Assignment E: Quantification of histochemical staining by color deconvolution

Course : Image Processing and Quantitative Data Analysis
Instructor : Marten Postma
Teaching assistants : Aaron Lin, Aoming Sun, Catherine Chia
Date: 21st June, 2023

Table of content

1.0 Introduction

In histology where biological tissues are microscopically studied, common histological stains such as the H&E stain (hematoxylin and eosin) and DAB (3,3'-Diaminobenzidine) and hematoxylin counterstain are routinely used under brightfield microscopy:

H&E

- hematoxylin - purple, which stains cell nuclei
- eosin - pink, stains cytoplasm and extracellular matrix

DAB and hematoxylin counterstain

- DAB - brown, stains nuclei that positively reacted to the target receptors (e.g. cancerous)
- hematoxylin counterstain - blue, stains portions that negatively reacted to the target receptors

These stainings should be analyzed individually without the presence of the partner stain. As the colors (purple, pink, or brown) in these multi-stains samples span across the red, green and blue (RGB) channels, color deconvolution is deployed to separate the stains into individual channels or images, each containing one of the stains (see Figure 2).

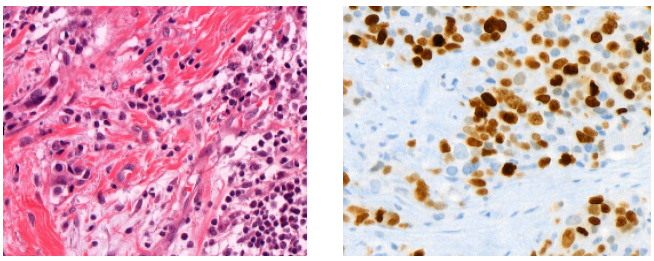

Figure 1 Histology images used in this assignment. (Left) H_E.tif - H&E stains showing nuclei (purple), cytoplasm and extracellular matrix (pink). **(Right) IHC2.tif** - DAB and hematoxylin counterstains show positive nuclei (brown) and negatively reacted components (blue)

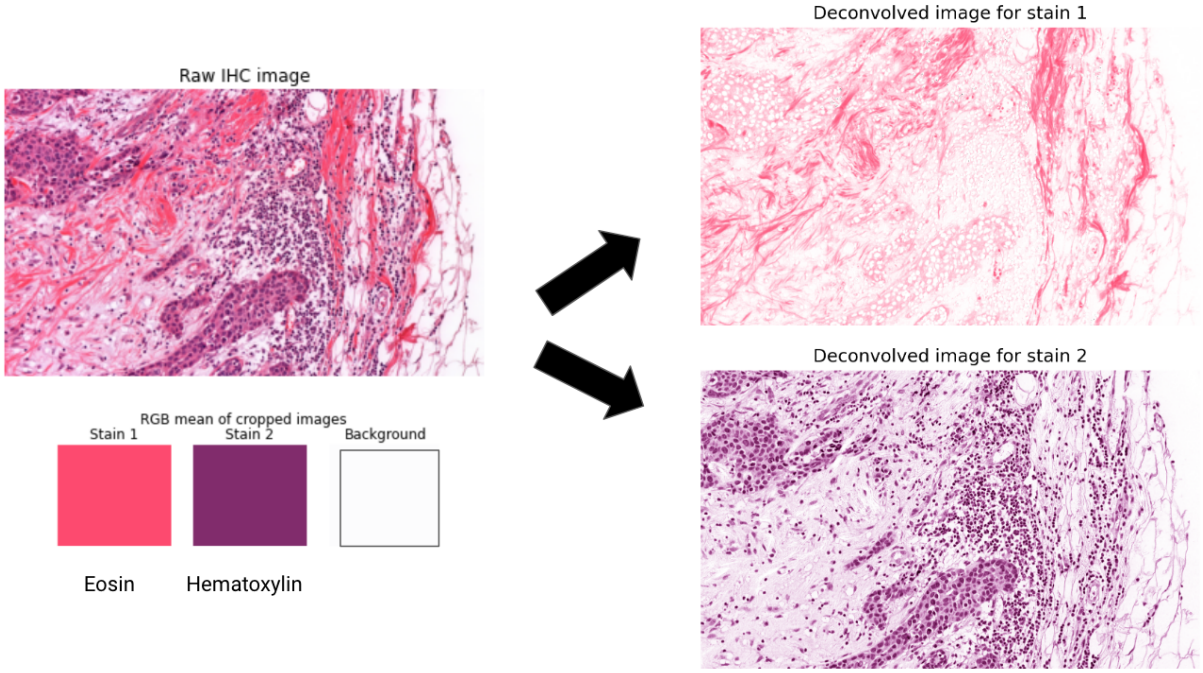

Figure 2 Color deconvole an IHC image resulting in two images, one for each stain. (Left) Input of the color deconvolution method. (Right) Output of the method.

The goal of this assignment is to reproduce a color deconvolution method from a frequently cited paper about color deconvolution (Ruifrok and Johnston, 2001), and apply the method to two provided brightfield histology images. We provide the necessary guidance in the following sections.

1.1 Guidance

1. We strongly recommend going through both the lecture for this topic and the paper (Ruifrok and Johnston, 2001) as they are complementary to each other. As there are differences in the symbols used in each of these materials, Table 1 is provided to aid in navigating the materials and Figure 3 illustrates the concept of transmitted and received light.

2. The color deconvolution method should separate each pixel (containing red, green, and blue (RGB) values) to two pixels, one for each selected stain (e.g. Eosin and Hematoxylin). The method performs on the absorbance values, hence the transmittance values (I) must be converted to absorbance (OD), and back to transmittance after applying the color deconvolution method.

3. Model each pixel in the raw measured image as a linear combination of two ODs, and solve for the coefficients using linear regression

$$\mathbf{OD} = \mathbf{OD}_1 + \mathbf{OD}_2 = c\mathbf{oeff}_1\mathbf{s'}_1 + c\mathbf{oeff}_2\mathbf{s'}_2$$

   where,
   OD = absorbance of the raw image
   OD_N = absorbance contribution of the used histochemical stains (here we have two stains for each raw image)
   $coeff_N$ = coefficient for each stain
   s'_N = normalized absorbance RGB vector of each stain

4. Solve for the coefficients using linear regression (least square) and matrix multiplication (x)

$$\mathbf{M}^T \times \mathbf{M} \times \mathbf{Coeff} = \mathbf{M}^T \times \mathbf{OD}$$

   where,
   M = OD matrix
   M^T = transposed M
   Coeff = coefficient matrix
   OD = absorbance of the raw image

1.2 General instruction

This assignment is Python-focused, but you are encouraged to use ImageJ to extract the RGB vectors over the stain and background areas (either exported in cropped images and extract the RGB in Python, or extract the RGB values (there are multiple ways. For e.g. using ROI, Image > Color > Make Composite, Analyze > Measure) and import these values via CSV to Python).

Design a color deconvolution workflow (draw a flowchart) and make sure you understand the key components in the workflow. You can learn about the workflow from the provided starting script. Then, complete the provided starting script at spaces between "*#Start coding here*" and "*#End coding here*".

Table 1 Comparison of symbols used in the lecture and the referred paper (Ruifrok and Johnston, 2001).

Definition	Lecture	(Ruifrok and Johnston, 2001)
Transmittance	T	(not available)
Absorbance (optical density)	A	OD
Transmitted light	ϕ_e^t	I_c
Received (background) light	ϕ_e^i	I_0
Beer-Lambert Law	$A = \varepsilon \ell c$	$OD_C = -\log_{10}(I_C/I_{0,C}) = A * c_C$
OD matrix	(not available)	M
Deconvolution matrix	(not available)	D

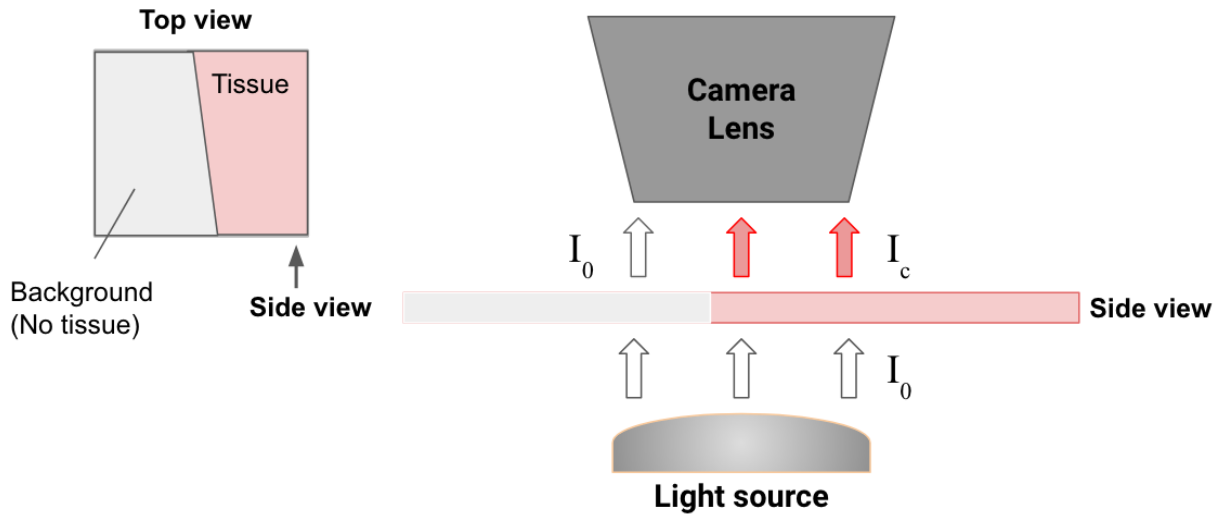

Figure 3 Transmitted and received light. (Left) Top view of a specimen that has an area covered with tissue (red) and a blank area known as background (grey). Arrow indicates side view direction. (Right) The light source emits light (I_0) through the specimen (side view). If the light passes through the blank area, the received light equals to I_0 (assuming no or negligible absorbance). Alternatively, if the light passes through the tissue, the transmitted light is labeled as I_c.

2.0 Tasks and questions (Total 10 points)

2.1 Image data and tracked data inspection (1 point)

- Question 1: Explain the choice of stains and background region of interest (ROI) in each IHC image. i.e. Which histological stains are used and what are the RGB values, and why are they selected at the specific locations?

2.2 Methodology (6 points)

- Question 2: Present a color deconvolution workflow using a flowchart and description about the key components.

- Question 3: Export the Jupyter Notebook as a PDF, one for each IHC image. Remember to include inline comments to demonstrate your understanding of the methods.

2.3 Data analysis and discussion (3 point)

- Question 4: Use the final 2 figures in the Jupyter Notebook for each IHC image, explain the observations on the color deconvolved images. i.e. cellular or tissue structure stained by the selected histological stains

- Question 5: Do you notice a difference in the output quality with IHC2.tiff compared to H_E.tiff, where DAB was used? Do some internet search to understand why applying color deconvolution on DAB signal can be more challenging.

2.4 Submit your work

1. One PDF that contains text-based answers, figures and/or tables. Don't forget about the references.
2. Two PDFs (one for each IHC image) generated from the Jupyter Notebook. Remember to place in-line comments
3. A zipped folder containing:
    a. T_img_ihc_stain1_norm.tif
    b. T_img_ihc_stain2_norm.tif

    One pair of images for each IHC image

3.0 Reference

Ruifrok, A C, and D A Johnston. Quantification of Histochemical Staining by Color Deconvolution. Analytical and Quantitative Cytology and Histology, 23 (4): 291–9, 2001.