

## A hybrid approach for stain normalisation in digital histopathological images

Faiza Bukenya

Received: 29 November 2018/Revised: 25 June 2019 / Accepted: 16 September 2019

**Abstract** Stain in-homogeneity adversely affects segmentation and quantification of tissues in histology images. Stain normalisation techniques have been used to standardise the appearance of images. However, most the available stain normalisation techniques only work on a particular kind of stain images. In addition, some of these techniques fail to utilise both the spatial and textural information in histology images, leading to image tissue distortion. In this paper, a hybrid approach has been developed, based on an octree colour quantisation algorithm combined with the Beer-Lambert law, a modified blind source separation algorithm, and a modified colour transfer approach. The hybrid method consists of two stages the stain separation stage and colour transfer stage. An octree colour quantisation algorithm combined with Beer-Lambert law, and a modified blind source separation algorithm are used during the stain separation stage to computationally estimate the amount of stain in an histology image based on its chromatic and luminous response. A modified colour transfer algorithm is used during the colour transfer stage to minimise the effect of varying staining and illumination. The hybrid method addresses the colour variation problem in both H&DAB (Haemotoxylin and Diaminobenzidine) and H&E (Haemotoxylin and Eosin) stain images. The stain normalisation method is validated against ground truth data. It is widely known that the Beer-Lambert law applies to only stains (such as haematoxylin, eosin) that absorb light. We demonstrate that the Beer-Lambert law applies is applicable to images containing a DAB stain. Better stain normalisation results are obtained in both H&E and H&DAB images.

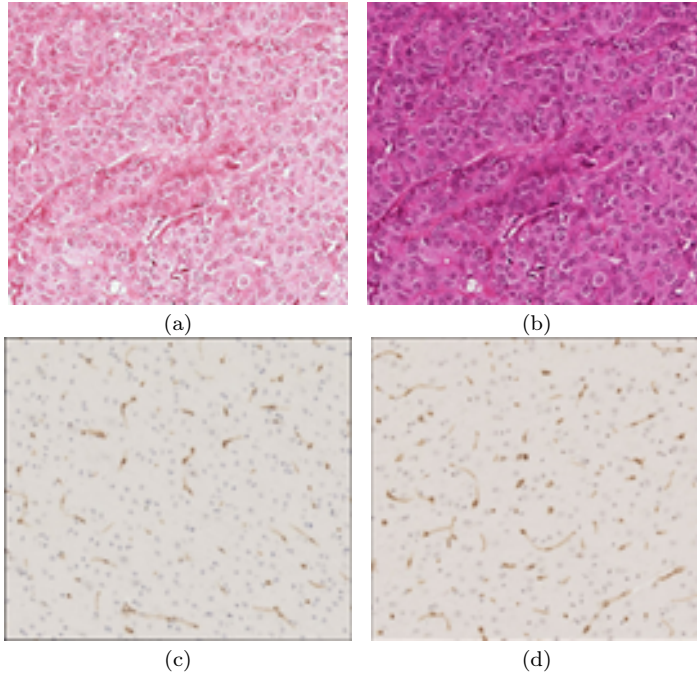
**Keywords** Histopathology images · Stain normalisation · Image processing.

---

F. Bukenya  
School of Computer Science, University of Nottingham, Nottingham, UK  
E-mail: faiza.bukenya@nottingham.ac.uk

## 1 Introduction

In histopathology, tissue samples are examined for disease diagnosis. Image staining is performed to increase the contrast of the image. In-homogeneity in histology images is still a major problem that impairs visual inspection as well as image analysis by computer systems [86, 41, 85, 66]. Figure 1 demonstrates the colour variation problem.



**Fig. 1** (a) and (b) are the same H&E histology images captured using an Aperio scanner and Hamamatsu scanner respectively. (c) and (d) are different H&DAB histology images from the same batch stained at different times.

Stain normalisation is performed to standardise the appearance of the images so as to improve image analysis during image segmentation and quantification[15]. However, serious tissue distortion may occur if stain normalisation is poorly performed leading to poor image interpretation [86, 55, 8, 85]. For example, several objects may appear to be linked as a single object after segmentation [30, 12]. For more accurate diagnosis, an automated stain normalisation method should preserve both tissue structure and texture in the image. Different stain normalisation approaches have been proposed. However, most of these stain normalisation techniques are based on histogram specification [34, 28], grey world normalisation [88], comprehensive colour normalisation [19], and histogram equalisation [46, 58]. These colour normalisation techniques work on natural images rather than histology images [86, 41, 85].

Stain normalisation techniques employ non-negative matrix factorisation (NMF) [86,41,85], singular value decomposition (SVD) [57], principal component analysis (PCA) [45], [33], independent component analysis (ICA) [84], statistical based algorithm ICA [1,51], and machine learning [97,78,35,98]. However, these methods are only applicable to certain histology images. In addition, some of these methods produce undesirable image tissue distortions and also introduce considerable visual artefacts in histology images [86,41]. Further more, some of these methods require expertise and are computationally expensive. For accurate stain normalisation, a precise estimation of each applied stain in both source image and target image is necessary to modify stain colour to match that of the target image [86,85,57]. For example in H&E stained image slides, purple and pink stains are estimated respectively based on the Beer-Lambert law. It is widely known that the Beer-Lambert law applies to only stains (such as haematoxylin, eosin) that absorb light [89]. We demonstrate that Beer-Lambert law is applicable to DAB images. In this paper, a hybrid approach has been developed, based on an octree colour quantisation algorithm combined with Beer-Lambert law, a modified blind source separation algorithm, and a modified colour transfer approach. Stain normalisation is carried out in similar way as existing approaches [86,57]. However, our method utilises blind source separation in [63] rather than NMF during stain separation. Note: we modified the blind source separation in [63] for a better stain normalisation.

The rest of the paper is structured as follows: Section 2 briefly reviews some stain normalisation algorithms, Section 3 presents the proposed methodology, Section 4 presents experimental results and the discussion of results, Section 6 concludes the paper.

## 2 Related Work

This section briefly reviews existing stain normalisation algorithms addressing the stain in-homogeneity problem in histology images. The commonly used stain normalisation methods include NMF, SVD, PCA, ICA. All of these methods are based on matrix factorisation (also known as matrix decomposition). Matrix decomposition methods reduce a complex matrix into constituent parts so as to make it easier to calculate. Although these methods are based on matrix factorisation, they use various kinds of assumptions translated into different constraints and/or objective functions during stain normalisation [87]. For-example, NMF aims at considering the physical constraint that each stain has a non-negative response, SVD aims at evaluating the decomposition that best approximates a data matrix, given some rank restriction [69], PCA aim at accounting for as much of the data variability while enforcing orthogonality between their components, and ICA is based on the assumption that each stain, stains the tissue independently from all the other stains [67,87]. In imaging, besides stain normalisation, matrix factorisation has been used to recover human pose [31], to extract muscle synergy [18], to segment objects [81,61],

to classify images [101], to monitor human motion [90], to reduce image dimension [4, 79], to extract features [83], and to retrieve images from the large database of images [96]. This is because of its effectiveness and efficiency. The four stain normalisation methods are described below.

## 2.1 Non-Negative Matrix Factorisation (NMF)

NMF was proposed by Paatero and Tapper in 1994 [62]. Non-negative matrix factorization (NMF) factors a data matrix into low-rank latent factor matrices with non-negativity constraints. Given a data matrix  $X$  and a target rank  $R$ , NMF seeks a factorization model [20].

$$X \approx WH^T, X \in \mathbf{R}^{M \times N}, H \in \mathbf{R}^M \quad (1)$$

where  $W \geq \mathbf{0}$ ,  $H \geq \mathbf{0}$ , and  $R \leq \min\{M, N\}$ .

Compared to other matrix factorisation methods, NMF is capable of learning parts of objects [47, 20]. In addition, NMF is able to identify the ground-truth-generative factors  $W$  and  $H$  up to certain trivial ambiguities, which results into a strong interpretability. NMF can be combined with other methods such as non negative least squares to normalise stain in FAST images as the FAST images consist of more objects than it had spectral dimensions [11]. The effectiveness of NMF in analyzing real-life non-negative data has sparked a substantial amount of research in image analysis. In image analysis, NMF and related techniques have been applied to different kinds of imaging (such as remote sensed images, medical images (such as histology images, CT images, MRI images), and natural images) during feature extraction [80, 56, 53, 92, 48], image classification [26, 76, 99, 5], image ranking [82, 14], image reconstruction [17, 16, 7], and image enhancement [43, 22, 42, 100, 85, 86, 54, 41, 42, 73, 87]. However, NMF, is less utilised during image enhancement of medical images, in particular histology imaging, which is affected by colour variation problem. This study focus on stain normalisation methods.

During stain normalisation, NMF-based algorithms attempt to factor the optical density (OD) matrix into  $W$  and  $H$  with the constraint that all elements must be non-negative, since stains cannot absorb a negative amount of light. It also forces all columns of the stain vectors to be unit length to keep it from having an infinite number of solutions that differ only by a constant. These methods work well on multispectral images, however they tend to inconsistently converge when only three spectral components are available [57]. The standard NMF is subject to local minimisation that is initialisation-dependent and thus requires a good approximation of the colour vectors to initialise the factorization process [87]. In addition, standard NMF is not capable of preserving structural and textual information during stain normalisation. Further more, NMF is not suitable for estimating the stain vector extraction [93, 87].

Following NMF, various analysis have been studied. For example, to help circumvent problems that can be found with NMF, Xu et al., [93] proposed



the use SNMF in place of NMF to extract the colour vectors. In SNMF, regularisation and sparsity terms are introduced in the NMF objective function. However, Xu et al., [93] does not provide information about the initialisation they use for the SNMF algorithm. In 2017, Van et al., [87] tested SNMF using the parameter in [52]. According to Van et al., [87], SNMF requires good initialisation to be efficient, which is not always an easy task. Vahadane et al., [86] developed a stain normalisation approach based on SNMF to preserve the structural information of the source image after color normalization. In this approach, sparsity constraints are added on to stain channels during stain separation. Parameters for stain normalization are calculated based on the pseudo maximum of the separated stains. Sparsity constraints are able to enhance the recognition ability of the model for separating independent stains. However, the weighting of stains is processed independently after stain separation, which limits the capacity for color normalization. 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 a single stain[102].

Amit et al., [54], developed a stain normalisation method based on graph regularised sparse NMF to preserve the textural information in the image. Like the method in [86], this method also add sparsity constraints on to stain channels during stain separation. This method also does not consider intensity and proportion of the separated stains in the model of stain separation. In [102], an adaptive color deconvolution (ACD) model for stain separation and color normalization of histological images is presented. In this method, the stain color appearance matrix is utilised as variables, the ACD model is solved through an integrated optimization. The normalization is achieved by a unified transformation for the pixels in the image. Unlike in [86,54], This method considers both intensity and proportion of the stains. Both intensity and proportion are embedded in the ACD model to effectively reduce the failure of stain separation and also to avoid color artifacts in the normalized images.

Recently Li et al., [50] proposed a majorisation-minimisation principle to develop a convergence-guaranteed algorithm NMF. However, the method was applied during communication analysis of signals.

## 2.2 Singular Value Decomposition (SVD)

The singular value decomposition of a matrix  $A$  is an attractive algebraic transform that factorises a matrix  $A$  into the product of three matrices( $U,D,V$ ): Let  $A$  be a  $m \times n$  matrix. The Singular Value Decomposition of  $A$  is as follows:

$$A = UDV^T \quad (2)$$

Where  $\mathbf{U}$  and  $\mathbf{V}$  are orthogonal matrices with singular values.  $\mathbf{V}^T$  is the conjugate transpose of  $\mathbf{V}$ .  $\mathbf{D}$  is a diagonal matrix with singular elements [39]. The process of SVD depends on the task at hand.

During stain normalisation, SVD is used for automated stain matrix estimation to address colour variation problems [57]. The SVD method works by calculating the plane from the two vectors corresponding to the two largest singular values of the SVD decomposition of the OD (optical density) transformed pixels and then later projecting OD transformed pixels onto the plane [84, 57, 1]. SVD easily adapts to variations in the local statistics of an image [95, 64, 27, 3, 38]. SVD is also stable and robust to noise [91, 94, 44, 44, 39, 6, 2, 71, 60, 37]. However, although SVD can effectively determine independent components, it changes the colour distribution of both source and reference images, which is not necessary [54]. Compared to NMF, SVD is fast. However, it tends to produce unrealistic colours in the normalised source image [86, 85].

### 2.3 Principal Component Analysis

Principal component analysis was first introduced by Karl Pearson in 1901 [65], and developed independently by Hotelling in 1933 [32]. Principal component analysis uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables known as principal components.

During stain normalisation, in PCA [40], a plane is constructed that optimally approximates variability between image pixels. The plane is spanned by the two first principal component vectors of the image pixel colour data. The deconvolution error is minimised, by making sure that the optimal colour deconvolution vectors of the two stains in the original image are part of the plane. Although PCA can identify an appropriate plane for describing the clouds of bi-coloured pixels for H&E staining, it is not good at representing variability in stain colours as it assumes fixed stain vectors for H and E images [67, 1, 87]. In comparison to PCA and NMF algorithms, ICA gives a superior projection of the data with respect to each stain [84].

### 2.4 Independent Component Analysis (ICA)

Independent Component Analysis is blind source separation method that obtains non-Gaussian signals from a Gaussian source. Independent Component Analysis is quite similar to PCA [84]. The only difference lies in the way PCA and ICA deal with data. PCA dispense the data in a way that best separates variance while ICA dispense the data along the axes in such a way that each independent component directly corresponds to a given stain [84]. ICA may not achieve an orthogonal distribution if the two stains in the image are of the same color. Although Trahearn et al., [84] introduced a correction step to estimate the true stain vector from within the ICA space, the correction step

involves two steps that include (a) the true stain vectors are a minor adjustment from the unit vectors in ICA space, (b) Most of the data in ICA space is close to one of the true stain vectors. However ICA is applicable to only images with more than two stains. Rabinovic et al.,[67] compared two stain deconvolution approaches, NMF and ICA. They showed that while NMF performed better however neither of the methods was sufficient to fully deconvolve the images [1], [13]. The study was performed on hyper-spectral images, rather than light microscopy images. Hyper-spectral imaging are different from microscopy image in such a way that they have higher number of input channels as compared to the three-channel RGB microscopy images, which may limit the comparison that can be made between methods for the two modalities.

## 2.5 Statistical Blind Source Separation Algorithm ICA

Alsubaie N et al., [1] chose to recover the original source signals blindly by using the statistical blind source separation algorithm ICA. In this method, independent component analysis is used in the wavelet domain where each colour channel of the input image is decomposed into a series of narrow sub-band images using a decimated wavelet transform. Then later each sub-band is statistically analysed in order to find the least Gaussian sub-bands. Finally, ICA is applied to the selected sub bands to estimate the stain matrix. The method cannot estimate weaker stains such as eosin.

## 3 Methods and material

This section presents the proposed methodology for automated stain normalisation of H&E and H&DAB stain histology images. The proposed methodology involves two stages: (a) stain separation, and (b) colour transfer.

### 3.1 Proposed automated stain normalisation

The hybrid method consist of two stages: the stain separation stage and colour transfer stage. An octree colour quantisation algorithm combined with the Beer-Lambert law, and a modified blind source separation algorithm are used during stain separation stage to computationally estimate the amount of stain in an histology image based on its chromatic and luminous response. A modified colour transfer algorithm is used during the colour transfer stage to minimise the effect of varying staining and illumination.

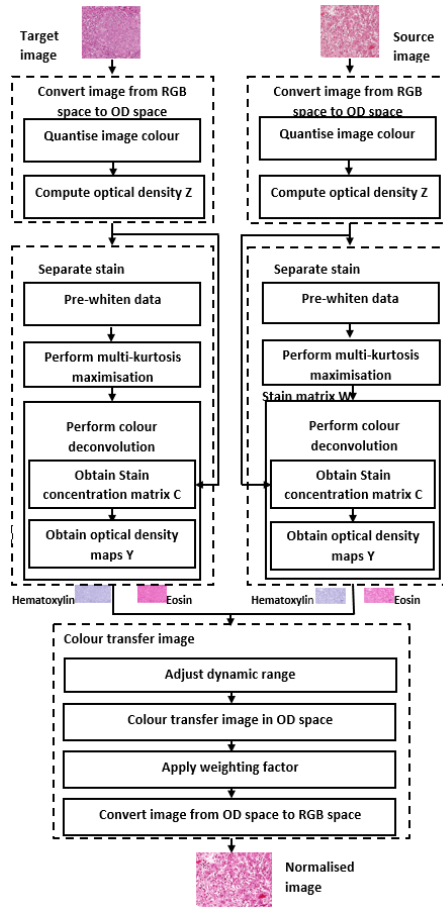


Fig. 2 Stain normalisation process

### 3.1.1 Stain separation

Stain separation is performed to computationally estimate the amount of stain in an histology image based on its chromatic and luminous response [70, 23, 67, 24]. Octree colour quantisation algorithm is combined with Beer-Lambert law, and a modified blind source separation algorithm are used during stain separation stage. Note: We modified the blind source separation [63]. As compared to blind source separation method in [63], our blind source separation method uses spatial-spectral Schrodinger-Eigen vectors and Eigen values during pre-whitening process to determine the set of stain vectors and stain matrix (3 X 3). By making use of Schrodinger-Eigen vectors, we are able to incorporate texture and spatial information [10]. Our Stain separation algorithm involves three steps that comprise, (a) Estimate stain, (b) Obtain the stain vectors, (c)

Perform colour deconvolution. The three steps are described in the section below:-

- (a) Estimate stain
  - (i) Colour quantisation

Colour quantisation is the process of choosing  $M$  colours from  $I$  to construct a quantised colour space,  $\bar{I}$ , where  $\bar{I} = \bar{I}_j, j = \{1, 2, \dots, M\}$  [77]. Colour quantisation is introduced due its ability to enhance the image [8, 29, 59, 77, 36]. Quantisation algorithms can be categorised into two classes: uniform quantisation and non-uniform quantisation algorithm. The difference between the two lies in the way they handle colour space. For example, in uniform colour quantisation, the colour space is broken into equal sized regions where the number of regions is less than or equal to  $M$ . Non-uniform colour quantisation breaks the colour space depending on distribution of colours in the image. In terms of performance, non-uniform algorithms outperforms uniform quantisation algorithms. Non-uniform algorithms are easy to implement and offer consistent results [77, 49]. We choose Octree algorithm that is highly non-uniform to extract the classes [72]. The colour classes are then labelled and the labelled image is converted into optical density. The octree quantisation algorithm quantises the colour following the steps below: given an image  $I$ , every colour in the image is stored in an octree of depth 8 (every leaf at depth 8 represents a distinct colour). A limit of  $M$  (in this case  $M = 256$ ) leaves are placed on the tree. The colours are filtered based on the limit of  $M$  leaves set. For example colour in the tree can result in two outcomes: (a) if there are less than  $M$  leaves the colour is filtered down the tree until either it reaches some leaf node that has an associated representative colour or it reaches the leaf node representing its unique colour, and (b) If there are greater than  $M$  leaves in the tree some set of leaves in the tree must be merged (their representative colours averaged) together and a new representative colour stored in their parent.

Leaves are selected and merged following two criteria: (a) Reducible nodes that have the largest depth in the tree should be chosen first. They represent colours that lie closest together, (b) If there is more than one group of leaves at the maximum depth the algorithm either merges leaves that represent the fewest number of pixels to keep the error small or reduce leaves that represent the most pixels. In this case large areas will be uniformly filled in a slightly wrong colour while maintaining detailed shading.

Once the entire image has been processed, the colour in the image is quantised by simply filtering each colour down the tree until a leaf is hit.

- (ii) Convert the colour quantised image to optical density space

$$Z = \text{LOG}(I_0) - \text{LOG}(I + 1) \quad (3)$$

where  $\mathbf{I}_0$  is the illuminating light associated with the quantised RGB Image I.

(b) Obtain stain vector

The stain vector is obtained following the steps below

(i) Pre-whiten data following the steps below [31].

Compute the correlation matrix using Eq. 4:-

$$\mathbf{R} = \mathbf{R}_s \times \frac{\mathbf{R}'_s}{L_s} \quad (4)$$

where  $\mathbf{R}_s$  is the received sigma.  $\mathbf{R}'_s$  is the conjugate transpose of  $\mathbf{R}_s$ .  $L_s$  is the length of the sigma.

Obtain Eigen values U and Eigen vectors V using Schroedinger Potential in [10,9], rather than the common Eigen values and vector. Note: As compared to blind source separation method in [63], our blind source separation method uses spatial-spectral Schroedinger-Eigen vectors and Eigen values during pre-whitening process to determine the set of stain vectors and stain matrix(3 X 3).

Perform Eigen values decomposition using shur decomposition method in Eq. 5. Shur decomposition method is expressed as [63,21,25,75]:

$$\tilde{\mathbf{A}}_{\mathbf{Y}\mathbf{Y}}^{(d)} = \mathbf{Q}\mathbf{U}\mathbf{Q}^{-1} \quad (5)$$

where Q is a unitary matrix,  $\mathbf{Q}^{-1}$  is a conjugate transpose of Q. U is the diagonal matrix with real entries. Denoising the estimate  $\tilde{\mathbf{A}}_{\mathbf{Y}\mathbf{Y}}$  of  $\mathbf{A}_{\mathbf{Y}\mathbf{Y}} = \mathbf{E}(\mathbf{Y}\mathbf{Y}^{-1})$  is performed by estimating the noise variance. The noise variance is estimated by obtaining the average of the smallest Eigen values. The denoised image is obtained using equation below:-

$$\tilde{\mathbf{A}}_{\mathbf{Y}\mathbf{Y}}^{(d)} = \tilde{\mathbf{A}}_{\mathbf{Y}\mathbf{Y}} - \sigma_n^2 \mathbf{I}_q \quad (6)$$

Given that ideally  $\tilde{\mathbf{A}}_{\mathbf{Y}\mathbf{Y}}^{(d)} = \mathbf{H}\mathbf{H}^{-1}$ , the non-zero part of the matrix  $\mathbf{L}\mathbf{D}^{\frac{1}{2}}$  equals H up to a unitary ambiguity matrix:

$$\tilde{\mathbf{L}} = \mathbf{H}\mathbf{U} \quad (7)$$

where U is a unitary matrix.  $\tilde{\mathbf{L}}$  is constructed as a matrix containing p large norm. Column in  $\mathbf{L}\mathbf{D}^{\frac{1}{2}}$ . Then, the received signal is pre-filtered using:

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{L}}\# \mathbf{Y} \quad (8)$$

where # denote matrix pseudo-inverse.

(ii) Perform multi-kurtosis maximisation

Multi user kurtosis (MUK) can be defined by  $\mathbf{M}_{q \times p}$  the following set of  $q \times p$  complex matrices:

$$\mathbf{M}_{q \times p} = \{\mathbf{D} \in \mathbf{C}^{q \times p} \text{ such that } \mathbf{D}^H \mathbf{D} = \mathbf{I}_p\}$$

We perform MUK following the steps below:-

Compute the gradient of  $F(\mathbf{G})$  using the formula in Eq. 9 with respect to  $\mathbf{W}$ . Then later obtain an adaptive stochastic-gradient algorithm. Let

$F(\mathbf{G}) = \sum_{j=1}^p \text{sign}(\mathbf{K}(z_j)) \mathbf{K}(z_j) = \text{sign}(\mathbf{K}_a) \sum_{j=1}^p \left( E|z_j|^4 - 2E^2|z_j|^2 - \left| E(z_j^2) \right|^2 \right)$  And by invoking C2 with the assumption that symmetrical inputs  $(E(\mathbf{a}^2(\mathbf{K})) = 0)$ , the gradient of  $F(\mathbf{G})$  with respect to  $\mathbf{W}$  is given by:

$$\nabla(F(\mathbf{G})) = 4\text{sign}(\mathbf{K}_a) \sum_{j=1}^p \left( E|z_j(k)|^2 z_j(k) \mathbf{Y}^*(k) \right) \quad (9)$$

Update  $\mathbf{W}(k)$  in the direction of the instantaneous gradient using Eq. 10.

$$\mathbf{W}'(k+1) = \mathbf{W}(k) + \mu \text{sign}(\mathbf{K}_a) \mathbf{Y}^*(k) \mathbf{Z}(k) \quad (10)$$

where  $\mu = 4$ .  $\mu$  is a constant.

$\mathbf{Z}(k) = \left[ |z_1(k)|^2 z_1(k) \dots |z_p(k)|^2 z_p(k) \right]$  Satisfy the orthogonality constraint at the next iteration of the algorithm by utilising the unitary constraint using Gram-Schmidt orthogonalisation in Eq. 11.

$$\mathbf{G} \in \mathbf{M}_{p \times p} \leftrightarrow \mathbf{G}^H(k+1) \mathbf{G}(k+1) = \mathbf{I}_p \quad (11)$$

Undo the remaining unitary mixture which makes the equaliser  $\mathbf{W}$  of size  $p \times p$ , hence saving computational cost. Then, to satisfy the constraint in Eq. 11, it suffices to satisfy  $\mathbf{W}(k+1) \in \mathbf{M}_{p \times p}$ , i.e.

$$\mathbf{W}^H(k+1) \mathbf{W}(k+1) = \mathbf{I}_p \quad (12)$$

Since there is no guarantee that  $\mathbf{W}'(k+1)$  will satisfy the constraint in Eq. 12, we have to transform  $\mathbf{W}'(k+1)$  into a unitary

$$\mathbf{W}(k+1) = f(\mathbf{W}'(k+1)) \quad (13)$$

By treating  $\mathbf{W}(k+1)$  as a  $p \times p$  matrix satisfying Eq. 12 that is as close as possible to  $\mathbf{W}'(k+1)$  in the Euclidean sense using the

iterative procedure that satisfies the following criterion successfully for  $j = 1, \dots, p$ :

$$\begin{cases} \min_{\mathbf{W}_j} \nabla(\mathbf{W}_j) = (\mathbf{W}_j - \mathbf{W}'_j)^H (\mathbf{W}_j - \mathbf{W}'_j) \\ \text{subject to : } \mathbf{W}_l^H \mathbf{W}_j = \sigma_{lj}, l = 1, \dots, j \end{cases} \quad (14)$$

where  $\mathbf{X}^2 = \mathbf{X}^H \mathbf{X}$  is squared Euclidean norm of vector  $\mathbf{X}$ ,  $\sigma_{lj}$  is kronecker delta,  $\mathbf{W}'_j$  is obtained from

$$\mathbf{W}' = [\mathbf{W}'_1, \dots, \mathbf{W}'_p] \quad (15)$$

To solve Eq. 14, we first construct the Lagrangian of  $\nabla(\mathbf{W}_j)$  using Eq. 16

$$\begin{aligned} L_{\nabla}(\mathbf{W}_j, \lambda_j, \mu, \nu) = & \nabla(\mathbf{W}_j) - \lambda_j (\mathbf{W}_j^H \mathbf{W}_j - 1) \\ & - \sum_{l=1}^{j-1} \mu_{lj} \text{Re}(\mathbf{W}_l^H \mathbf{W}_j) - \sum_{l=1}^{j-1} \nu_{lj} \text{Im}(\mathbf{W}_l^H \mathbf{W}_j) \end{aligned} \quad (16)$$

Where  $\lambda_j$ ,  $\mu_{lj}$ , and  $\nu_{lj}$  are real scalar parameters. Setting the gradient of  $L_{\nabla}(\mathbf{W}_j, \lambda_j, \mu, \nu)$  with respect to  $\mathbf{W}_j$  to zero  $\frac{\partial L_{\nabla}}{\partial \mathbf{W}_j} = 0$ , we obtain the following equations for each  $j$ :

$$\mathbf{W}_j - \mathbf{W}'_j - \lambda_j \mathbf{W}_j - \sum_{l=1}^{j-1} \mathbf{B}_{lj} \mathbf{W}_j = \mathbf{0} \quad (17)$$

Where  $\mathbf{B}_{lj} = \frac{1}{2}(\mu_{lj} + i\nu_{lj})$ . From Eq. 17, we obtain

$$\begin{cases} \lambda_j = 1 - \mathbf{W}_j^H \mathbf{W}'_j \\ \mathbf{B}_{lj} = -\mathbf{W}_l^H \mathbf{W}'_j \end{cases} \text{ which gives} \\ (1 - \lambda_j) \mathbf{W}_j = \mathbf{W}'_j - \sum_{l=1}^{j-1} (\mathbf{W}_l^H \mathbf{W}'_j) \mathbf{W}_l \quad (18)$$

According to Eq. 18

$$\mathbf{W}_j \left( \mathbf{W}'_j - \sum_{l=1}^{j-1} (\mathbf{W}_l^H \mathbf{W}'_j) \mathbf{W}_l \right)$$

(c) Perform colour deconvolution

Colour deconvolution scheme is used to transform quantised RGB channels of a histology image into a density map for each stain.

(i) Determine the stain concentration of individual stains:

$$\mathbf{C} = \mathbf{W} \times \mathbf{Z} \quad (19)$$

where  $\mathbf{Z}$  is the optical density map obtained in Eq. 3,  $\mathbf{W}$  is the stain matrix obtained in Eq. 18.



The optical density maps of each stain are obtained by taking the columns of stain matrix  $\mathbf{W}$  and taking rows of the stain concentration matrix using the equation below:

$$\mathbf{Y} = \mathbf{W}\mathbf{C} \quad (20)$$

### 3.1.2 Colour transfer

Colour transfer is performed to minimise the effect of varying staining and illumination [74]. We modified the colour transfer method in [86,57,54] to achieve better stain normalisation results. We also introduce a weighting factor in the colour to easily adjust the contrast and brightness of the image. Our colour transfer involves 6 steps as follows:

- (a) Let  $\mathbf{W}$  be the stain matrix and  $\mathbf{C}$  be the coefficient matrix from Eq. 20

$$\mathbf{Y}_{source} = \mathbf{W}_{source}\mathbf{C}_{source}, \mathbf{Y}_{target} = \mathbf{W}_{target}\mathbf{C}_{target} \quad (21)$$

- (b) Adjust the dynamic range of  $\mathbf{C}_{source}$  so as it matches that of  $\mathbf{C}_{target}$  to obtain a normalized source stains  $\mathbf{C}_{snorm}$ , by Calculating the 95 percentile of each row of the  $\mathbf{C}_{source}$  and  $\mathbf{C}_{target}$  as  $\mathbf{C}_{s-extreme}(i)$  and  $\mathbf{C}_{t-extreme}(i)$  to obtain a scale matrix. The dynamic range is estimated with a pseudo maximum (95%) as it is robust. We used Eq. 22 instead of Eq. 23.

$$\mathbf{D}_{scale(i,1)} = \frac{\mathbf{C}_{s-extreme}(i)}{\mathbf{C}_{T-extreme}(i)} \quad (22)$$

Instead of

$$\mathbf{D}_{scale(i,1)} = \frac{\mathbf{C}_{t-extreme}(i)}{\mathbf{C}_{s-extreme}(i)} \quad (23)$$

where  $i = 1, \dots, r$  which is the row number in a coefficient matrix  $\mathbf{C}$ .

- (c) Obtain the scale row of  $\mathbf{C}_{source}$  using Eq. 24 instead of Eq. 25.

$$\mathbf{C}_{s-scaled}(i,:) = \mathbf{C}_{target}(i,:) \mathbf{D}_{scale(i,1)} \quad (24)$$

Instead of

$$\mathbf{C}_{s-scaled}(i,:) = \mathbf{C}_{source}(i,:) \mathbf{D}_{scale(i,1)} \quad (25)$$

- (d) Colour transfer image in optical density space.

$$\mathbf{Y}_{transferred} = \mathbf{W}_{target}\mathbf{C}_{s-scaled} \quad (26)$$

- (e) Apply a weighting factor to a colour transferred image

$$NewY_{transferred} = T \times Y_{transferred} \quad (27)$$

where  $T \geq 1$ , T can be adjusted to improve the brightness and contrast of the normalised image. The default values of T for both H&DAB and H&E stain images is 1.

- (f) Convert the normalised image from optical density space to RGB space

$$I_{normalised} = I_0 \exp(-NewY_{transferred}) \quad (28)$$

where  $I_0 = 255$

## 4 Experiments

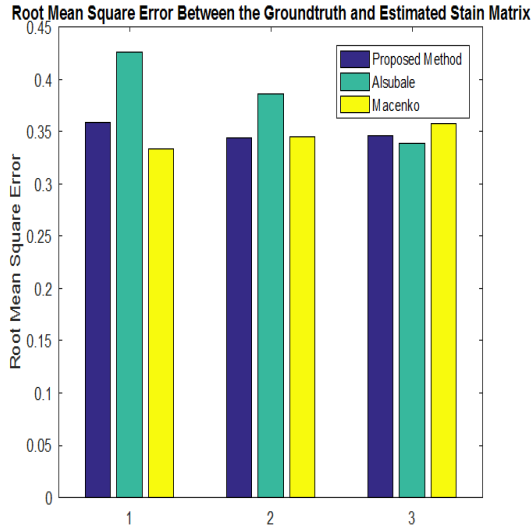
We tested the proposed method using 48 H&DAB image slides with dimensions of  $1920 \times 1080$  pixels from the mouse cortex, and 100 H&E stain images with dimensions of  $1539 \times 1376$  pixels provided by Dr Abhishek Vahadane. We also used 22 H&E histology images with the dimension of  $2000 \times 2000$  pixels of different datasets (i.e. breast dataset, colon dataset and lung dataset) provided by Alsubaie N et al. [1] under the support information section. H&E images came with ground truth. Please refer to [1] for more information about the ground truth. H&DAB images were stained on different dates. H&E images are from different scanners (Hamamatsu scanner and Aperio scanner). To increase reliability, more than one sample of slides were extracted from each mouse. 48 H&DAB image slides were extracted from different mice. As compared to other colour transfer methods, a weighting factor is introduced in the colour transfer stage. The weighting factor improves the brightness and contrast of the image as shown in Figure 5

### 4.1 Validation and evaluation of stain normalisation

We evaluated results from stain separation stage and the colour transfer stage both qualitatively and quantitatively. The accuracy of stain separation is evaluated using both estimated stain matrix and density maps. This is because stain matrix represents the principal colour of each stain and each vector in the stain matrix represents RGB values within an optical density space for each stain colour [86, 1]. An accurate stain matrix leads to better stain estimation. In this paper, ground truth data is used during the evaluation of the estimated stain matrix. During this stage, the root mean square error between the estimated stain vector and the ground truth is used. Bar charts are used during the comparison of our method with existing methods. The density maps represent the amount of stain absorbed at each pixel. A similarity quality metric between the results obtained using the proposed method and the ground truth is used to evaluate the accuracy of the proposed method. Box plots and a bar

graph are used to compare stain separation obtained using our methods with that of other methods (Figure 3). We used the median for easy comparison. The quantitative results shown in Figure 4 and Figure 3, show the superior performance of the proposed method over others.

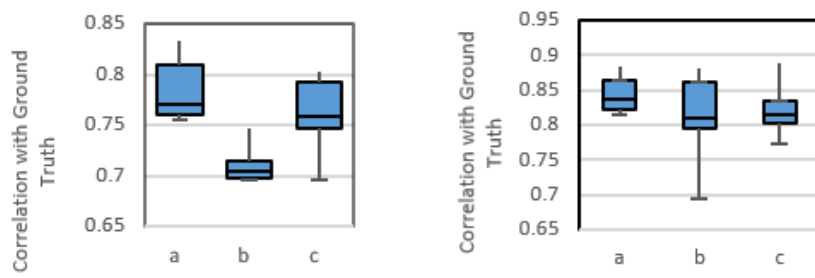
The accuracy of colour transfer is evaluated both qualitatively as shown in Figure 6(f), Figure 7(f). Our stain normalisation method produces better results compared to other methods.



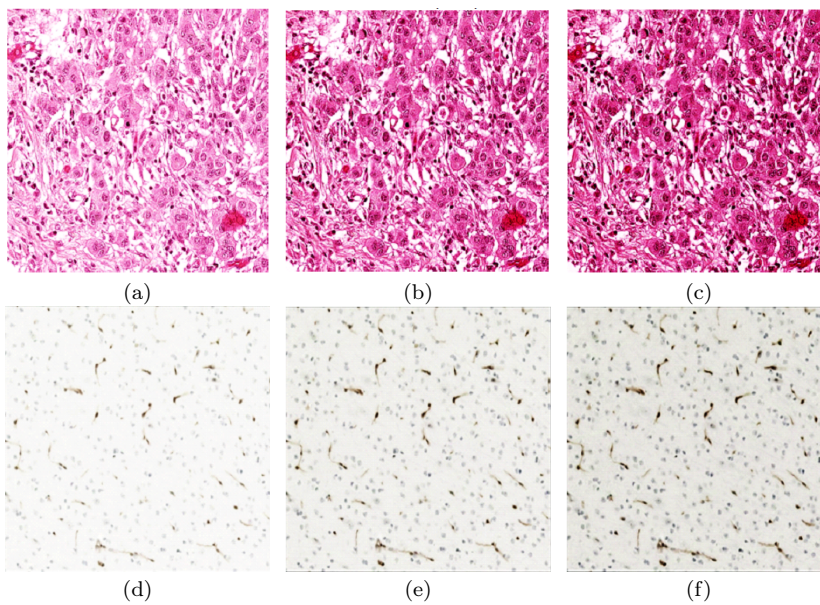
**Fig. 3** Root mean square error between the estimated stain matrices and the ground truth data for all methods on all datasets. (1) breast dataset, (2) Colon dataset, and (3) Lungs dataset

**Table 1** SSIM Index between the stain matrices and the ground truth. The last two columns show the median of SSIM index for all methods.

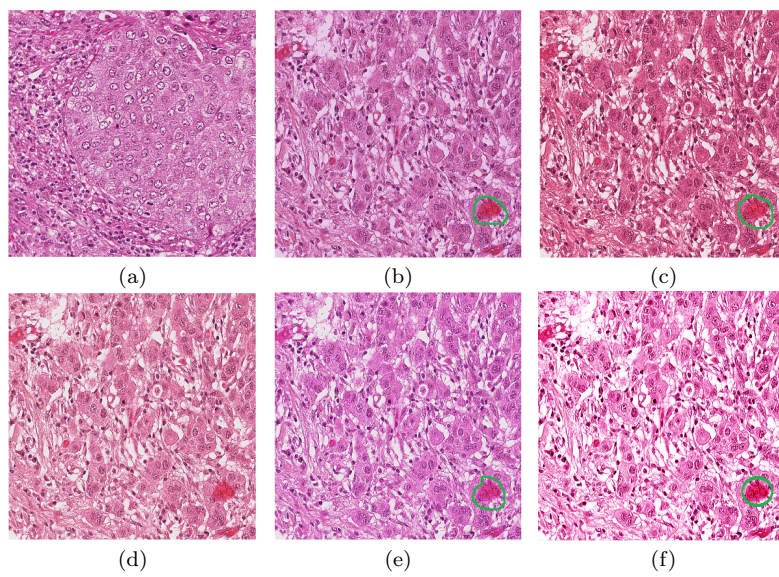
Method	Breast Dataset		Lung Dataset		Colon Dataset		Median	
	H	E	H	E	H	E	H	E
Alsubaie N et al., [1]	0.697	0.879	0.722	0.797	0.714	0.758	0.705	0.811
Macenko et al., [57]	0.803	0.830	0.719	0.847	0.759	0.787	0.759	0.814
Proposed Method	0.827	0.878	0.764	0.825	0.763	0.829	0.770	0.838



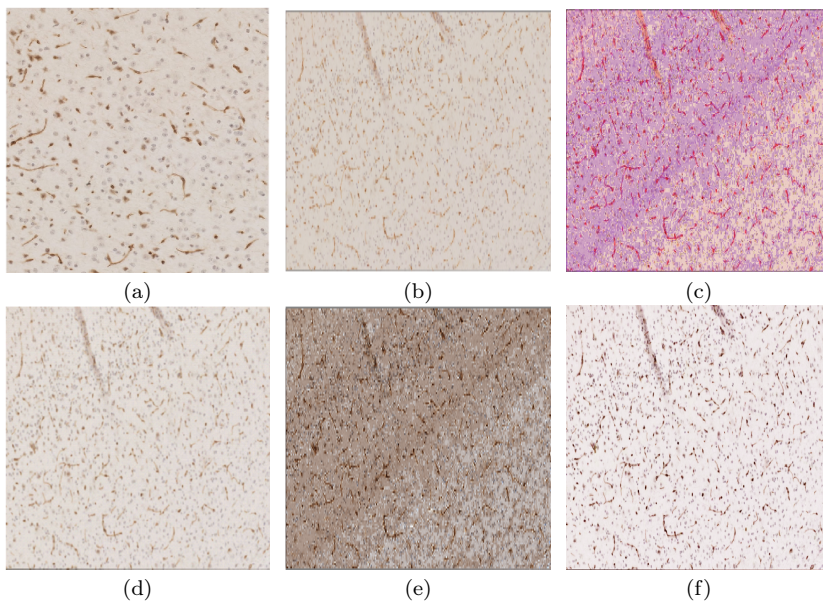
**Fig. 4** Correlation between the density maps and the ground truth with associated SSIM Index for each method with respect to H(Left) and E (right) stains in all three datasets (a) proposed method, (b) Alsubaie N et al. [1] and (c) Macenko et al. [57]



**Fig. 5** H&E and H&DAB normalised histology images at different weighting factors. First row are H&E stain images normalised at weighting factor 1, 2 and 3 respectively, and second row are H&DAB stain images normalised at weighting factors 1, 2, 3 respectively

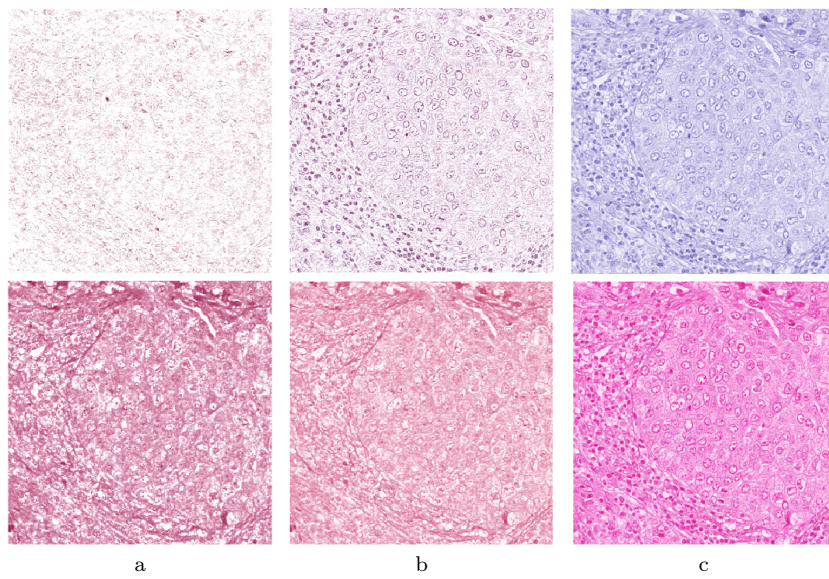


**Fig. 6** Comparison of different stain normalisation methods on H&E stain images.(a) target image,(d) source image, (b) Reinhard et al., [68],(c) Khan et al., [41], (e) Macenko et al., [57], (f) Proposed Method

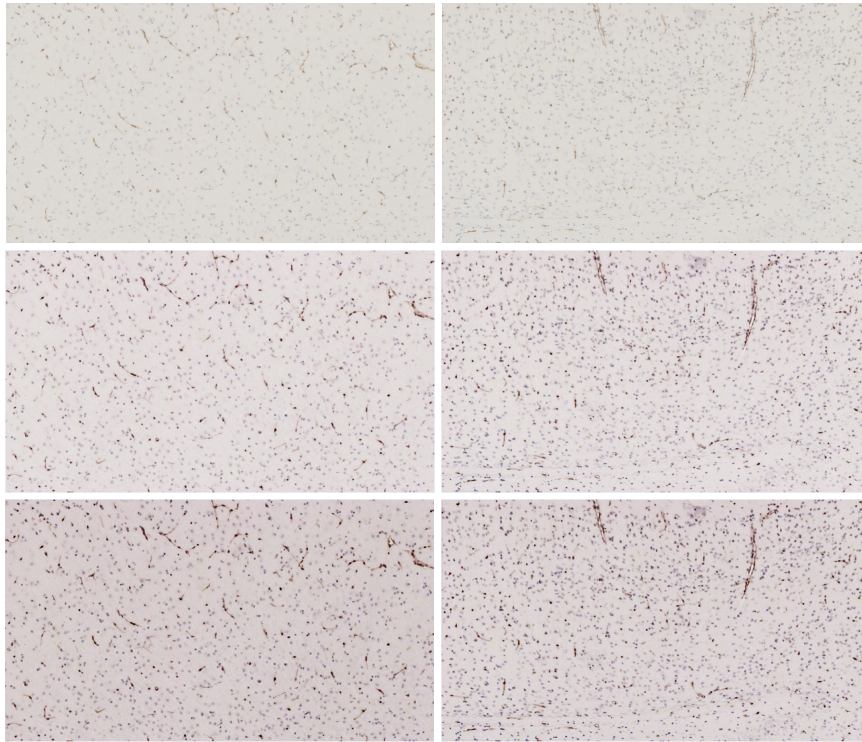


**Fig. 7** Comparison of different stain normalisation methods on H&DAB stain images.(a) target image, (d) source image, (b) Reinhard et al., [68],(c) Khan et al., [41], (e) Macenko et al., [57], (f) Proposed Method

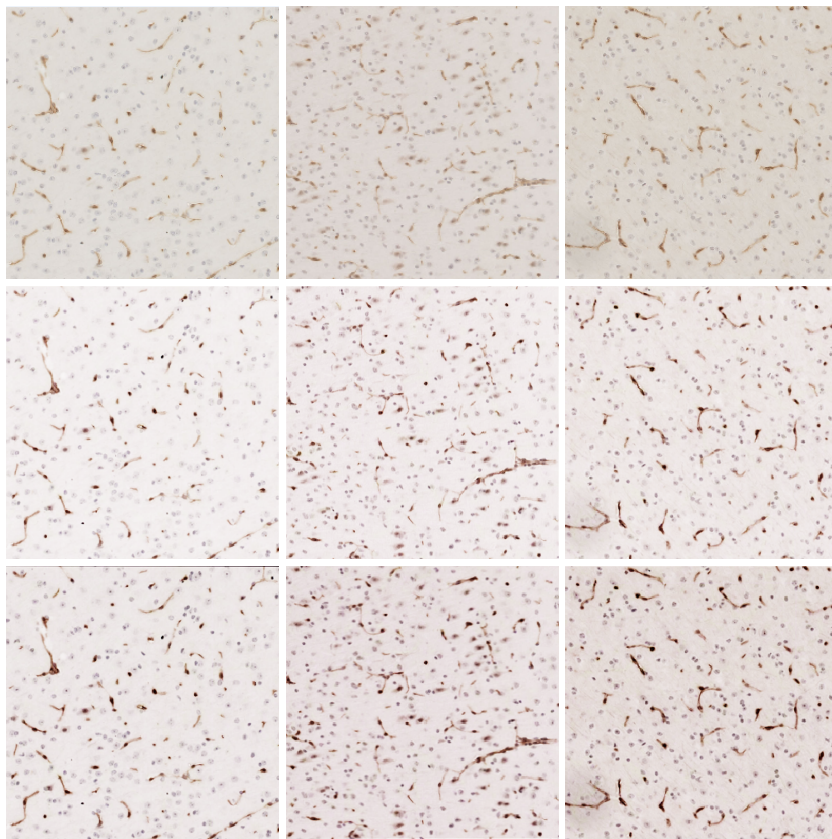




**Fig. 8** Stain separation after application of SVD, statistical blind source separation ICA and proposed method on H&E images. (a) statistical blind source separation ICA H&E images. (b) SVD H&E stains. (c) Proposed method



**Fig. 9** Different image slides normalized at different factors. First row are original image slides from cortex, second row are images normalized at factor 4, and third row are images normalized at factor 5



**Fig. 10** Different image slides normalized at different factors. First row are original image slides from cortex, second row are images normalized at factor 2, and third row are images normalized at factor 3.



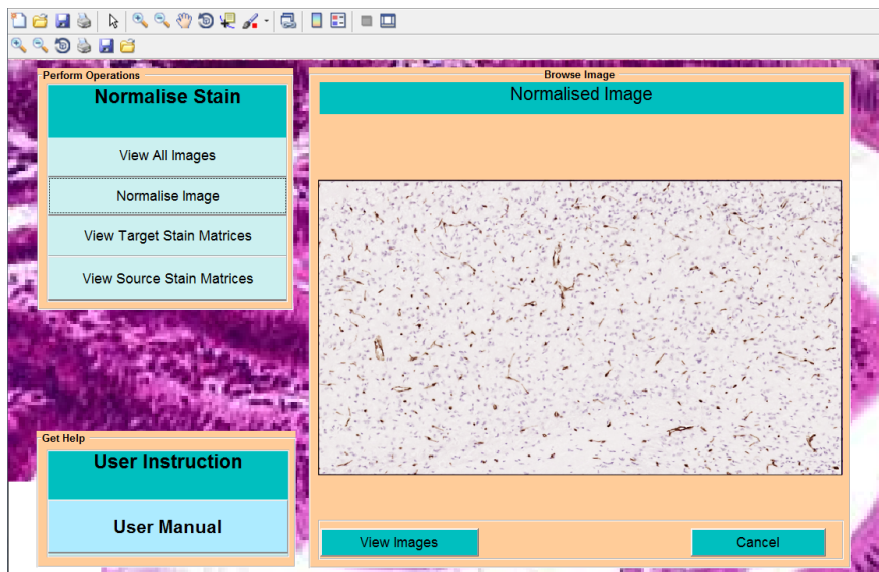


Fig. 11 System interface with normalised H&DAB source image

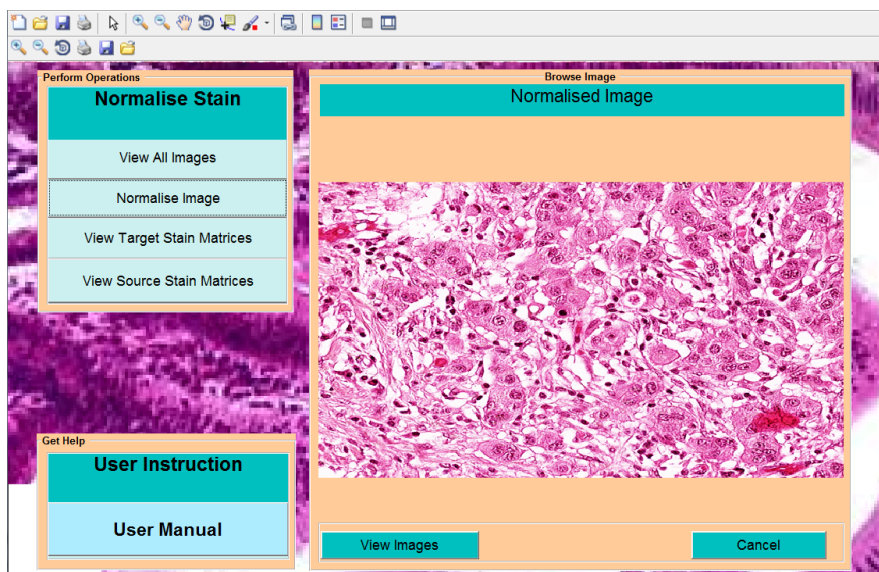


Fig. 12 System interface with normalised H&E source image

## 4.2 Discussion

The qualitative and the overall quantitative results of stain separation in Figure 8, Figure 4 and Figure 3, show that the proposed method outperforms all the other methods. The similarity metric in Figure 4 above shows that the proposed method outperforms Alsubaie et al., [1] by 4-13% for H for breast dataset, colon dataset, and the lung dataset. However Alsubaie's method [1] outperforms the proposed method by 0.1% for E in breast datasets. The proposed method still outperforms Alsubaie's method [1] by 3-7% for E in colon and lung dataset. In comparison to Macenko's method [57], the proposed method outperforms Macenko's method [57] by 2-4% and 4-5% for H&E in all three datasets (breast and colon dataset). Like Alsubaie's method [1], Macenko's method [57] also outperforms the proposed method in estimating Eosin stain. Macenko's method [57] outperforms the proposed method by 2.2% for E in the lung dataset. This is due to usage of low weighting factor. Figure 3 shows that proposed method outperforms Alsubaie's method [1] for estimated stain matrix in both the breast dataset and lung dataset. The proposed method also outperforms Macenko's method [57] for estimated stain matrix in the colon dataset but achieves the same result for estimated matrix as Macenko's method [57] in the lung dataset. Macenko's method [57] achieves better results for estimated stain matrix in breast datasets as compared to all methods.

The colour transfer stage is the final stage of stain normalisation process. The qualitative results of the color transfer stage show that the proposed method can be applied to both H&E stain images and H&DAB images (Figure 6 and Figure 7). Khan's NMF method [41] and Macenko's SVD method [57] produces unrealistic stains when applied to H&DAB images. Khan's NMF method [41] and Macenko's SVD method [57] also introduce artefacts when applied to H&DAB images. Reinhard's method produces slightly better results when applied on H&DAB but still not better than the proposed method. Compared to all methods, our method produces satisfactory results when applied on both H&DAB and H&E images (Figure 10, Figure 9, Figure 12, Figure 11). With the introduction of a weighting factor in the colour transfer stage, the contrast and brightness of the image can be adjusted. An increase in the weighting factor results into increase in brightness of the image (Figure 5). Further work is needed to evaluate the proposed method at different weighting factors.

## 5 Conclusion

Stain normalisation is an essential step to remove inherent in-homogeneity in histopathological images for diagnosing disease and its progression. A potential drawback of such a technique is that tissue structures and texture in the original image could be distorted after stain normalisation. In this paper, a hybrid approach for stain normalisation has been developed, based on an octree colour quantisation algorithm combined with Beer-Lambert law, a

modified blind source separation algorithm, and a modified colour transfer approach. The hybrid method consists of two stages: the stain separation stage and colour transfer stage. An octree colour quantisation algorithm combined with Beer-Lambert law, and a modified blind source separation algorithm are used during stain separation stage to computationally estimate the amount of stain in an histology image based on its chromatic and luminous response. A modified colour transfer algorithm is used during the colour transfer stage to minimise the effect of varying staining and illumination. The method is capable of addressing colour variation in both H&DAB (Haemotoxylin and Diaminobenzidine) and H&E (Haemotoxylin and Eosin) stain image. Unlike other colour transfer, our normalisation algorithm introduces weighting factor which allows one to adjust the brightness and contrast of the normalised image. Both qualitative and quantitative validation demonstrate the superior performance of the proposed stain separation and colour normalisation technique.

### Acknowledgements

Faiza Bukenya acknowledges the Islamic Development Bank (IDB) Merit scholarship scheme and University of Nottingham for their generous support. My sincere thanks goes out to Dr. Marie-Christine Pardon and Culi Nerissa (School of Life Science, University of Nottingham) for providing H&DAB stained image slides. My warm thanks goes out to Dr. Abhishek Vahadane for providing H&E stained image slides.

### Conflict of Interest

There is no competing interests or exclusive licenses used in preparing this manuscript.

### References

1. Najah Alsubaie, Nicholas Trahearn, Shan E. Ahmed Raza, David Snead, and Nasir M. Rajpoot. Stain deconvolution using statistical analysis of multi-resolution stain colour representation. *PLOS ONE*, 12(1):1–15, 01 2017.
2. H. Andrews and C. Patterson. Singular value decompositions and digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(1):26–53, February 1976.
3. Louis Asiedu, Atinuke Adebajji, Francis T Oduro, and Felix O Mettle. Statistical assessment of pca/svd and fft-pca/svd on variable facial expressions. 2016.
4. Mohammadreza Babae, Stefanos Tsoukalas, Maryam Babae, Gerhard Rigoll, and Mihai Datcu. Discriminative nonnegative matrix factorization for dimensionality reduction. *Neurocomputing*, 173:212–223, 2016.
5. Michael W Berry, Nicolas Gillis, and François Glineur. Document classification using nonnegative matrix factorization and underapproximation. In *2009 IEEE International Symposium on Circuits and Systems*, pages 2782–2785. IEEE, 2009.

6. Laurens Bogaardt, Romulo Goncalves, Raul Zurita-Milla, and Emma Izquierdo-Verdiguier. Dataset reduction techniques to speed up svd analyses on big geo-datasets. *ISPRS International Journal of Geo-Information*, 8(2):55, 2019.
7. Aymen Bougacha, Ines Njeh, Jihene Boughariou, Omar Kammoun, Kheireddine Ben Mahfoudh, Mariem Dammak, Chokri Mhiri, and Ahmed Ben Hamida. Rank-two nmf clustering for glioblastoma characterization. *Journal of healthcare engineering*, 2018, 2018.
8. Peter D Burns and Roy S Berns. Quantization in multispectral color image acquisition. In *Color and Imaging Conference*, volume 1999, pages 32–35. Society for Imaging Science and Technology, 1999.
9. Nathan D Cahill, Selene E Chew, and Paul S Wenger. Spatial-spectral dimensionality reduction of hyperspectral imagery with partial knowledge of class labels. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI*, volume 9472, page 94720S. International Society for Optics and Photonics, 2015.
10. Nathan D. Cahill, Wojciech Czaja, and David W. Messinger. Schroedinger eigenmaps with nondiagonal potentials for spatial-spectral clustering of hyperspectral imagery. 2014.
11. Duane Carey, VN Wijayathunga, Andrew J Bulpitt, and Darren Treanor. A novel approach for the colour deconvolution of multiple histological stains. In *Proceedings of the 19th Conference of Medical Image Understanding and Analysis*, pages 156–162. BMVA, 2015.
12. REYESAÑRALDASORO C.C., WILLIAMS L.J., AKERMAN S., KANTHOU C., and TOZER G.M. An automatic algorithm for the segmentation and morphological analysis of microvessels in immunostained histological tumour sections. *Journal of Microscopy*, 242(3):262–278.
13. R Celis, D Romo, and E Romero. Blind colour separation of h&e stained histological images by linearly transforming the colour space. *Journal of microscopy*, 260(3):377–388, 2015.
14. Richard M Charles, Kye M Taylor, and James H Curry. Nonnegative matrix factorization applied to reordered pixels of single images based on patches to achieve structured nonnegative dictionaries. *arXiv preprint arXiv:1506.08110*, 2015.
15. Francesco Ciompi, Oscar Geessink, Babak Ehteshami Bejnordi, Gabriel Silva de Souza, Alexi Baidoshvili, Geert Litjens, Bram van Ginneken, Iris Nagtegaal, and Jeroen van der Laak. The importance of stain normalization in colorectal tissue classification with convolutional networks. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 160–163. IEEE, 2017.
16. Dimah Dera, Nidhal Bouaynaya, and Hassan M Fathallah-Shaykh. Automated robust image segmentation: level set method using nonnegative matrix factorization with application to brain mri. *Bulletin of mathematical biology*, 78(7):1450–1476, 2016.
17. Viet-Hang Duong, Yuan-Shan Lee, Bach-Tung Pham, Pham The Bao, and Jia-Ching Wang. Nmf-based image segmentation. In *2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, pages 1–2. IEEE, 2016.
18. Ahmed Ebied, Eli Kinney-Lang, Loukianos Spyrou, and Javier Escudero. Evaluation of matrix factorisation approaches for muscle synergy extraction. *Medical engineering & physics*, 57:51–60, 2018.
19. Graham D Finlayson, Bernt Schiele, and James L Crowley. Comprehensive colour image normalization. In *European conference on computer vision*, pages 475–490. Springer, 1998.
20. Xiao Fu, Kejun Huang, Nicholas D Sidiropoulos, and Wing-Kin Ma. Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications. *arXiv preprint arXiv:1803.01257*, 2018.
21. Zielke G. Horn, r. a.; johnson, c. r., matrix analysis. cambridge etc., cambridge university press 1985. xiii, 561 s., Åč 35.00. isbn 0âĀR521âĀR30586âĀR1. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift fĀĳr Angewandte Mathematik und Mechanik*, 67(3):212–212.
22. Luis Garcia-Torres, Juan J Caballero-Novella, David Gomez-Candon, and Ana Isabel De-Castro. Semi-automatic normalization of multitemporal remote images based on vegetative pseudo-invariant features. *PloS one*, 9(3):e91275, 2014.

23. Milan Gavrilovic, Jimmy C Azar, Joakim Lindblad, Carolina Wählby, Ewert Bengtsson, Christer Busch, and Ingrid B Carlbom. Blind color decomposition of histological images. 2013.
24. Biswajoy Ghosh, Sri Phani Krishna Karri, Debdoot Sheet, Hrushikesh Garud, Arindam Ghosh, Ajoy K Ray, and Jyotirmoy Chatterjee. A generalized framework for stain separation in digital pathology applications. In *India Conference (INDICON), 2016 IEEE Annual*, pages 1–4. IEEE, 2016.
25. G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 2013.
26. David Guillaumet, Bernt Schiele, and Jordi Vitria. Analyzing non-negative matrix factorization for image classification. In *Object recognition supported by user interaction for service robots*, volume 2, pages 116–119. IEEE, 2002.
27. Lili Guo, Dan Xu, and Zhenping Qiang. Background subtraction using local svd binary pattern. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 86–94, 2016.
28. Azam Hamidinekoo and Reyer Zwiggelaar. Stain colour normalisation to improve mitosis detection on breast histology images. In M. Jorge Cardoso, Tal Arbel, Gustavo Carneiro, Tanveer Syeda-Mahmood, João Manuel R.S. Tavares, Mehdi Moradi, Andrew Bradley, Hayit Greenspan, João Paulo Papa, Anant Madabhushi, Jacinto C. Nascimento, Jaime S. Cardoso, Vasileios Belagiannis, and Zhi Lu, editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 213–221, Cham, 2017. Springer International Publishing.
29. Markku Hauta-Kasari, Jussi Parkkinen, Timo Jääskeläinen, and Reiner Lenz. Multi-spectral texture segmentation based on the spectral cooccurrence matrix. *Pattern Analysis Applications*, 2:275–284, 1999.
30. R. A. Hoffman, S. Kothari, and M. D. Wang. Comparison of normalization algorithms for cross-batch color segmentation of histopathological images. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 194–197, Aug 2014.
31. Chaoqun Hong, Jun Yu, Dacheng Tao, and Meng Wang. Image-based three-dimensional human pose recovery by multiview locality-sensitive sparse retrieval. *IEEE Transactions on Industrial Electronics*, 62(6):3742–3751, 2014.
32. Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933.
33. A. Hyvärinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Netw.*, 13(4-5):411–430, May 2000.
34. Anil K Jain. *Fundamentals of digital image processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
35. Andrew Janowczyk, Ajay Basavanahally, and Anant Madabhushi. Stain normalization using sparse autoencoders (stanosa): Application to digital pathology. *Computerized Medical Imaging and Graphics*, 57:50 – 61, 2017. Recent Developments in Machine Learning for Medical Imaging Applications.
36. Ellen C Jensen. Quantitative analysis of histological staining and fluorescence using imagej. *The Anatomical Record*, 296(3):378–381, 2013.
37. Zhongxiao Jia and Yanfei Yang. Modified truncated randomized singular value decomposition (mtrsvd) algorithms for large scale discrete ill-posed problems with general-form regularization. *Inverse Problems*, 34(5):055013, 2018.
38. Ehsanhosein Kalatehjari and Farzin Yaghmaee. A new reduced-reference image quality assessment based on the svd signal projection. *Multimedia Tools and Applications*, 77(19):25053–25076, 2018.
39. Ram Kumar Karsh, Rabul Hussain Laskar, et al. Robust image hashing through dwt-svd and spectral residual method. *EURASIP Journal on Image and Video Processing*, 2017(1):31, 2017.
40. Jakob Nikolas Kather, Cleo-Aron Weis, Alexander Marx, Alexander K Schuster, Lothar R Schad, and Frank Gerrit Zöllner. New colors for histology: optimized bivariate color maps increase perceptual contrast in histological images. *PloS one*, 10(12):e0145572, 2015.

41. Adnan Mujahid Khan, Nasir Rajpoot, Darren Treanor, and Derek Magee. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. *IEEE Transactions on Biomedical Engineering*, 61(6):1729–1738, 2014.
42. Haris Khan, Sofiane Mihoubi, Benjamin Mathon, Jean-Baptiste Thomas, and Jon Hardeberg. Hytexila: high resolution visible and near infrared hyperspectral texture images. *Sensors*, 18(7):2045, 2018.
43. Hongsuk H Kim and Gregory C Elman. Normalization of satellite imagery. *International Journal of Remote Sensing*, 11(8):1331–1347, 1990.
44. Konstantinos Konstantinides, Balas Natarajan, and Gregory S Yovanof. Noise estimation and filtering using block-based singular value decomposition. *IEEE Transactions on Image Processing*, 6(3):479–483, 1997.
45. Sonal Kothari, John H Phan, Richard A Moffitt, Todd H Stokes, Shelby E Hassberger, Qaiser Chaudry, Andrew N Young, and May D Wang. Automatic batch-invariant color segmentation of histological cancer images. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 657–660. IEEE, 2011.
46. Robert Krutsch and David Tenorio. Histogram equalization. 2011.
47. Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788, 1999.
48. Johannes Leuschner, Maximilian Schmidt, Pascal Fernsel, Delf Lachmund, Tobias Boskamp, and Peter Maass. Supervised non-negative matrix factorization methods for MALDI imaging applications. *Bioinformatics*, 11 2018.
49. Kang Li. Advanced image processing homework 1 color quantization : A median cut approach. 2004.
50. Li Li, Hirokazu Kameoka, and Shoji Makino. Discriminative non-negative matrix factorization with majorization-minimization. In *Hands-free Speech Communications and Microphone Arrays (HSCMA), 2017*, pages 141–145. IEEE, 2017.
51. X. Li and 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):1862–1873, July 2015.
52. Xingyu Li and Konstantinos 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):1862–1873, 2015.
53. Lin Liang, Lei Shan, Fei Liu, Ben Niu, and Guanghua Xu. Sparse envelope spectra for feature extraction of bearing faults based on nmf. *Applied Sciences*, 9(4):755, 2019.
54. Amit Sethi Lingdao Sha, Dan Schonfeld. Color normalization of histology slides using graph regularized sparse nmf, 2017.
55. Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen AWM van der Laak, Bram van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.
56. Weixiang Liu, Fei Peng, Shu Feng, Jiangsheng You, Ziqiang Chen, Jian Wu, Kehong Yuan, and Datian Ye. Semantic feature extraction for brain ct image clustering using nonnegative matrix factorization. In *International Conference on Medical Biometrics*, pages 41–48. Springer, 2008.
57. Marc Macenko, Marc Niethammer, James S Marron, David Borland, John T Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E Thomas. A method for normalizing histology slides for quantitative analysis. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 1107–1110. IEEE, 2009.
58. Izzati Muhimmah, Dhina Puspasari Wijaya, and Indrayanti. Color swapping to enhance breast cancer digital images qualities using stain normalization. *IOP Conference Series: Materials Science and Engineering*, 185(1):012029, 2017.
59. P. Naylor, M. LaÁf, F. Reyat, and T. Walter. Nuclei segmentation in histopathology images using deep neural networks. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 933–936, April 2017.
60. Jørgen Ødegård, Ulf Indahl, Ismo Strandén, and Theo HE Meuwissen. Large-scale genomic prediction using singular value decomposition of the genotype matrix. *Genetics Selection Evolution*, 50(1):6, 2018.

61. Sandra Ortega-Martorell, Paulo JG Lisboa, Alfredo Vellido, Rui V Simões, Martí Pumarola, Margarida Julià-Sapé, and Carles Arús. Convex non-negative matrix factorization for brain tumor delimitation from mrsi data. *PLoS One*, 7(10):e47824, 2012.
62. Pentti Paatero and Unto Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994.
63. C. B. Papadias. Globally convergent blind source separation based on a multiuser kurtosis maximization criterion. *IEEE Transactions on Signal Processing*, 48(12):3508–3519, Dec 2000.
64. LOCAL BINARY PATTERN. Content based image retrieval using gray.
65. Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.
66. Loic Peter, Diana Mateus, Pierre Chatelain, Noemi Schworm, Stefan Stangl, Gabriele Multhoff, and Nassir Navab. Leveraging random forests for interactive exploration of large histological images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1–8. Springer, 2014.
67. Andrew Rabinovich, Sameer Agarwal, Casey Laris, Jeffrey H Price, and Serge J Belongie. Unsupervised color decomposition of histologically stained tissue samples. In *Advances in neural information processing systems*, pages 667–674, 2004.
68. Erik Reinhard, Michael Adhikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer graphics and applications*, 21(5):34–41, 2001.
69. William Rey. Total singular value decomposition. robust svd, regression and location-scale. *arXiv preprint arXiv:0706.0096*, 2007.
70. Arnout C Ruifrok, Dennis A Johnston, et al. Quantification of histochemical staining by color deconvolution. *Analytical and quantitative cytology and histology*, 23(4):291–299, 2001.
71. Rowayda A Sadek. Svd based image processing applications: state of the art, contributions and research challenges. *arXiv preprint arXiv:1211.7102*, 2012.
72. R. Sampath and G. Biros. A parallel geometric multigrid method for finite elements on octree meshes. *SIAM Journal on Scientific Computing*, 32(3):1361–1392, 2010.
73. Mukesh Saraswat and KV Arya. Colour normalisation of histopathological images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1(4):185–197, 2013.
74. Mukesh Saraswat and K.V. Arya. Automated microscopic image analysis for leukocytes identification: A survey. *Micron*, 65:20 – 33, 2014.
75. James R Schott. *Matrix analysis for statistics*. John Wiley & Sons, 2016.
76. I Selim, Arabi E Keshk, and Bassant M El Shourbugy. Galaxy image classification using non-negative matrix factorization. *International Journal of Computer Applications*, 137(5):4–8.
77. Olcay Sertel, Jun Kong, Gerard Lozanski, Arwa Shana’ah, Umit Catalyurek, Joel Saltz, and Metin Gurcan. Texture classification using nonlinear color quantization: Application to histopathological image analysis. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 597–600. IEEE, 2008.
78. M Tarek Shaban, Christoph Baur, Nassir Navab, and Shadi Albarqouni. Staingan: Stain style transfer for digital histological images. *arXiv preprint arXiv:1804.01601*, 2018.
79. Dong Shan, Xinzhen Xu, Tianming Liang, and Shifei Ding. Rank-adaptive non-negative matrix factorization. *Cognitive Computation*, 10(3):506–515, 2018.
80. Li Shang, Yan Zhou, Jie Chen, and Wen-jun Huai. Nature image feature extraction using several sparse variants of non-negative matrix factorization algorithm. In *International Symposium on Neural Networks*, pages 274–281. Springer, 2012.
81. Jan Soelter, Jan Schumacher, Hartwig Spors, and Michael Schmuker. Automatic segmentation of odor maps in the mouse olfactory bulb using regularized non-negative matrix factorization. *NeuroImage*, 98:279–288, 2014.
82. Steven Squires, Adam Prügel-Bennett, and Mahesan Niranjan. Rank selection in non-negative matrix factorization using minimum description length. *Neural computation*, 29(8):2164–2176, 2017.

83. Weiwei Sun, Gang Yang, Bo Du, Lefei Zhang, and Liangpei Zhang. A sparse and low-rank near-isometric linear embedding method for feature extraction in hyperspectral imagery classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7):4032–4046, 2017.
84. Nicholas Trahearn, David Snead, Ian Cree, and Nasir Rajpoot. Multi-class stain separation using independent component analysis. In *Medical Imaging 2015: Digital Pathology*, volume 9420, page 94200J. International Society for Optics and Photonics, 2015.
85. Abhishek Vahadane, Tingying Peng, Shadi Albarqouni, Maximilian Baust, Katja Steiger, Anna Melissa Schlitter, Amit Sethi, Irene Esposito, and Nassir Navab. Structure-preserved color normalization for histological images. In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pages 1012–1015. IEEE, 2015.
86. Abhishek Vahadane, Tingying Peng, Amit Sethi, Shadi Albarqouni, Lichao Wang, Maximilian Baust, Katja Steiger, Anna Melissa Schlitter, Irene Esposito, and Nassir Navab. Structure-preserving color normalization and sparse stain separation for histological images. *IEEE transactions on medical imaging*, 35(8):1962–1971, 2016.
87. Yves-Rémi Van Eycke, Justine Allard, Isabelle Salmon, Olivier Debeir, and Christine Decaestecker. Image processing in digital pathology: an opportunity to solve inter-batch variability of immunohistochemical staining. *Scientific reports*, 7:42964, 2017.
88. Maria Vanrell, Felipe Lumbreras, Albert Pujol, Ramón Baldrich, Josep Lladós, and Juan José Villanueva. Colour normalisation based on background information. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 1, pages 874–877. IEEE, 2001.
89. Jared Vicory, Heather D Couture, Nancy E Thomas, David Borland, JS Marron, John Woosley, and Marc Niethammer. Appearance normalization of histology slides. *Computerized Medical Imaging and Graphics*, 43:89–98, 2015.
90. Christian Vollmer, Sven Hellbach, Julian Eggert, and Horst-Michael Gross. Sparse coding of human motion trajectories with non-negative matrix factorization. *Neurocomputing*, 124:22–32, 2014.
91. Michael E Wall, Andreas Rechtsteiner, and Luis M Rocha. Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis*, pages 91–109. Springer, 2003.
92. Z. Wang, X. Kong, H. Fu, M. Li, and Y. Zhang. Feature extraction via multi-view non-negative matrix factorization with local graph regularization. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 3500–3504, Sep. 2015.
93. Jun Xu, Lei Xiang, Guan hao Wang, Shridar Ganesan, Michael Feldman, Natalie NC Shih, Hannah Gilmore, and Anant Madabhushi. Sparse non-negative matrix factorization (snmf) based color unmixing for breast histopathological image analysis. *Computerized Medical Imaging and Graphics*, 46:20–29, 2015.
94. Xiaowei Xu, Scott D Dexter, and Ahmet M Eskicioglu. A hybrid scheme for encryption and watermarking. In *Security, Steganography, and Watermarking of Multimedia Contents VI*, volume 5306, pages 725–737. International Society for Optics and Photonics, 2004.
95. Jar-Ferr Yang and Chiou-Liang Lu. Combined techniques of singular value decomposition and vector quantization for image coding. *IEEE Transactions on Image Processing*, 4(8):1141–1146, 1995.
96. Jun Yu, Dacheng Tao, Meng Wang, and Yong Rui. Learning to rank using user clicks and visual features for image retrieval. *IEEE transactions on cybernetics*, 45(4):767–779, 2014.
97. F. G. Zanjani, S. Zinger, B. E. Bejnordi, J. A. W. M. van der Laak, and 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)*, pages 573–577, April 2018.
98. Farhad Ghazvinian Zanjani, Svitlana Zinger, Babak Ehteshami Bejnordi, Jeroen AWM van der Laak, and Peter HN de With. Stain normalization of histopathology images using generative adversarial networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 573–577. IEEE, 2018.
99. Rafał Zdunek, Anh Huy Phan, and Andrzej Cichocki. Image classification with non-negative matrix factorization based on spectral projected gradient. In *Artificial Neural Networks*, pages 31–50. Springer, 2015.



100. Junyi Dai Tao He Zhanzhuang Zhang, Juli Zhang. Exploring weighted dual graph regularized non-negative matrix tri-factorization based collaborative filtering framework for multi-label annotation of remote sensing images. *Remote Sens*, 11(8), 2019.
101. Yungang Zhang, Tianwei Xu, and Jieming Ma. Image categorization using non-negative kernel sparse representation. *Neurocomputing*, 269:21–28, 2017.
102. Yushan Zheng, Zhiguo Jiang, Haopeng Zhang, Fengying Xie, Jun Shi, and Chenghai Xue. Adaptive color deconvolution for histological wsi normalization. *Computer methods and programs in biomedicine*, 170:107–120, 2019.