



# HuBMAP + HPA – Hacking the Human Body

Progress Meeting 7 Group A

# CONTENTS

01

Stain Normalization

02

Practice

03

Further improvements

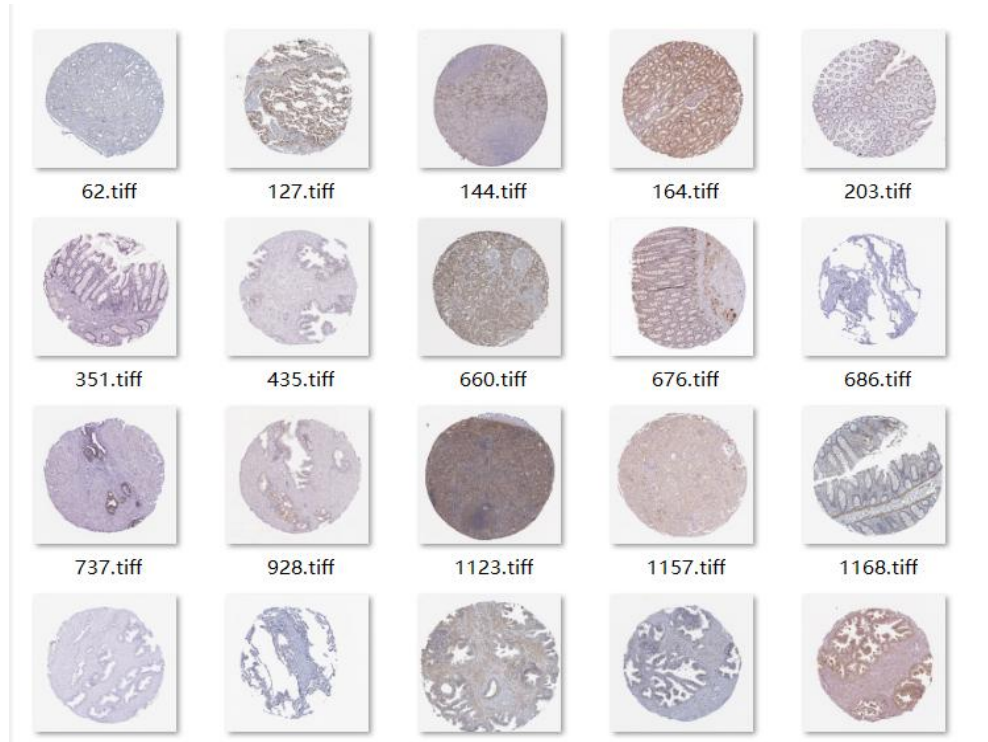
## ■ Stains

In histopathology, stains are used to increase the contrast between different structures for their manual examination. These two chemicals typically stain: the nuclei a dark purple (Hematoxylin) and the cytoplasm a light pink (Eosin). Thus all pixels in an idealized histology image are principally composed of two colors.

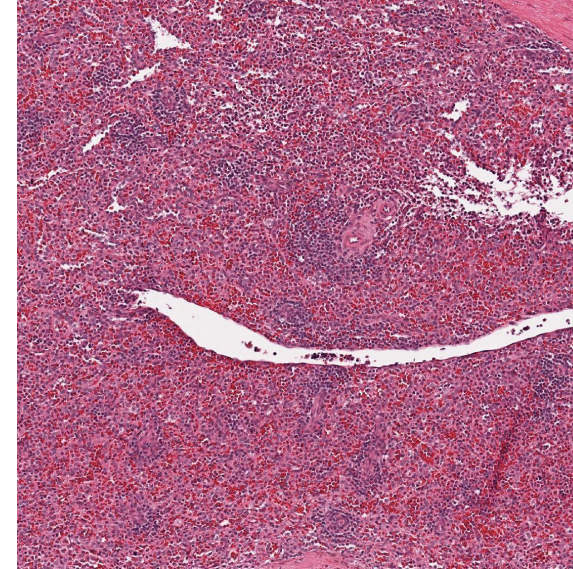
## ■ common problem

One common problem in analysis of tissue samples is undesirable variation in color due to differences in color responses of slide scanners (as shown in Fig. 1), raw materials and manufacturing techniques of stain vendors, and staining protocols across different pathology labs.

This competition uses data from two different consortia, the **Human Protein Atlas (HPA)** and **Human BioMolecular Atlas Program (HuBMAP)**.



Training-bluish-gray



Testing-red or purple

## A METHOD FOR NORMALIZING HISTOLOGY SLIDES FOR QUANTITATIVE ANALYSIS

*Marc Macenko<sup>1</sup>, Marc Niethammer<sup>1,2</sup>, J. S. Marron<sup>3,4</sup>, David Borland<sup>5</sup>, John T. Woosley<sup>6</sup>,  
Xiaojun Guan<sup>5</sup>, Charles Schmitt<sup>5</sup>, and Nancy E. Thomas<sup>4,7</sup>*

Departments of <sup>1</sup>Computer Science, <sup>2</sup>Biomedical Research Imaging Center,  
<sup>3</sup>Statistics and Operations Research, <sup>4</sup>Lineberger Comprehensive Cancer Center,  
<sup>5</sup>Renaissance Computing Institute, <sup>6</sup>Pathology and Laboratory Medicine, <sup>7</sup>Dermatology  
University of North Carolina, Chapel Hill, NC

## Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images

*Abhishek Vahadane\*, Tingying Peng\*, Amit Sethi, Shadi Albarqouni, Lichao Wang, Maximilian Baust,  
Katja Steiger, Anna Melissa Schlitter, Irene Esposito, and Nassir Navab*



### ***Macenko method***

- 1. Convert RGB to OD

$$OD = -\log\left(\frac{I}{255}\right)$$

- 2. Two vertical color matrices are obtained by SVD decomposition

stain matrix S

concentration matrix C

- 3. Normalize the color matrix to Target (testing)

### ***Vahadane method***

- *SNMF*—*sparse non-negative matrix factorization*  
(function: separate the stains)

sparse: a given location(pixel) is likely to be occupied with just one material, not a mixture of two

non-negativity: make the optical density be zero or positive, optical density is meaningless if the value is negative

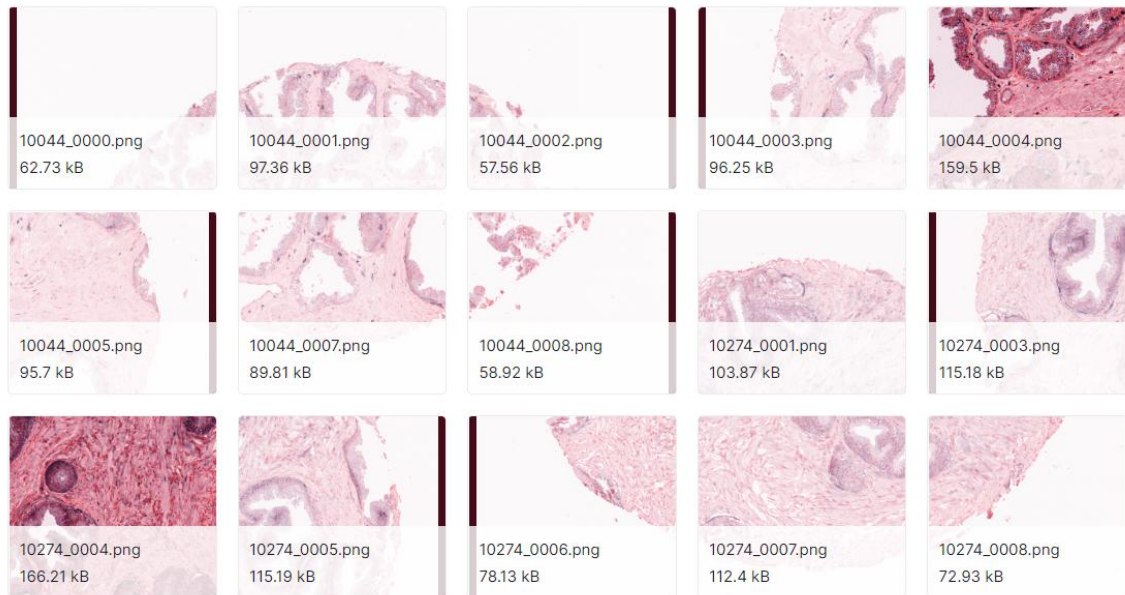
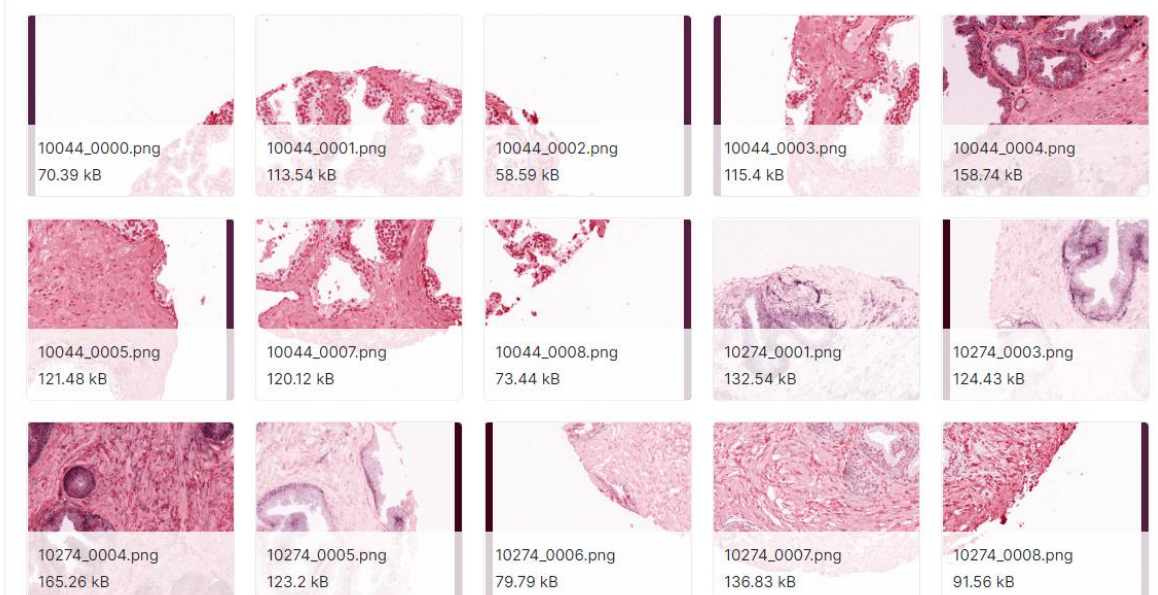
- *SPCN*—*structure-preserving color normalization*

changes color of one image (source) to match that of another (target) while reliably keeping source structural information intact

## StainTools

```
!pip install staintools  
!pip install spams
```

```
# Read data  
target = staintools.read_image("./data/my_target_image.png")  
to_transform = staintools.read_image("./data/my_image_to_transform.png")  
  
# Standardize brightness (This step is optional but can improve the tissue mask calculation)  
target = staintools.LuminosityStandardizer.standardize(target)  
to_transform = staintools.LuminosityStandardizer.standardize(to_transform)  
  
# Stain normalize  
normalizer = staintools.StainNormalizer(method='vahadane')  
normalizer.fit(target)  
transformed = normalizer.transform(to_transform)
```

***Macenko method******Vahadane method***



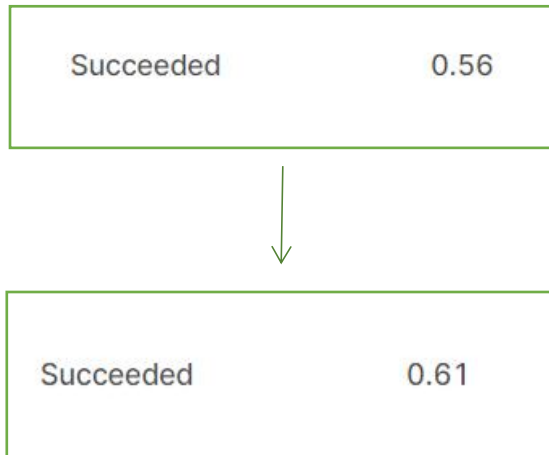
## *Macenko method*



## *Vahadane method*



## *Macenko method*



## *Vahadane method*



Only did colour normalization for the training sets

If we want to do colour normalization to testing sets, several problems...

1. Fail to pip install ***spams*** without Internet
2. We try to use ***torchstain*** instead of ***staintools***, it has good effect as well, and it can be installed without Internet, while some errors happen during scoring

Succeeded 0.56



Succeeded 0.62

Notebook Threw  
Exception Error ⓘ

1. The cause of Vahadane's poor performance is being investigated: Is Vahadane a bad fit for this dataset or is it a problem with the dataset handling.
2. Try to resolve the error message
3. GAN-stain normalization for testing image
4. Find the final effect of combining the previous improvements with the normalization method that is known to be useful

**Thanks for listening**