

Adaptive Color Deconvolution for Histological WSI Normalization

Yushan Zheng^{a,b,c}, Zhiguo Jiang^{a,b,c,*}, Haopeng Zhang^{a,b,c}, Fengying Xie^{a,b,c},
Jun Shi^d, Chenghai Xue^e

^a*Image Processing Center, School of Astronautics, Beihang University, Beijing, 100191, China, {yszheng, jiangzg, zhanghaopeng, xfy_73@buaa.edu.cn}*

^b*Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing, 100191, China*

^c*Beijing Key Laboratory of Digital Media, Beihang University, Beijing, 100191, China*

^d*School of Software, Hefei University of Technology, Hefei 230601, China*

^e*Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China*

Abstract

Background and objective: Color consistency of histological images is significant for developing reliable computer-aided diagnosis (CAD) systems. However, the color appearance of digital histological images varies across different specimen preparations, staining, and scanning situations. This variability affects the diagnosis and decreases the accuracy of CAD approaches. It is important and challenging to develop effective color normalization methods for digital histological images.

Methods: We proposed a novel adaptive color deconvolution (ACD) algorithm for stain separation and color normalization of hematoxylin-eosin-stained whole slide images (WSIs). To avoid artifacts and reduce the failure rate of normalization, the multiple prior knowledge of staining is considered and embedded in the ACD model. To improve the capacity of color normalization for various WSIs, an integrated optimization is designed to simultaneously estimate the parameters of the stain separation and normalization. The solving of ACD model and application of the proposed method involves only pixel-wise operation, which makes it very efficient and applicable to WSIs.

*Corresponding author: Zhiguo Jiang, Tel.:+86 10 82316173.

Results: The proposed method was evaluated on four WSI-datasets including breast, lung and cervix cancers and was compared with 6 state-of-the-art methods. The proposed method achieved the most consistent performance in color normalization according to the quantitative metrics. Through a qualitative assessment for 500 WSIs, the failure rate of normalization was 0.4 % and the structure and color artifacts were effectively avoided. Applied to CAD methods, the area under receiver operating characteristic curve for cancer image classification was improved from 0.842 to 0.914. The average time of solving the ACD model is 2.97 s.

Conclusions: The proposed ACD model has proven effective for color normalization of hematoxylin-eosin-stained WSIs in various color appearances. The model is robust and can be applied to WSIs containing different lesions. The proposed model can be efficiently solved and is effective to improve the performance of cancer image recognition, which is adequate for developing automatic CAD programs and systems based on WSIs.

Keywords: Color normalization, digital pathology, stain separation, WSI, CAD

1. Introduction

Cancer diagnosis still relies on histopathology [1], which involves the microscopic examination of tissue to study the manifestations of the disease. Histological slides are stained with multiple dyes to color different types of tissues [2]. With the development of digital pathology, histological slides can be scanned rapidly using advanced micro-scanners and stored as digital whole slide images (WSIs). This enables pathologists to view slides on a screen. Based on digital WSIs, an increasing number of computer-aided-diagnosis (CAD) approaches have emerged in the last two decades [3, 4, 5, 6]. A competent WSI-based CAD system can help pathologists locate diagnostically relevant regions from the WSIs [7, 8], which improves the efficiency and reliability of diagnoses.

Color consistency of digital WSIs is quite important for CAD based on WSI

analysis [9, 10]. In practical applications, the appearance of WSIs varies due to different specimen preparations, staining situations, and section scanners [11]. This variability affects the diagnosis and decreases the accuracy of CAD approaches. To overcome the variability, many color normalization methods for histological images have been proposed. [12]

A group of methods realized the normalization via histogram transformation in different color space [13, 8, 14]. These methods were designed with reference to the normalization of natural scene images and ignored the prior knowledge that histological images are colored by multiple stains. The performance of normalization is limited.

In contrast, more methods are based on the prior knowledge that the color of WSIs is the combination of several independent stains. A set of these methods [15, 16, 11] proposed normalizing histological images by separate transformations, where pixels belonging to different stains were classified and then normalized through stain-specific transformations. The color of histological images can be effectively transformed to a template image by these methods. However, the normalization via multiple transformations occasionally causes structural artifacts in the normalized images.

In comparison, a number of methods proposed establishing a unified transformation for all the pixels [17, 18, 19, 20, 21]. Instead of directly classifying pixels for different stains, the independent stain components were estimated and separated from the pixel values. Next, the separated stains were aligned to that of the template image, and then recombined to achieve the normalization. The color of the normalized images is relatively smoother than that obtained by separate-transformation-based methods, and thus the structure can be effectively preserved [19]. While, the constraints for stain proportion and overall intensity are relatively weak compared to the methods based on separated transformation, for which color artifact occasionally appears in the normalized images.

In this study, we propose a novel adaptive color deconvolution (ACD) model for color normalization of hematoxylin-eosin-stained (H&E-stained) WSIs. The

normalization is achieved through a unified color transformation for pixels from the source image to the template image. Compared to the separate-transformation-based methods [16, 11], our approach does not rely on the classification of stains. The parameters for color normalization are estimated from the distribution of pixel values through an integrated optimization. Therefore, the structural information of histological images can be well preserved. Different from current methods based on unified transformation [19, 22, 20], the parameters for stain separation and normalization are simultaneously estimated in the integrated optimization, for which the color consistency of the normalization is improved. Moreover, the proportion of stains and the overall staining intensity are considered and embedded in the ACD model, which effectively reduces the failure rate of stain separation and thus delivers a more robust color normalization. Compared to the transformation-based methods of [11, 20, 23], both the solution and application stages of the proposed method only contain pixel-wise operation and involve no pixel interaction. It determines the proposed method is light, fast, and applicable for developing efficient CAD programs and systems based on WSIs.

The proposed method was evaluated on aspects of color normalization, stain separation, computational complexity and effectiveness for CAD approaches on four histopathological image datasets, and was compared with the state-of-the-art methods [8, 14, 16, 11, 22, 19]. The experimental results have shown that the proposed method is effective for H&E-stained histological image normalization and the overall performance is superior to the compared methods in quantitative and qualitative assessments.

The remainder of this paper is organized as follows. Section 2 reviews the related studies. Section 3 introduces the methodology of the proposed method. The experiment is presented in Section 4. Section 5 provides necessary discussions and suggests directions for future work. Finally, Section 6 summarizes the contributions.

2. Related work

Most of the works related to our method are reviewed in this section. The histogram-transformation-based methods are first introduced. Then, the methods based on separate transformation and unified transformation are reviewed. A briefly comparison for different category of methods are summarized in Table 1.

2.1. Histogram-transformation-based methods

Wang et al. [13] introduced a linear color transform method [24] into histological images. The normalization was achieved by transforming the color to a template image by a linear projection in $l\alpha\beta$ color space. Zheng et al. [8] proposed normalizing WSIs in HSV-space, in which the saturation and value channels were stretched. The method can normalize global illumination and saturation of WSIs and has proven effective in improving the performance of WSI analysis. Janowczyk et al. [14] realized stain normalization using sparse auto-encoders. Pixels in histological images were clustered based on the features generated by the sparse auto-encoders and the pixels belonging to the same cluster were transformed to the template image using a specific histogram projection.

These methods derive from nature scene image processing and barely utilize the staining characteristic of histological images. As a result, tissue area and background are occasionally confused after the normalization [15].

2.2. Separate-transformation-based methods

Based on the prior knowledge that histological images are colored by independent stains, the methods based on multiple transformations were proposed [15, 16, 7]. Typically, Khan et al. [16] proposed a specific color deconvolution (SCD) algorithm. In this method, pixels belonging to the same stain were extracted through a classification model and a staining vector was estimated by analyzing the distribution of the pixels belonging to the stain. Based on the

staining vector, a non-linear mapping approach was designed to transfer the image color to the template image. For more stable performance, the prior knowledge of nuclei structure was considered for the pixel classification process. In [11], the nuclei were detected by using the Hough transform [25]. Then, the pixels belonging to the type of nuclei, cytoplasm and background were accurately classified. Therefore, the color of different types of pixels can be sufficiently transformed to the template image. While, since the pixels in the images are normalized through different transformations, it probably brings color discontinuity into normalized images and causes improper structural changes on the border of different stains. Moreover, the Hough transform method requires additional computation for the normalization, which may become the bottleneck of an efficient CAD approach.

2.3. Unified-transformation-based methods

Compared to the aforementioned methods, An increasing number of methods proposed to establish a unified transformation model for pixels in the images. Instead of classifying the pixels as different stains, these methods proposed extracting independent stain components from pixels and then established a unified transformation for all the pixels based on the separated stain information.

Color deconvolution (CD) [26] has proven effective in stain separation for histological images [27, 28, 29, 30, 31, 32]. Based on CD, the color normalization methods were emerged [33, 18], in which transformations between the source image and the template image were established through CD and its inverse operation. These methods commonly deliver a more reasonable visual performance and help preserve the local contrast in histological images. Specifically, Zhou et al. [22] proposed modifying the CD matrix through an optimization for H&E-stained histological images. The variable of the optimization is the exact CD matrix. The objective function enforces the third channel (i.e., the background channel) of the deconvolution result to be zero. However, the specificity of H&E-staining is not considered in the optimization, for which the normalization performance is limited. Referring to CD [26], the items in the CD matrix

Table 1: A summary of the consideration and ability of models in different categories of color normalization methods for histological images.

Category	Staining Specificity	Structure Preservation	Color Consistency
Histogram transformation	No	No	Yes
Separated transformation	Yes	Weak	Yes
Unified transformation	Yes	Yes	Weak

are not independent variables but decided by a stain color appearance (SCA) matrix. Therefore, the SCA matrix is more reasonable than CD matrix to be the variables of normalization models. Li et al. [17, 34] regarded the SCA matrix as the model variables and applied non-negative matrix factorization (NMF) to solving the corresponding model. Furthermore, sparse constraint was considered in NMF model for the assumption that most of the pixels contain one type of stain [35, 19, 20]. Typically, Vahadane [19] designed a normalization approach for WSIs based on sparse NMF (SNMF). The sparse constraint enhanced the recognition ability of the model for separating independent stains. In these methods, the weighting of stains is processed independently after stain separation, for which the capacity of color normalization is limited. Furthermore, the overall intensity and proportion of the separated stains are not considered in the model of stain separation. It risks a bias in stain separation, for which most of the pixels would be normalized to share the appearance of single stain.

Building on these methods, we propose an adaptive color deconvolution (ACD) model for stain separation and color normalization of histological images. Taking the SCA matrix as the variables, the ACD model is solved through an integrated optimization. The normalization is achieved by a unified transformation for the pixels in the image. The proposed method inherits the advantage of the unified-transformation-based methods [19, 20]. The structural information of histological images is well preserved.

The contribution of our work to the problem is two-fold:

- Besides the consideration for the specificity of H&E staining, the overall intensity and proportion of the stains are additionally considered and em-

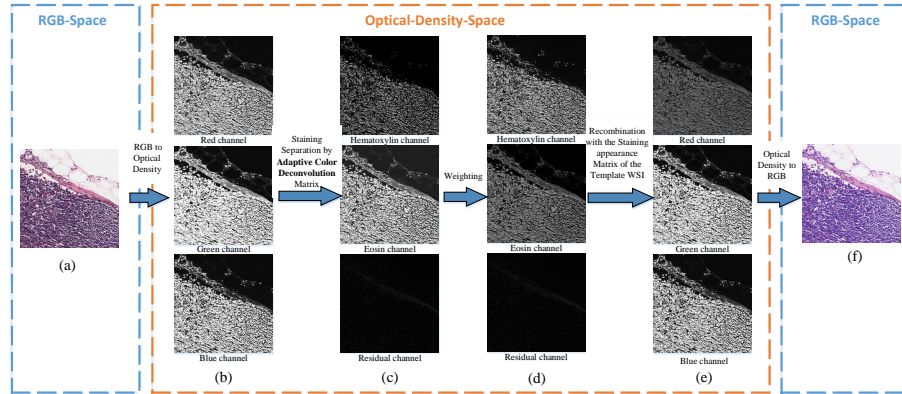


Figure 1: Flowchart of the proposed normalization method, where (a) denotes the original WSI (A featured sub-region of the WSI is chosen to display.), (b) is the visualization of R/G/B channel in the optical-density-space, (c) shows the density of hematoxylin and the eosin stains separated by an adaptive color deconvolution matrix, (d) displays the weighted stains, (e) shows the R/G/B channel recombined with the stain parameters of the template WSI, and (f) shows the result of the normalization.

bedded in the ACD model, which effectively reduces the failure of stain separation and then avoids color artifacts in the normalized images.

- The parameters for stain intensity normalization are simultaneously estimated with the parameters for stain separation based all the pixels involved in an integrated optimization. Compared to the present models of [19, 20] where the parameters for stain normalization are calculated based on the pseudo maximum of the separated stains, the capacity of our model to handle different color variance is extended and thereby the color consistency of the normalized images is improved.

3. Method

Fig. 1 presents the flowchart of the proposed normalization. For a certain WSI, a group of pixels are sampled from the tissue region and converted into optical density (OD) space. The normalized H&E components are obtained based on an ACD matrix and a stain-weight matrix for the WSI. Finally, the

H and E components are recombined with the SCA matrix of a template WSI, achieving the color normalization. The approach to obtain the ACD matrix and the stain-weight matrix are the essential of our method. In this section, the ACD model is first introduced and then the normalization method based on ACD is described.

3.1. Color deconvolution

The theory of color deconvolution (CD) [26] is the basis of ACD. CD is proposed based on Beer-Lambert law. Letting $\mathbf{x}_i \in \mathbb{R}^{3 \times 1}$ denote the value in RGB color space for the i -th pixel in a WSI, CD can be briefly represented with the following equations

$$\begin{aligned}\mathbf{o}_i &= -\ln(\mathbf{x}_i/I_{max}) \\ \mathbf{s}_i &= \mathbf{D} \cdot \mathbf{o}_i\end{aligned}\tag{1}$$

where $\mathbf{o}_i \in \mathbb{R}^{3 \times 1}$ denotes the optical density (OD) of RGB channels, $\mathbf{D} \in \mathbb{R}^{3 \times 3}$ is the so-called color deconvolution matrix, and $\mathbf{s}_i \in \mathbb{R}^{3 \times 1}$ is the output that contains stain densities. I_{max} denotes the intensity of background, i.e. the value of pixel when no stained tissue is present. The exact value of I_{max} varies with different section scanners. Generally, I_{max} approximates the maximum of digital image intensity (255 for 8-bit data format). For H&E-stained WSIs, the separated densities of stains can be represented as $\mathbf{s}_i = (h_i, e_i, d_i)^T$, where h_i and e_i are the values for hematoxylin and eosin stains, respectively, and d_i represents the residual of the separation. The deconvolution matrix \mathbf{D} is determined by a SCA matrix \mathbf{M} with an inverse operation $\mathbf{D} = \mathbf{M}^{-1}$. Further, \mathbf{M} can be manually measured using a designed experiment [26].

3.2. Adaptive color deconvolution

The ACD parameters are obtained by optimization. The variables, objective and solving of the optimization are presented in the following sub-sections.

3.2.1. Variables

Considering that the deconvolution matrix \mathbf{D} is determined by the SCA matrix \mathbf{M} , we propose directly optimizing \mathbf{M} and then calculating \mathbf{D} . Mean-

while, the parameters for normalizing stain intensities are also solved in the optimization. Specifically, a stain-weight matrix $\mathbf{W} = \text{diag}(w_h, w_e, 1)$ is defined to modify the CD algorithm (Eq. 1) as

$$\begin{aligned}\mathbf{o}_i &= -\ln(\mathbf{x}_i/I_{max}) \\ \mathbf{s}_i &= \mathbf{W} \cdot \mathbf{D} \cdot \mathbf{o}_i.\end{aligned}\tag{2}$$

\mathbf{W} is also regarded as the variable of ACD model and obtained in the optimization.

The SCA matrix can be decomposed as $\mathbf{M} = (\mathbf{m}_h, \mathbf{m}_e, \mathbf{m}_d)$, where $\mathbf{m}_j \in \mathbb{R}^{3 \times 1} (j = h, e, d)$. In general, \mathbf{m}_j is a unit vector [19, 20], which describes the contributions of the j -th stain to the intensities in red, green, and blue channels. To ensure $\mathbf{m}_j \equiv 1$ throughout the optimization, we propose representing \mathbf{m}_j using two degree variables as

$$\mathbf{m}_j = (\cos \alpha_j \sin \beta_j, \cos \alpha_j \cos \beta_j, \sin \alpha_j)^T, j = h, e, d.$$

Then, the SCA matrix \mathbf{M} can be represented by six independent degree variables. For convenience, the six degree variables are represented by a collection

$$\varphi = \{\alpha_h, \beta_h, \alpha_e, \beta_e, \alpha_d, \beta_d\},$$

the SCA matrix decided by φ is represented as $\mathbf{M}(\varphi)$, and the corresponding CD matrix is $\mathbf{D}(\varphi)$.

3.2.2. Objective

An objective function about variables φ and \mathbf{W} is defined. By resolving the function, the optimized set of variables $\hat{\varphi}$ and $\hat{\mathbf{W}}$ are obtained, and then the adaptive matrices $\mathbf{M}(\hat{\varphi})$ and $\mathbf{D}(\hat{\varphi})$ for the WSI are determined. For brevity, $\mathbf{M}(\hat{\varphi})$ and $\mathbf{D}(\hat{\varphi})$ are also represented as $\hat{\mathbf{M}}$ and $\hat{\mathbf{D}}$ in this paper.

The objective function for ACD is designed primarily on the basis of the following prior knowledge: (1) There are two types of stains in H&E-stained WSIs. Therefore, the third channel of the separated result (d_i) should be zero in ideal situation. (2) H&E staining has high specificity. Hematoxylin mainly

stains nuclei and eosin mainly stains the cytoplasm and stroma. Therefore, the majority of pixels in images alternatively contain H or E stain. Based on the prior knowledge, the objective function is defined as

$$L_p = \frac{1}{N} \sum_{i=1}^N d_i^2 + \lambda_p \frac{1}{N} \sum_{i=1}^N \frac{2h_i e_i}{h_i^2 + e_i^2}, \quad (3)$$

where the first item of the function minimizes the residual of the separation, the second item enforces the value of a pixel being assigned to one stain (H or E) after the separation, λ_p is the weight of the two items, and N is the number of pixels used for the optimization.

Besides the features considered above, the proportion of the two stains and the overall intensity of staining are equally important for the normalization. Therefore, the two factors are embedded in the objective function. First, a function to control the ratio of H and E components is defined:

$$L_b = \left[(1 - \eta) \frac{1}{N} \sum_{i=1}^N h_i - \eta \frac{1}{N} \sum_{i=1}^N e_i \right]^2, \quad (4)$$

where $\eta \in (0, 1)$ is defined as the balance parameter. Similarly, a function to control the overall energy of stains is defined:

$$L_e = \left[\gamma - \frac{1}{N} \sum_{i=1}^N (h_i + e_i) \right]^2, \quad (5)$$

where γ controls the desired intensity of staining.

Finally, the objective function is modified as

$$L = L_p + \lambda_b L_b + \lambda_e L_e, \quad (6)$$

where λ_b and λ_e are the weights.

3.2.3. Solution

The objective is a function of variables φ and \mathbf{W} , and thus the optimization is described as

$$(\hat{\varphi}, \hat{\mathbf{W}}) = \arg \min_{(\varphi, \mathbf{W})} L(\varphi, \mathbf{W})$$

$L(\varphi, \mathbf{W})$ is continuous and differentiable for variables φ and \mathbf{W} . Therefore, we utilized a gradient descent algorithm to solve it. The derivatives of the objective function on the variables of the model are given in the appendix.

In the optimization, only the pixels located on the tissue area are used. In WSIs, the regions that are devoid of stain are approximately white, and the optical densities of the region pixels are close to zero. Therefore, the background pixels can be easily filtered by a threshold [15, 11, 19]. Specifically, the pixels within $\mathbf{o}_i < T_{back}$ are recognized as background. T_{back} was tuned in the interval of [0.2, 0.5] and determined as 0.28 for the most robust normalization performance in the statistical assessment. Then, a binary tissue mask for the WSI can be obtained. The pixels used in the optimization are randomly sampled from the WSI based on the tissue mask.

3.3. Color normalization

After the optimization, the adaptive variables for stain separation $\hat{\mathbf{D}}$ and stain intensity normalization $\hat{\mathbf{W}}$ are simultaneously obtained. With $\hat{\mathbf{D}}$, the stain components of a WSI can be separated. Next, the separated stains are weighted by $\hat{\mathbf{W}}$. Finally, the normalization is completed by recombining the weighted stain components with the SCA matrix of a template WSI $\overline{\mathbf{M}}$. Specifically, for the i -th pixel \mathbf{x}_i of the WSI, the normalization can be formulated by equations

$$\begin{aligned}\mathbf{o}_i &= -\ln(\mathbf{x}_i/I_{max}), \\ \bar{\mathbf{o}}_i &= \overline{\mathbf{M}} \cdot \hat{\mathbf{W}}\hat{\mathbf{D}} \cdot \mathbf{o}_i, \\ \bar{\mathbf{x}}_i &= \exp(-\bar{\mathbf{o}}_i) \cdot I_{max},\end{aligned}\tag{7}$$

where $\bar{\mathbf{x}}_i$ is the normalized result for \mathbf{x}_i . Because, $\overline{\mathbf{M}}$, $\hat{\mathbf{W}}$ and $\hat{\mathbf{D}}$ are constant after the solving of ACD, the three matrices can be combined as a transform matrix $\mathbf{T} = \overline{\mathbf{M}}\hat{\mathbf{W}}\hat{\mathbf{D}}$. Then, the normalization can be efficiently achieved through linear transformation of pixel values in optical density space, which is formulated as

$$\bar{\mathbf{o}}_i = \mathbf{T} \cdot \mathbf{o}_i.\tag{8}$$

4. Experiments and results

4.1. Setup

Four dataset, Camelyon-16, Camelyon-17, Motic-cervix, and Motic-lung, were used in the experiments. Camelyon-16 and Camelyon-17 were obtained from the Camelyon challenge¹ for cancer metastasis detection in the lymph node [36, 37]. Motic-cervix and Motic-lung were supplied by Motic (Xiamen) Medical Diagnostic Systems Co. Ltd. The profiles are provided as follows.

- *Camelyon-16* contains 400 H&E-stained lymph node WSIs, in which 270 WSIs are used for training and the remainder are used for testing. Regions with cancer in these WSIs are annotated by pathologists. All the annotations for Camelyon-16 are available.
- *Camelyon-17* contains 1000 WSIs from 5 medical centers, in which 500 WSIs are used for training and the remainder are used for testing. The annotations of testing WSIs are not yet available.
- *Motic-cervix* contains 47 WSIs from 47 patient with cervical cancer (including adenocarcinoma and squamous cell carcinoma), in which regions with cancer are annotated by pathologists.
- *Motic-lung* contains 39 WSIs from 39 patient with lung cancer (including adenocarcinoma and squamous cell carcinoma, large cell carcinoma and small cell carcinoma), in which regions with cancer are annotated by pathologists.

The quantitative and qualitative assessments were processed on the Camelyon-17 dataset, since it consists of WSIs from 5 medical centers and contains rich color variations. The Camelyon-16 dataset is used to evaluate the normalization performance for the CAD method, because the labels for both the training and testing set are available. The experiments were also conducted on Motic-cervix

¹<https://camelyon17.grand-challenge.org/>

and Motic-lung datasets to verify the applicable ability of our method to other lesions.

The normalized median intensity (NMI) measure [38] is used to quantitatively assess the consistency of normalization, NMI is defined as

$$NMI(I) = \frac{Med(u_i)}{P_{95}(u_i)}, \quad (9)$$

where I denotes a WSI, u_i denotes the mean value of R, G and B channels of the i -th pixel in the WSI. $Med()$ denotes the median value, and $P_{95}()$ denotes the 95th percentile [11]. The standard deviation of the NMI values (NMI SD) and coefficient of the variation (i.e., standard deviation divided by mean) of the NMI values (NMI CV) for all WSIs were calculated and used as the metrics. The lower the values of NMI SD and NMI CV, the more consistent the normalization. To avoid the influence of extensive background regions in WSIs and limit the amount of computation, sub-images were sampled from the tissue regions of a WSI to substitute the WSI and the NMI for the WSI was calculated based on all the pixels in the sub-images. Specifically, sub-images within 2048×2048 pixels under the $40\times$ lenses were sampled and the percentage of tissue pixels in each sub-image was controlled at more than 70% (according to the tissue mask defined in section 3.2.3). The number of sub-images used for each WSI was evaluated from 5 to 35 and found 20 was sufficient to deliver a stable assessment. Therefore, 20 sub-images within 2048×2048 pixels were used to calculate the NMI for each WSI. Notice that some WSIs in Camleyon17 dataset contain pure black background. These regions were filtered beforehand and were not considered in the optimization and assessment.

The ACD model was solved using gradient descent algorithm. The optimizer was selected from SGD, AdaGrad, AdaDelta and Adam and was determined as AdaGrad [39] as it achieved the lowest NMI SD. The variable φ is initialized based on the SCA matrix suggested in [26] and, \mathbf{W} was initialized as a unit matrix.

The algorithm was implemented in python with tensorflow [40] and was processed on a computer with an Intel Core i7-7700k CPU of 4.2 GHz and a

RAM of 32GB. All the experiments were conducted on the same computer.

In this section, the ACD model are first validated on the training set of Camleyon-17. Then, the normalization performance of the proposed method is evaluated and compared with the state-of-the-art methods within the test set of Camelyon-17.

4.2. Validation of ACD model

The structure of ACD model and the settings in the model solving are validated in this section. The experiments were conducted on the training set of Camelyon-17 dataset. The NMI SD for the normalized WSIs is used as the metric.

4.2.1. Hyper-parameters

There are five hyper-parameters $\lambda_p, \lambda_b, \lambda_e, \gamma, \eta$ involved in our model. These parameters were adjusted in large ranges. The curve of NMI SDs for different setting of the hyper-parameters are presented in Fig. 2. A low NMI SD indicates a good normalization performance. Note that the other hyper-parameters were fixed when tuning a specific hyper-parameter.

$\lambda_p, \lambda_b,$ and λ_e are weights of different items involved in the cost function. The results for different settings of the three parameters are shown in Fig. 2 (a-c). The three hyper-parameters were selected for relatively low NMI SDs. Specifically, the three parameters were set as $\lambda_p = 0.002, \lambda_b = 10,$ and $\lambda_e = 1$ in the following experiments.

In the ACD model, η constrains the ratio of the staining components, and γ constrains the staining density, respectively. The setting of the two parameters influences the visual performance. The normalization performance for the different settings of the two parameters are visualized in Fig. 4. According to the statistical metrics (Fig. 2(d-e)), we suggested $\gamma \in [0.25, 0.4]$ and $\eta \in [0.55, 0.7]$ in application for a consistent normalization performance. In the following experiments, γ is set to 0.3 and η is set to 0.6 for relatively low NMI SDs.

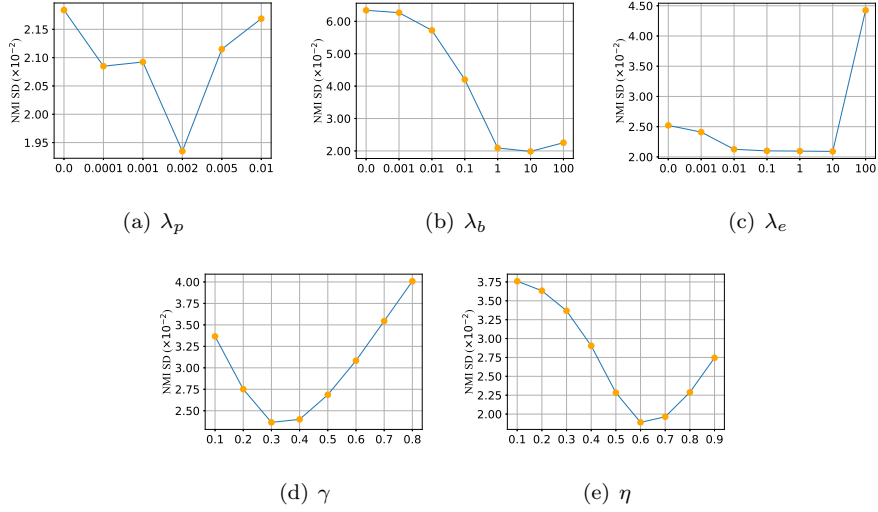


Figure 2: NMI SDs for hyper-parameters of ACD model.

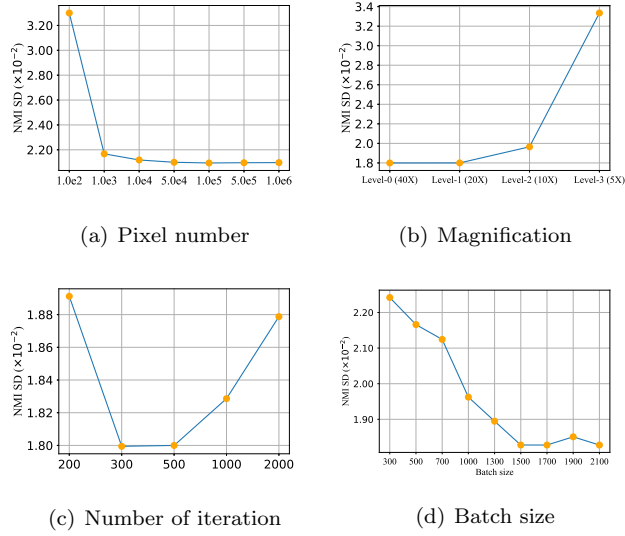


Figure 3: NMI SDs for different training settings of ACD model.

4.2.2. Optimization settings

The normalization performance is also influenced by the settings of the optimization, including the number of pixels, the magnification of pixel sampling, the pixel number used in each interaction (batch size), and the number of in-

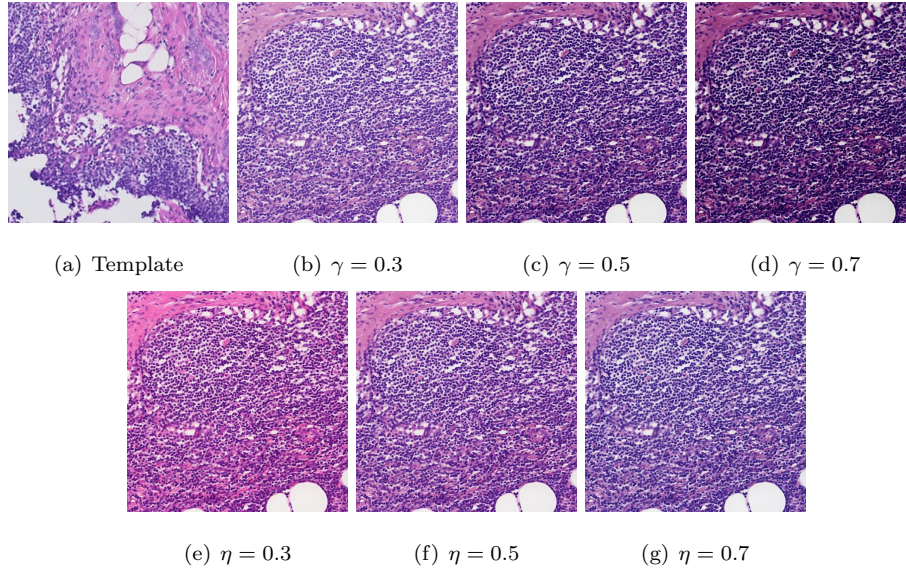


Figure 4: Visual performance of the normalized image varied with the control parameters γ and η , where (a) is a region from the template WSI, (b-d) display the results for different γ , and (e-g) present the results for different η .

interactions in the optimization. The curves of NMI SD for different settings of these factors are given in Fig. 3. The number of pixels used in the optimization is ranged from 10^2 to 10^6 . Fig. 3 (a) shows that the model trained with 1,000 pixels can achieve a desirable normalization consistency with an NMI SD of 0.0217. This indicates that the proposed model does not rely on massive number of pixels. The metric improves from 0.0217 to 0.0210 as the training pixels are increased from 1,000 to 100,000, and then changes little when further increasing the training pixels. Hence, the number of training pixels is set 100,000 for consistent normalization performance. The magnification of pixel sampling is also important in the optimization. According to Fig. 3 (b), the proposed model has a certain robustness to decrease in magnification. For reasonable normalization results, the pixels used in the optimization are sampled from WSIs under $20\times$ lenses. It can be seen from Fig. 3 (c-d) that the normalization is stable when the training step is set between 300 and 500 and the batch size is above 1500. To limit the calculation amount of optimization, the training step

Table 2: Results of the ablation experiments for the proposed ACD model

Model	NMI SD	NMI CV
ACD w/o \mathbf{W}	0.022	0.028
ACD w/ $\lambda_p = 0$	0.021	0.027
ACD w/ $\lambda_b = 0$	0.063	0.101
ACD w/ $\lambda_e = 0$	0.025	0.034
ACD	0.018	0.023

is set to 300 and the batch size is set to 1500 in the following experiments.

4.2.3. Ablation experiments

In the proposed method, the parameters for stain weighting are simultaneously optimized with the parameters of stain separation in our model. To verify the necessity of the simultaneous optimization, we implemented a control approach (Abbreviated as ACD w/o \mathbf{W}), where the stain-weight matrix \mathbf{W} was not the variable in the optimization but was estimated according to the pseudo maximum of separated stains [19, 20] after the optimization. Meanwhile, the optimization without the constraints of stain specificity, stain proportion and stain intensity were validated by setting the hyper-parameters λ_p , λ_b , and λ_e in Eq. 6 to be zero. The results are presented in Table 2.

When the estimation of \mathbf{W} is independent from the parameters for stain separation, the result of ACD w/o \mathbf{W} deteriorates. It has indicated that the simultaneous optimization designed in the ACD model can effectively extend the model capacity of color transformation from source images to the template image. Therefore, the consistency of the normalization is improved. Obviously, when the three hyper-parameters λ_p , λ_b , and λ_e are set to zero, the performance of normalization deteriorates. It has demonstrated that the items defined based on the prior knowledge (Eq. 3), the proportion (Eq. 4), and the intensity (Eq. 5) of stains are all necessary for a consistent normalization performance.

4.3. Comparison with the state-of-the-art

4.3.1. Methods for comparison

The color normalization methods developed from different aspects of the histological slides are compared. Specifically, two methods introduced from nature scene image processing proposed by Zheng et al. [8] and Janowczyk et al. [14], the separate-transformation-based methods proposed by Khan et al. [16] and Bejnordi et al. [11], and other two unified-transformation-based methods developed by Vahadane et al.[19] and Zhou et al.[22] are involved in the comparison. The methodologies for these approaches are introduced in the related works (Section 2).

Table 3: The comparisons of NMI SD and NMI CV for different normalization methods.

Method	NMI SD	NMI CV
Original	0.139	0.210
Zheng et al. [8]	0.077	0.117
Janowczyk et al.[14]	0.027	0.037
Khan et al. [16]	0.049	0.067
Bejnordi et al. [11]	0.028	0.045
Vahadane et al. [19]	0.042	0.062
Zhou et al. [22]	0.054	0.095
The proposed	0.025	0.034

Table 4: The comparisons of NMI SD and NMI CV for different normalization methods, where NMI_h and NMI_e represent the NMI for hematoxylin and eosin stains, respectively.

Method	NMI_h SD	NMI_h CV	NMI_e SD	NMI_e CV
Original	0.166	0.582	0.144	0.438
Zheng et al. [8]	0.092	0.321	0.081	0.242
Janowczyk et al. [14]	0.017	0.071	0.184	0.392
Khan et al. [16]	0.055	0.203	0.089	0.214
Bejnordi et al. [11]	0.042	0.117	0.028	0.070
Vahadane et al. [19]	0.043	0.109	0.036	0.103
Zhou et al. [22]	0.160	0.362	0.142	0.363
The proposed	0.029	0.067	0.027	0.087

4.3.2. Quantitative comparison

The test set of Camelyon-17 was used in this experiment. The NMI SD and NMI CV calculated based on all testing WSIs were used as metrics. The results of the proposed method were obtained under the hyper-parameters determined in the training set.

The results of the compared methods are presented on Table 3. It is shown that the proposed method achieves the best performance in NMI SD and NMI CV assessment. To intuitively present the distribution of NMI values for the normalized images, the violin plots [41] for different methods are utilized (Fig. 5). The NMI values of the proposed method are the most clustered. It indicates that the normalization of the proposed method is most consistent.

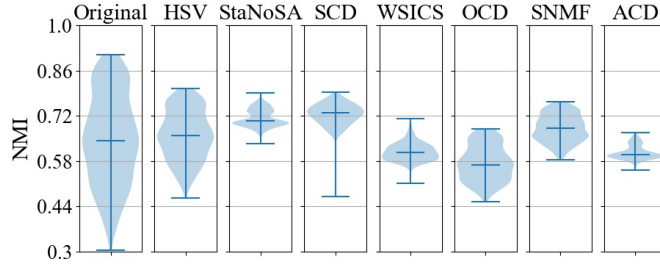


Figure 5: Violin plots of NMIs for the compared methods, where the blue shadow presents the allocation of NMIs for each plot, and the maximum, median, and minimum values for each plot are signed with bars.

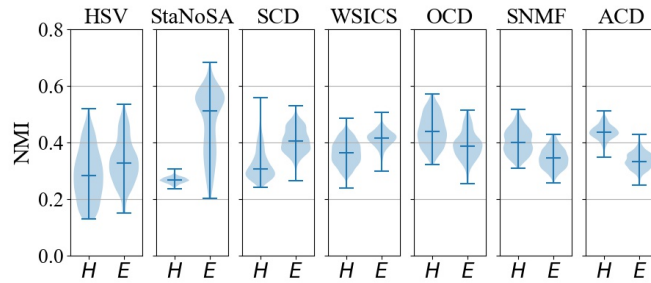


Figure 6: Violin plots of NMIs for independent stains, where H represents the hematoxylin stain, E represents the eosin stain, the blue shadow presents the allocation of NMIs for each plot, and the maximum, median, and minimum values for each plot are signed with bars

The stability for staining separation of the compared methods was also evaluated. The normalization results were separated using CD with the parameters of the template WSIs. The NMI metrics for independent staining components are presented in Table 4. Correspondingly, violin plots of NMIs for independent stains are given in Fig. 6. Janowczyk et al. [14] achieves the best NMI SD in the hematoxylin stain component, but the metrics for eosin are inferior to other methods. ACD and Bejnordi et al. [11] achieve an equally good quantitative performance for eosin component. While, the performance of ACD for hematoxylin is better than Bejnordi et al. [11]. Overall, ACD is the most consistent for stain separation among all the compared methods.

4.3.3. Qualitative comparison

To evaluate the quality of normalization, three pathologists were invited to inspect the normalized results and the WSIs containing apparent artifacts in normalization were annotated. Specifically, a WSI was considered as failure once any of the 20 sub-images used in the quantitative assessment contained structure or color artifacts. The failure rate is calculated by a division of the number of failures WSI to the number of test WSIs (500) in Camelyon-17. The results of the assessment are given in Fig. 7. Correspondingly, the visual performances of the compared methods for 6 challenging WSIs are visualized in Fig. 8, where the results that contain typical artifacts are framed by red boxes.

The method Zheng et al. [8] was designed to eliminate the variances of illumination and saturation and Zhou et al. [22] did not consider the specificity of stains in the optimization. The normalization performances of the two methods are limited. In contrast, the other compared methods have effectively transformed the color to the template WSI. However, various artifacts appear in the normalized images. Specifically, the eosin stain and the background are occasionally confused in the results of Janowczyk et al. [14]. In Fig. 8(a), A certain amount of eosin stain surrounding the nuclei is eliminated, which has changed the environment of nuclei in the image. The result obtained by Bejnordi et al. [11] exhibits ringing artifacts around nuclei (Fig. 8(c)). The results of Khan et

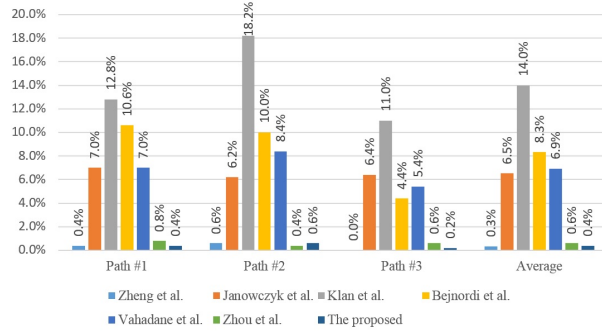


Figure 7: Statistical results in the visual assessment, where the percentage of WSIs annotated by the three pathologists as containing structural or color artifacts are compared, and the average results are presented on the right.

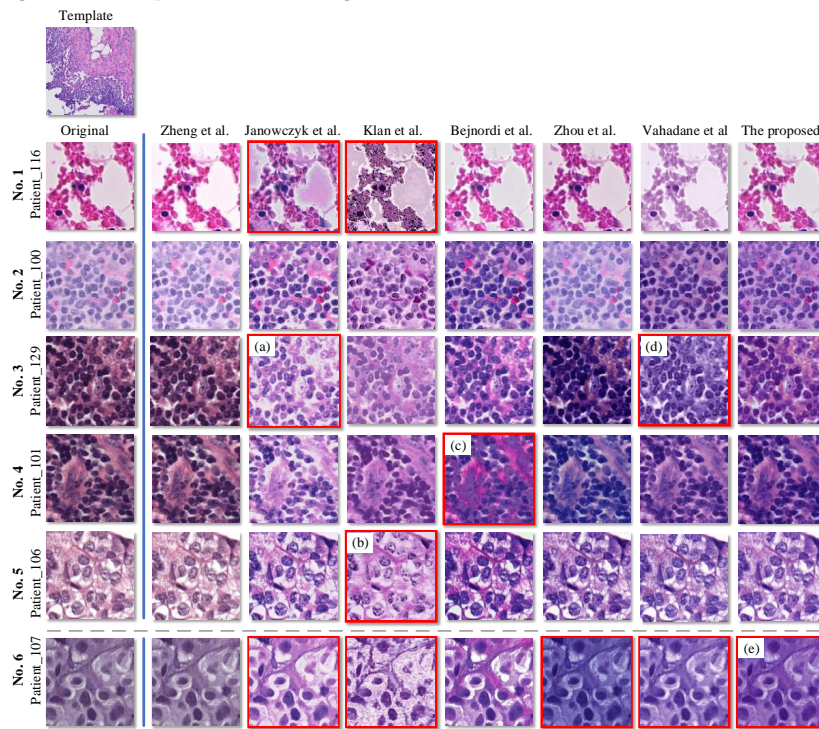


Figure 8: Visual performance of color normalization, in which ROIs cropped from challenging WSIs are displayed, the original ROIs are represented in the first column, the names of these WSIs are given on the left, the normalization results of the compared methods are presented on the right, and the results that have apparent artifacts are framed with red boxes.

al. [16] also have the similar problem. In Fig. 8(b), the area of nuclei apparently decreases, which will affect the performance of CAD approaches developed based on nuclei statistics. Compared to the methods above, the results obtained by unified-transformation-based methods, Vahadane et al. [19] has preserved the structure of tissue in the images. Therefore, the failure rate is lower than that of separate-transformation-based methods. On the other hand, color artifacts remain in the results of normalization, for instance the result in Fig. 8(d).

Overall, the failure rate of the proposed method is 0.4 %, which is relatively low in the compared methods. Through visual assessment, most of the structure and color artifacts appeared in the other methods were effectively avoided by our method. Therefore, the proposed method is the most robust in the color normalization compared to the compared methods. Typically, the failure occurred when the color distribution of the WSI was very monochromatic, as shown on the last row of Fig. 8. This type of color variance is challenging that all the methods except Bejnordi et al. [11] failed to process it.

4.3.4. Time complexity

The time complexity of color normalization methods is equally important in application. Especially for an automatic CAD approach based on WSIs, the running time of the normalization module should not become a bottleneck for the entire system. For the compared methods, the difference in computational complexity mainly derives from the estimation of model parameters. The time complexity for the pixel number n is given in Table 5. The methods proposed by Khan et al. [16] and Bejnordi et al. [11] depend on pixel classification, which is individually considered as pre-processing in Table 5. Khan et al. applied a pixel-wise classifier and the computation is linearly related to the pixel number n . Bejnordi et al. utilized Hough transform to detect nuclei, for which the time complexity is $\mathcal{O}(n^3)$. And in the stage of parameter estimation, the method of Zheng et al. [8] involved sorting algorithm of pixel values, for which the complexity is $\mathcal{O}(n \log_2 n)$. Other methods estimated the parameters based on pixel-wise operation, and the complexity is $\mathcal{O}(n)$.

Table 5: The comparison of time complexity for pixel number n used in the estimation of model parameters and the average running time for the estimation.

Method	Time complexity		Running time
	Pre-processing	Model solution	
Zheng et al. [8]	-	$\mathcal{O}(n \log_2 n)$	62.1 s
Janowczyk et al. [14]	-	$\mathcal{O}(n)$	93.1 s
Khan et al. [16]	$\mathcal{O}(n)$	$\mathcal{O}(n)$	1420 s
Bejnordi et al. [11]	$\mathcal{O}(n^3)$	$\mathcal{O}(n)$	370 s
Vahadane et al. [19]	-	$\mathcal{O}(n)$	65.4 s
Zhou et al. [22]	-	$\mathcal{O}(n)$	2.57 s
The proposed	-	$\mathcal{O}(n)$	2.97 s

The average times required for model parameter estimation are given in Table 5, where all the methods are implemented on CPU. Furthermore, the joint results of the quantitative metric and running time of the compared methods are presented in Fig. 9. For the proposed ACD model, the average time to estimate the normalization parameters is 2.97 s (including a pixel sampling time of 2.10 s and a model solving time of 0.87 s), which is relatively short compared to a WSI-based CAD approach.

Janowczyk et al. [14] and Bejnordi et al. [11], which are competitive in the evaluation of NMI SD, respectively takes 65 s and 370 s in average. Janowczyk et al. [14] utilized sparse auto-encoders to cluster pixels in different textures, and Bejnordi et al. [11] applied Hough transform to detect the nuclei for pixel classification. These operations extended the running time of the two methods. In contrast, the proposed method estimates the stain parameters using an integrated optimization and involves no additional operation to identify different types of pixels. Therefore, the proposed method is much faster than the two methods. Zhou et al. [22] is simpler than our model, thereby taking less time. However, the color consistency cannot surpass that of ours. Overall, the proposed method is both effective and efficient for histological image normalization.

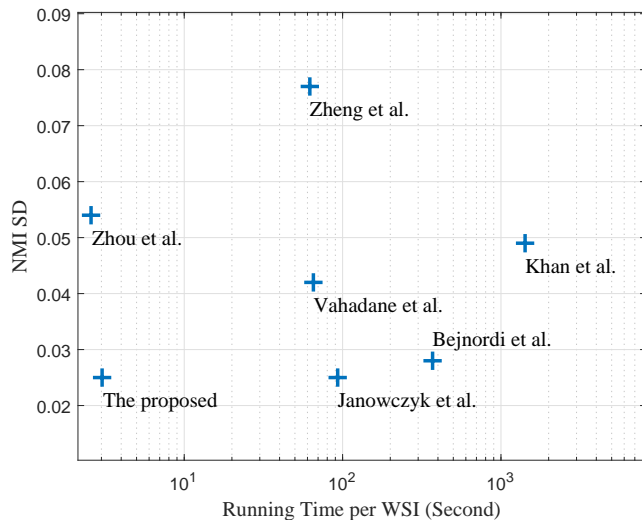


Figure 9: Joint plot of NMI SD and average running time for computing the model parameters for a WSI, where the time is counted in second and presented in logarithmic coordinate.

4.3.5. Effectiveness in CAD

In recent years, the convolutional neural network (CNN) has become a popular algorithm for medical image analysis [42]. An increasing number of studies have applied CNNs to histological image analysis [43, 36, 44, 45, 36]. The recent studies [9, 10] have shown that the normalization of histological images can improve the performance of CNN-based CAD methods. We conducted experiments to evaluate the improvement of the normalization methods for CNN model. Camelyon-16 dataset was used in this experiment.

The classification of image patches is the major action of CNNs in CAD systems for WSI analysis. Hence, a CNN model for histological image patch classification is established as the benchmark of this experiment. Specifically, a *ResNeXt-50* [46] structure was employed. To fit the input of the network, patches in size of 224×224 pixels were sampled from the WSIs in Camelyon-16 dataset. In the 270 training WSIs, there are 110 WSIs containing cancer regions that the pathologists annotated. To ensure the balance, 500 images were randomly sampled from the cancer regions in each WSI. Therefore, a total 55,000

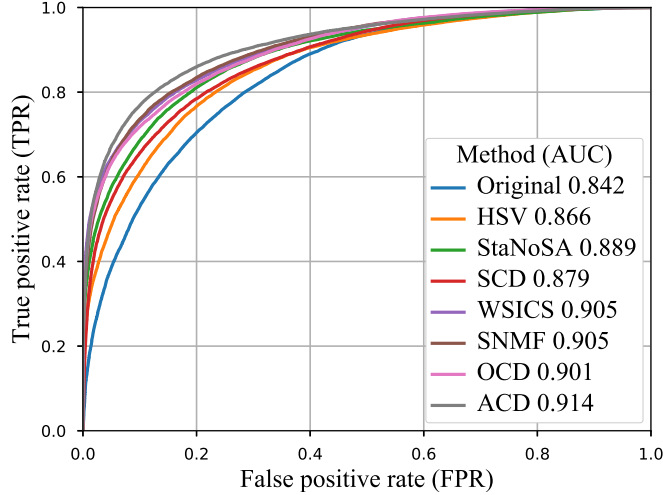


Figure 10: Comparison of ROC curve and AUC in application of breast histological image classification using different normalization methods.

positive samples were generated. In corresponding, 80,000 negative samples were generated (500 images per WSI) from the 160 WSIs without cancer area. To make a balance of the positive and negative samples, the 80,000 negative images were reduced to 55,000 through a randomly sampling. Therefore, a total 110,000 images in size of 224×224 pixels were sampled to train the CNN model. Following the same paradigm, 65,000 images were sampled from the 130 testing WSIs to evaluate the trained CNN model in cancer image classification. The receiver operating characteristic (ROC) curves and the corresponding area under curve (AUC) values for the testing set were calculated after 60,000 steps of training, where the CNN models for the compared color normalization methods were converged. The results are shown in Fig. 10.

Overall, the proposed method achieves the most evident outcome, which increases the AUC from 0.842 (the bench mark) to 0.914. The normalized images by Bejnordi et al. [11], Janowczyk et al. [14] and Khan et al. [16] changed structure of tissue. The results by Vahadane et al. [19] produced color distortion. These artifacts are detrimental for the CNN-model to distinguish patterns

Table 6: Comparison of metrics before and after the normalization on Motic-cervix dataset.

Metrics	NMI		NMI _h		NMI _e		Classification
	SD	CV	SD	CV	SD	CV	AUC
Original	0.113	0.142	0.107	0.551	0.188	0.583	0.831
Normalized	0.027	0.039	0.032	0.104	0.047	0.132	0.894

Table 7: Comparison of metrics before and after the normalization on Motic-lung dataset.

Metrics	NMI		NMI _h		NMI _e		Classification
	SD	CV	SD	CV	SD	CV	AUC
Original	0.112	0.139	0.142	0.711	0.183	0.879	0.886
Normalized	0.035	0.043	0.031	0.097	0.046	0.151	0.911

in histological images. In comparison, our method performs more robust and can effectively avoid artifacts in the normalization, and therefore achieves better performance than other methods in the classification of histological images.

4.4. Application performance on Motic datasets

To further evaluate the application performance of the proposed method on WSIs containing other lesions, we conducted experiments on Motic-cervix and Motic-lung datasets. In this experiment, the hyper-parameters of ACD model were the same as that determined on Camelyon dataset. The metrics for quantitative assessment was calculated before and after the normalization. Furthermore, the performance for improving CNN classification model (cancer versus non-cancer patches) was also evaluated. The paradigm of the evaluation was the same as that on Camelyon dataset.

The results before and after the color normalization are compared in Tables 6 and 7. It presents that the improvements in color consistency and for cancer image classification are significant after the normalization in the two datasets. Fig. 11 and Fig. 12 are the normalization results in motic-lung and motic-cervix datasets. Through visual assessment, the color normalization is consistent and no apparent artifacts are found in normalized WSIs on the two datasets. The experimental results have indicated that the proposed normalization method is robust and applicable to H&E-stained WSIs from other lesions.

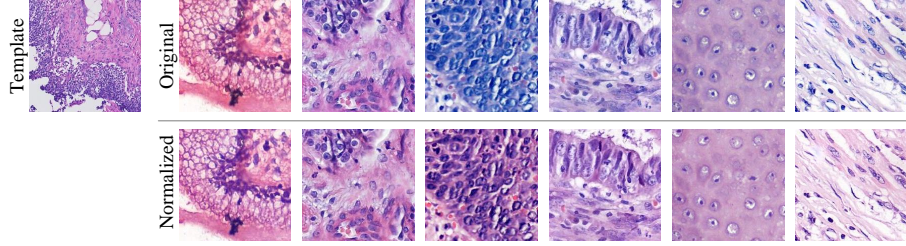


Figure 11: Normalized regions from different WSIs in Motic-cervix dataset.

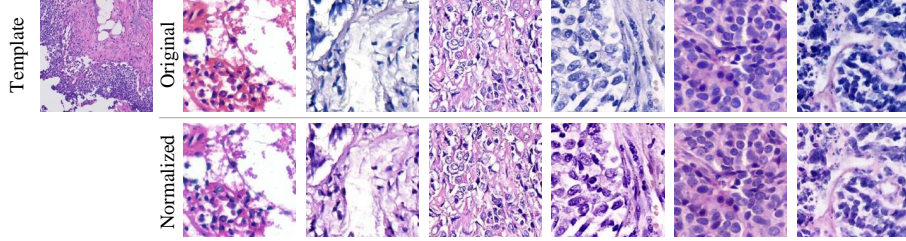


Figure 12: Normalized regions from different WSIs in Motic-lung dataset.

5. Discussion

The ACD model can also be used for color transformation between any two WSIs based on Eq. 8. The only modification is the transform matrix \mathbf{T} . For example, the color transform from WSI p to WSI q can be achieved with a matrix

$$\mathbf{T}_{pq} = (\hat{\mathbf{W}}_q \hat{\mathbf{D}}_q)^{-1} \hat{\mathbf{W}}_p \hat{\mathbf{D}}_p, \quad (10)$$

where $\hat{\mathbf{W}}_p$ and $\hat{\mathbf{D}}_p$ are the adaptive variables for WSI p , $\hat{\mathbf{W}}_q$ and $\hat{\mathbf{D}}_q$ are the adaptive variables for WSI q . The adaptive variables $\hat{\mathbf{W}}$ and $\hat{\mathbf{D}}$ for a WSI needs to be solved only once after the digitalization of the WSI, and can be stored along with the WSI. When the color transform between any two WSIs is required, the transform matrix can be directly obtained using Eq. 10. Therefore, it is very convenient to develop online color transformation applications for digital pathology systems (e.g. MoticGallery²) using the proposed model.

Based on the transform matrix \mathbf{T} , the normalization of the proposed method is a pixel-wise transformation. It can be completed efficiently by parallel com-

²<https://med.motic.com/MoticGallery/>

puting on CPU or GPU. For further acceleration, a look-up-table (LUT) from the original pixel values to the normalized pixel values can be established for the transformation. Then, the normalization of the WSI can be efficiently achieved through the LUT.

In the evaluation for number of pixels used in the optimization, the results show that the thousands of pixels can deliver a consistent performance of normalization (Fig. 2(b)). The main reason is that the pixels used in the optimization are sampled from the tissue area, and the staining in Camelyon dataset is relatively homogeneous in the WSI. Hence, thousands of pixels can cover the staining condition of the WSI. For a robust normalization performance, number of pixels is set to 100,000 in the experiments. It makes the ACD model successfully estimate the stain appearance matrix. We also tuned the number in the experiment on Motic dataset and found that 100,000 pixels were sufficient to the ACD model.

In the experiment on Motic datasets, the hyper-parameters of ACD model were the same with those used in Camelyon-17 dataset. The results have shown that the proposed method achieves a consistent normalization performance and avoids apparent artifacts. The reason is that the hyper-parameters of ACD model have robust intervals, and the Camelyon-17 dataset contains rich color variances in WSIs. Therefore, the hyper-parameters determined in section 4.2 can be robustly applied to other H&E-stained WSIs. When the ACD model is applied to WSIs in other type of stains, the hyper-parameters should be tuned following the paradigm provided in section 4.2.

Recently, color normalization methods based on convolutional neural networks [47, 23] are emerged. While, the deep convolutional structure in this type of method brings high computational cost into the normalization, which limits the applications in efficient CAD systems based on WSIs.

The model of Vahadane et al. [19] was solved by alternating between the staining matrix \mathbf{M} and the stain density of pixels (s_i). The computational complexity of this method is similar with our method. To maintain a robust performance of normalization, the pixels used in the optimization were about

10 million as suggested in [19]. Therefore, the time to solve the model longer than our method. The optimization in Zhou et al. [22] is similar with our model and therefore efficient in solution. However, Zhou et al. used the deconvolution matrix \mathbf{D} as the variable of the optimization, and objective function did not consider the specificity, the proportion, and the intensity of H and E stains. It makes the optimization mainly adjust the third row of matrix \mathbf{D} , and pay little attention to the first two rows in \mathbf{D} (i.e. the deconvolution parameters for H and E stains). Therefore, the color appearance cannot be sufficiently transformed to the template WSI.

The method proposed by Bejnordi et al. [11] identified different type of stains by detecting nuclei in histological images, and established specific transformation for each stain. Therefore, the normalized color of each stain is consistent with the template image (especially the eosin stain as shown in Fig. 8). On the other hand, once pixels sharing the similar color in the original images are classified into different classes, the color of these pixels may be quite different after the normalization. It will cause apparent color discontinuity in the normalized WSIs. This type of color discontinuity mainly appears on the boundary of eosin and hematoxylin stains (as shown Fig. 8 (c)). In contrast, the unified-transformation-based methods, e.g. [19, 22] and the proposed method, embedded the prior knowledge of H&E-staining into the model and estimated a unified transformation for all the pixels. The structural artifacts are effectively avoided. However, these methods depend on the assumption that the color distributions of different stains are discriminative in the original WSIs. When the WSI is very monochromatic, the model is difficult to identify different stains from the distribution of pixel values. A typical failure instance is presented on the last row of Fig. 8. This is one of the limitations of our method at present. One direction of the future work is to introduce the structural knowledge of histological images into the ACD model to assist in the identification of different stains.

In the ACD model, stain separation is achieved based on Beer-Lamber law. Hence, the ACD model is potential to be applied to other stains that satisfy

Beer-Lamber law. While, for stains that do not satisfy Beer-Lamber law (e.g. some immunohistochemical stains), the application of ACD model is limited. Another direction of future work is to utilize more general stain separation approaches in the model, for extending the application scope of the proposed color normalization method.

6. Conclusion

In this study, we have proposed a novel adaptive color deconvolution (ACD) model for color normalization of H&E-stained histological WSIs. The normalization is achieved through a unified transformation of pixels from the source WSI to the template WSI. The prior knowledge involving the specificity, the proportion and the overall intensity of stains are jointly considered and embedded in the ACD model, which has effectively reduced the failure rate of color normalization. The adaptive parameters for stain separation and stain normalization are simultaneously solved through an integrated optimization, for which the consistency of color appearance for the normalized images has been improved. In terms of computation, both the solution and application of the proposed method only involve pixel-wise operation, which determines the proposed method is light and applicable to WSI normalization. The experiment has demonstrated that the proposed method is effective in color normalization of H&E-stained WSIs in various color situations and from different lesions. The normalization results are consistent in color appearance and contain few structure or color artifacts. The average running time for parameter estimation is 2.97s. When using pre-processing method, performance of cancer image classification is significantly improved. Therefore, the proposed method is robust, effective, and efficient in color normalization for histological images and is adequate for developing efficient CAD programs and systems based on WSIs. The future work will focus on 1) using structural knowledge of histological images to improve the identification of different stains, 2) utilizing more general stain separation approaches to extend the application scope of the proposed color

normalization.

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 61771031, 61871011, 61371134, and 61501009).

Appendix

The derivatives of the objective function (Eq.6) of the ACD model on the variables are given in this section.

For a variable $\theta \in \{\alpha_h, \beta_h, \alpha_e, \beta_e, \alpha_d, \beta_d, w_h, w_e\}$, the partial derivatives of the variable can be calculated based on Eq.6 in the body of the paper as

$$\frac{\partial L}{\partial \theta} = \frac{\partial L_p}{\partial \theta} + \lambda_b \frac{\partial L_p}{\partial \theta} + \lambda_e \frac{\partial L_e}{\partial \theta}, \quad (11)$$

where the items of $\partial L_p/\partial \theta$, $\partial L_p/\partial \theta$, and $\partial L_e/\partial \theta$ are given as follows

$$\begin{aligned} \frac{\partial L_p}{\partial \theta} &= \frac{\partial}{\partial \theta} \left[\frac{1}{N} \sum_{i=1}^N d_i^2 + \lambda_p \frac{1}{N} \sum_{i=1}^N \frac{2h_i e_i}{h_i^2 + e_i^2} \right] \\ &= \frac{2}{N} \sum_{i=1}^N d_i \frac{\partial d_i}{\partial \theta} \\ &\quad + \frac{2\lambda_p}{N} \sum_{i=1}^N \left[\frac{(e_i \frac{\partial h_i}{\partial \theta} + h_i \frac{\partial e_i}{\partial \theta})(h_i^2 + e_i^2)}{(h_i^2 + e_i^2)^2} \right. \\ &\quad \left. - \frac{h_i e_i (2h_i \frac{\partial h_i}{\partial \theta} + 2e_i \frac{\partial e_i}{\partial \theta})}{(h_i^2 + e_i^2)^2} \right] \\ &= -\frac{4\lambda_p}{N} \sum_{i=1}^N \frac{h_i^2 e_i}{(h_i^2 + e_i^2)^2} \frac{\partial h_i}{\partial \theta} \\ &\quad - \frac{4\lambda_p}{N} \sum_{i=1}^N \frac{h_i e_i^2}{(h_i^2 + e_i^2)^2} \frac{\partial e_i}{\partial \theta} \\ &\quad + \frac{2}{N} \sum_{i=1}^N d_i \frac{\partial d_i}{\partial \theta} \end{aligned} \quad (12)$$

$$\begin{aligned}
\frac{\partial L_b}{\partial \theta} &= \frac{\partial}{\partial \theta} \left[(1-\eta) \frac{1}{N} \sum_{i=1}^N h_i - \eta \frac{1}{N} \sum_{i=1}^N e_i \right]^2 \\
&= 2\sqrt{L_b} \left[\frac{(1-\eta)}{N} \sum_{i=1}^N \frac{\partial h_i}{\partial \theta} - \frac{\eta}{N} \sum_{i=1}^N \frac{\partial e_i}{\partial \theta} \right] \\
&= 2\sqrt{L_b} \frac{(1-\eta)}{N} \sum_{i=1}^N \frac{\partial h_i}{\partial \theta} - 2\sqrt{L_b} \frac{\eta}{N} \sum_{i=1}^N \frac{\partial e_i}{\partial \theta}
\end{aligned} \tag{13}$$

$$\begin{aligned}
\frac{\partial L_e}{\partial \theta} &= \frac{\partial}{\partial \theta} \left[\gamma - \frac{1}{N} \sum_{i=1}^N (h_i + e_i) \right]^2 \\
&= -2\sqrt{L_e} \frac{1}{N} \sum_{i=1}^N \left(\frac{\partial h_i}{\partial \theta} + \frac{\partial e_i}{\partial \theta} \right) \\
&= -2\sqrt{L_e} \frac{1}{N} \sum_{i=1}^N \frac{\partial h_i}{\partial \theta} - 2\sqrt{L_e} \frac{1}{N} \sum_{i=1}^N \frac{\partial e_i}{\partial \theta}
\end{aligned} \tag{14}$$

From the definition of stains $\mathbf{s}_i = (h_i, e_i, d_i)^\top$, the partial derivatives of h_i, e_i, d_i on θ can be represented by a vector

$$\frac{\partial \mathbf{s}_i}{\partial \theta} = \left(\frac{\partial h_i}{\partial \theta}, \frac{\partial e_i}{\partial \theta}, \frac{\partial d_i}{\partial \theta} \right)^\top.$$

Then, the calculation of $\partial L / \partial \theta$ can be written as

$$\frac{\partial L}{\partial \theta} = \frac{1}{N} \sum_{i=1}^N \mathbf{c}^\top \cdot \frac{\partial \mathbf{s}_i}{\partial \theta}, \tag{15}$$

where \mathbf{c} is a vector that consists of coefficients summarized from Equations 11-14:

$$\mathbf{c} = \begin{pmatrix} -4\lambda_p \frac{h_i^2 e_i}{(h_i^2 + e_i^2)^2} + 2(1-\eta)\lambda_b \sqrt{L_b} - 2\lambda_e \sqrt{L_e} \\ -4\lambda_p \frac{h_i e_i^2}{(h_i^2 + e_i^2)^2} - 2\eta\lambda_b \sqrt{L_b} - 2\lambda_e \sqrt{L_e} \\ 2d_i \end{pmatrix}$$

Next, the calculation of $\partial \mathbf{s}_i / \partial \theta$ is presented. From Eq.2 in the body of the paper, it is

$$\frac{\partial \mathbf{s}_i}{\partial \theta} = \frac{\partial}{\partial \theta} (\mathbf{W} \mathbf{D} \mathbf{o}_i).$$

Here, \mathbf{W} consists of the weighting variables w_h and w_e , and \mathbf{D} is a function of degree variables φ . Thus, the partial derivatives of \mathbf{s}_i on $\theta \in \{w_h, w_e\}$ and $\theta \in \varphi$ are discussed separately.

The partial derivatives of \mathbf{s}_i on w_h and w_e are

$$\frac{\partial \mathbf{s}_i}{\partial w_h} = \text{diag}(1, 0, 0) \mathbf{D} \mathbf{o}_i, \quad \frac{\partial \mathbf{s}_i}{\partial w_e} = \text{diag}(0, 1, 0) \mathbf{D} \mathbf{o}_i. \quad (16)$$

And, for variables $\theta \in \varphi$,

$$\begin{aligned} \frac{\partial \mathbf{s}_i}{\partial \theta} &= \frac{\partial}{\partial \theta} (\mathbf{W} \mathbf{D} \mathbf{o}_i) = \mathbf{W} \frac{\partial \mathbf{D}}{\partial \theta} \mathbf{o}_i \\ &= \mathbf{W} \frac{\partial \mathbf{M}^{-1}}{\partial \theta} \mathbf{o}_i \\ &= \mathbf{W} \mathbf{M}^{-1} \frac{\partial \mathbf{M}}{\partial \theta} \mathbf{M}^{-1} \mathbf{o}_i \\ &= \mathbf{W} \mathbf{D} \frac{\partial \mathbf{M}}{\partial \theta} \mathbf{D} \mathbf{o}_i. \end{aligned}$$

Specifically, the derivatives of \mathbf{M} on each degree variables are

$$\begin{aligned} \frac{\partial \mathbf{M}}{\partial \alpha_h} &= \left(\frac{\partial \mathbf{m}_h}{\partial \alpha_h}, \mathbf{0}, \mathbf{0} \right), & \frac{\partial \mathbf{M}}{\partial \beta_h} &= \left(\frac{\partial \mathbf{m}_h}{\partial \beta_h}, \mathbf{0}, \mathbf{0} \right), \\ \frac{\partial \mathbf{M}}{\partial \alpha_e} &= \left(\mathbf{0}, \frac{\partial \mathbf{m}_e}{\partial \alpha_e}, \mathbf{0} \right), & \frac{\partial \mathbf{M}}{\partial \beta_e} &= \left(\mathbf{0}, \frac{\partial \mathbf{m}_e}{\partial \beta_e}, \mathbf{0} \right), \\ \frac{\partial \mathbf{M}}{\partial \alpha_d} &= \left(\mathbf{0}, \mathbf{0}, \frac{\partial \mathbf{m}_d}{\partial \alpha_d} \right), & \frac{\partial \mathbf{M}}{\partial \beta_d} &= \left(\mathbf{0}, \mathbf{0}, \frac{\partial \mathbf{m}_d}{\partial \beta_d} \right), \end{aligned} \quad (17)$$

and

$$\begin{aligned} \frac{\partial \mathbf{m}_j}{\partial \alpha_j} &= (-\sin \alpha_j \sin \beta_j, -\sin \alpha_j \cos \beta_j, \cos \alpha_j)^T, \\ \frac{\partial \mathbf{m}_j}{\partial \beta_j} &= (\cos \alpha_j \cos \beta_j, -\cos \alpha_j \sin \beta_j, \sin \alpha_j)^T, \\ j &= h, e, d \end{aligned}$$

References

- [1] S. L. Robbins, V. Kumar, A. K. Abbas, N. Fausto, J. C. Aster, Robbins and Cotran pathologic basis of disease, Saunders/Elsevier, 2010.
- [2] F. Ghaznavi, A. Evans, A. Madabhushi, M. Feldman, Digital imaging in pathology: Whole-slide imaging and beyond, Annual Review of Pathology-mechanisms of Disease 8 (1) (2013) 331–359.

- [3] J. S. Duncan, N. Ayache, Medical image analysis: Progress over two decades and the challenges ahead, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (1) (2000) 85–106. doi:10.1109/34.824822.
- [4] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, B. Yener, Histopathological image analysis: a review, *IEEE Reviews in Biomedical Engineering* 2 (2009) 147–171.
- [5] S. Zhang, D. Metaxas, Large-scale medical image analytics: Recent methodologies, applications and future directions, *Medical Image Analysis* 33 (2016) 98–101.
- [6] Z. Li, X. Zhang, H. Miller, S. Zhang, Large-scale retrieval for medical image analytics: A comprehensive review., *Medical Image Analysis* 43 (2018) 66–84.
- [7] B. E. Bejnordi, M. Balkenhol, G. Litjens, R. Holland, P. Bult, N. Karssemeijer, J. A. van der Laak, Automated detection of dcis in whole-slide h&e stained breast histopathology images, *IEEE Transactions on Medical Imaging* 35 (9) (2016) 2141–2150. doi:10.1109/tmi.2016.2550620.
- [8] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Y. Ma, H. Shi, Y. Zhao, Histopathological whole slide image analysis using context-based cbir, *IEEE Transactions on Medical Imaging* 37 (7) (2018) 1641–1652. doi:10.1109/TMI.2018.2796130.
- [9] F. Ciompi, O. Geessink, B. E. Bejnordi, G. S. De Souza, A. Baidoshvili, G. J. S. Litjens, B. Van Ginneken, I. D. Nagtegaal, J. A. W. M. V. Der Laak, The importance of stain normalization in colorectal tissue classification with convolutional networks, in: *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2017, pp. 160–163.
- [10] B. E. Bejnordi, J. Lin, B. Glass, M. Mullooly, G. L. Gierach, M. E. Sherman, N. Karssemeijer, J. V. D. Laak, A. H. Beck, Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in

- histopathology images, in: IEEE International Symposium on Biomedical Imaging, 2017, pp. 929–932.
- [11] B. E. Bejnordi, G. Litjens, N. Timofeeva, I. Otte-Hller, A. Homeyer, N. Karssemeijer, J. A. V. D. Laak, Stain specific standardization of whole-slide histopathological images, *IEEE Transactions on Medical Imaging* 35 (2) (2016) 404–415.
- [12] D. Onder, S. Zengin, S. Sarioglu, A review on color normalization and color deconvolution methods in histopathology, *Applied Immunohistochemistry & Molecular Morphology* 22 (10) (2014) 713–719.
- [13] Y. Y. Wang, S. C. Chang, L. W. Wu, S. T. Tsai, Y. N. Sun, A color-based approach for automated segmentation in tumor tissue classification., in: *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, p. 6577.
- [14] A. Janowczyk, A. Basavanhally, A. Madabhushi, Stain normalization using sparse autoencoders (stanosa): Application to digital pathology, *Computerized Medical Imaging and Graphics* 57 (2017) 50–61.
- [15] D. Magee, D. Treanor, D. Crellin, M. Shires, K. Smith, K. Mohee, P. Quirke, Colour normalisation in digital histopathology images, in: *Proc Optical Tissue Image analysis in Microscopy, Histopathology and Endoscopy (MICCAI Workshop)*, 2009, pp. 100–111.
- [16] A. M. Khan, N. M. Rajpoot, D. Treanor, D. R. Magee, A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution, *IEEE Transactions on Biomedical Engineering* 61 (6) (2014) 1729–1738.
- [17] X. Li, K. N. Plataniotis, A complete color normalization approach to histopathology images using color cues computed from saturation-weighted statistics, *IEEE Transactions on Biomedical Engineering* 62 (7) (2015) 1862–1873. doi:10.1109/TBME.2015.2405791.

- [18] J. Vicory, H. D. Couture, N. E. Thomas, D. Borland, J. S. Marron, J. T. Woosley, M. Niethammer, Appearance normalization of histology slides, *Computerized Medical Imaging and Graphics* 43 (2015) 89–98.
- [19] A. Vahadane, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A. M. Schlitter, I. Esposito, N. Navab, Structure-preserving color normalization and sparse stain separation for histological images, *IEEE Transactions on Medical Imaging* 35 (8) (2016) 1962–1971.
- [20] L. Sha, D. Schonfeld, Sethi, Color normalization of histology slides using graph regularized sparse nmf, in: *Proceeding of SPIE Medical Imaging*, Vol. 10140, 2017, p. 1014010.
- [21] N. Hidalgogavira, J. Mateos, M. Vega, R. Molina, A. K. Katsaggelos, Fully automated blind color deconvolution of histopathological images., in: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2018, pp. 183–191.
- [22] N. Zhou, Y. Gao, Optimized color decomposition of localized whole slide images and convolutional neural network for intermediate prostate cancer classification, in: *Proceeding of SPIE Medical Imaging*, Vol. 10140, 2017, p. 101400W.
- [23] F. G. Zanjani, S. Zinger, B. E. Bejnordi, J. A. van der Laak, P. H. N. de With, Stain normalization of histopathology images using generative adversarial networks, in: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 573–577. doi:10.1109/ISBI.2018.8363641.
- [24] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images, *IEEE Computer Graphics and Applications* 21 (5) (2001) 34–41.
- [25] J. Illingworth, J. Kittler, A survey of the hough transform, *Graphical Models and Image Processing* 44 (1) (1988) 87–116.

- [26] A. C. Ruifrok, D. A. Johnston, Quantification of histochemical staining by color deconvolution, *Analytical and quantitative cytology and histology* 23 (4) (2001) 291–299.
- [27] Y. Ma, S. Jun, Z. Jiang, F. Hao, Plsa-based pathological image retrieval for breast cancer with color deconvolution, *Proceedings of SPIE Medical Imaging* 8920 (8) (2013) 89200L–89200L–7.
- [28] Y. Zheng, Z. Jiang, J. Shi, Y. Ma, Retrieval of pathology image for breast cancer using plsa model based on texture and pathological features, in: *IEEE International Conference on Image Processing*, 2014, pp. 2304–2308.
- [29] Y. Ma, Z. Jiang, H. Zhang, F. Xie, Y. Zheng, H. Shi, Y. Zhao, Breast histopathological image retrieval based on latent dirichlet allocation, *IEEE Journal of Biomedical and Health Informatics* 21 (4) (2017) 1114–1123. doi:10.1109/jbhi.2016.2611615.
- [30] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Y. Ma, H. Shi, Y. Zhao, Size-scalable content-based histopathological image retrieval from database that consists of wsis, *IEEE journal of biomedical and health informatics* 22 (4) (2018) 1278–1287. doi:10.1109/jbhi.2017.2723014.
- [31] Y. Ma, Z. Jiang, H. Zhang, F. Xie, Y. Zheng, H. Shi, Y. Zhao, Proposing regions from histopathological whole slide image for retrieval using selective search, in: *IEEE International Symposium of Biomedical imaging*, 2017.
- [32] Y. Ma, Z. Jiang, H. Zhang, F. Xie, Y. Zheng, H. Shi, Y. Zhao, J. Shi, Generating region proposals for histopathological whole slide image retrieval, *Computer Methods and Programs in Biomedicine* 159 (2018) 1–10. doi:{10.1016/j.cmpb.2018.02.020}.
- [33] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, N. E. Thomas, A method for normalizing histology slides for quantitative analysis, in: *IEEE International Conference on*

Symposium on Biomedical Imaging: From Nano To Macro, 2009, pp. 1107–1110.

- [34] X. Li, K. N. Plataniotis, Circular mixture modeling of color distribution for blind stain separation in pathology images, *IEEE Journal of Biomedical and Health Informatics* 21 (1) (2017) 150–161. doi:10.1109/JBHI.2015.2503720.
- [35] J. Xu, L. Xiang, G. Wang, S. Ganesan, M. Feldman, N. N. Shih, H. Gilmore, A. Madabhushi, Sparse non-negative matrix factorization (snmf) based color unmixing for breast histopathological image analysis., *Computerized Medical Imaging and Graphics* 46 Part 1 (2015) 20–29.
- [36] B. E. Bejnordi, M. Veta, P. J. Van Diest, B. Van Ginneken, N. Karssemeijer, G. J. S. Litjens, J. A. W. M. V. Der Laak, M. Hermsen, Q. F. Manson, M. Balkenhol, et al., Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer, *JAMA* 318 (22) (2017) 2199–2210.
- [37] G. J. S. Litjens, P. Bandi, B. E. Bejnordi, O. Geessink, M. Balkenhol, P. Bult, A. Halilovic, M. Hermsen, R. J. M. V. De Loo, R. Vogels, et al., 1399 h&e-stained sentinel lymph node sections of breast cancer patients: the camelyon dataset, *GigaScience* 7 (6).
- [38] A. Basavanthally, A. Madabhushi, Em-based segmentation-driven color standardization of digitized histopathology, in: *Proceeding of SPIE Medical Imaging*, Vol. 8676, 2013, p. 86760G.
- [39] J. C. Duchi, E. Hazan, Y. Singer, Adaptive subgradient methods for on-line learning and stochastic optimization, *Journal of Machine Learning Research* 12 (2011) 2121–2159.
- [40] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. J. Goodfellow,

- A. Harp, G. Irving, M. Isard, Y. Jia, R. Józefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. G. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. A. Tucker, V. Vanhoucke, V. Vasudevan, F. B. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, Tensorflow: Large-scale machine learning on heterogeneous distributed systems, arXiv abs/1603.04467.
- [41] J. L. Hintze, R. D. Nelson, Violin plots: A box plot-density trace synergism, *The American Statistician* 52 (2) (1998) 181–184.
- [42] G. Litjens, T. Kooi, B. E. Bejnordi, S. Aaa, F. Ciompi, M. Ghafoorian, V. D. L. Jawm, G. B. Van, C. I. Snchez, A survey on deep learning in medical image analysis, *Medical Image Analysis* 42 (9) (2017) 60–88.
- [43] Y. Xu, Z. Jia, L. B. Wang, Y. Ai, F. Zhang, M. Lai, E. I. Chang, Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features, *Bmc Bioinformatics* 18 (1) (2017) 281.
- [44] Y. Zheng, Z. Jiang, F. Xie, H. Zhang, Y. Ma, H. Shi, Y. Zhao, Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification, *Pattern Recognition* 71 (2017) 14–25. doi:10.1016/j.patcog.2017.05.010.
- [45] J. Xu, X. Luo, G. Wang, H. Gilmore, A. Madabhushi, A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images, *Neurocomputing* 191 (2016) 214–223. doi:10.1016/j.neucom.2016.01.034.
- [46] S. Xie, R. B. Girshick, P. Dollar, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [47] D. Bug, S. Schneider, A. Grote, E. Oswald, F. Feuerhake, J. Schler, D. Merhof, Context-based normalization of histological stains using deep convo-

lutional features, in: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2017, pp. 135–142.