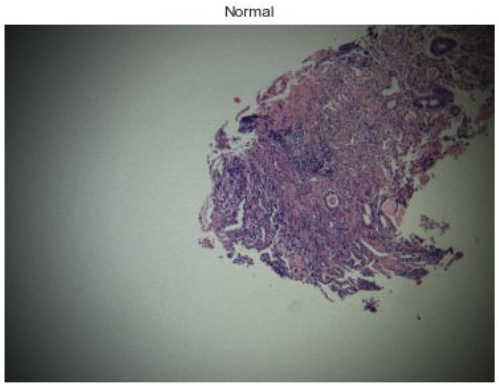
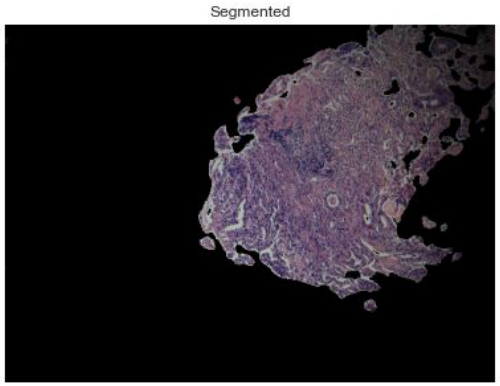
Microscopic Images Binary Classification

🔬 *SeeGene Project Report* 🔬

Willmer R. Quinones R.
2021.03.05

> Background is Enough?



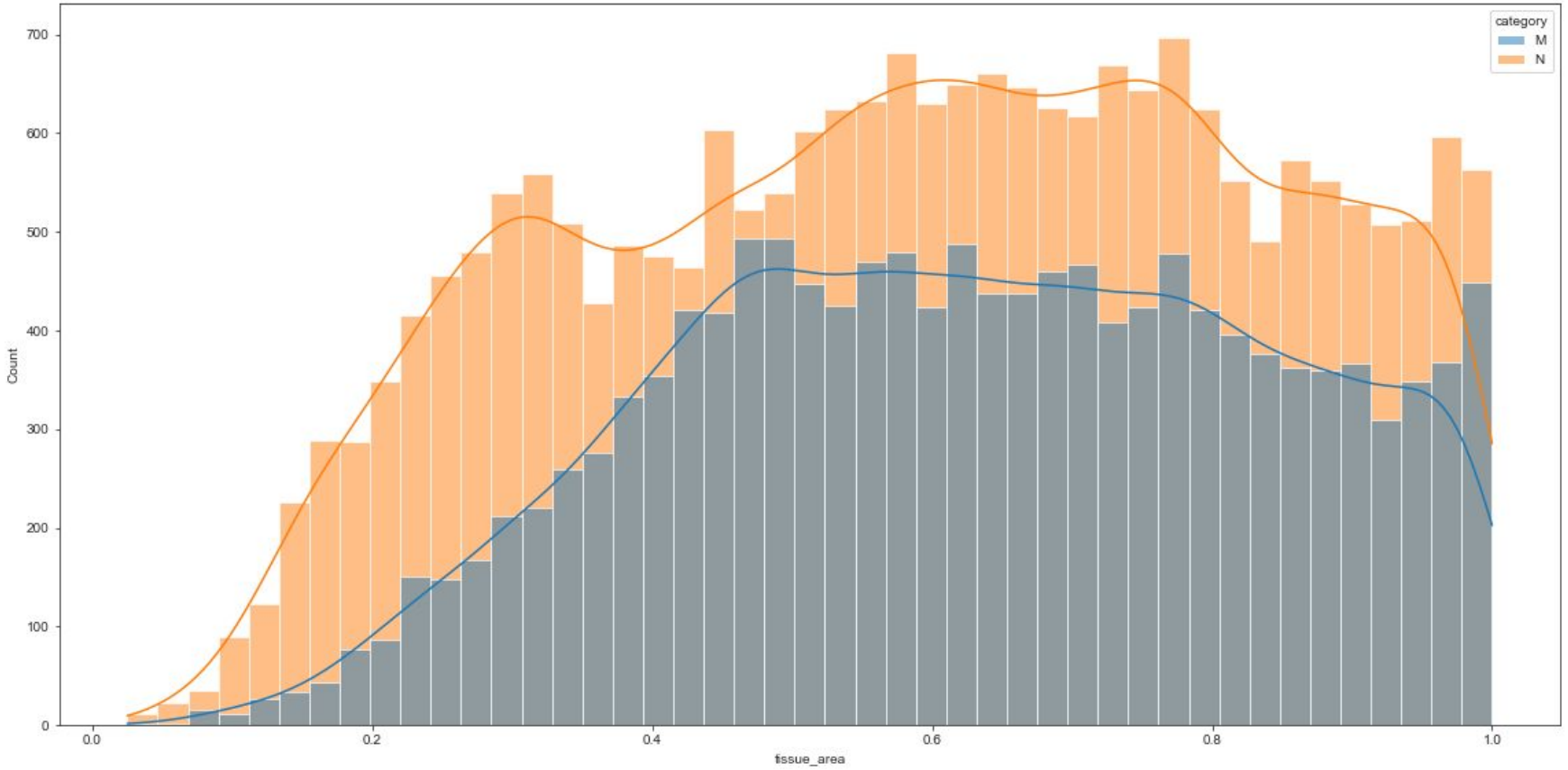


Accuracy	**0.971 ± 0.015**	0.888 ± 0.011	**0.969 ± 0.022**
TPR	**0.948 ± 0.051**	0.854 ± 0.020	**0.943 ± 0.063**
TNR	**0.981 ± 0.001**	0.900 ± 0.009	**0.979 ± 0.007**

A model is able to accurately distinguish between malignant and benign without looking at the tissue

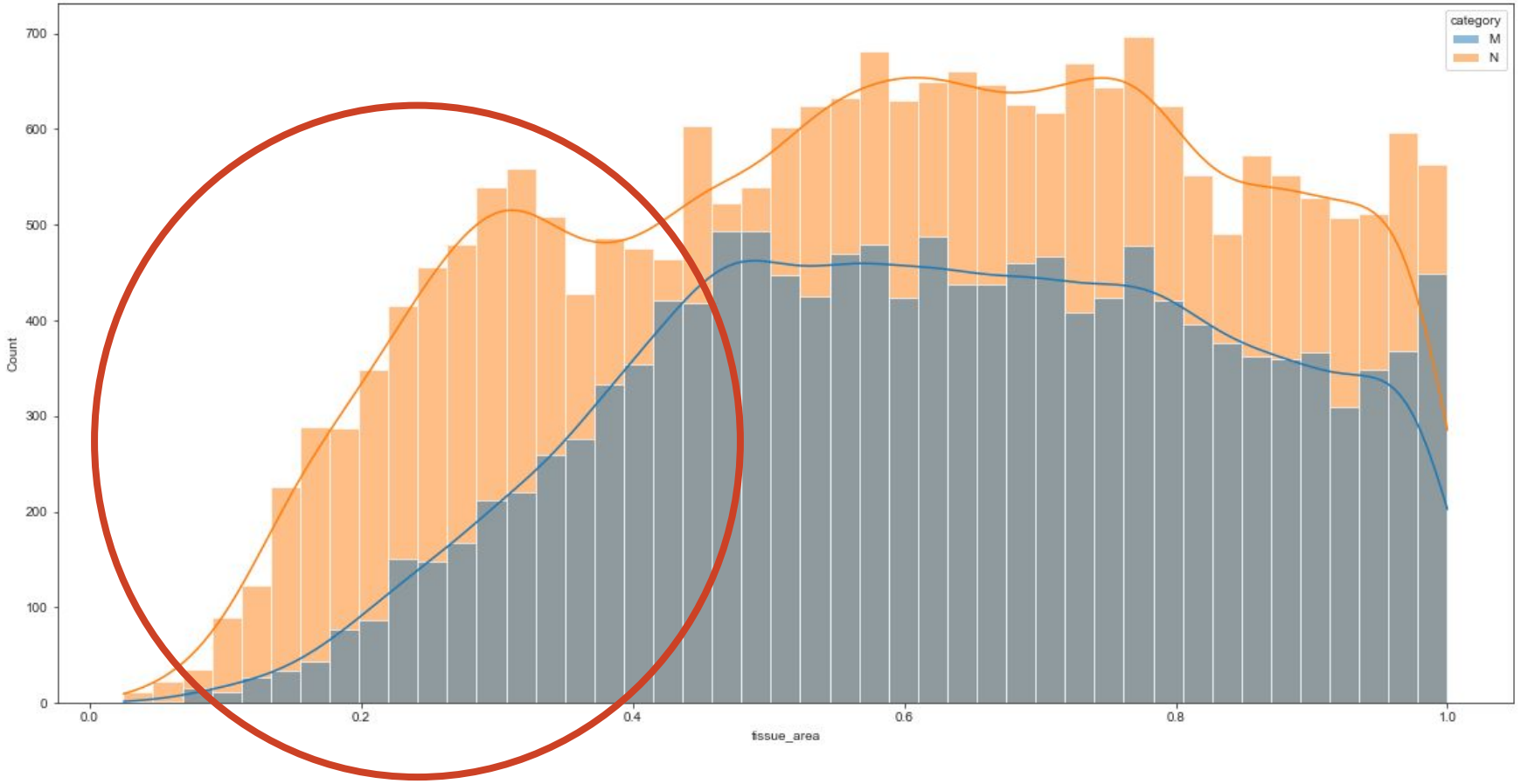
> Image Analysis

Tissue Area



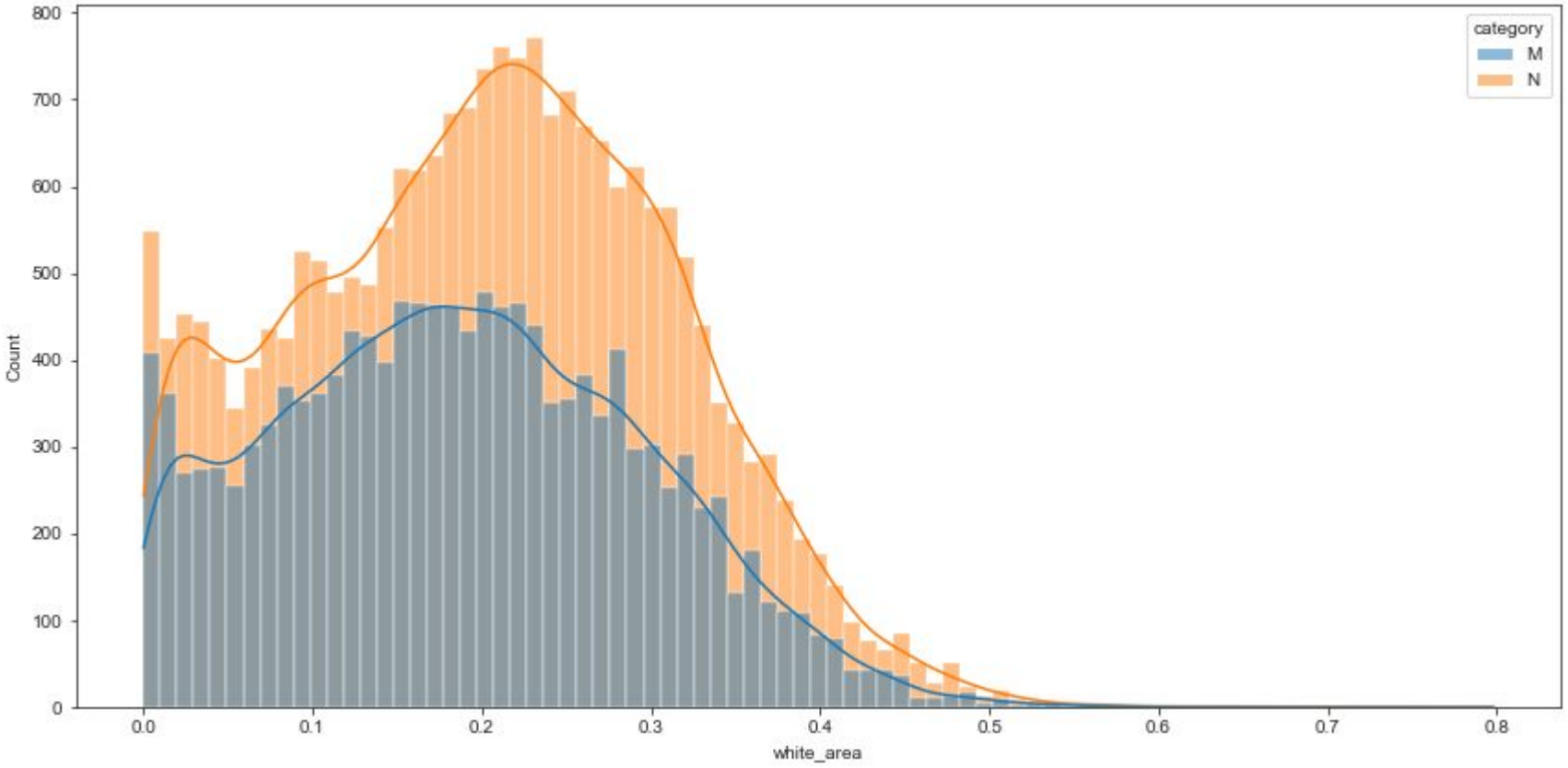
> Image Analysis

Tissue Area



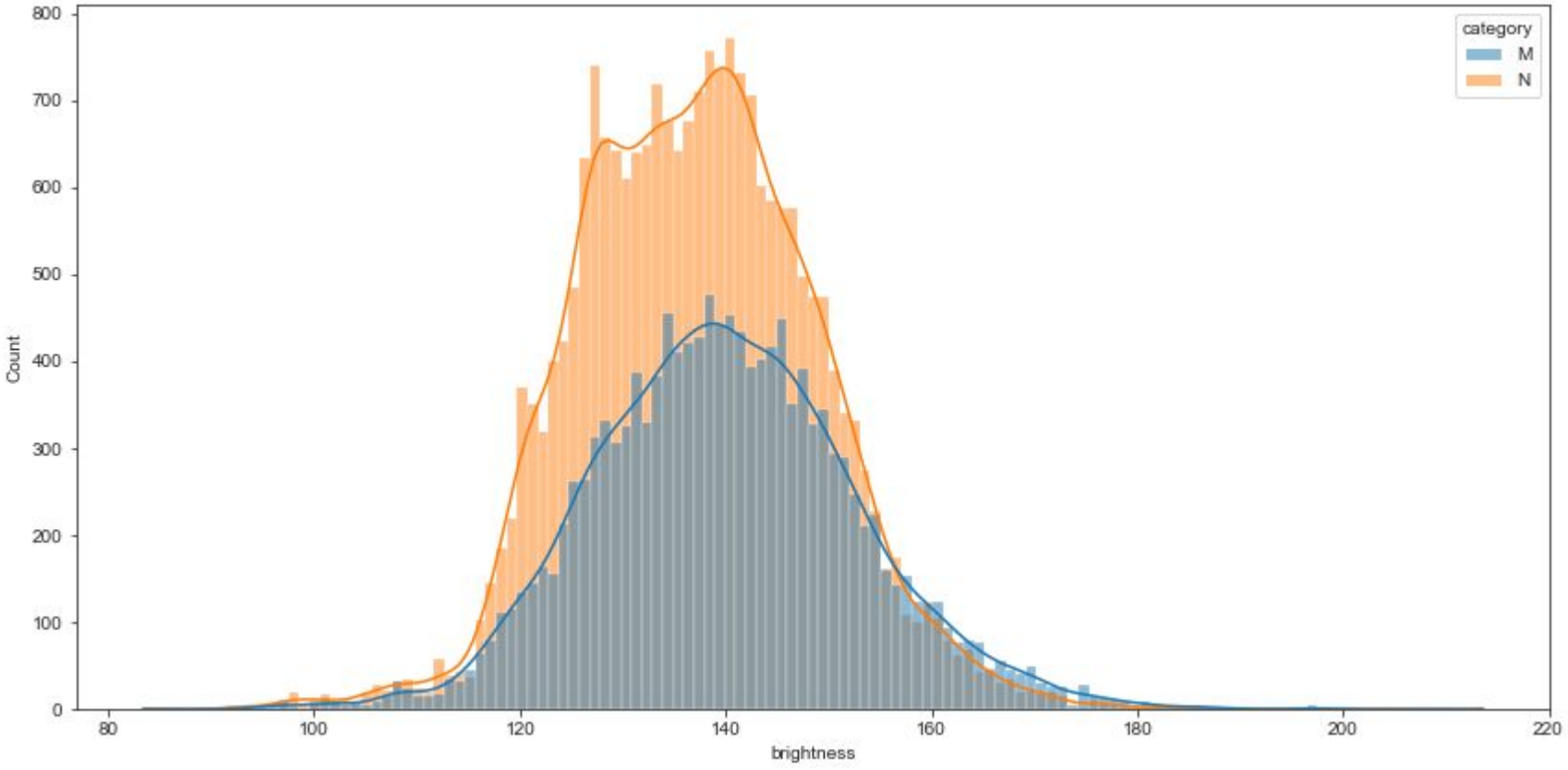
> Image Analysis

White-ish



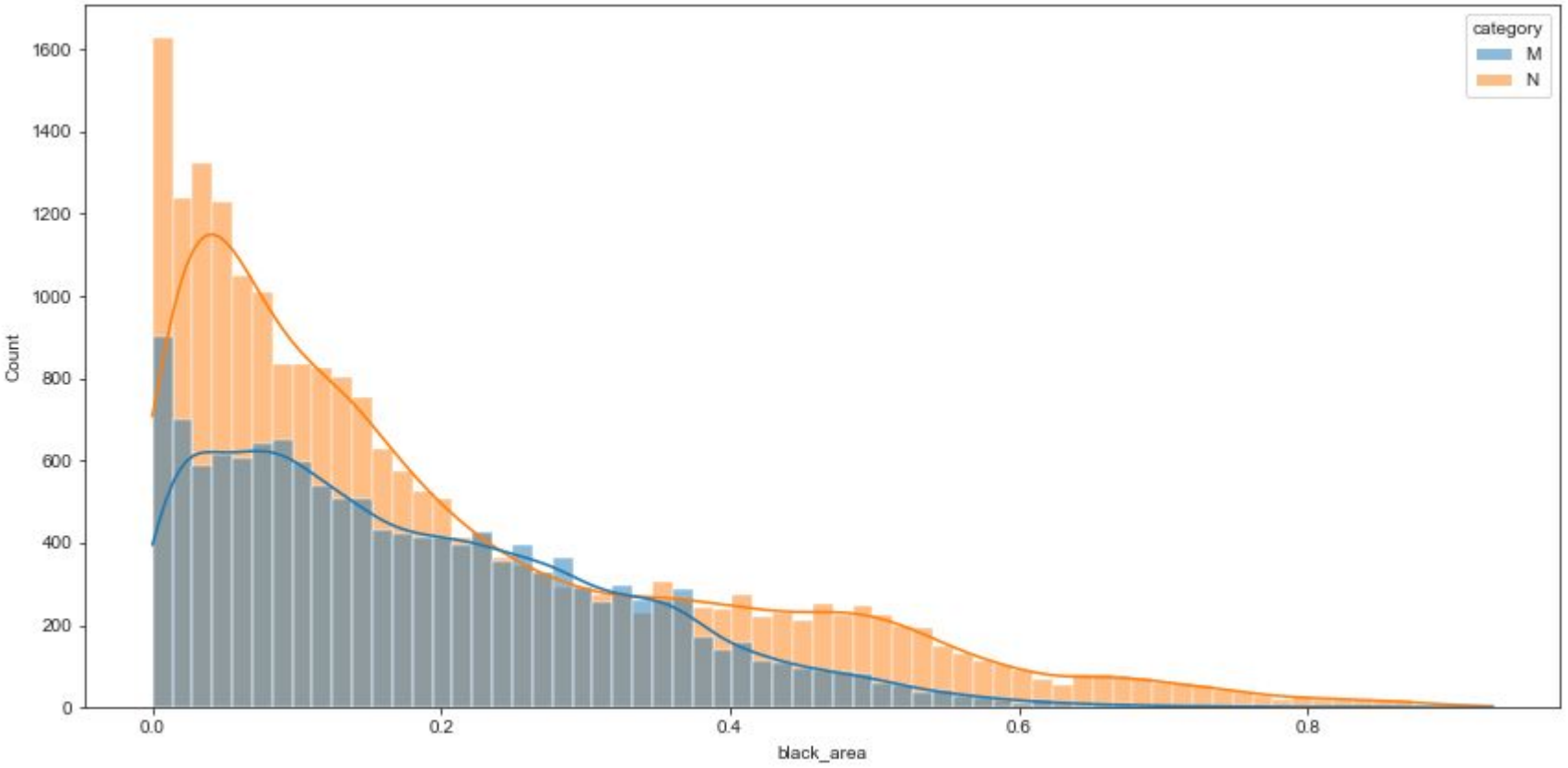
> Image Analysis

Brightness

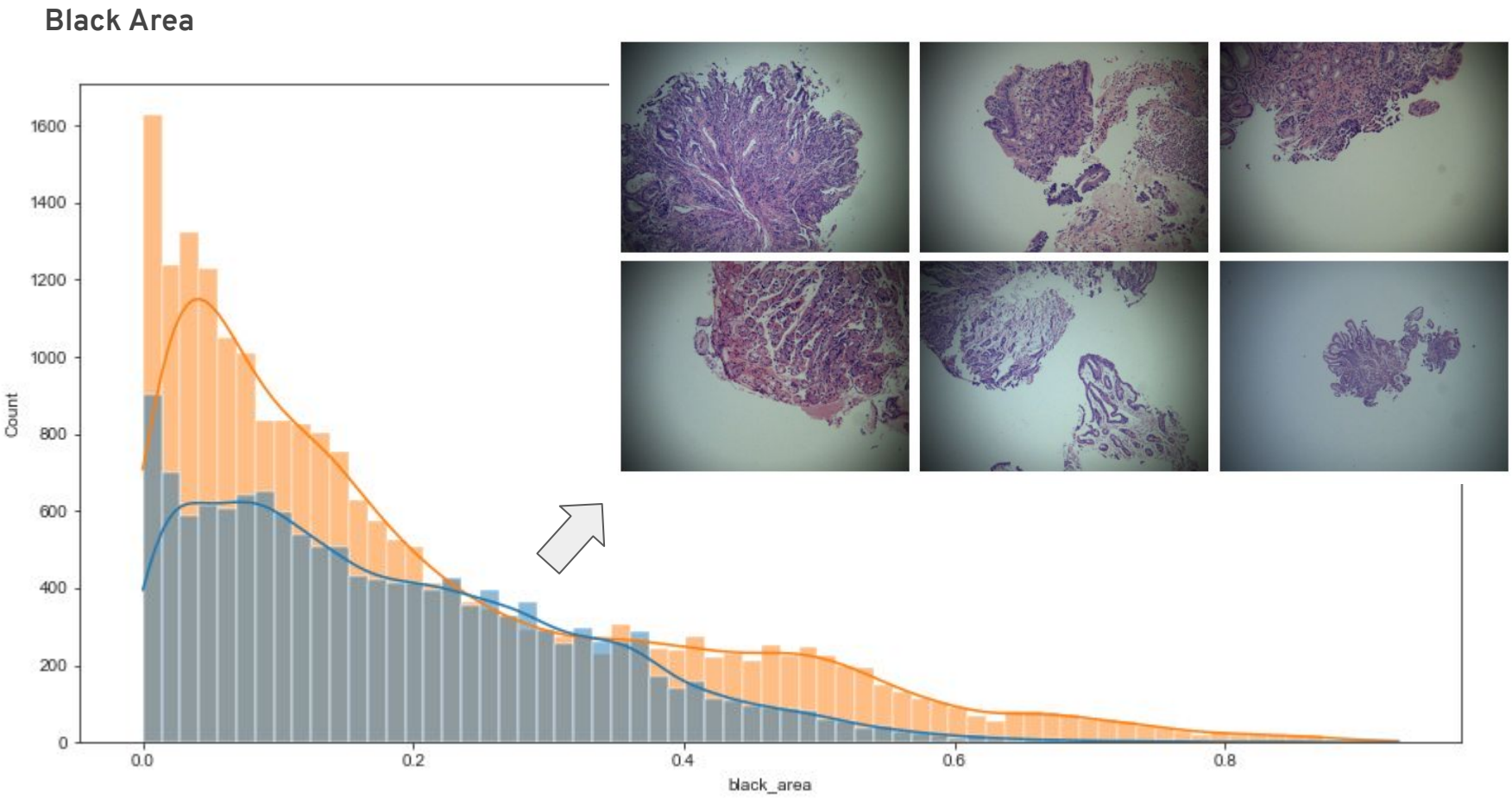


> Image Analysis

Black Area

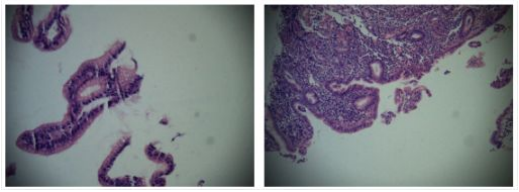


> Image Analysis



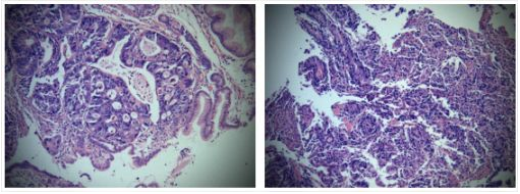
> Image Analysis

- N images tend to have less tissue than M images
- N images tend to be less bright than M images
- N dataset contains more ‘zoom-out’ images than the M dataset
- The model seems to pick spurious features from the background



N

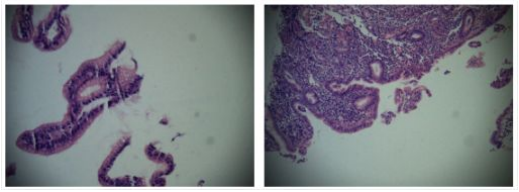
vs.



M

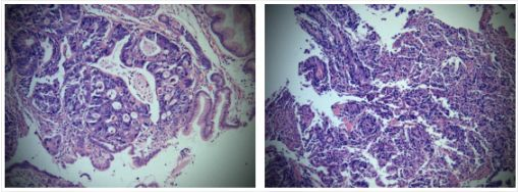
> Image Analysis

- ~~N images tend to have less tissue than M images~~ ⇒ No really
- ~~N images tend to be less bright than M images~~ ⇒ No really
- N dataset contains more ‘zoom-out’ images than the M dataset ⇒ Maybe
- The model seems to pick spurious features from the background ⇒ Yes



N

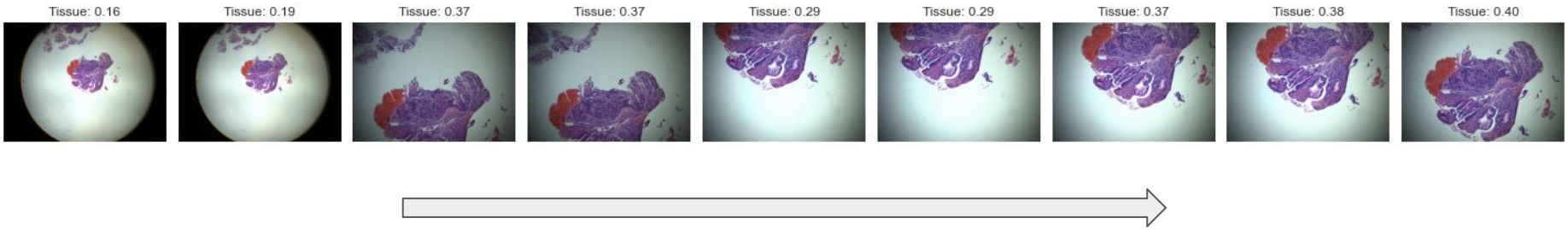
vs.



M

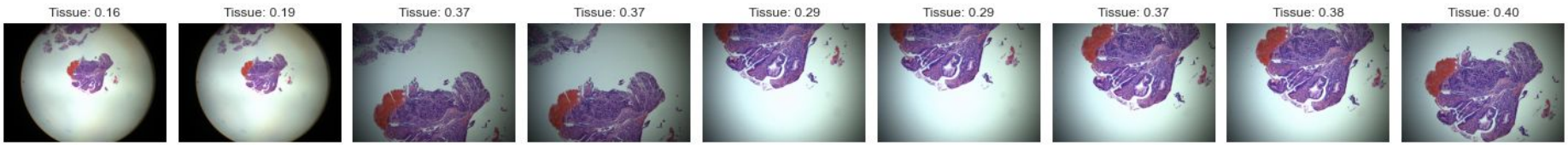
> Pathologist's footprint

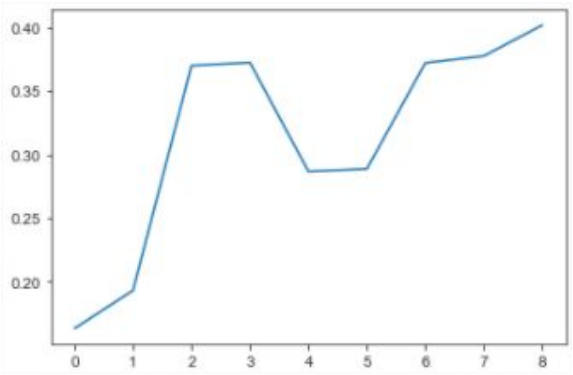
- Pathologists could transfer their **inherent bias** to the dataset through the way they examine the samples.
    - Do pathologists move differently on Normal samples and Malignant samples?



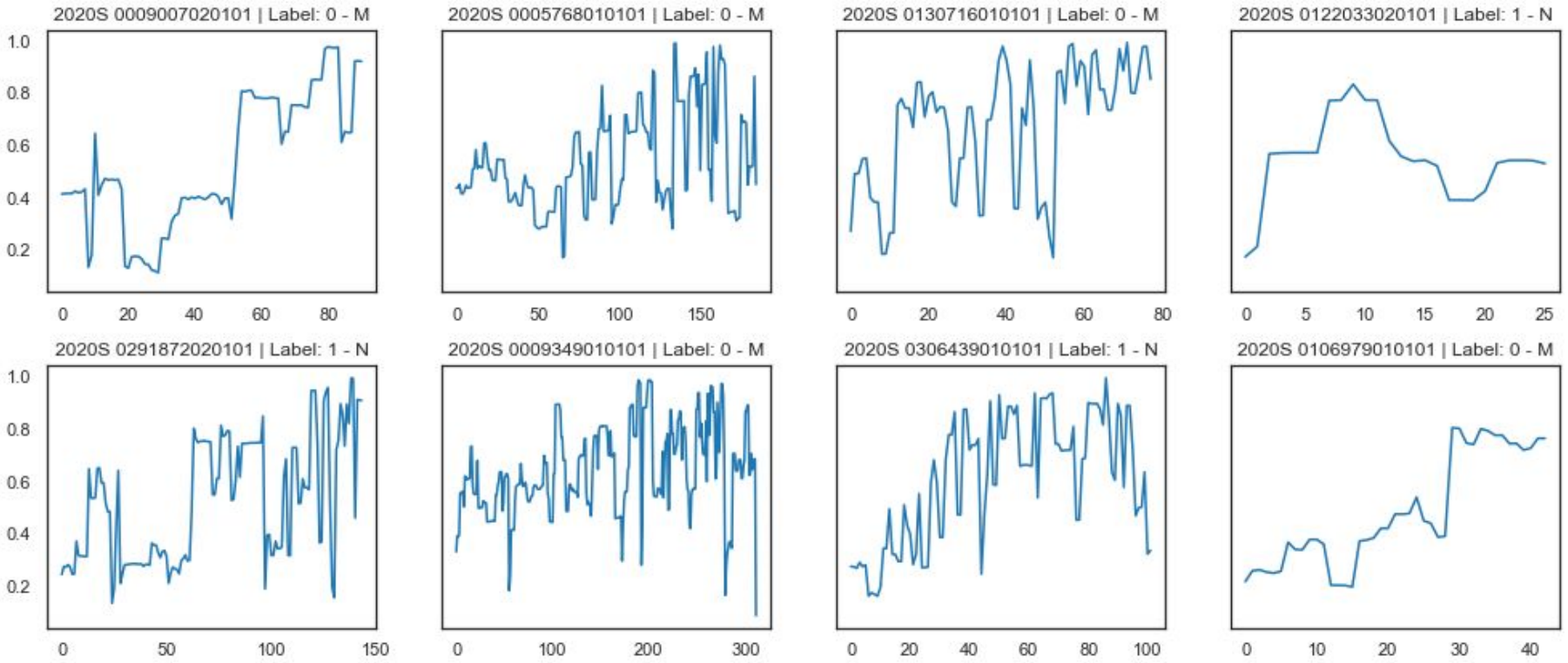
> Pathologist's footprint

- Pathologists could transfer their **inherent bias** to the dataset through the way they examine the samples.
    - Do pathologists move differently on Normal samples and Malignant samples?

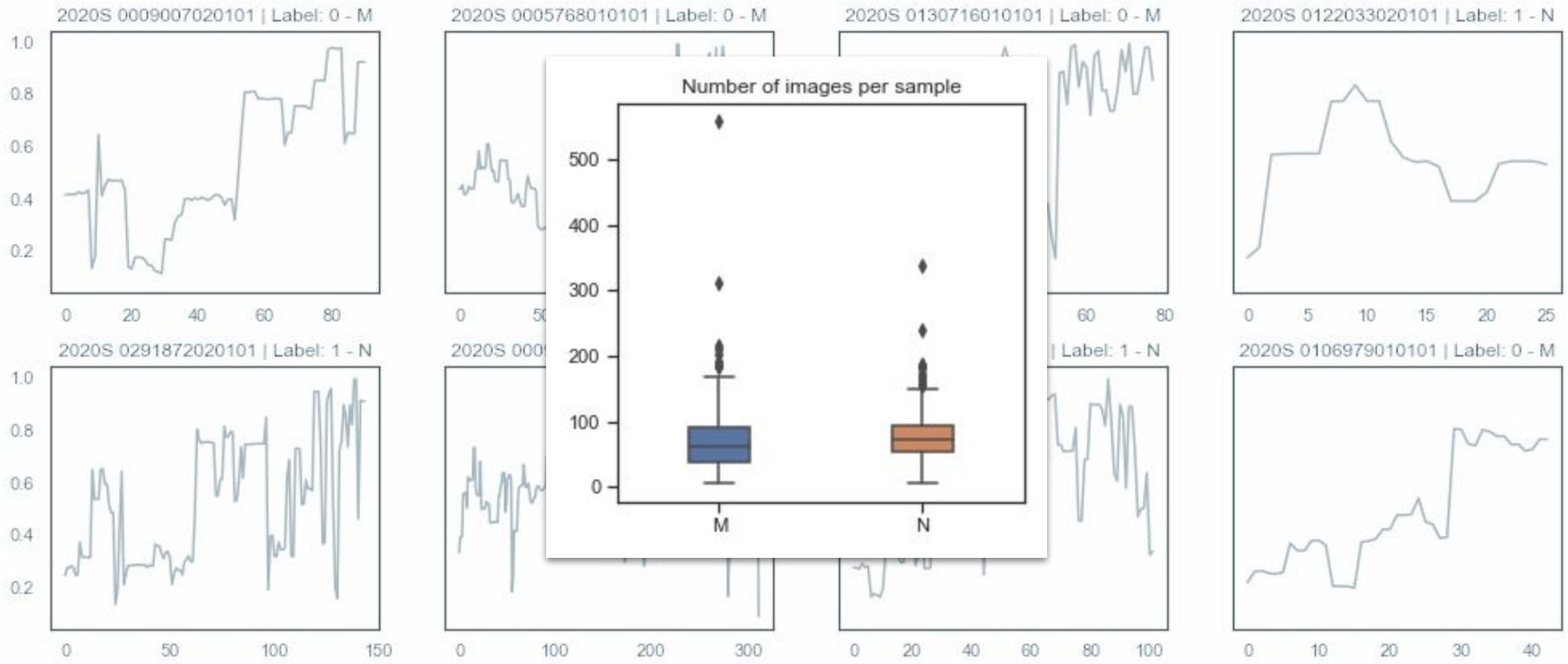




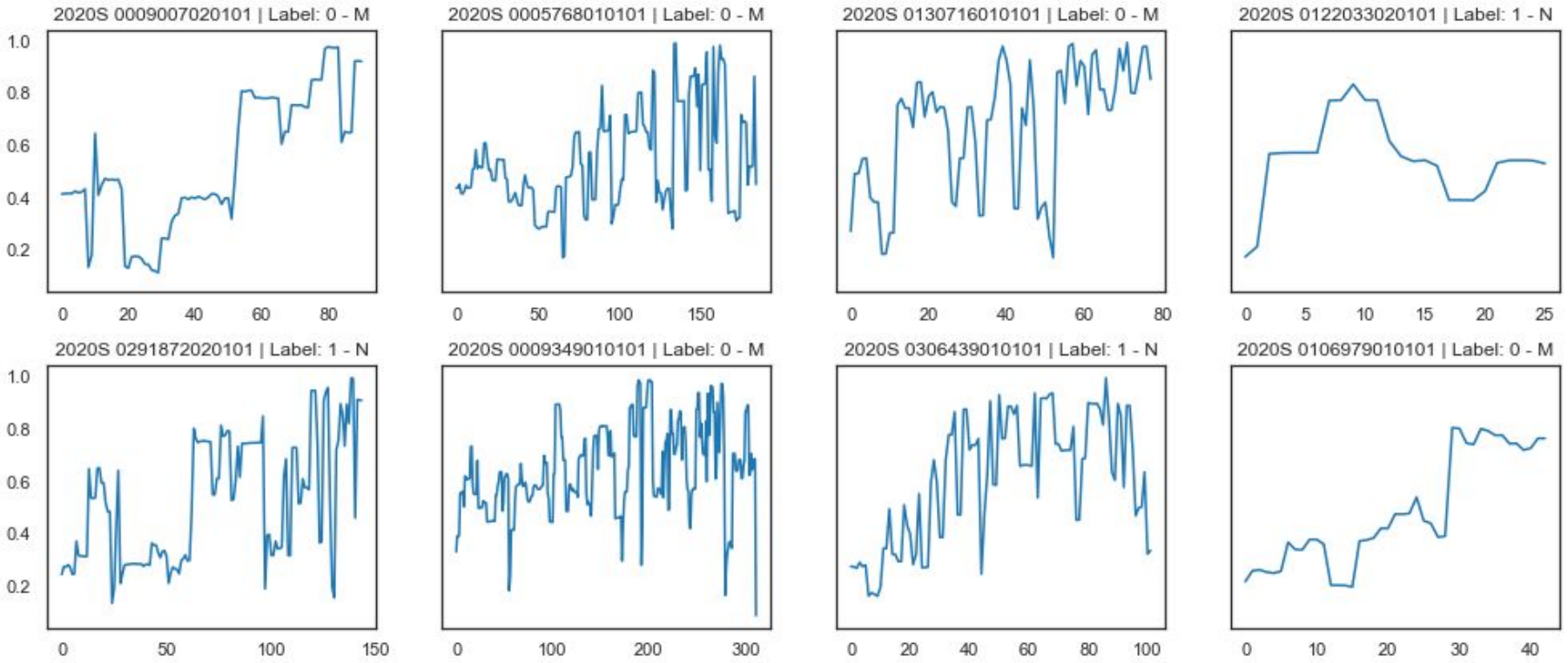
> Pathologist’s footprint



> Pathologist's footprint

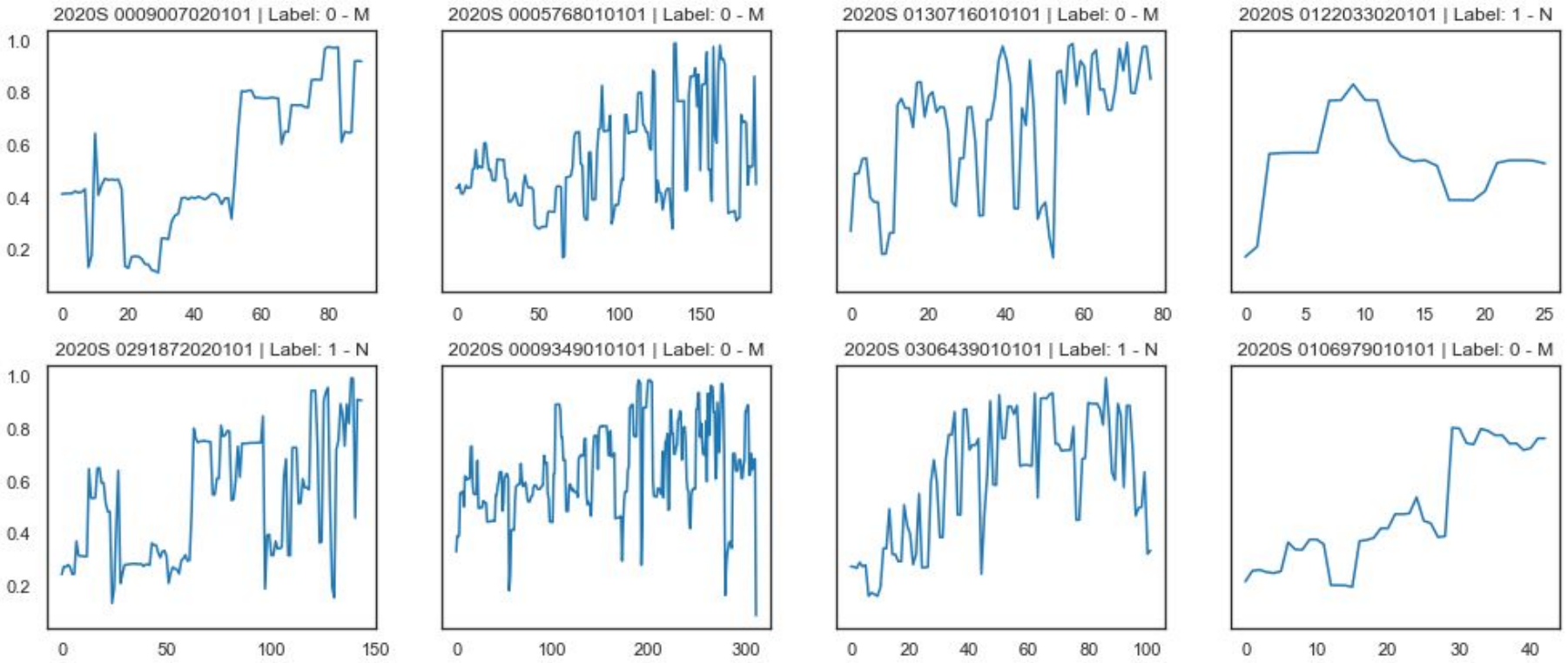


> Pathologist's footprint



GOAL: Find differences between N and M using this data

> Pathologist's footprint



Is there any difference?

> Pathologist's footprint

Is there any difference?

- Vanilla RNN, lr = 0.01, epochs = 50

	Pred M	Pred N
True M	28	30
True N	61	85

- LSTM, lr = 0.01, epochs = 50

	Pred M	Pred N
True M	58	0
True N	146	0

> Pathologist’s footprint

More data?

Another approach?

~ THANK YOU ~