

Retinex model based stain normalization technique for whole slide image analysis



Md. Ziaul Hoque ^{a,b,*}, Anja Keskinarkaus ^{a,b}, Pia Nyberg ^{c,d}, Tapio Seppänen ^{a,b}

^a Physiological Signal Analysis Group, Center for Machine Vision and Signal Analysis, University of Oulu, Finland

^b Faculty of Information Technology and Electrical Engineering, University of Oulu, Finland

^c Biobank Borealis of Northern Finland, Oulu University Hospital, Finland

^d Translational & Cancer Research Unit, Medical Research Center Oulu, Faculty of Medicine, University of Oulu, Finland

ARTICLE INFO

Keywords:

Medical image analysis
Computer aided diagnosis
Color normalization
Stain separation
Illumination estimation

ABSTRACT

Medical imaging provides the means for diagnosing many of the medical phenomena currently studied in clinical medicine and pathology. The variations of color and intensity in stained histological slides affect the quantitative analysis of the histopathological images. Moreover, stain normalization utilizing color for the classification of pixels into different stain components is challenging. The staining also suffers from variability, which complicates the automatization of tissue area segmentation with different staining and the analysis of whole slide images. We have developed a Retinex model based stain normalization technique in terms of area segmentation from stained tissue images to quantify the individual stain components of the histochemical stains for the ideal removal of variability. The performance was experimentally compared to reference methods and tested on organotypic carcinoma model based on myoma tissue and our method consistently has the smallest standard deviation, skewness value, and coefficient of variation in normalized median intensity measurements. Our method also achieved better quality performance in terms of Quaternion Structure Similarity Index Metric (QSSIM), Structural Similarity Index Metric (SSIM), and Pearson Correlation Coefficient (PCC) by improving robustness against variability and reproducibility. The proposed method could potentially be used in the development of novel research as well as diagnostic tools with the potential improvement of accuracy and consistency in computer aided diagnosis in biobank applications.

1. Introduction

Cancer is a major life-threatening health problem around the world and biomedical imaging is one of the crucial data modalities in cancer research. Radiology and Radiomics provide means of quantitative study of cancer properties at the macroscopic scale. The histopathologic examination of tissue, on the other hand, reveals the effects and properties of cancer at the microscopic level (Kurc et al., 2020). The new digital technologies can provide means for attaining digital images from biological tissue samples in a faster and more cost-effective way without compromising quality or patient care. A trend prevalent in many areas of healthcare over the past decade has also entered into the area of pathology. In digital pathology, the histology slides are digitized to produce whole slide images with microscopic scanners (Gurcan et al., 2009a). The new scanning technologies are enabling researchers to capture various types of morphological information at very high

resolutions. Whole slide image analysis technology provides a way to diagnose huge numbers of medical conditions and phenomena currently studied in clinical medicine and pathology (Stathonikos et al., 2013; Al Janabi et al., 2011). There are different possibilities to apply image analysis techniques into whole slide images. Common tasks, where pathologists locate and recognize tissue components can be automated with the aid of image analysis algorithms, such as image segmentation and recognition algorithms. Mapping and quantification of cell/nuclei in the tumor area is another significant direction.

In histopathology, different color of stains or dyes are used to identify and visualize the tissue components and structures as originally the tissue looks transparent under the microscope (Ghaznavi et al., 2013). In absence of the stain the entire light passes through, as a result, it appears bright white (McCann et al., 2014a,b). Accordingly, mixtures of stains with different spectral absorption characteristics are used. Two or three different forms of colored stains are commonly used (Thaina et al.,

* Corresponding author at: Physiological Signal Analysis Group, Center for Machine Vision and Signal Analysis, University of Oulu, P.O. BOX 8000, FI-90014, Finland.

E-mail address: mdziaul.hoque@oulu.fi (Md.Z. Hoque).

2019). Hematoxylin and eosin are the most common routine stains to evaluate disease morphology where the hematoxylin specifies as a purple or blue color appearance (e.g. cells, nuclei) and eosin provides eosinophilic structures (e.g., muscle fibers, cytoplasm, and collagen) as a pinkish hue appearance (Altsubaie et al., 2017).

However, the concentration of the stain present in the pixel determines the intensity of each pixel. But the use of different scanner, slide preparation, different stain coloring, concentration of the stain, specimen thickness, temperature, batches of stains, and activities from different manufacture causes color variation in whole slide images. In general, a whole slide image implies scanning sections of tissue that are mounted on a glass slide with high resolution (normally 20X or 40X resolution). It has often a pyramidal structure, with many levels of varying resolution. Therefore, handling whole slide images with the highest resolution is challenging for the lack of available computer memory and display characteristics (Gurcan et al., 2009b).

The Biological Stain Commission in USA regulates the synthesis of biological stains that promoted the awareness of the issue of staining variability and encouraged efforts to reduce it. Also, the expert assessors (histology professionals and pathologists) in UK survey slides submitted for quantitative analysis, visually looking for staining characteristics. They examined each histochemical stain and assigning a semi-quantitative score to the slide (Gray et al., 2015). Such semiquantitative scoring technique is well established for the assessment of stain quality and intensity in every aspect of pathology and histotechnology. However, it is relatively expensive and time consuming in routine daily practice due to the manual assessment of slides by specialists.

From the existing research, it can be stated that image quality is highly important. The resulting image quality is a sum of several operations, starting from tissue preparation, and how the sections are handled in the laboratories to the selection of scanner devices and selected parameters. Management of image quality consistency is particularly important in biobank research where cumulative data combined from several biobanks will enable large enough sample sizes, multidisciplinary research, and innovations. Sharing knowledge is important in evaluating diseases and digital pathology provides a means for sharing high resolution digitalized information about the samples. The challenges in histological image analysis include the need for a wide variety of different methods to identify medical problems. Different tissues have different biological properties (biological factors) and staining protocol inconsistencies causing a high visual appearance variability (Arevalo and Cruz-Roa, 2014; Kothari et al., 2013; Hoque, 2019; Wei and Simpson, 2014).

It is important to study how color variability affects image analysis in computer-aided diagnostics in histopathology (Azevedo et al., 2019). Color normalization techniques for stained and digitalized pathology samples have been proposed to minimize the color variations. Color normalization is a process to realize color transformation from one image to another for the minimization of color variations. In this paper, a novel method with key elements of stain normalization is described. Color variability in H&E (hematoxylin and eosin) stained and digitalized pathology samples of 3D organotypic model for carcinoma cell invasion studies (Nurmenniemi et al., 2009) are considered for examination, and color deconvolution under illumination estimation of sample images is considered as a test case for the developed color normalization algorithm.

Section 2 describes the related work. Section 3 describes the proposed method. Section 4 shows the simulation results and the performance comparison of the developed method with selected state of the art methods. Finally, Section 5 and Section 6 summarize our main findings.

2. Related work

Stain normalization is a significant research topic with immense literature (Onder et al., 2014). Currently, there are various color normalization methods based on stain specific color deconvolution (Ruifrok and Johnson, 2001) which requires prior reference stain vectors for the dyes present in whole slide image. Macenko et al. (2009) introduced an algorithm to find the fringe of pixel distribution in the optical density space and determine the stain vectors in presence of strong staining variations. The pathologist needs to identify not only the stain components but also the incorporation of the spatial dependency of tissue structures (e.g. find the near elliptic shape of cell nuclei from stained hematoxylin and eosin tissue structures). The negligence of spatial dependency of tissue structures limits the robustness of the method. A linear color transform method introduced by (Wang et al., 2007) only works with transforming the color to a template image using linear projection in color space.

Zheng et al. (2018) introduced a normalization method that works in HSV space. Janowczyk et al. (2017) introduced a stain normalization method that clusters the pixels in histogram by sparse autoencoders. These histogram transformation based methods can utilize staining characteristics of histological images but have difficulties in analyzing different complex tissue areas due to confusion after normalization (Magee et al., 2009). Different multiple transformation methods were proposed to analyze different colored independent stains (Bejnordi et al., 2016; Khan et al., 2014). Typically, Khan et al. (2014) introduced a nonlinear mapping approach to utilize a pretrained relevance vector machine classifier to classify the stain pixel distributions which is comparatively good under noise. Zheng et al. (2019) proposed an adaptive color deconvolution method that avoids artifacts, but the overall failure rate of this method is 4 %.

Bejnordi et al. (2016) proposed hue saturation density-based color normalization method which has good performance in normal stain variations but comparatively bad performance in strong stain variations. Reinhard et al. (2001a, 2001b) introduced a technique that aligns the color channels with LAB color model (Hunter, 1948). However, dye has its specific reaction pattern with an independent contribution of color so that a single transformation function leads to improper color mapping for the stained components. To overcome this problem separate transformations are needed in different stained variations (Magee et al., 2009; Bejnordi et al., 2014; Basavanhally and Madabhushi, 2013). The Gaussian mixture model introduced in Basavanhally and Madabhushi (2013) has less robustness in strong stain variations.

The method described in Bejnordi et al. (2016) has also less accuracy due to the considerable imbalance in stained components in tissue images. The stain variations caused by different staining protocols have various problems (Bejnordi et al., 2014) where most of the published algorithms were enhanced the performance not typically evaluated in an existing computer-aided system.

3. Methods

Our proposed algorithm first takes the whole slide image, with multi resolution pyramid structure as input, and a group of pixels is separated and converted into optical density space. The source image is transformed into optical density space to determine the color appearance matrix from the measurement of relative color for each pixel in RGB channels. Color stains act linearly in optical density space according to Beer's Law (Ruifrok and Johnson, 2001) and color deconvolution method is applied. Depth matrix of stain is evaluated by taking the inverse of the color appearance matrix. The normalized hematoxylin and eosin components are obtained with multiscale Retinex model that

estimates the illumination map of hematoxylin and eosin components.

This Retinex method is different from the traditional multiscale Retinex model based method which partitions a digital image into reflectance and illumination factors (Fu et al., 2016). Our method estimates illumination factors individually by finding the maximum intensity of pixel in three color channels. Further, our Retinex model refines the illumination map by exploiting the structure of the illumination. We manipulate the illumination map with the help of Gamma correction in the range of [0,1] (Zheng et al., 2019). Finally, we can get back to the normalized original image by the multiplication of optical density space with color appearance matrix and the recombination of hematoxylin and eosin components. The flow chart of the proposed

illumination estimation and Retinex model based technique is introduced in Fig. 1.

3.1. Stain separation

The tissue samples are generally stained with multiple dyes, where the first preprocessing step is to separate the areas stained with different dyes. Ruifrok and Johnson (2001) developed a color deconvolution method to separate a combination of two or three colors. To make this possible separation the color formation is presented as

$$I_{Input} = I_0 \exp(-Ac) \quad (1)$$

where, I_0 represents the intensity of light entering the specimen and I is the light detected both represented in the RGB color space. Moreover, A is the staining intensity and c is the matrix related to the absorption factors of the stains. The color deconvolution transforms RGB color space to another color space defined by the stains used for staining the tissue. The method needs an estimation of a stain matrix.

In practice, the color conversion matrixes can be worked out from single-stained control slides or ROIs.

In here, our proposed algorithm estimates the color deconvolution vectors based on selecting representative sample pixels for each stain class which is tricky as the stain concentration varies over the sample. It works on image-specific color deconvolution (Ruifrok and Johnson, 2001) for unmixing different stains from that specific image. The absorption factors (Ruifrok and Johnson, 2001) of the stains affect the unmixing accuracy represented by the following equation,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 255 \times \prod_3^1 n = 1e^{-AC} \\ 255 \times \prod_3^2 n = 1e^{-AC} \\ 255 \times \prod_3^3 n = 1e^{-AC} \end{bmatrix} \quad (2)$$

where, A is the amount of stain and n is the pixel, and C is the color deconvolution matrix. We calculated the mean color of regions of interest for each stain of n (eq. 3).

$$\begin{bmatrix} c_r \\ c_g \\ c_b \end{bmatrix}_n = \left[\begin{bmatrix} -\log(\frac{r_n + 1}{256}) \\ -\log(\frac{g_n + 1}{256}) \\ -\log(\frac{b_n + 1}{256}) \end{bmatrix} \right] / \left[\begin{bmatrix} -\log(\frac{r_n + 1}{256}) \\ -\log(\frac{g_n + 1}{256}) \\ -\log(\frac{b_n + 1}{256}) \end{bmatrix} \right] \quad (3)$$

Finally, we estimated colors of interest for each stain of three color channels. In here we used the approximate vectors forming where the matrixes have been specified for different stain combinations like hematoxylin and eosin, and hematoxylin and DAB.

3.2. Pixel distribution in optical density space

The intensities I_{Red} , I_{Green} , and I_{Blue} are obtained with the help of the scanner for each pixel in RGB model. The relative intensity depends on the stain concentration which is non-linear and cannot directly be used in stain separation calculation. Whereas, the optical density of different channels is linear with the concentration of absorbing factors. For more accurate quantitative analysis and measurement we used the optical density space (OD) as following,

$$OD = -\log_{10}(I_{Input}) \quad (4)$$

where, I_{Input} is the RGB color vector. Our proposed algorithm separates a group of pixels of whole slide image and converts it into optical density space to find the fringe of pixel distribution in the optical density space.

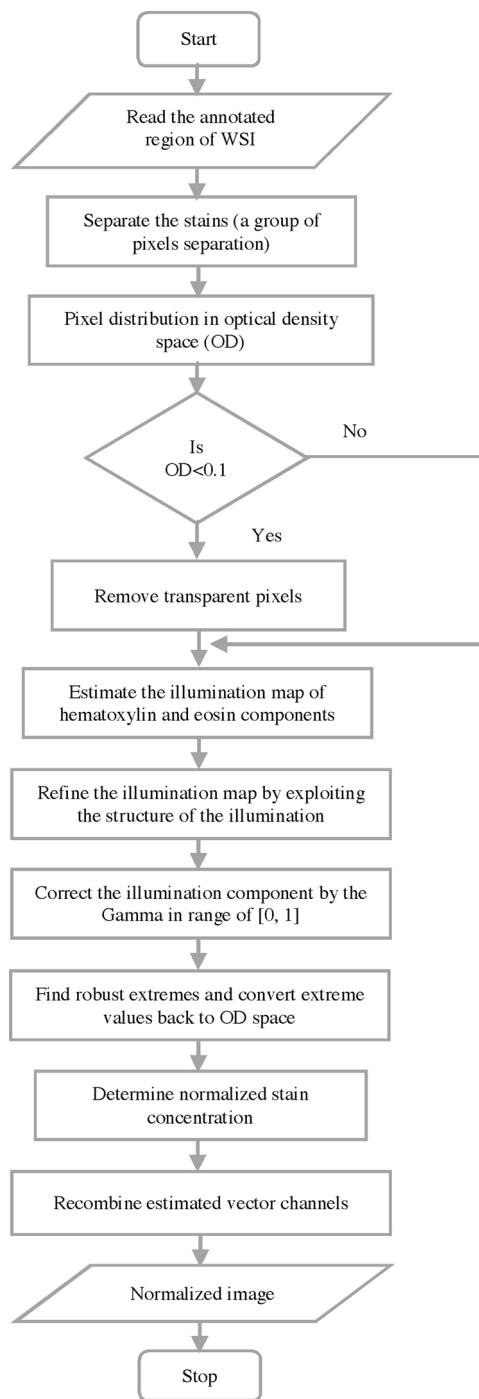


Fig. 1. Flow chart of the proposed Retinex model based stain normalization technique.

3.3. Multiscale retinex model

Multiscale Retinex model is used to estimate illumination map constructed by finding maximum intensity of RGB color channels (Fu et al., 2016; Jozé et al., 2012). The formula of the multiscale Retinex model is following,

$$I_{In} = I_{Out} M_E T_{Map} \quad (5)$$

where, I_{In} is the input image, and I_{Out} is the desired output result, M_E is the operator of element-wise multiplication and T_{Map} is the illumination map. Our designed method only estimates global illumination. For the non-uniform illumination, we adopt the following estimation formulation,

$$T_{Map}(x) \leftarrow \text{Max}_{R,G,B} I_{In} \quad (6)$$

where, x is the individual pixel. This operation helps to illuminate the maximum value of the three channels. For the structural refinement of the illumination map, we set the weights as following,

$$W_h(x) \leftarrow 1; W_v(x) \leftarrow 1 \quad (7)$$

where, W_h and W_v are the horizontal and vertical weights. We used the gradient of the illumination map as the weights of the following sequel formulations,

$$W_h(x) \leftarrow \frac{1}{|\nabla_h T_{Map}| + \epsilon}; W_v(x) \leftarrow \frac{1}{|\nabla_v T_{Map}| + \epsilon} \quad (8)$$

where, $\nabla_h T_{Map}$ (horizontal) and $\nabla_v T_{Map}$ (vertical) are the first-order derivative filters.

To characterize the relative total variation, we set the weights as following,

$$W_h(x) \leftarrow \sum_{y \in \Omega(x)} \frac{G_\sigma(x, y)}{G_\sigma(x, y) |\nabla_h T_{Map}| + \epsilon} \quad (9)$$

$$W_v(x) \leftarrow \sum_{y \in \Omega(x)} \frac{G_\sigma(x, y)}{G_\sigma(x, y) |\nabla_v T_{Map}| + \epsilon} \quad (10)$$

where, $G(x, y, \sigma)$ is the Gaussian factor and σ is the standard deviation. Here, $G(x, y, \sigma)$ is denoted as,

$$G(x, y, \sigma_i) = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{D(x, y)}{2\sigma_i^2}\right) \quad (11)$$

where, $D(x, y)$ is the measurement of the spatial Euclidean distance. This convolution with Gaussian factors can be helpful to reduce the noise in the images and emphasize the contrast of color structures. Finally, the multiscale Retinex (Starck et al., 1998) estimates the component at

Table 1

Information of an example image (.svs format) scanned by Leica Aperio AT2 whole slide scanner.

Image Format	SVS
Description	Aperio image library v12.0.15 11,988 × 89,547
Apparent Magnification	40X
Image Width	117 528 pixels
Image Height	89 447 pixels
Image Depth	1 pixel
Image Channels	3
Image Bit Depth	8 bits
File Size	3,653,670,000 bytes
Compression Type	JPEG using libjpeg
Compression Quality	60
Compression Ratio	8.63
Tile Width	240 pixels
Tile Height	240 pixels

scale, s with Gaussian low pass filter. The image component R is obtained with the subtraction of the log of estimated illumination from the log of the original image I as the following equation,

$$R(x, y, \sigma_i) = \log \{I(x, y)\} - \log \{G(x, y, \sigma_i)^* I(x, y)\} \quad (12)$$

The refined illumination map T_{Map} can be recovered with the help of the (eq. 5) and manipulated through gamma transformation as following,

$$T_{Map} \leftarrow T_{Map}^\gamma \quad (13)$$

For the avoidance of unbalance processing e.g. dark, brighter, or over smoothed tissue area, we can use the following operation,

$$I_f \leftarrow I_{In} M_E T_{Map} + I_d M_E (1 - T_{Map}) \quad (14)$$

where, I_d is the result after the estimation process and I_f is the result after recombination.

The whole process of our Retinex model based stain normalization algorithm is described in Algorithm 1.

Algorithm 1. Retinex Model Based Stain Normalization

Input: Read annotated whole slide image as I_{Input} .

Step 1: A group of pixels separated and converted into optical density space.

Step 2: Taking separated hematoxylin and eosin components as I_{In} ,
Gamma transformation parameter γ .

Step 3: Initialization by constructing weight matrixes with equations 7, 8, 9, and 10.

Processing:

1. Estimate initial illumination map T_{Map} through eq. 6.
2. Refine illumination map T_{Map} based on initial illumination map.
3. Gamma correction via $T_{Map} \leftarrow T_{Map}^\gamma$.
4. Estimate I_{In} using T_{Map} through equation 5.
5. If unbalance processing needs to be avoided, then refine I_{In} by I_d through eq. 14.

Output: Final normalization result as I_{Output} .

4. Results

4.1. Tissue collection and organotypic culturing

The collection of tissues was approved by the ethical committee of the Northern Ostrobothnia Hospital District. The oral tongue squamous carcinoma cells were cultured on top of the myoma tissues as described earlier (Nurmenniemi et al., 2009). Samples were digitalized using Leica Aperio AT2 whole slide scanner.

Descriptive information of an example whole slide image (.svs format) scanned by Leica Aperio AT2 scanner are shown in Table 1.

4.2. Experimental setup

For our experiment, we used ImageScope (Aperio ImageScope) to annotate the whole slide image, and our method analyzed by taking that annotated image automatically. We developed our algorithm using Matlab and Python separately to provide both Matlab and Python users in support of the digital pathology and cancer analysis.

We used ImageJ (Schneider et al., 2012) software for the ground truth labeling, which is the manual stain normalization and segmentation method, e.g. (Marek et al., 2019) used for their analysis. Based on selected regions from the datasets the pixel characteristics were labeled including hematoxylin, eosin, and background.

Details of the implementation parameters including their values are listed in Table 2.

Table 2

Implementation parameters and their values.

Parameter	Value
Density space (OD)	OD<0.1
The number of stains, A	A = 2
Threshold, Th	Th = 0.95
Hyperparameter, σ	$\sigma = 2$ and 4
Gamma transformation parameter, γ	$\gamma = 0.6$ and 0.8 in the range of [0,1]
Hyperparameter, ϵ	$\epsilon = 1e-4$
Hyperparameter, μ	$\mu > 0$
Hyperparameter, α , β , and λ	$\alpha=\beta=\lambda = 1$ where, $\alpha > 0$, $\beta > 0$, and $\lambda > 0$
Constant, c_1	$c_1 = (0.01*255)^2$
Constant, c_2	$c_2 = (0.03*255)^2$
Constant, c_3	$c_3 = \frac{c_2}{2}$
P_{95}	95 th percentile (Bejnordi et al., 2016)
$u(i)$	Average of the RGB channels
QSSIM, SSIM, and PCC	In the range of [0,1]
P-value	≤ 0.05

4.3. Performance evaluation criterion

In this experiment, the normalized median intensity (NMI) (Basavanhally and Madabhushi, 2013) is used to evaluate color constancy and compare the statistics of intensity over whole slide image. We used the following formula for the normalized median intensity measurement to utilize the color constancy of normalization (Basavanhally and Madabhushi, 2013).

$$NMI (I_M) = \frac{\frac{Median}{i \in I_M} \{u(i)\}}{\frac{P_{95}}{i \in I_M} \{u(i)\}} \quad (15)$$

where, I_M is the processed whole slide image, $u(i)$ is the average of the RGB channels, Median () is the median value and P_{95} is the 95th percentile (Bejnordi et al., 2016). The image quality metrics including Quaternion Structure Similarity Index Metric (QSSIM) (Kolaman and Yadid-Pecht, 2011), Structural Similarity Index Metric (SSIM) (Wang et al., 2004), and Pearson Correlation Coefficient (PCC) (Wang and Bovik, 2002) are calculated in order to measure performance. For the

quaternion structural similarity index metric (QSSIM), numerical value lies between 0–1 where the value closer to 1, indicates the better color normalization method. The mathematical formula is given in eq. 16.

$$QSSIM_{ref,norm} = \left| \left(\frac{2\mu_{qref}\mu_{qnorm}}{\mu_{qref}^2 + \mu_{qnorm}^2} \right) \left(\frac{2\sigma_{qref,qnorm}}{\sigma_{qref}^2 + \sigma_{qnorm}^2} \right) \right| \quad (16)$$

where, μ_{qref} and μ_{qnorm} are the sample mean and normalized image; σ_{qref} and σ_{qnorm} are the standard deviation of source and normalized images. $\sigma_{qref,qnorm}$ is the correlation coefficient between source and normalized images. Similarly, the numerical value lies between 0–1 for structural similarity index metric (SSIM), where the value closer to 1, indicates the better color normalization method. The SSIM can be defined as follows,

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\lambda \quad (17)$$

From the eq. 17 we can see that SSIM depends on three factors such as,

$$(a) Luminance, l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (18)$$

$$(b) Contrast, c(x, y) = \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (19)$$

$$(c) Structure, s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (20)$$

After combining eq. 18, 19, and 20 the mathematical formula of SSIM is derived as follows,

$$SSIM (x, y) = \left\{ \left(\frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \right) \left(\frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \right) \right\} \quad (21)$$

where, μ_x and μ_y are the means of the source and normalized images; σ_x and σ_y are the standard deviation of source and normalized images; c_1 , c_2 and c_3 are the constants. For the pearson correlation coefficient (PCC), we measured the linear correlation between source and normalized images in a range, 0–1 where 0 denotes not any similarity between two images. However, a value greater than 0 indicates the correlation be-

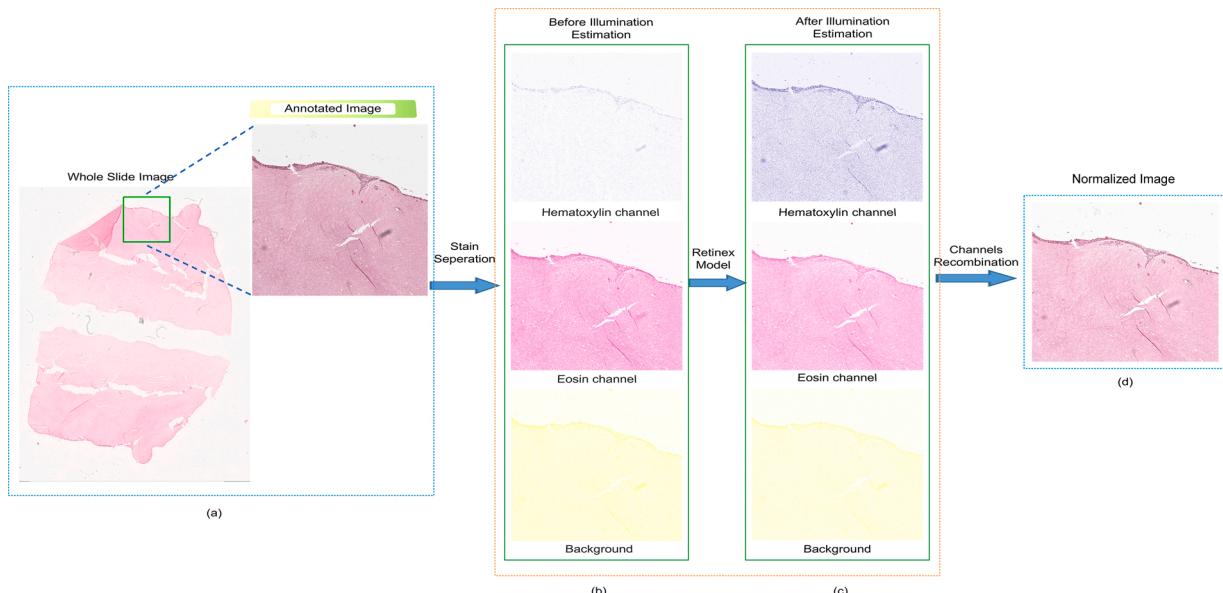


Fig. 2. The detailed architecture of the proposed Retinex model based stain normalization method. (a) Annotated image from the original whole slide image with artifacts and noise, (b) Hematoxylin, Eosin, and Background channels separation before illumination estimation, (c) Normalized hematoxylin, eosin, and background channels using proposed method, (d) Normalized image with channels recombination using proposed method.

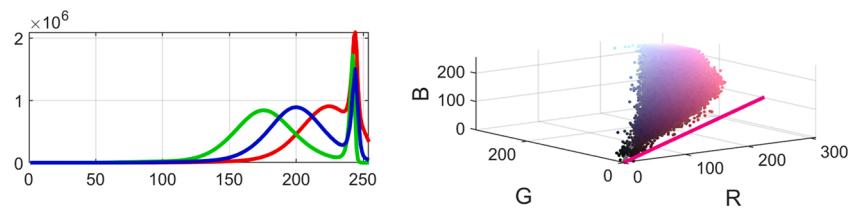


Fig. 3. Stain vectors using 3D color gamut as a point cloud in RGB color space (Right) and Histogram (Left).

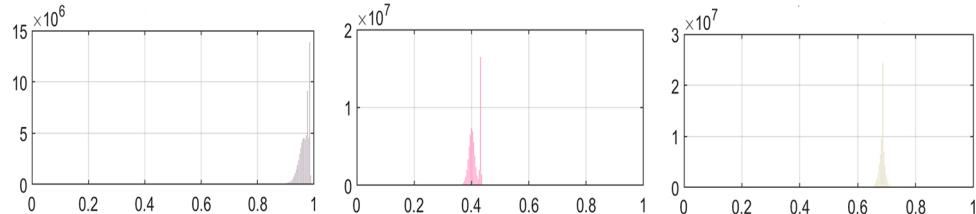


Fig. 4. Histograms of the three normalized stain density maps.

tween two images (Roy et al., 2018a). The mathematical formula is given in eq. 22.

$$PCC = \frac{\sum_i (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_i (x_i - \mu_x)^2} \sqrt{\sum_i (y_i - \mu_y)^2}} \quad (22)$$

where, x_i and y_i are the source and normalized image; μ_x and μ_y are the means of source and normalized images. We calculated the relative square error (rSE) using the formula (Salvi et al., 2020) given in eq. 23.

$$rSE = \frac{\|(W - W_{GT})^T (W - W_{GT})\|}{\sqrt{\|W^T \cdot W\| \| (W_{GT})^T \cdot W_{GT} \|}} = \frac{\|W - W_{GT}\|_F^2}{\|W\|_F \cdot \|W_{GT}\|_F} \quad (23)$$

where, W_{GT} is the ground truth stain color appearance matrix, F is the Frobenius norm of the matrix (Salvi et al., 2020) and W is the estimated color appearance matrix. Finally, we calculated standard deviation and

coefficient variation of normalized median intensity for all myoma samples where the lower value of these denotes more consistent normalization.

4.4. Experimental results

The aim of the experiment was to study stain normalization by utilizing color from the stained whole slide images. Currently developed algorithms and software e.g. SPCN (Vahadane et al., 2016), do not have support to read whole slide image formats directly. In majority of cases, for example Aperio (.svs) and Hamamatsu (.ndpi) formats need to be converted into tagged image file format (.tiff) in where we lose some important information. Sometimes we need to use an extra library which is expensive and makes the system more complex and comparatively slow in performance e.g. OpenSlide (Goode et al., 2013a,b). So, another aim was to develop such an algorithm that can directly read these types of whole slide image formats without any extra library requirements,

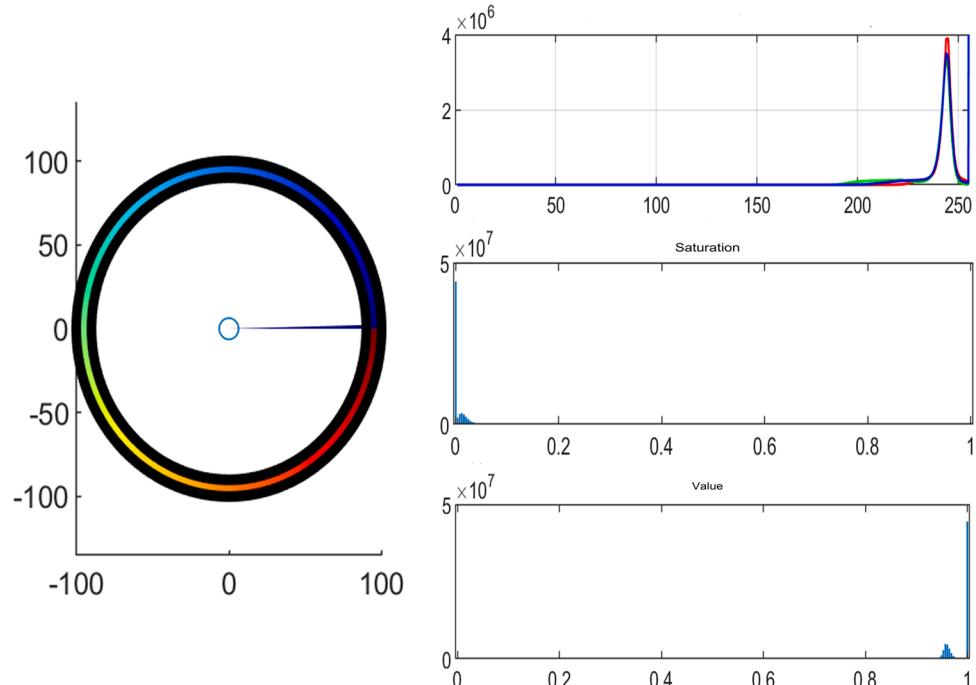


Fig. 5. Histogram of RGB color space (Right) and hue channel (Left).

Table 3Comparisons based on average Euclidean distance for hematoxylin and eosin vector (mean \pm STD).

Method	Average Euclidean distance for hematoxylin vector (mean \pm STD)	Average Euclidean distance for eosin vector (mean \pm STD)
Macenko et al., 2009	0.040 \pm 0.011	0.038 \pm 0.015
Khan et al., 2014	0.192 \pm 0.027	0.183 \pm 0.030
Bejnordi et al., 2016	0.127 \pm 0.092	0.022 \pm 0.021
Janowczyk et al., 2017	0.053 \pm 0.041	0.097 \pm 0.037
Zheng et al., 2018	0.079 \pm 0.022	0.145 \pm 0.042
Zheng et al., 2019	0.036 \pm 0.004	0.040 \pm 0.007
Proposed method	0.010 \pm 0.001	0.020 \pm 0.004

where there is not any issue to lose the information. In our experiment, we first utilized the staining difference and color normalization of hematoxylin and eosin (HE) stained whole slide images. Then we examined the effectiveness and the fluctuation of the results by calculating hematoxylin and eosin stain vectors in the presence of various color intensities.

Finally, we compared the effectiveness and robustness of our proposed Retinex model based method with state of art methods. The detailed architecture of our proposed algorithm is shown in Fig. 2.

It shows that the individual stain components of the histochemical stains are effectively quantified and separated. Moreover, the normalized output image is free from artifacts and apparent color discontinuity because of less difficulties to identify the color distributions from different stains (Zheng et al., 2019). In here, the variability was examined more closely, with normalized density histograms and visual inspection.

The normalized stain vectors using 3D color gamut as a point cloud in RGB color space are shown in Fig. 3 and three normalized stain density maps are shown in Fig. 4.

The histogram of RGB color space and the hue channel of the normalized image is shown in Fig. 5.

It is seen from Fig. 3 that, the 3D color gamut as a point cloud in RGB color space represents the existence of hematoxylin and eosin staining where the blue color denotes hematoxylin staining and the purple color denotes the eosin staining. Figs. 4 and 5 help to visualize the variations of staining vectors using density maps, histogram of RGB color space, and hue channel.

4.5. Comparison with different methods

We studied the means for objective measurement of the differences in color deconvolution results. To evaluate the performance of our proposed algorithm, comparative evaluation results are shown in calculating stain vectors and achieving color constancy. To compare the efficiency and robustness of different algorithms the correct stain vectors are measured based on the mean and standard deviation of the Euclidean distances.

Table 3 represents the mean and standard deviation of the Euclidean distance between stain vectors from the annotated image and derived by each algorithm for the comparison of efficiency and robustness. Comparatively high mean value and low standard deviation of the Euclidean distances indicate better results.

It is seen that our proposed method has comparatively low standard deviation of the Euclidean distances which is considerably better compared to other methods.

Table 4 represents the normalized median intensity measure to evaluate the color constancy of different regions.

The overall performance of our proposed method is considerably better compared to other methods in calculating normalized median intensity for the hematoxylin and eosin stains. Statistically, the standard deviation (SD) and the coefficient of variation (CV) of normalized median intensity results of different methods are utilized in Figs. 6 and 7.

We used boxplot representation for comparison of the distributions e.g. shape of the distribution, central value, and variability.

We also used violin plot (a combination of boxplot and density plot) to visualize the distribution and probability density of the data. The performance of color normalization is shown in Fig. 8, which represents different normalization results of the state of art methods in terms of comparison.

It is clear from Figs. 6, 7, and 8 that, Zheng et al. (2019) and Macenko et al. (2009) methods are good in color normalization but failure occurred due to color artifacts and variance. The result of Janowczyk et al. (2017) method is good but the results are not smooth due to various artifacts that appeared in normalized images. These methods have a typical failure rate in color distribution of the monochromatic whole slide image. Whereas, compared to all seven methods our proposed method removes the color artifacts with superior performance and effectively estimates the hematoxylin, eosin, and background distributions and classifies the pixels with better performance in color distribution of the monochromatic whole slide image. Our method provides effective normalization results that help to overcome the color variance challenges. Table 5 represents the detailed evaluation results of staining separation (mean \pm standard deviation), relative square error of hematoxylin and eosin respectively, time complexity, and average running time of the reference methods.

Here, our algorithm consistently has the smallest standard deviation and low relative square error. The time complexity is another important application where Table 5 also represents the comparison of time complexity for the pixel number, n , and the average running time. It is specified that the pixel-wise classification based method, Khan et al. (2014) is linearly related to the number of pixels, n so that it has the time complexity as $\sigma(n)$. Bejnordi et al. (2016) has the time complexity as $\sigma(n^3)$ due to the utilization of Hough transform.

Zheng et al. (2018) is the parameter estimation method that involved sorting algorithm of n value so that this method has time complexity as $\sigma(n \log_2 n)$. Other methods estimate the parameters in terms of pixel-wise operation where the time complexity is $\sigma(n)$. The average normalization parameters estimation running time of our proposed method and (Zheng et al., 2019) method are 2.11 s and 2.89 s which are comparatively short time compare to the other reference methods.

We evaluated and compared the quality performance of the reference

Table 4

Comparison based on Standard Deviation (SD) and Coefficient of Variation (CV) of Normalized Median Intensity (NMI).

Method	Hematoxylin		Eosin	
	(NMI) SD	(NMI) CV	(NMI) SD	(NMI) CV
Macenko et al., 2009	0.011	0.15	0.020	0.04
Khan et al., 2014	0.025	0.39	0.040	0.07
Bejnordi et al., 2016	0.040	0.33	0.030	0.10
Janowczyk et al., 2017	0.050	0.21	0.030	0.19
Zheng et al., 2018	0.064	0.27	0.050	0.20
Zheng et al., 2019	0.030	0.12	0.020	0.03
Proposed method	0.010	0.02	0.015	0.02

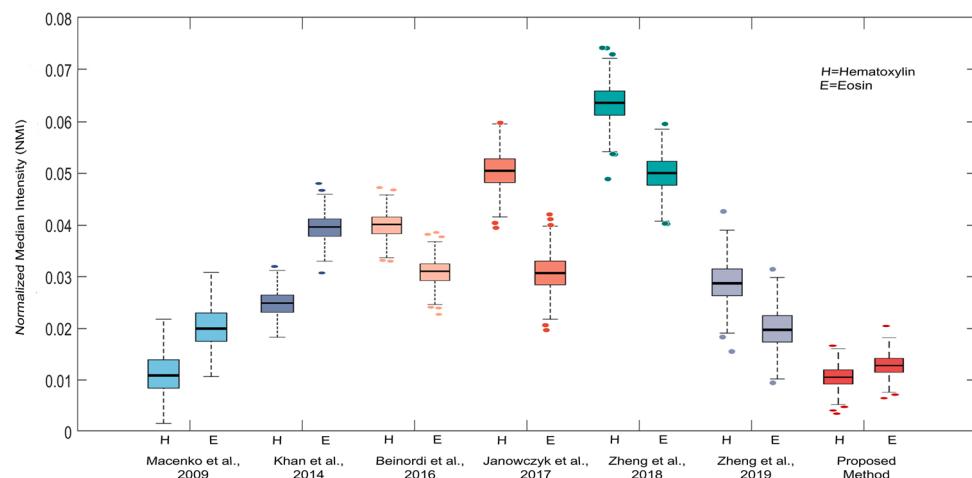


Fig. 6. Boxplots of the standard deviation of normalized median intensity to compare distributions e.g. shape of distribution, central value, and variability of different state of the art methods.

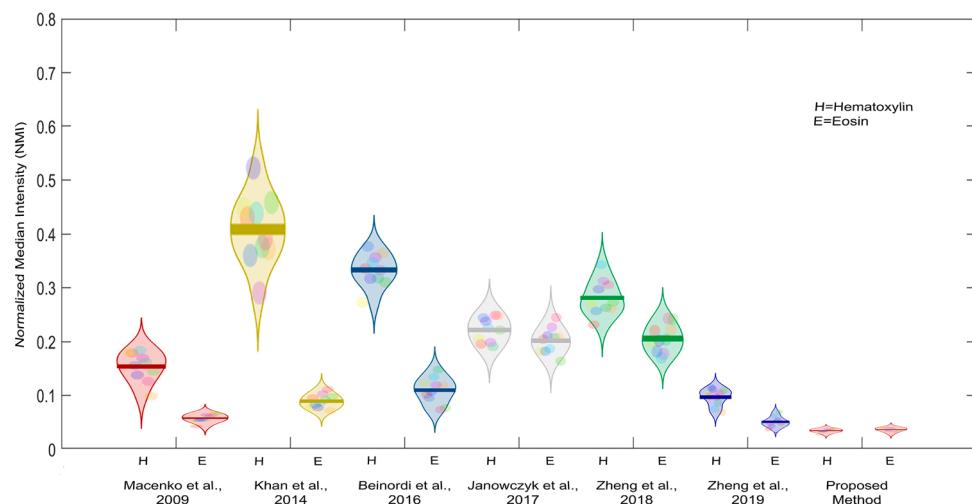


Fig. 7. Violin plots of the coefficient variation of normalized median intensity results of different methods where the dot points inside the shadow represent the normalized median intensity allocation. The thick bar at the center highlights the interquartile range and the extended thin points represent the maximum (upper), and minimum (lower) values of data.

methods using quaternion structure similarity index metric (QSSIM), structure similarity index metric (SSIM), and pearson correlation coefficient (PCC) (Roy et al., 2018b) on different myoma samples (results shown in Table 6). The mean, standard deviation (STD), minimum (min), maximum (max) along *p*-value are shown in Table 7.

The QSSIM, SSIM, and PCC of our proposed method on different myoma samples are comparatively higher than other reference methods (Table 6). On the other hand, a small *p*-value (*p*-value ≤ 0.01 means highly significant) presents the proof against the null hypothesis testing (Altman and Bland, 2011). Table 7 shows that the *p*-values of our method and (Zheng et al., 2019) method are lower than other reference methods that means the results are highly significant.

The experimental results (Tables 3, 4, 5, 6 and 7) indicate that our proposed algorithm is consistently better for the stain normalization with less apparent artifacts in normalized whole slide images.

5. Discussion

Our proposed algorithm effectively removes the unwanted variability and makes the analysis more consistent. The scope of possibilities and alternatives to histopathological image analysis is wide (Arevalo and Cruz-Roa, 2014). Additionally, there has been an increasing interest in quantifying the tumor microenvironment and correlating the structural information with tumor invasion and patient prognosis (Wang et al., 2019).

In particular cases, the image analysis techniques can provide the supports to help the pathologists in disease grading and decision making. These techniques can be used to visualize the tissue morphology, stromal or fatty adipose tissue findings, cancer epithelium area detection, tubule counting, mitosis quantification, tissue segmentation, and classification (Aeffner et al., 2019). In that context, the myoma organotypic model (Nurmenniemi et al., 2009) is a particularly attractive tool combining real human stromal tissue with the possibility for various

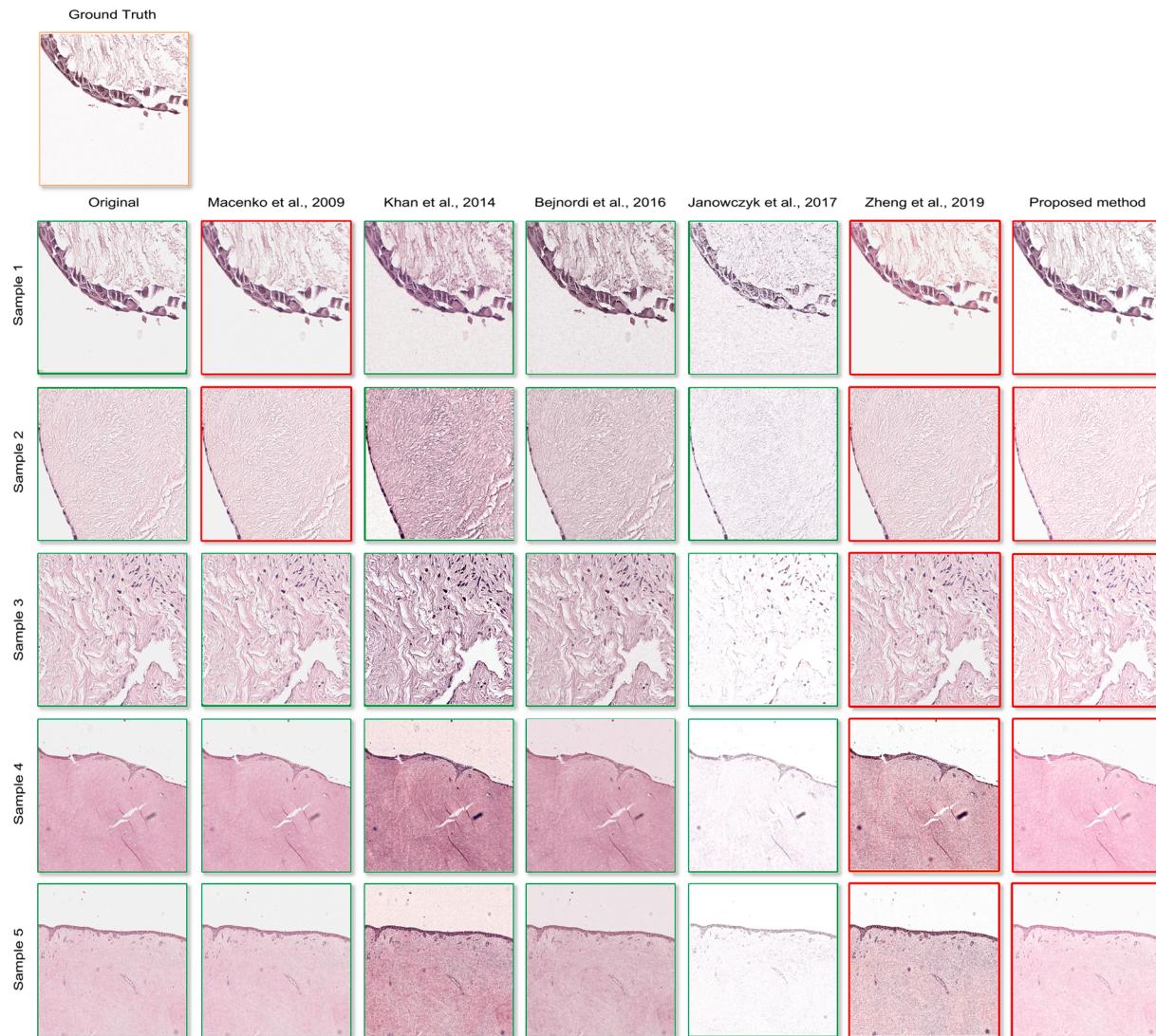


Fig. 8. Normalization results (five myoma samples) of compared methods. Here the red boxes represent the results without artifacts or apparent color discontinuity (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 5

Evaluation results of staining separation (mean \pm standard deviation), relative square error ($rSE_{Hematoxylin}$ and rSE_{Eosin} are the stain estimation relative square error of hematoxylin and eosin respectively), time complexity, and average running time of the reference methods.

Method	Staining Separation (mean \pm standard deviation)		Relative Square Error (rSE)		Time Complexity		Average running time (S)
	$S_{Hematoxylin}$	S_{Eosin}	$rSE_{Hematoxylin}$	rSE_{Eosin}	Pre-processing	Model solution	
Macenko et al., 2009	0.9149 ± 0.0545	0.3505 ± 0.0545	0.0086	0.0051	–	$\sigma(n)$	14.22
Khan et al., 2014	0.8723 ± 0.0815	0.3118 ± 0.0630	0.0053	0.0039	$\sigma(n)$	$\sigma(n)$	1210
Bejnordi et al., 2016	0.8538 ± 0.0975	0.3059 ± 0.0570	0.0081	0.0168	$\sigma(n^3)$	$\sigma(n)$	489
Janowczyk et al., 2017	0.9036 ± 0.0745	0.2785 ± 0.0720	0.0064	0.0073	–	$\sigma(n)$	73.16
Zheng et al., 2018	0.8481 ± 0.0810	0.3811 ± 0.0830	0.0092	0.0115	–	$\sigma(n \log_2 n)$	42.01
Zheng et al., 2019	0.9597 ± 0.0340	0.4515 ± 0.0365	0.0025	0.0031	–	$\sigma(n)$	2.89
Proposed method	0.9761 ± 0.0160	0.4852 ± 0.0135	0.0011	0.0017	–	$\sigma(n)$	2.11

experimental cell culture set-ups with e.g. anti-invasive drug candidates.

Recently different convolutional neural networks and deep learning based color normalization methods have emerged (Shafiei et al., 2020; Zanjani et al., 2018; Bug et al., 2017). They come with high computational cost which limits the applicability of these methods in practical whole slide image analysis systems. Our proposed method has low computational cost and is easy to use in effective computer aided whole

slide image analysis systems. Most recently developed adaptive color deconvolution method proposed by Zheng et al. (2019) is effective in stain normalization with 4% failure rate due to color artifacts and variance which is relatively low. Our proposed method is more effective in stain normalization without any specific failure rate due to color artifacts. On the other hand, it is difficult to analyze the color distribution of the monochromatic whole slide image (Zheng et al., 2019).

Table 6

Quality performance of the reference methods in terms of QSSIM, SSIM, and PCC.

Method	QSSIM	SSIM	PCC
Macenko et al., 2009	0.93865	0.93350	0.94184
Khan et al., 2014	0.90215	0.84692	0.87913
Bejnordi et al., 2016	0.82327	0.79114	0.83690
Janowczyk et al., 2017	0.88105	0.87225	0.89576
Zheng et al., 2018	0.87289	0.85531	0.91223
Zheng et al., 2019	0.95950	0.95180	0.96105
Proposed method	0.97591	0.98221	0.98431

Table 7

The quantitative comparisons (mean, standard deviation, min, max, and *p*-value) of the reference methods.

Method	Mean	STD	Min	Max	<i>P</i> -value
Macenko et al., 2009	0.92	0.05	0.90	0.94	0.009
Khan et al., 2014	0.88	0.08	0.84	0.91	0.02
Bejnordi et al., 2016	0.86	0.09	0.81	0.87	0.04
Janowczyk et al., 2017	0.90	0.07	0.87	0.93	0.01
Zheng et al., 2018	0.85	0.08	0.82	0.90	0.05
Zheng et al., 2019	0.96	0.03	0.93	0.97	0.007
Proposed method	0.97	0.02	0.95	0.99	0.004

Hence, our proposed method ideally removes the color artifacts and the normalization result is consistent with color appearance in analyzing the color distribution of the monochromatic whole slide image. The evaluation of the state of art methods showed color discontinuity in normalized whole slide images. This is one of the major issues that appear on hematoxylin and eosin boundary. However, our proposed method only estimates illumination map of the whole slide image and utilizes more general color pixels without any apparent color discontinuity. So, the normalized whole slide images are consistent in color appearance containing almost zero percent color artifacts.

Besides, the measure of hematoxylin and eosin (HE)-ratio is vulnerable to being dependent on image content (the actual amount of hematoxylin and eosin stained areas) and our method can be effective also in this case. Therefore, our Retinex model based method is more robust and effective compared to the state of art methods to quantify the individual stain components of the histochemical stains for the perfect expulsion of variability and area segmentation from hematoxylin and eosin stained tissue images.

6. Conclusion

It is evident that digital pathology is taking vast steps towards realization in clinical work in addition to being part of research methodology. Applications, both to replace conventional time-consuming work routines and the new methods are appealing to the community of pathologists as well as researchers. In this work, we studied the staining difference and color normalization of hematoxylin and eosin stained whole slide images. We examined the performance by calculating the stain vectors. Our Retinex model based stain normalization method separates the individual stain components of the histochemical stains and the normalized images are achieved via illumination estimation utilizing more general color pixels without any apparent color

discontinuity. It helps to remove the color artifacts with superior performance and effectively estimates the hematoxylin, eosin, and background distributions of whole slide images.

Experimental evaluation showed that compared to the selected state of art methods, our method consistently has the smallest standard deviation, skewness value, and coefficient of variation in normalized median intensity measurements. The quality performance of our method in terms of QSSIM, SSIM, and PCC is considerably better for the stain normalization with less apparent artifacts in normalized whole slide images. Many recent studies are focused on medical image standardization. Moreover, the effect of staining variation has been noticed as a challenge when developing advanced tools in the area of computational pathology. Ideally, computational pathology would provide new innovations and tools for diagnostic purposes, as well as methods for automated analysis in biobank and other multi-center studies where the large database of digitalized samples are retrieved from multiple laboratories.

Our proposed method can provide a potential tool to remove the color variability to overcome the heterogeneity in the digitalized slides towards automating and comparable analysis in histopathology.

Credit author statement

Md. Ziaul Hoque: Conceptualization; Methodology; Software Implementation; Data Analysis; Visualization; Writing-Original Draft. **Anja Keskinarkaus:** Supervision; Data Curation; Investigation; Writing Review. **Pia Nyberg:** Resources; Writing Review. **Tapio Seppänen:** Principal Supervision; Project Administration; Resources; Funding Acquisition; Writing Review.

Ethical approval

The Ethical Committee of Northern Ostrobothnia Hospital District approved the research and the use of myoma tissue samples.

Declaration of Competing Interest

The authors have no conflict of interest that could have appeared to influence the work reported in this paper.

Acknowledgments

The research work of this paper was conducted with Physiological Signal Analysis Group at Center for Machine Vision and Signal Analysis (CMVS) in the Faculty of Information Technology and Electrical Engineering (ITEE) at University of Oulu, Finland. This research has been financially supported by Academy of Finland 6 Genesis Flagship (Grant 318927) and Academy of Finland Identifying trajectories of healthy aging via integration of birth cohorts and biobank data (Grant 309112). We are thankful to Professor Tuula Salo group (University of Oulu and University of Helsinki) for selection and access to the myoma organotypic slides. We thank Sanna Juntunen, Eeva-Maija Kiljander, Maija-Leena Lehtonen, and Merja Tyynismaa for technical assistance in preparing the organotypic cultures and slides.

Appendix A

Table A1 and **Table A2** represent the detailed statistics derived from the separated channels of twenty different myoma samples.

Table A1
Comparisons with different optimal mean and standard deviation (STD) values for hematoxylin channel.

Method	Hematoxylin			Khan et al., 2014			Bejnordi et al., 2016			Janowczyk et al., 2017			Zheng et al., 2019			Proposed method		
	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew
11642	0.90	0.05	-3.93	0.91	0.06	-3.82	0.87	0.07	-4.23	0.89	0.08	-4.99	0.95	0.03	-3.22	0.97	0.01	-2.63
11643	0.93	0.06	-3.98	0.90	0.06	-3.85	0.87	0.07	-4.70	0.88	0.08	-4.97	0.95	0.04	-3.27	0.98	0.02	-2.31
11644	0.92	0.05	-3.92	0.91	0.06	-3.81	0.86	0.09	-4.53	0.88	0.07	-4.93	0.94	0.04	-3.49	0.97	0.02	-2.55
11645	0.91	0.06	-3.92	0.91	0.05	-3.80	0.84	0.07	-4.67	0.89	0.07	-4.91	0.94	0.03	-3.50	0.96	0.01	-2.36
11646	0.90	0.06	-3.90	0.92	0.05	-3.80	0.83	0.07	-4.58	0.91	0.07	-4.98	0.95	0.03	-3.66	0.99	0.01	-2.27
11647	0.93	0.05	-3.88	0.90	0.06	-3.81	0.80	0.09	-4.88	0.93	0.08	-4.98	0.93	0.04	-3.72	0.98	0.02	-2.22
11648	0.92	0.05	-3.82	0.91	0.06	-3.77	0.81	0.08	-4.81	0.87	0.07	-4.92	0.95	0.04	-3.80	0.96	0.02	-2.50
11649	0.91	0.05	-3.83	0.89	0.08	-3.78	0.84	0.09	-4.87	0.88	0.08	-4.98	0.95	0.04	-3.91	0.96	0.02	-2.30
11650	0.94	0.05	-3.87	0.86	0.08	-3.89	0.84	0.09	-4.74	0.87	0.07	-4.97	0.93	0.03	-3.77	0.95	0.02	-2.19
11651	0.90	0.06	-3.69	0.87	0.07	-3.79	0.84	0.09	-4.73	0.88	0.07	-4.97	0.92	0.04	-3.57	0.97	0.02	-2.66
11652	0.91	0.06	-2.81	0.87	0.08	-2.71	0.85	0.07	-3.70	0.88	0.08	-4.98	0.92	0.04	-2.86	0.95	0.02	-1.81
11653	0.91	0.06	-3.80	0.87	0.07	-3.69	0.85	0.09	-4.69	0.88	0.07	-4.99	0.92	0.04	-3.44	0.96	0.01	-2.14
11654	0.94	0.06	-3.77	0.87	0.07	-3.82	0.83	0.08	-4.33	0.89	0.07	-4.97	0.93	0.04	-3.86	0.96	0.01	-3.92
11655	0.93	0.06	-3.79	0.87	0.07	-3.76	0.85	0.09	-4.37	0.87	0.07	-4.95	0.94	0.03	-3.56	0.97	0.02	-2.35
11656	0.91	0.05	-3.73	0.89	0.08	-3.70	0.85	0.09	-4.38	0.88	0.08	-4.98	0.93	0.03	-3.59	0.96	0.01	-2.15
11657	0.90	0.05	-2.98	0.85	0.07	-2.76	0.83	0.08	-4.43	0.88	0.07	-4.97	0.93	0.03	-2.97	0.97	0.02	-1.77
11658	0.90	0.05	-3.89	0.87	0.06	-3.78	0.86	0.08	-4.47	0.87	0.07	-4.97	0.95	0.04	-3.92	0.99	0.01	-2.02
11659	0.90	0.05	-3.89	0.87	0.08	-3.82	0.87	0.09	-4.79	0.88	0.07	-4.99	0.95	0.04	-3.88	0.98	0.02	-2.52
11660	0.90	0.06	-3.81	0.87	0.07	-3.79	0.85	0.09	-4.42	0.93	0.07	-4.98	0.95	0.04	-3.21	0.99	0.02	-2.00
11661	0.91	0.05	-3.83	0.84	0.08	-3.77	0.86	0.10	-4.84	0.87	0.08	-4.93	0.94	0.03	-3.81	0.98	0.01	-2.55

Table A2
Comparisons with different optimal mean and standard deviation (STD) values for eosin channel.

Method	Eosin			Khan et al., 2014			Bejnordi et al., 2016			Janowczyk et al., 2017			Zheng et al., 2019			Proposed method		
	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew	Mean	STD	Skew
11642	0.41	0.05	-1.22	0.40	0.06	-1.01	0.33	0.06	-1.31	0.30	0.07	-1.47	0.47	0.04	-1.10	0.50	0.01	-0.90
11643	0.39	0.05	-0.97	0.39	0.07	-1.02	0.36	0.06	-1.00	0.31	0.07	-1.09	0.43	0.04	-0.93	0.62	0.02	-0.58
11644	0.45	0.05	-0.95	0.38	0.06	-1.15	0.37	0.05	-1.01	0.30	0.07	-1.81	0.41	0.04	-0.91	0.48	0.02	-0.70
11645	0.38	0.05	-0.91	0.38	0.07	-1.07	0.31	0.06	-1.00	0.29	0.07	-1.12	0.41	0.04	-0.90	0.45	0.02	-0.96
11646	0.39	0.05	-0.82	0.39	0.07	-1.07	0.32	0.06	-1.09	0.30	0.08	-1.11	0.41	0.03	-0.89	0.46	0.01	-0.77
11647	0.33	0.06	-0.85	0.33	0.07	-1.02	0.30	0.06	-1.02	0.28	0.07	-1.15	0.37	0.04	-0.88	0.39	0.01	-0.76
11648	0.34	0.06	-0.65	0.34	0.07	-1.02	0.31	0.06	-0.97	0.29	0.07	-1.07	0.33	0.03	-0.69	0.47	0.01	-0.40
11649	0.40	0.06	-0.63	0.34	0.07	-1.01	0.31	0.05	-0.91	0.28	0.07	-1.17	0.42	0.03	-0.60	0.66	0.01	-0.40
11650	0.39	0.05	-1.12	0.39	0.05	-1.02	0.31	0.06	-1.13	0.29	0.07	-1.57	0.41	0.04	-1.00	0.53	0.02	-0.82
11651	0.39	0.06	-0.92	0.39	0.05	-1.15	0.31	0.06	-1.07	0.29	0.07	-1.23	0.41	0.04	-0.91	0.53	0.01	-0.53
11652	0.38	0.05	-0.90	0.38	0.06	-1.13	0.33	0.06	-1.18	0.30	0.07	-1.37	0.45	0.04	-0.87	0.68	0.01	-0.61
11653	0.33	0.06	-0.85	0.33	0.07	-1.17	0.29	0.06	-0.97	0.22	0.07	-1.49	0.37	0.04	-0.81	0.41	0.01	-0.90
11654	0.38	0.05	-1.80	0.33	0.07	-1.11	0.30	0.06	-1.96	0.27	0.08	-2.01	0.41	0.03	-1.83	0.68	0.01	-1.52
11655	0.29	0.06	-1.01	0.29	0.07	-1.13	0.30	0.05	-1.01	0.23	0.08	-1.81	0.47	0.04	-0.95	0.54	0.02	-0.42
11656	0.39	0.05	-0.92	0.39	0.06	-1.13	0.31	0.05	-0.98	0.27	0.08	-1.13	0.43	0.03	-0.67	0.41	0.01	-0.13
11657	0.39	0.06	-0.88	0.38	0.05	-1.19	0.32	0.06	-0.88	0.29	0.07	-1.11	0.40	0.03	-0.58	0.46	0.01	-0.22
11658	0.33	0.06	-0.85	0.33	0.05	-1.02	0.29	0.06	-0.83	0.24	0.07	-1.02	0.47	0.04	-0.69	0.62	0.02	-0.52
11659	0.30	0.06	-0.45	0.30	0.05	-1.05	0.27	0.06	-0.51	0.21	0.07	-0.86	0.41	0.04	-0.28	0.41	0.02	-0.05
11660	0.39	0.05	-1.91	0.38	0.07	-1.07	0.31	0.06	-1.93	0.33	0.07	-1.96	0.45	0.04	-1.86	0.50	0.01	-1.55
11661	0.34	0.05	-1.67	0.31	0.07	-1.09	0.30	0.05	-1.59	0.30	0.07	-1.77	0.41	0.03	-1.51	0.43	0.01	-1.00

References

- Aeffner, F., Zarella, M.D., Buchbinder, N., Bui, M.M., Goodman, M.R., Hartman, D.J., Lujan, G.M., Molani, M.A., Parwani, A.V., Lillard, K., Turner, O.C., 2019. Introduction to digital image analysis in whole-slide imaging: a white paper from the digital pathology association. *J. Pathol. Inform.* 10.
- Al Janabi, S., Huisman, A., Van Diest, P.J., 2011. Digital pathology: current status and future perspectives. *Histopathology* 61 (1), 1–9.
- Altman, D.G., Bland, J.M., 2011. How to obtain the P value from a confidence interval. *Bmj* 343, d2304.
- Altsubaie, N., Trahearn, N., Raza, S.E.A., Snead, D., Rajpoot, N.M., 2017. Stain deconvolution using statistical analysis of multi-resolution stain colour representation. *PLoS One* 12 (1), e0169875.
- Arevalo, J., Cruz-Roa, A., Fabio, 2014. Histopathology image representation for automatic analysis: a state of the art review. *Rev. Med* 22 (2), 79–91.
- Azevedo, T.A., Faria, P.R., Neves, L.A., Nascimento, M.Z., 2019. Computational normalization of H&E-stained histological images: progress, challenges and future potential. *Artif. Intell. Med.* 95, 118–132. PMID: 30420242.
- Basavanahally, A., Madabhushi, A., 2013. EM-based segmentation-driven color standardization of digitized histopathology. *Proc. SPIE Med. Imag.* 86, 760G-86 760G.
- Bejnordi, B.E., Timofeeva, N., Otte-Höller, I., Karssemeijer, N., van der Laak, J.A., 2014. Quantitative analysis of stain variability in histology slides and an algorithm for standardization. *Proc. SPIE Med. Imag.* 108-904 (108), 904.
- Bejnordi, B.E., Litjens, G., Timofeeva, N., Otte-Höller, I., Homeyer, A., Karssemeijer, N., van der Laak, J., 2016. Stain specific standardization of whole-slide histopathological images. *IEEE Trans. Med. Imaging* 35 (2), 404–415.
- Bug, D., Schneider, S., Grote, A., Oswald, E., Feuerhake, F., Schuler, J., Merhof, D., 2017. Context-based Normalization of Histological Stains Using Deep Convolutional Features. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 135–142.
- Fu, X., Zeng, D., Huang, Y., Zhang, X., Ding, X., 2016. A weighted variational model for simultaneous reflectance and illumination estimation. *CVPR* 2782–2790.
- Ghaznavi, F., Evans, A., Madabhushi, A., Feldman, M., 2013. Digital imaging in pathology: whole-slide imaging and beyond. *Annu. Rev. Pathol. Biol.* 8, 331–359.
- Goode, A., Gilbert, B., Harkes, J., Jukic, D., Satyanarayanan, M., 2013a. OpenSlide: a vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* 4, 27.
- Goode, A., Gilbert, B., Harkes, J., Jukic, D., Satyanarayanan, M., 2013b. OpenSlide: a vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* 4, 27.
- Gray, A., Wright, A., Jackson, O., Hale, M., Treanor, D., 2015. Quantification of histochemical stains using whole slide imaging: development of a method and demonstration its usefulness in laboratory quality control. *J. Clin. Pathol.* 68 (3), 192.
- Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009a. Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* 2, 147–171.
- Gurcan, M.N., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., et al., 2009b. Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* 2, 147–171.
- Hoque, M.Z., 2019. Stained Tissue Sample Image Analysis for Digital Pathology. e-thesis Workstations. University of Oulu. <http://urn.fi/URN:NBN:fi:oulu-201904131492>.
- Hunter, R.S., 1948. Accuracy precision and stability of new photoelectric color-difference meter. *J. Opt. Soc. Am.* 38 (12), pp. 1094–1094.
- Janowczyk, A., Basavanahally, A., Madabhushi, A., 2017. Stain normalization usingsparse autoencoders: application to digital pathology. *Comput. Med. Imaging Graph.* 57, 50–61.
- Joze, H., Drew, M., Finlayson, G., Rey, P., 2012. The role of bright pixels in illumination estimation. In: Proceedings of European Conference on Colour in Graphics. Imaging and Vision.
- Khan, A., Rajpoot, N., Treanor, D., Magee, D., 2014. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. *IEEE Trans. Biomed. Eng.* 61 (6), 1729–1738.
- Kolaman, A., Yadid-Pecht, O., 2011. Quaternion structural similarity: a new quality index for color images. *IEEE Transactions on Image Processing. Image Process* 21 (4), 1526–1536.
- Kothari, S., Phan, J.H., Stokes, T.H., Wang, M.D., 2013. Pathology imaging informatics for quantitative analysis of whole-slide images. *J. Am. Med. Inform. Assoc.* 20 (6), 1099–1108.
- Kurc, T., Bakas, S., Ren, X., Bagari, A., Momeni, A., Huang, Y., Zhang, L., Kumar, A., Thibault, M., Qi, Q., Wang, Q., 2020. Segmentation and classification in digital pathology for glioma research: challenges and deep learning approaches. *Front. Neurosci.* 14.
- Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., Thomas, N.E., 2009. A method for normalizing histology slides for quantitative analysis. *Proc. IEEE Int. Symp. Biomed. Imag. From Nano to Macro* 1107–1110.
- Magee, D., Treanor, D., Crellin, D., Shires, M., Smith, K., Mohee, K., Quirke, P., 2009. Colour normalisation in digital histopathology images. *Proc. Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy* 100–111.
- Marek, K., Michal, Z., Marcin, S., Jozef, K., Roman, M., 2019. Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm. *J. Digit. Imaging* 33 (2020), 231–242.
- McCann, M.T., Majumdar, J., Peng, C., Castro, C.A., Kovačević, J., 2014a. Algorithm and benchmark dataset for stain separation in histology images. *IEEE International Conference on Image Processing* 3953–3957.
- McCann, M.T., Majumdar, J., Peng, C., Castro, C.A., Kovačević, J., 2014b. Algorithm and benchmark dataset for stain separation in histology images. *IEEE International Conference on Image Processing* 3953–3957.
- Nurmenniemi, S., Sinikumpu, T., Alahuhta, I., Salo, S., Sutinen, M., Santala, M., Risteli, J., Nyberg, P., Salo, T., 2009. A novel organotypic model mimics the tumor microenvironment. *Am. J. Pathol.* Sep 175 (3), 1281–1291.
- Onder, D., Zengin, S., Sarıoglu, S., 2014. A review on color normalization and color deconvolution methods in histopathology. *Appl. Immunohistochem. Mol. Morphol.* 22 (Nov-Dec 10), 713–9. PMID: 24897076.
- Reinhard, E., Adhikmin, M., Gooch, B., Shirley, P., 2001a. Color transfer between images. *IEEE Comput. Graph. Appl.* 21 (5), 34–41.
- Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P., 2001b. Color transfer between images. *IEEE Comput. Graph. Appl.* 21 (5), 34–41.
- Roy, S., Kumar Jain, A., Lal, S., Kini, J., 2018a. A study about color normalization methods for histopathology images. *Micron* 114, 42–61. PMID: 30096632.
- Roy, S., Jain, A.K., Lal, S., Kini, J., 2018b. A study about color normalization methods for histopathology images. *Micron* 114, 42–61.
- Ruifrok, A.C., Johnson, D.A., 2001. Quantification of histochemical staining by color deconvolution. *Analyt. Quant. Cytol. Histol. Int. Acac. Cytol. Am. Soc. Cytol.* 23 (4), 291–299.
- Salvi, M., Michielli, N., Molinari, F., 2020. Stain Color Adaptive Normalization (SCAN) algorithm: Separation and standardization of histological stains in digital pathology. *Computer methods and programs in biomedicine* 193, 105506.
- Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9 (7), 671–675.
- Shafiei, S., Safarpoor, A., Jamalizadeh, A., Tizhoosh, H.R., 2020. Class-agnostic weighted normalization of staining in histopathology images using a spatially constrained mixture model. *IEEE Trans. Med. Imaging*. <https://doi.org/10.1109/TMI.2020.2992108>.
- Starck, J.L., Murtagh, F.D., Bijaoui, A., 1998. *Image Processing and Data Analysis: the Multiscale Approach*. Cambridge University Press.
- Stathonikos, N., Veta, M., Huisman, A., van Diest, P.J., 2013. Going fully digital: perspective of a Dutch academic pathology lab. *J. Pathol. Inform.* 4 (1), 15.
- Thaina, A., Tosta, A., Faria, P.R., Neves, L.A., Nascimento, M.Z., 2019. Computational normalization of H&E-stained histological images: progress, challenges and future potential. *Artif. Intell. Med.* 95, 118–132.
- Vahadane, A., Peng, T., Sethi, A., Albarqouni, S., Wang, L., Baust, M., Steiger, K., Schlitter, A.M., Esposito, I., Navab, N., 2016. Structure-preserving color normalization and sparse stain separation for histological images. *IEEE Trans. Med. Imaging* 35 (8), 1962–1971.
- Wang, Z., Bovik, A.C., 2002. A universal image quality index. *IEEE Signal Process* 9 (3), 81–84.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wang, Y.Y., Chang, S.C., Wu, L.W., Tsai, S.T., Sun, Y.N., 2007. A color-based approach for automated segmentation in tumor tissue classification. *Proc. Conf. of the IEEE Engineering in Medicine and Biology Society* 6577–6580.
- Wang, S., Yang, D.M., Rong, R., Zhan, X., Xiao, G., 2019. Pathology image analysis using segmentation deep learning algorithms. *Am. J. Pathol.* 189 (9), 1686–1698.
- Wei, B.R., Simpson, R.M., 2014. Digital pathology and image analysis augment biospecimen annotation and biobank quality assurance harmonization. *Clinical Biochem.* 47 (4), 274–279.
- Zanjani, F.G., Zinger, S., Bejnordi, B.E., van der Laak, J.A., de With, P.H.N., 2018. Stain normalization of histopathology images using generative adversarial networks. *IEEE 15th International Symposium on Biomedical Imaging (ISBI)* 573–577.
- Zheng, Y., Jiang, Z., Zhang, H., Xie, F., Ma, Y., Shi, H., Zhao, Y., 2018. Histopathological whole slide image analysis using context-based cbir. *IEEE Trans. Med. Imaging* 37 (7), 1641–1652.
- Zheng, Y., Jiang, Z., Zhang, H., Xie, F., Shi, J., Chenghai, 2019. Adaptive color deconvolution for histological WSI normalization. *Comput. Methods Programs Biomed.* 170–120.