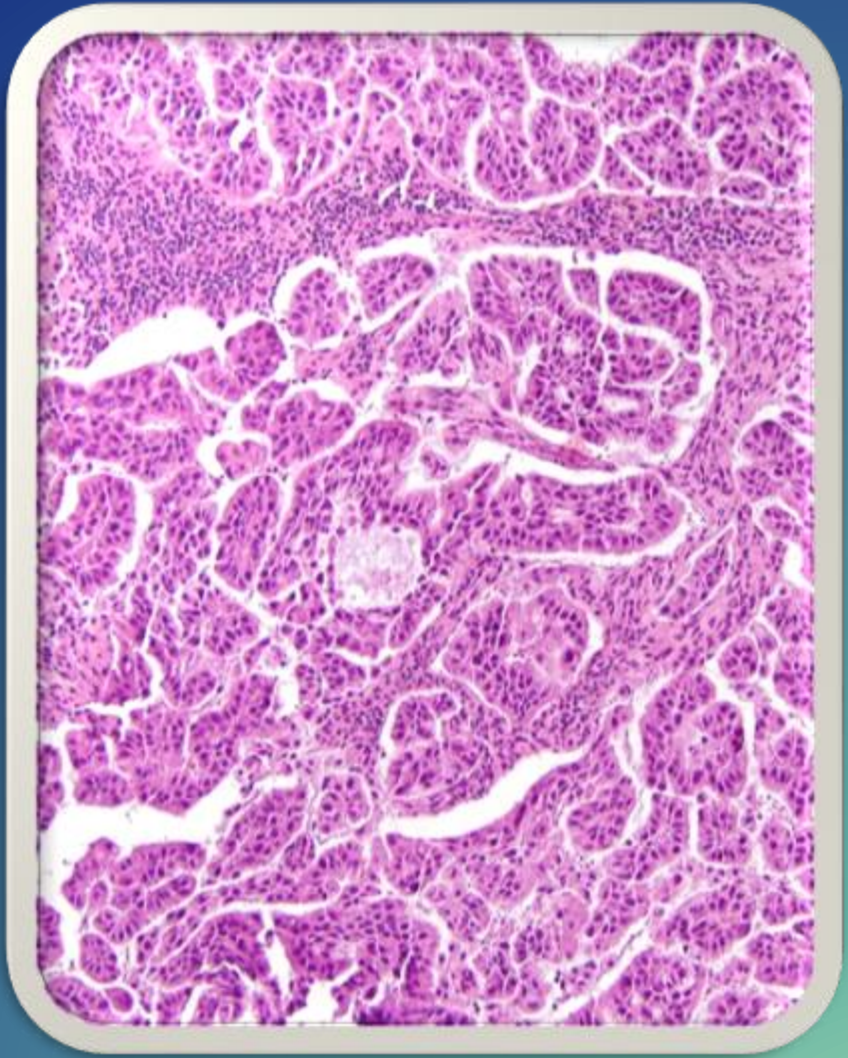
Improving Domain Generalization for Deep Learning-Powered Cancer Cell Detection

Raphaël Attias, supervised by Eric Cosatto
NECLA, ML Department



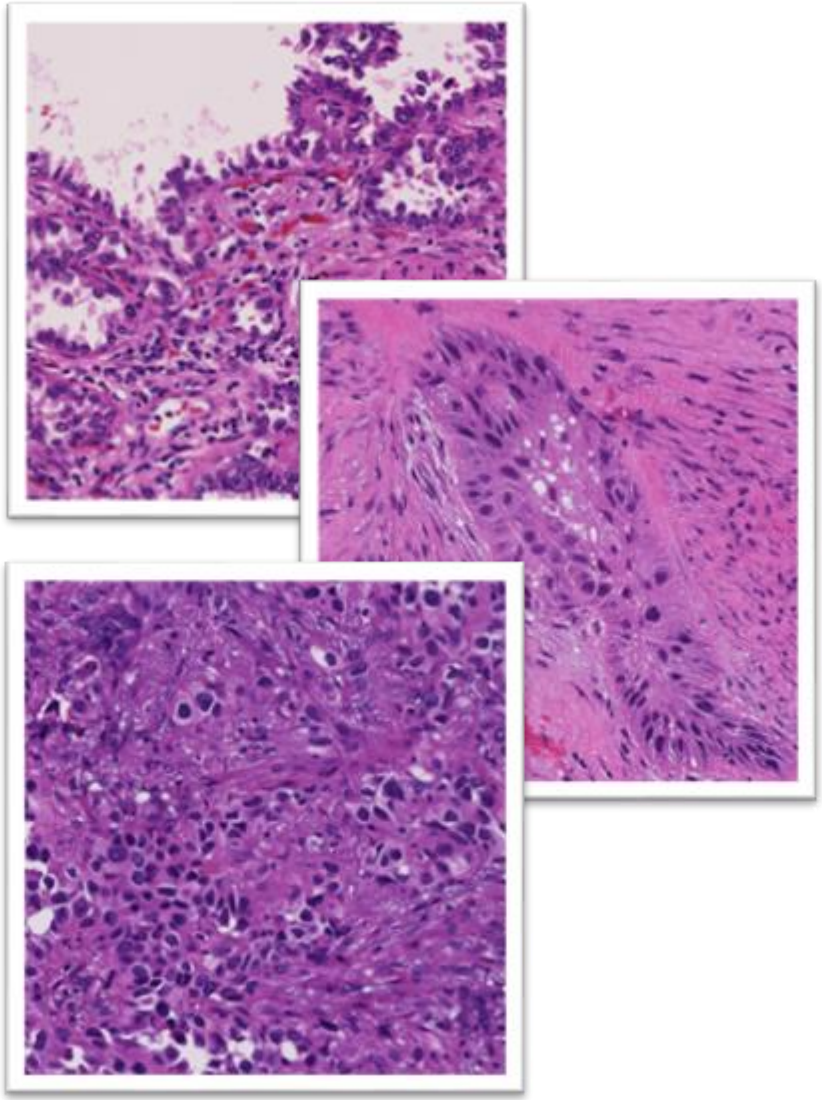
Goals and Challenges

- Goal: Segment and count number of healthy and tumor cells, in slides from patients with lung cancer.

- Real-world application: Perform genetical testing on cells that have cancer

- Challenge: Model decreases in accuracy when tested on different hospitals or slides scanner.

- Idea: Review ways to counter overfitting, such as data augmentation, Out of Distribution (OoD) estimation, representation learning etc.



Generalize cancer cell detection model to new hospitals and scanners

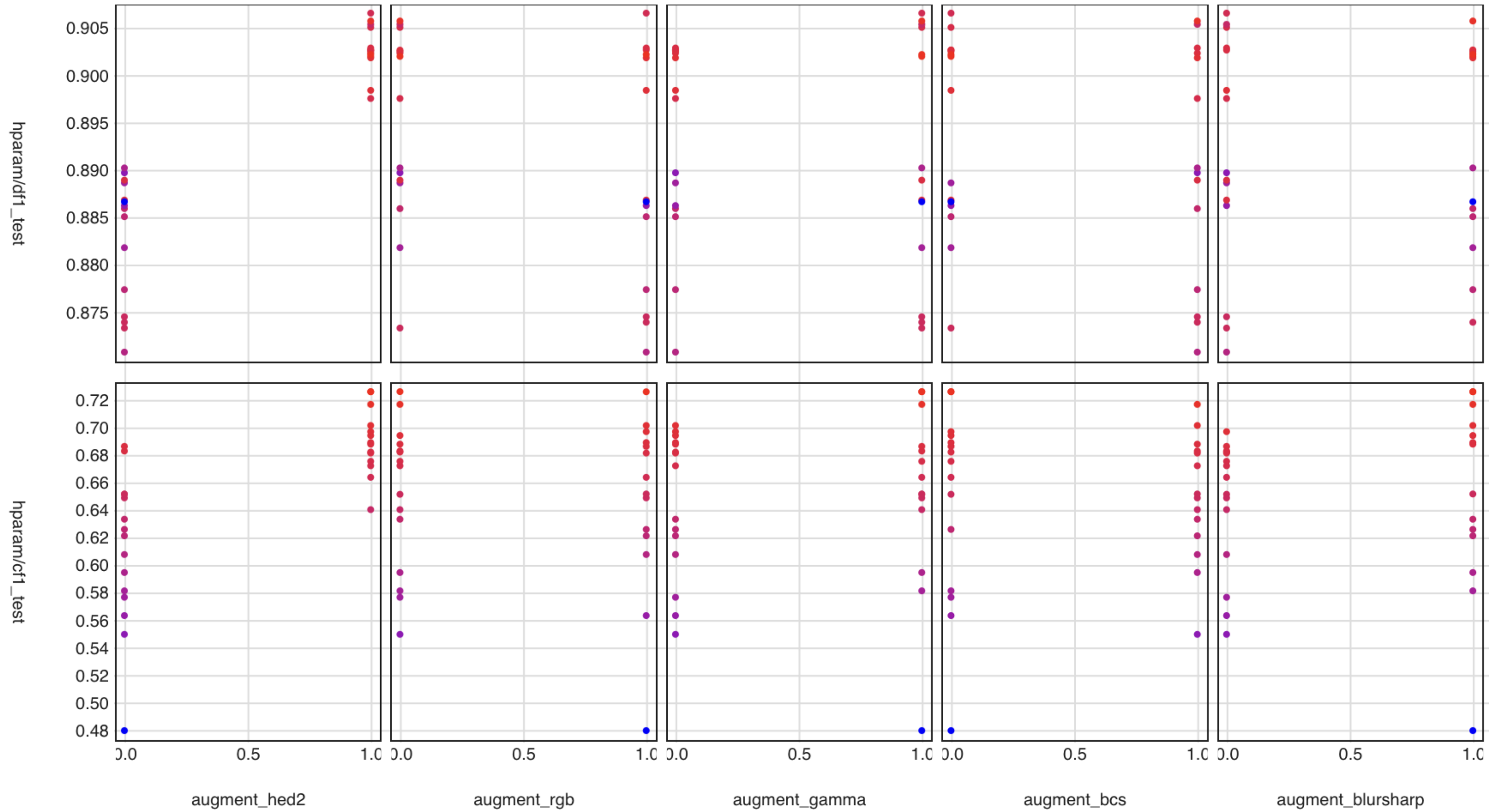
Experiments and hyperparameter search

- Slides are obtained from real-world patients, labeled by doctors and experts in Japan.

- Hyperparameter tunning and positive augmentations were identified usings loggers such as weight & biases and tensorboard hyperparameter search.

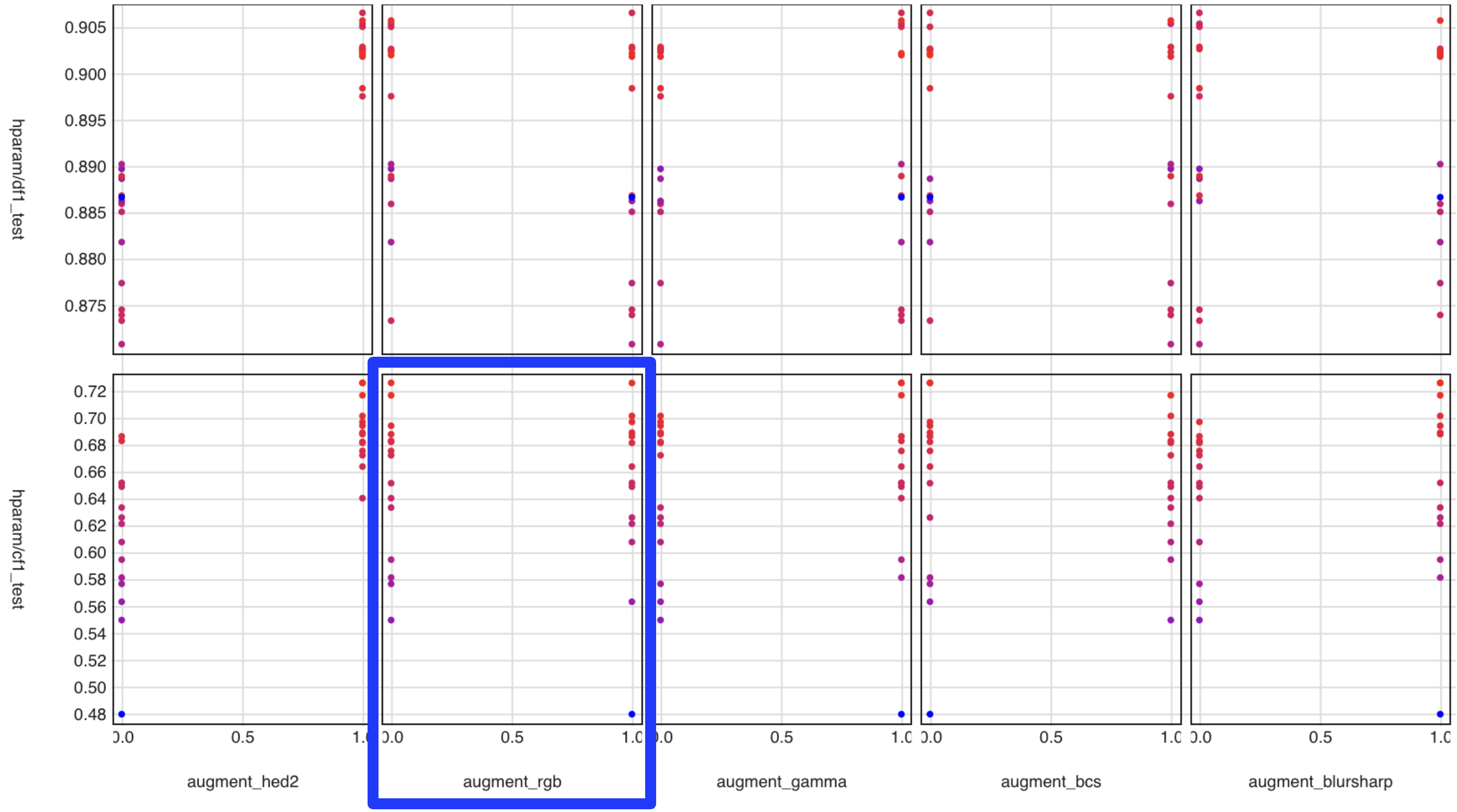
Hundreds runs done to validate and find hyperparameters

Data Augmentation Binary Search



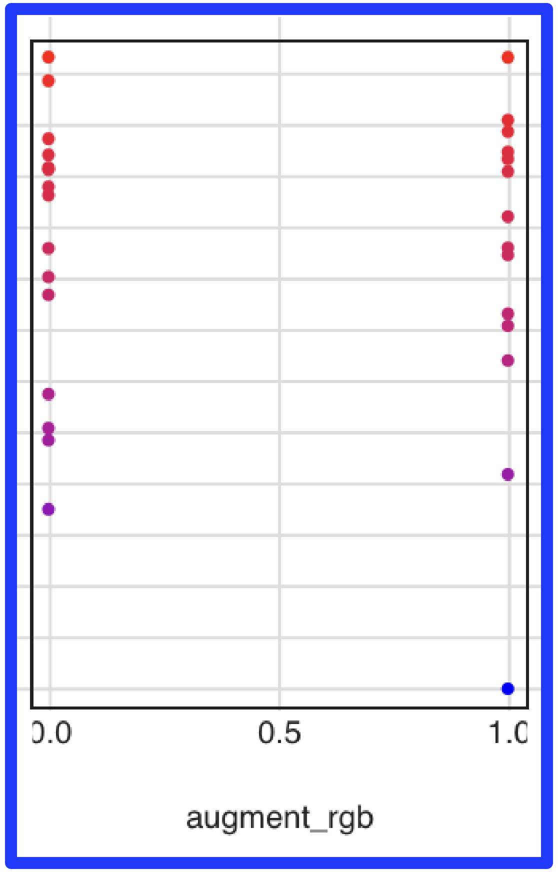
Understanding which augmentations improved the model

Data Augmentation Binary Search



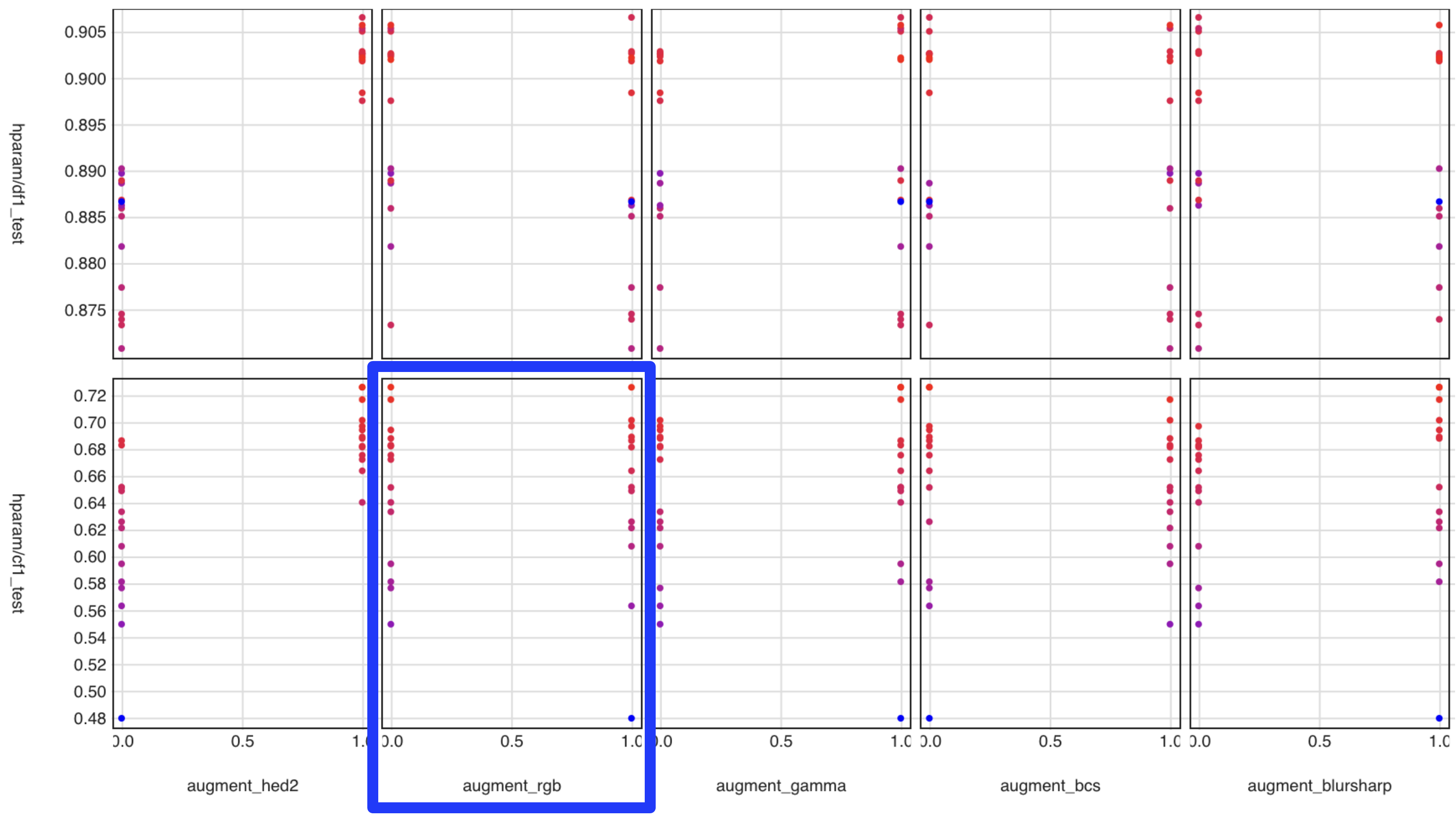
Understanding which augmentations improved the model

Data Augmentation Binary Search



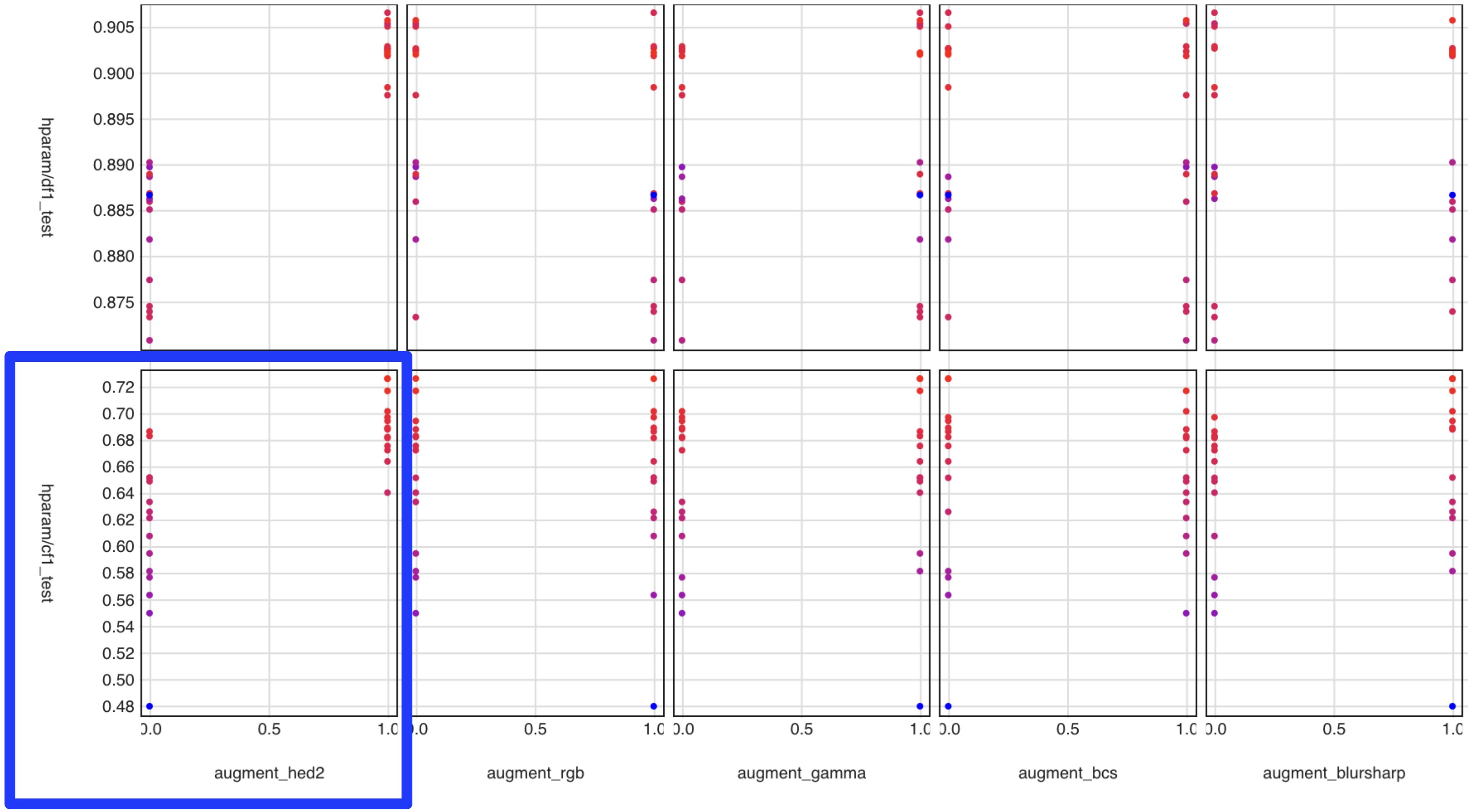
Understanding which augmentations improved the model

Data Augmentation Binary Search



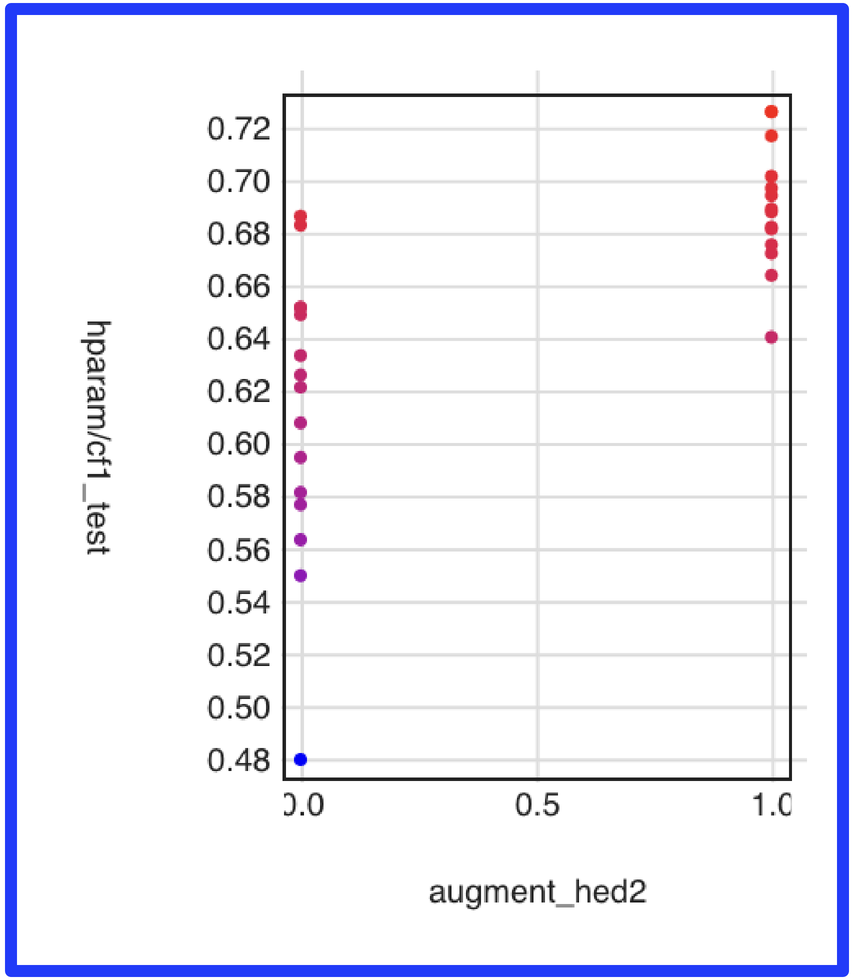
Understanding which augmentations improved the model

Data Augmentation Binary Search



Understanding which augmentations improved the model

Data Augmentation Binary Search



Understanding which augmentations improved the model

1 / 5 – Augmentation with changes in histological stains

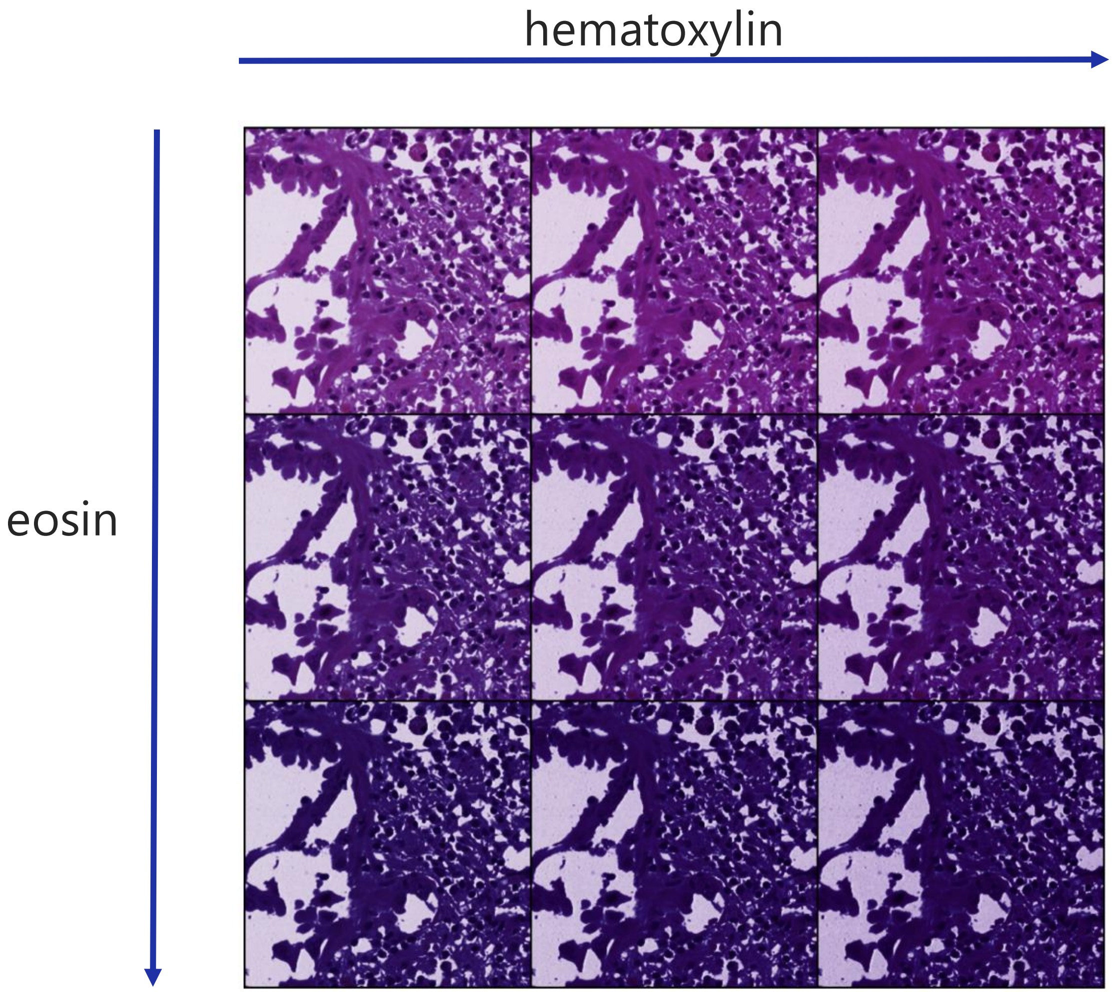
- Context: Cell slides are obtained by using a combination of two histological stains: hematoxylin and eosin.

- Idea: Make the model more robust against different amounts of histological stains, which may vary among datasets. The intuition is to replicate the change in light absorption.

- Technical: Changing the quantity of histological stains in slides ⇔ Change of coordinates in the vector space directed by hematoxylin and eosin.

Augmentation by varying the amount of "ink" applied to slides

1 / 5 – Augmentation with changes in histological stains



Increase of F1 accuracy.
Classification: 0.62 → 0.71
Detection: 0.88 → 0.90

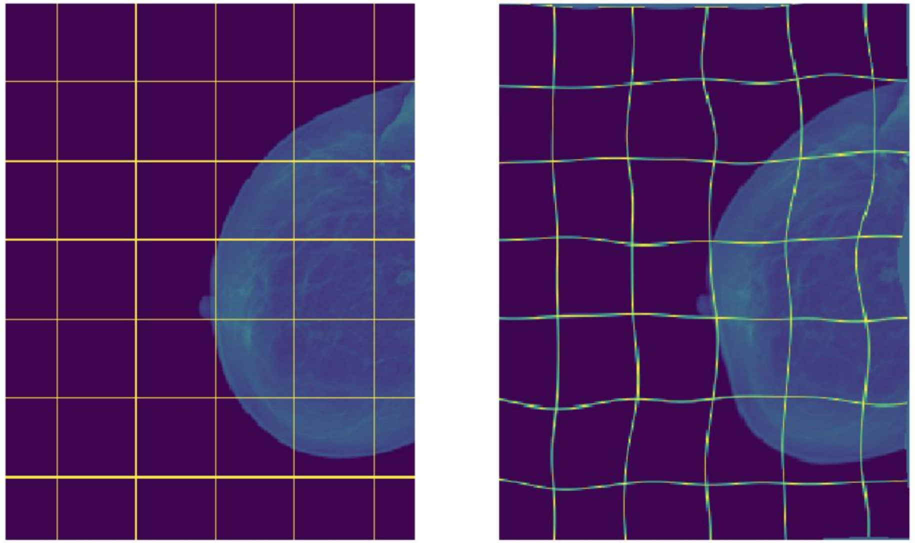
Augmentation by varying the amount of "ink" applied to slides

2 / 5 – Augmentation with elastic transformations

- Context: Elastic transformations have been shown to be an useful augmentation techniques for medical segmentation tasks. [1]

- Idea: Variation in shape and angles of the cells may need to be accounted for during training.

- Results: Overall the model performs worse with such transformations. An explanation is that cells are not affected by such distortions.

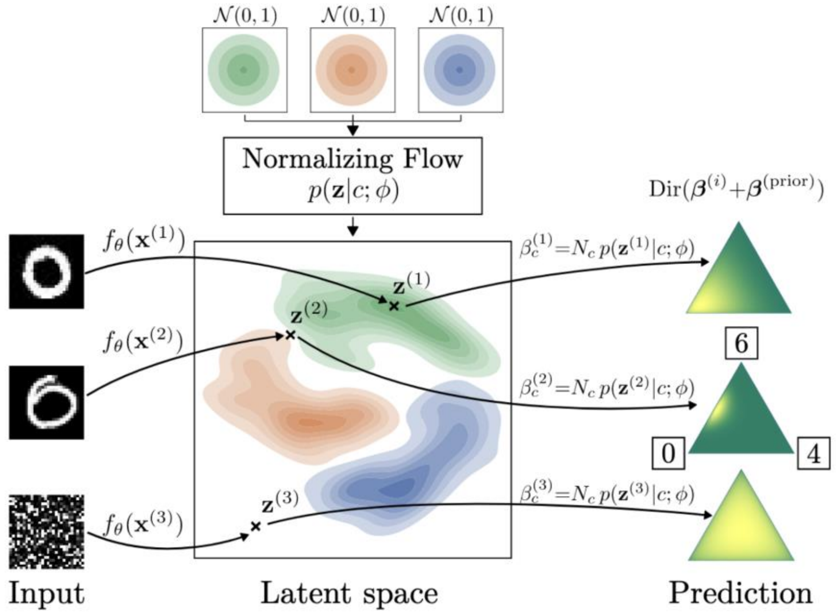


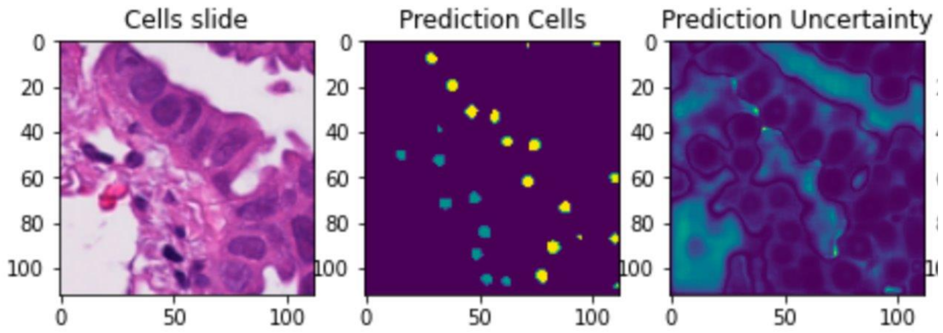
(a) Original (b) Deformed

Augmentation by "stretching" the slides randomly

3 / 5 – Posterior Network for detecting out of distribution samples

- Context: The model performs worse on samples or datasets not seen a train time.
- Idea: Use the Posterior Network model [2] for uncertainty estimation and detecting out of distribution samples, that may require further tunning.
- Results: Model correctly predicts higher uncertainty for out of training examples.
- Future work: Think of a way to exclude samples that have high uncertainty.





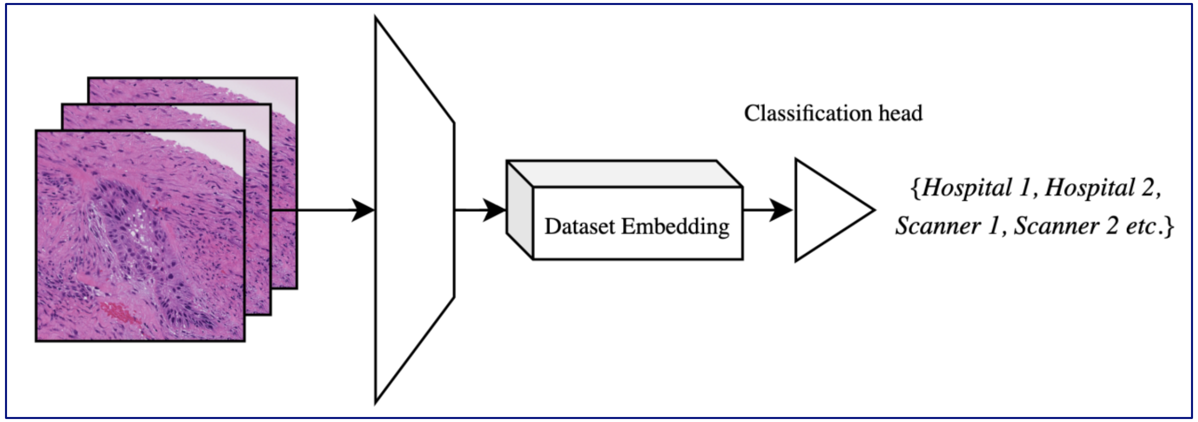
Detecting samples that look "unusual"

4 / 5 – Build a dataset representation encoder

- Context: Datasets have different characteristics due to external factors and machine calibrations.

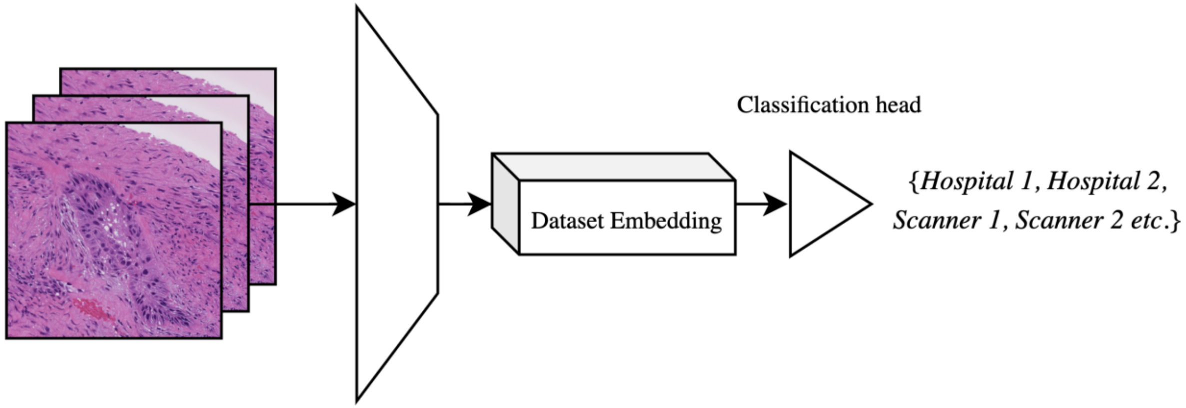
- Idea: Use those variations to our advantage by training a representation learner with a classification task.

- Technical: The representation encoder is obtained by training a ConvNet or LeNet on a classification task, then stripping out the classification layer.



Building a representation of the dataset features

4 / 5 – Build a dataset representation encoder

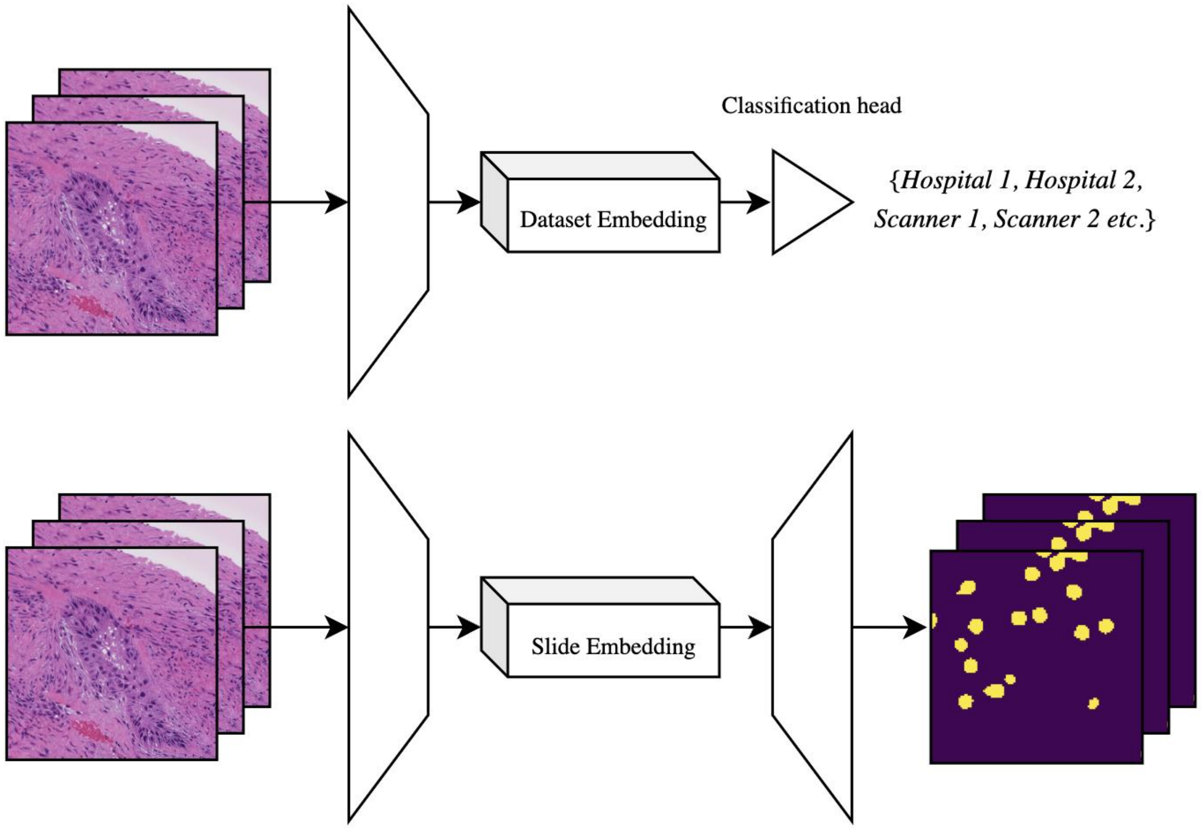


Classifier F1 accuracy of 0.92 on test

Building a representation of the dataset features

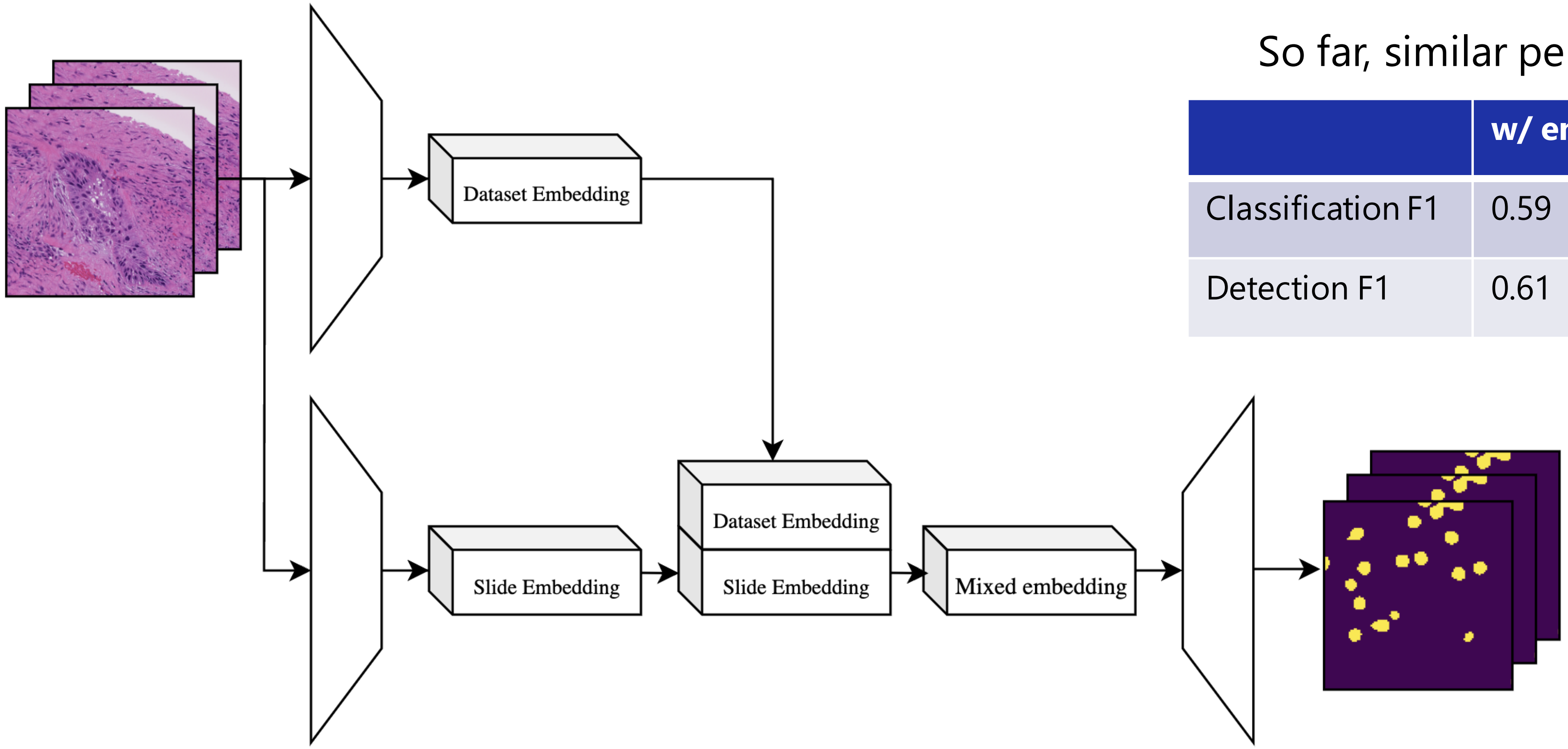
\Orchestrating a brighter world

4 / 5 – Build a dataset representation encoder



Building a representation of the dataset features

4 / 5 – Build a dataset representation encoder



So far, similar performance on test set

	w/ encoder	w/o encoder
Classification F1	0.59	0.71
Detection F1	0.61	0.69

Building a representation of the dataset features

5 / 5 – Conclusions and future research

- Medical ML rises new challenges in term of robustness and data generalization.

- Domain expertise brings critical knowledge from their field, see stain augmentation.

- Building dataset representation through self-supervised or weakly-supervised learning may be a solution.

- Recent works on image-to-image translation seems to go in that direction.

Model generalization is hard!







Thank you!




Q&A

\Orchestrating a brighter world

NEC

NEC Laboratories **America**