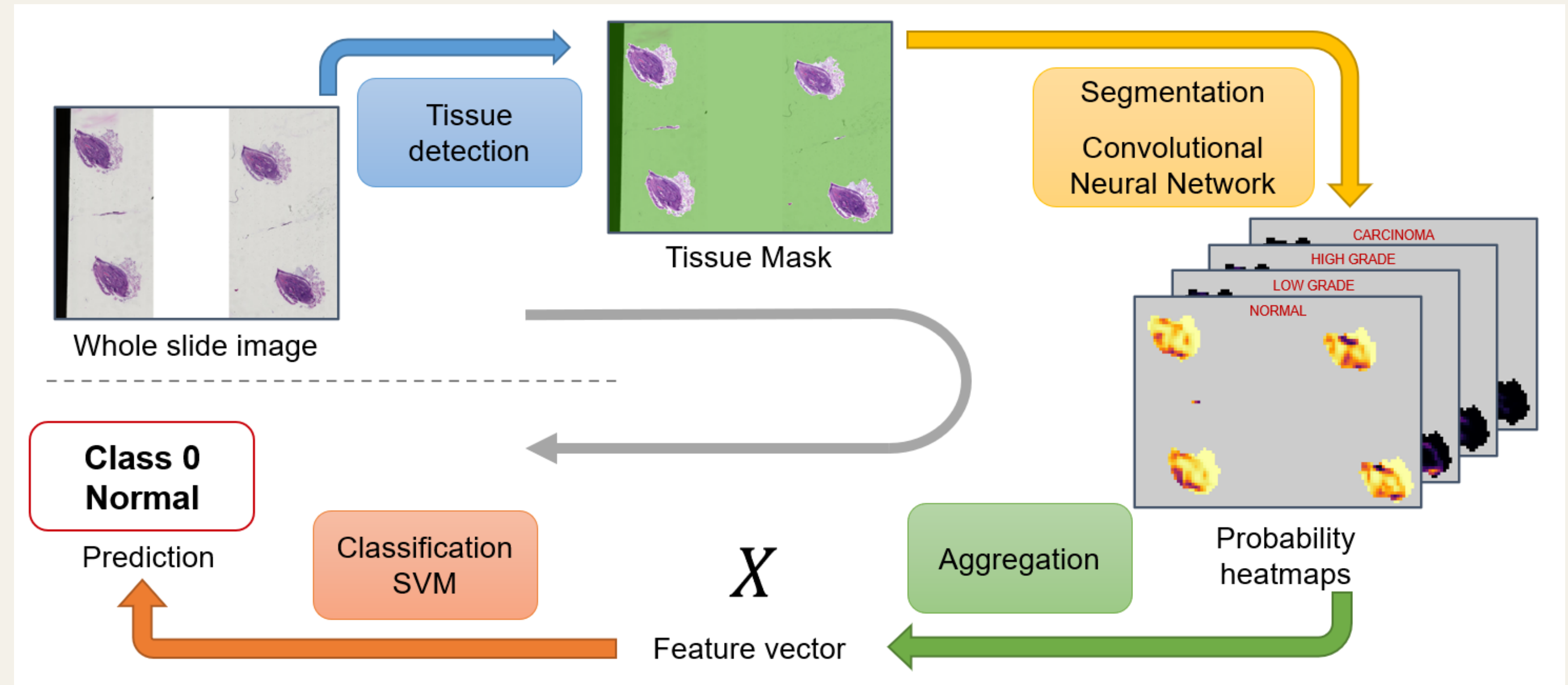
The total score is the average across all predictions.

#### Computing time constraint

- 16 h for 1500 slides (around 38 s per slide)

## Method

### Algorithm outline:



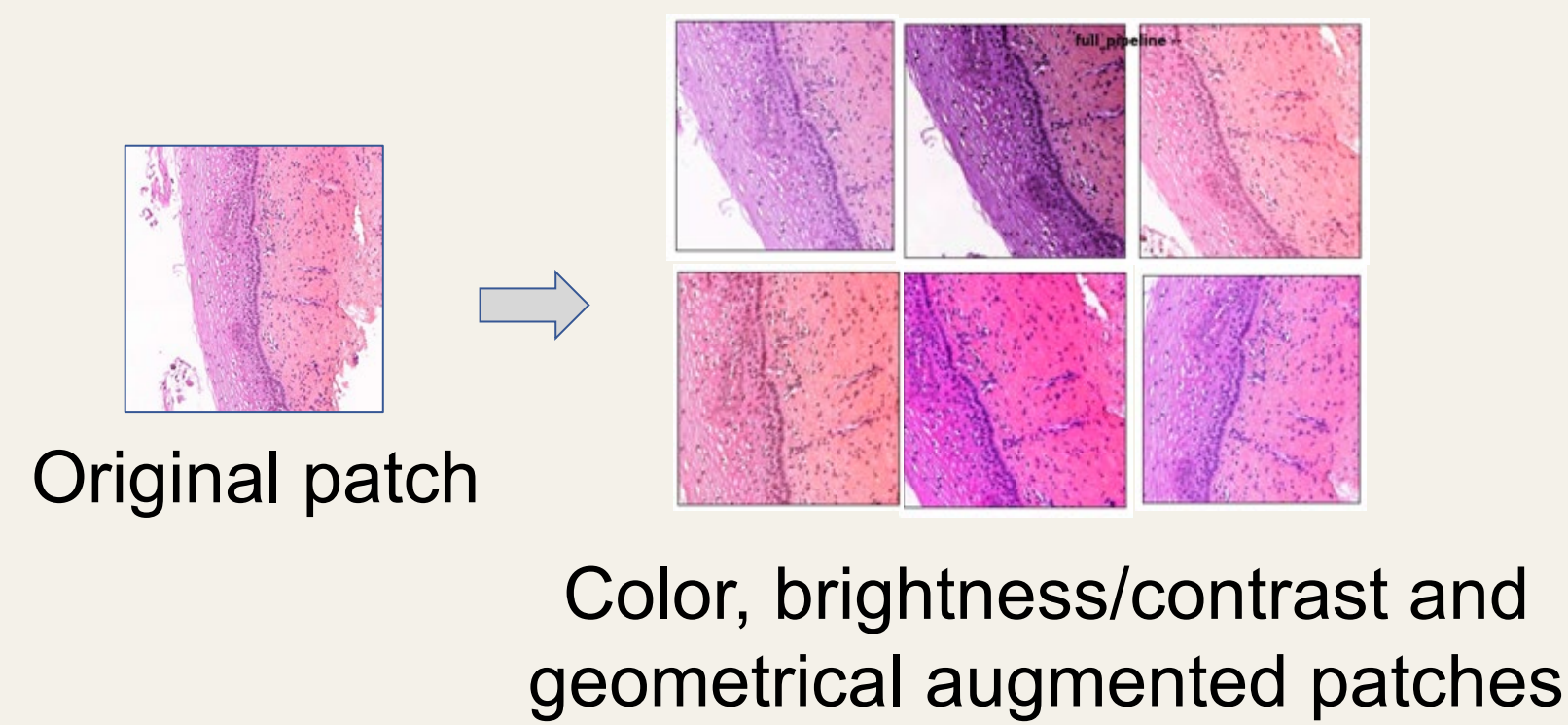
### Training:

- **DenseNet model pretrained on ImageNet** (database of millions of small images)

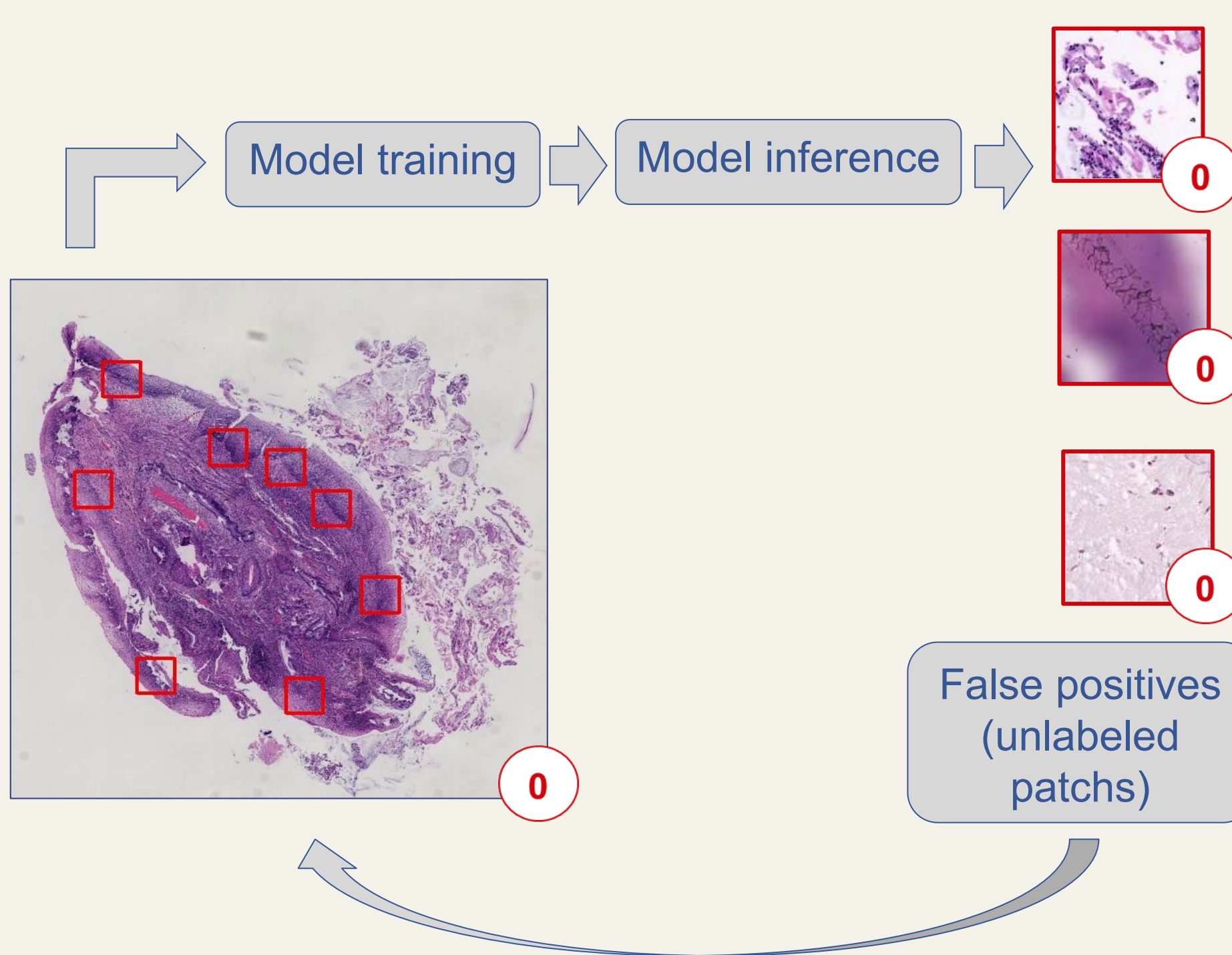


Some of the techniques that we used to improve our training focused on getting more usable data:

- **Data augmentation**

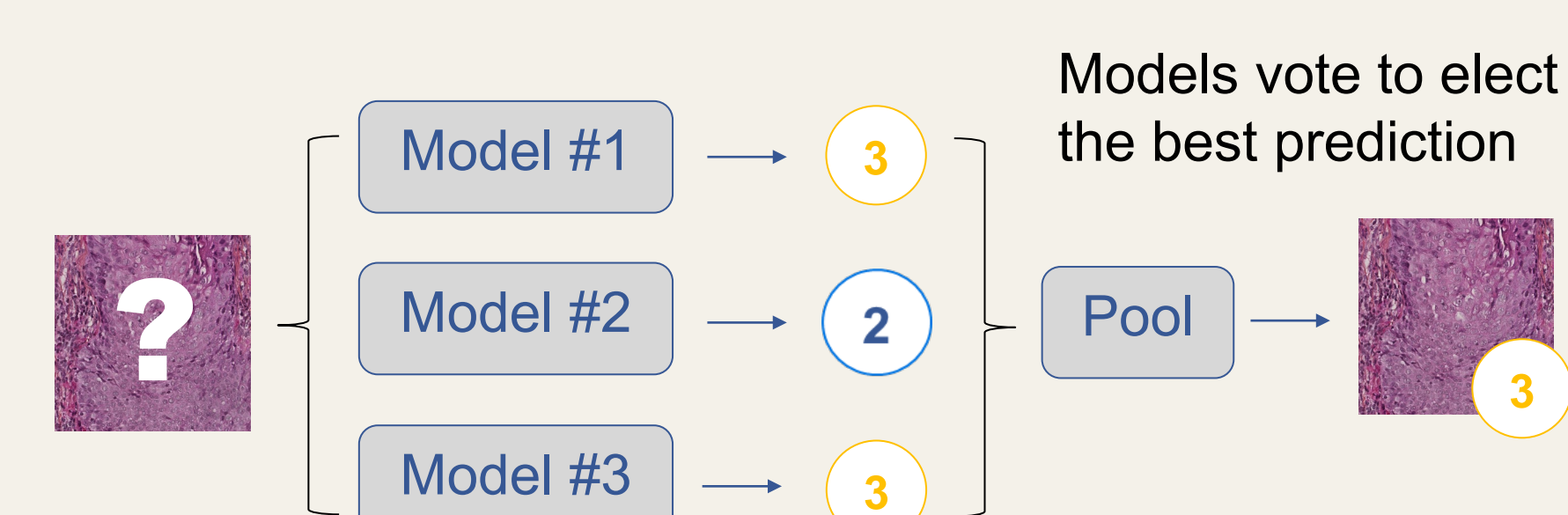


- **Hard negative mining (HNM)**

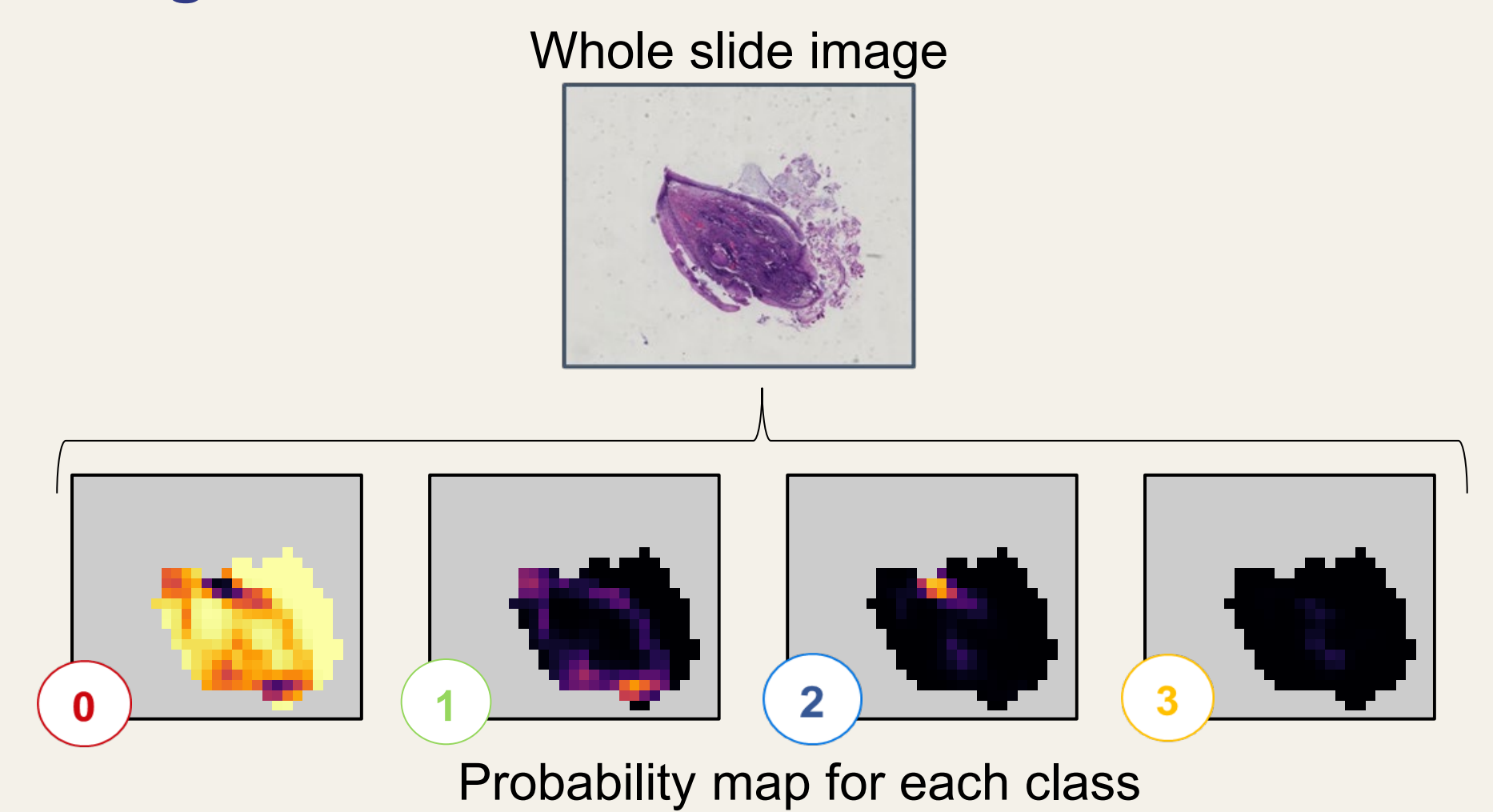


The segmentation model, already trained, is applied to slides with a *normal* global label, that are known not to contain lesions. The model sometimes predicts that there is a lesion (class 1-3) with a high probability. These false positives are added to the training set, as examples of images hard to predict. It enhances training, as the model will see images drawn from a **wider distribution**.

- **Model ensemble**

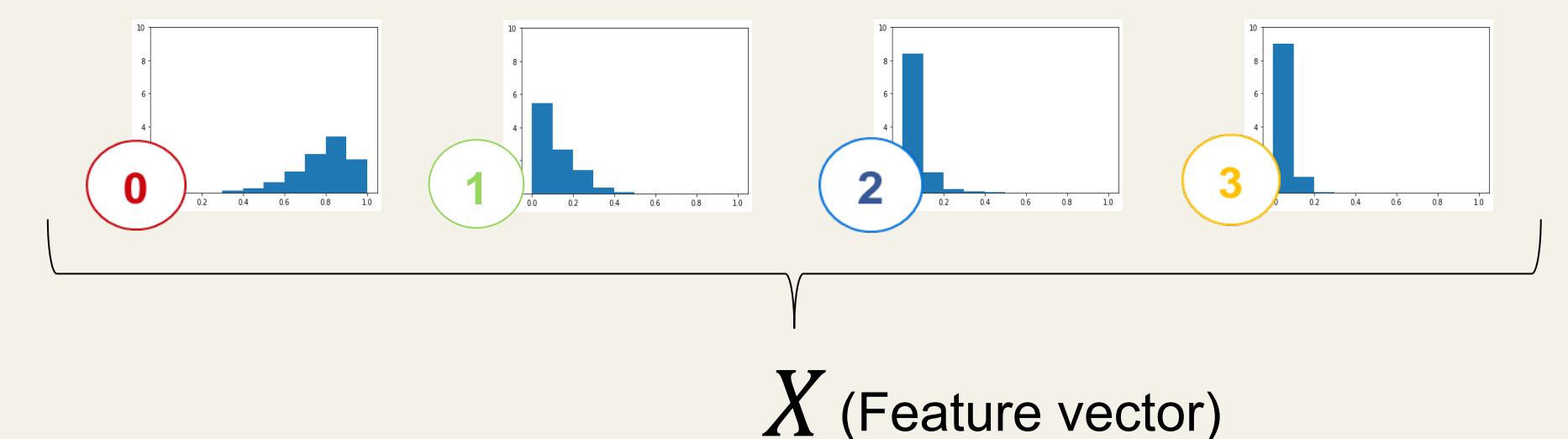


### Segmentation:



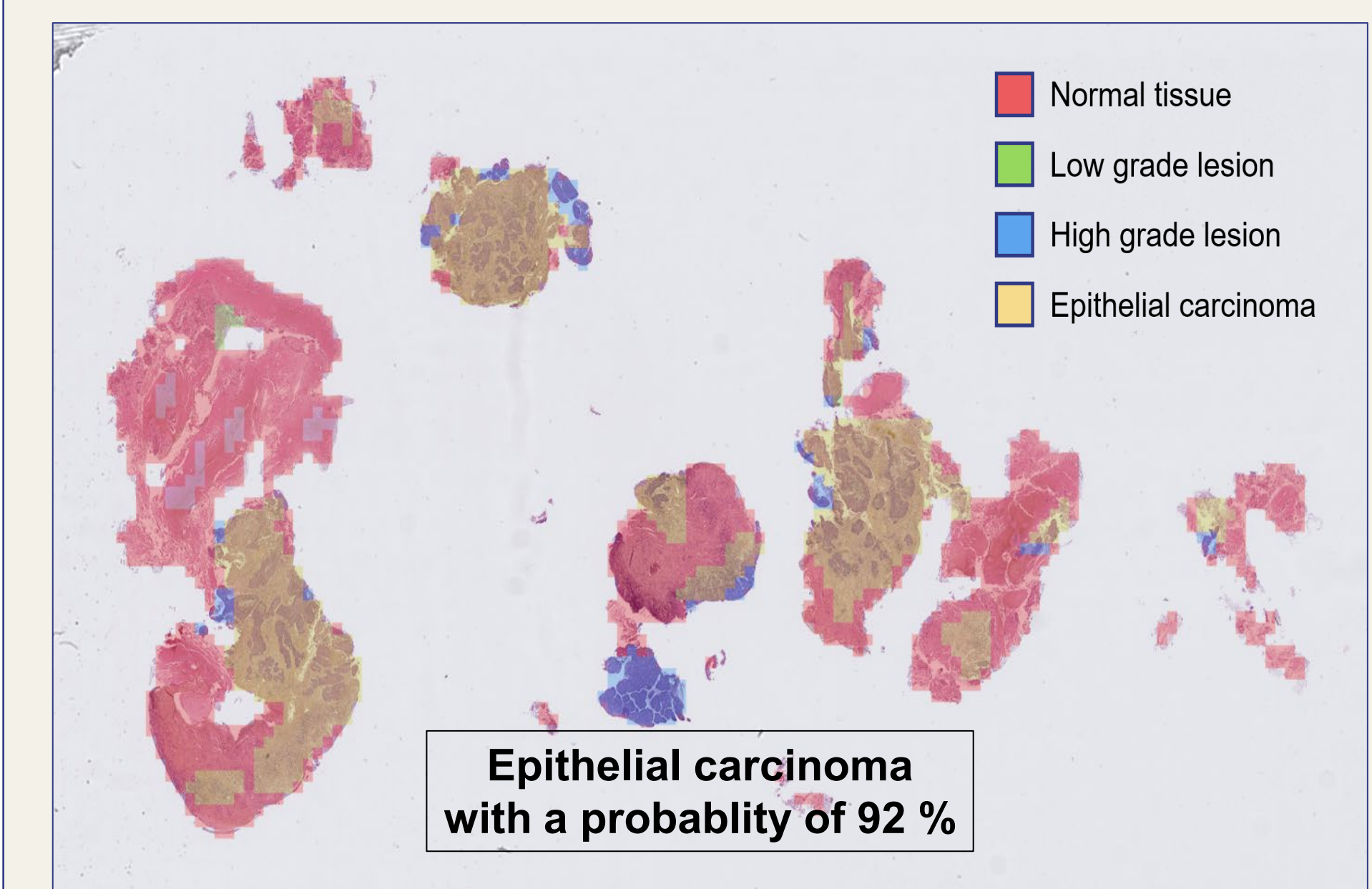
### Aggregation:

After the segmentation, we obtain four probability maps. To summarize this information, we compute distribution statistics for each map that are aggregated into a unique feature vector.



## Results

**Winner of the data challenge with a score of 0,9475.**  
Example on a whole slide image:



## Acknowledgement

This work has been performed on data that were available during the 2020 Data Challenge of the French Society of Pathology and the Health Data Hub, with the support of Grand Defi for A.I. in Health. We also thanks our collaborators : Dr. Farré and Dr Tilmant, external pathologists and F. Ciompi, our scientific chair.

## Author contact

Saïma Ben Hadj

2, rue du Capitaine  
Scott - 75015 Paris,  
France

Tel: +33 1 78 96 72 75  
Email: [sbenhadj@tribvn-hc.com](mailto:sbenhadj@tribvn-hc.com)  
Web: [www.tribvn-hc.com](http://www.tribvn-hc.com)