Robust blind color deconvolution and blood detection on histological images using Bayesian K-SVD*

Fernando Pérez-Bueno[a,b,*], Kjersti Engan[c], Rafael Molina[a]

[a]*Dpto. Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, Spain.*
[b]*Research Center for Information and Communication Technologies (CITIC-UGR)*
[c]*Department of Electrical Engineering and Computer Science, University of Stavanger, Norway.*

Abstract

Hematoxylin and Eosin (H&E) color variation among histological images from different laboratories can significantly degrade the performance of Computer-Aided Diagnosis systems. The staining procedure is the primary factor responsible for color variation, and consequently, the methods designed to reduce such variations are designed in concordance with this procedure.

In particular, Blind Color Deconvolution (BCD) methods aim to identify the true underlying colors in the image and to separate the tissue structure from the color information. Unfortunately, BCD methods often assume that images are stained solely with pure staining colors (e.g., blue and pink for H&E). This assumption does not hold true when common artifacts such as blood are present, requiring an additional color component to represent them. This is a challenge for color standardization algorithms, which are unable to correctly identify the stains in the image, leading to unexpected results.

In this work, we propose a Blood-Robust Bayesian K-Singular Value Decomposition model designed to simultaneously detect blood and extract color from histological images while preserving structural details. We evaluate our method using both synthetic and real images, which contain varying amounts of blood pixels.

Keywords: Stain Separation, Blood Detection, Histological Images

1. Introduction

The development of Computer-Aided Diagnosis (CAD) systems for the analysis of Whole Slide Images (WSIs) poses important challenges [1], with one of the most

*This work was supported by project PID2022-140189OB-C22 funded by MCIN / AEI / 10.13039 / 501100011033, Spain. The work by Fernando Pérez-Bueno was supported in part by JDC2022-048784-I, funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/PRTR.

*Corresponding author

Email addresses: fpb@ugr.es (Fernando Pérez-Bueno), kjersti.engan@uis.no (Kjersti Engan), rms@decsai.ugr.es (Rafael Molina)

significant being that their performance is adversely affected when using data from different hospitals [2]. This decrease in performance is mainly due to color variation and artifacts in the image [3], which arise from a wide range of sources such as different chemicals, scanners, or laboratory conditions. Therefore, preprocessing is often a key step [3] for reliable CAD systems.

Blind Color Deconvolution (BCD) methods [4] estimate both the image-specific stain colors and structure (concentration). The separation itself can reduce the impact of color variation [5] and it is often included as a crucial step for other approaches such as color normalization [6] or color augmentation [7]. See [3] for a complete survey.

However, a critical assumption underlying BCD methods is that the image contains only the expected stains (e.g. Blue for nuclei and pink for cytoplasm and connective tissue in H&E images). This assumption does not hold true when artifacts are present. Artifacts degrade the quality and diagnosis value of a WSI and introduce additional sources of color variation [3]. In this work, we focus on blood in the images. It can appear due to vessels in the tissue, hemorrhages, or contamination of the slide. Once stained, blood exhibits metachromasia [8]. It reacts differently to H&E staining and acquires an entirely different color (usually red) [9]. This distinct color is commonly used for its detection [9] but also hampers the performance of BCD methods [10]. We refer to areas with blood, as blood artifacts.

1.1. Contributions

While the relationship between artifacts and color is noteworthy, the intersection of artifact detection and color variation has hardly been explored. In most cases, the presence of blood and other artifacts is ignored by BCD methods. Similarly, there is a scarcity of works using BCD for artifact detection. Our work bridges the gap between the fields of artifact detection and color variation, with a focus on blood and its impact on BCD methods applied to H&E images. We propose the use of BCD for blood detection, leveraging the color difference between blood and H&E stains. For this goal, we use the recently proposed Bayesian K-Singular Value Decomposition (BKSVD) [4] for BCD. While BKSVD excels at identifying the H&E stains, it is negatively affected by the presence of blood. We extend the model to detect the presence of blood, enabling both blood detection and robust estimation of the H&E stains.

This paper builds upon our recent work presented at the international conference of Artificial Intelligence in Medicine (AIME 2023) [11]. Four major enhancements are included in this extended work:

- In [11] we solely rely on the blood channel to detect anomalies. We now introduce a fully Bayesian detection approach that relies on the posterior distribution of the stain concentrations. This novel approach is more accurate and robust. We also provide a complete and clearer mathematical derivation of the model.

- With this novel approach, the need for a manual threshold is eliminated. Instead, we introduce a new thresholding method based on the estimated variance.

- We expand the experimental validation to include additional experiments and more extensive datasets. These experiments include evaluating blood detection, accuracy of color-vector matrix estimation, and tissue fidelity.

- Furthermore, we evaluate the impact of the proposed approach in the context of breast cancer classification.

The paper is organized as follows. Section 1.2 describes related works on BCD and blood detection. Section 2 provides an overview of the BKSVD method in [4], discusses its limitations, and provides the necessary improvements for its application to robust blood detection and BCD. In section 3 we experimentally evaluate the proposed method. Finally, section 5 concludes the paper.

1.2. Related Work

1.2.1. Blind Color Deconvolution

Most BCD methods use the Beer-Lambert law [12], which establishes a linear combination between the stains in the *optical density* (OD) space. Let I be a RGB image $\mathbf{I} \in \mathbb{R}^{3\times Q}$, where each value $i_{cq} \in \mathbf{I}$ correspond to pixel q and channel $c \in$ RGB. Then, the OD is defined as $y_{cq} = -\log(i_{cq}/i^0_{cq})$, where $i^0_{cq} = 255$ denotes the incident light. Then, a WSI $\mathbf{Y}$ stained with S stains follows the equation

$$\mathbf{Y} = \mathbf{MC} + \mathbf{N}\,, \tag{1}$$

where $\mathbf{M} = [\mathbf{m}_1, \cdots, \mathbf{m}_S] \in \mathbb{R}^{3\times S}$ is the normalized stains' specific color-vector matrix; $\mathbf{C} \in \mathbb{R}^{S\times Q}$ is the stain concentration matrix, its q-th column, $\mathbf{c}_q = [c_{1,q}, \ldots, c_{S,q}]^{\mathrm{T}}$, represents the contribution of each stain to the q-th pixel value in $\mathbf{Y}$; and, finally, $\mathbf{N} \in \mathbb{R}^{3\times Q}$ is a Gaussian noise matrix with independent components of variance β^{-1}.

The goal of BCD is to estimate $\mathbf{C}$ and $\mathbf{M}$ from $\mathbf{Y}$. Here we summarize the most relevant approaches in the literature. See [3] for a complete survey. Ruifrok *et al.* [12] use a given color-vector matrix to separate the stains, which is widely used as a standard. However, the actual color is usually considered to be unknown due to color variation. In the work by Macenko *et al.* [13] the H&E channels are estimated using Singular Value Decomposition (SVD). Vahadane *et al.* [6] estimate the color-vector matrix using Non-Negative Matrix Factorization (NMF) and the assumption that most pixels in the image are stained by a single stain. Alsubaie *et al.* [14] apply Independent Component Analysis (ICA) in the wavelet domain, under the assumption that stains might not be independent.

A Bayesian approach is followed by Hidalgo *et al.* [15] by defining a smoothness prior on the concentrations and a similarity prior on the color-vectors. To improve the quality of the concentration obtained, this work was extended by Pérez-Bueno *et al.* in [16] by using a Total Variation (TV) prior and in [5] with general super-Gaussian priors. These Bayesian methods share a need for a reference color-vector matrix for the similarity prior. The use of a prior on the color [16, 5] can reduce the effect of noise and artifacts, but limits the adaptability to different color distributions. To solve this issue,

the work in [4] proposes the use of Bayesian K-SVD (BKSVD) to find the color-vector matrix as a dictionary learning problem.

Deep Learning (DL) techniques are not widely used in BCD, mostly because of the lack of ground truth for stain separation. Most works proposed for color variation using DL avoid the BCD step and directly produce color normalization [17, 18, 19]. However, some works have applied DL for BCD. Duggal *et al.* [20] proposed the use of [13] as a CNN stain deconvolution layer to feed a stain-separated input to CNN-classifiers. Zheng *et al.* [21] produced multiple stain separation candidates using a Capsule Network and then assembled them with a sparse constraint. An Autoencoder was proposed by Abousamra *et al.* [22] to separate multiplex immunohistochemistry images using weak labels. Finally, the work by Yang *et al.* [23] proposed a deep Bayesian neural network that jointly estimates the stains and concentrations without ground truth, using the similarity prior proposed in [15].

From the above works, few consider the effect of noise and artifacts. The use of the method by Macenko *et al.* is widely extended, although it is known to be affected by noise and artifacts [24, 10]. To speed up the method, Vahadane *et al.* perform a patch sampling of the WSI. They take advantage of their patch-wise stain estimation to calculate the median color-vector and provide a robust estimation against small artifacts. Alsubaie *et al.* [14] include linear filtering to reduce the noise contamination when estimating the stain matrix but did not consider large artifacts. The presence of artifacts and its effect on the color estimation is acknowledged in [10]. The estimation of the color-vector and the robust maximum (99th percentile) of the H&E concentrations are used to identify low-quality images, substituting the color-vector matrix with average estimates from other images when poor quality is detected. The use of a prior on the color [25, 16, 5] can reduce the effect of the noise but limits the adaptability to different color distributions.

1.2.2. Blood Detection

The Detection of blood is frequently formulated as a color-related problem. In [9], Bukenya *et al.* classify blood segmentation techniques into (i) RGB segmentation, (ii) segmentation using other color space (such as HSV, Lab, and LUV) and additional techniques, (iii) segmentation using one or two channels of a non-RGB color space. Kim *et al.* [26] approach the detection of blood with a combination of staining protocol and image processing, using mathematical morphology and thresholding the RGB channels of the image. Sertel *et al.* [27] use the color to distinguish five major components in the H&E images (i.e. nuclei, cytoplasm, background, blood, and extracellular material). First they threshold the RGB channels to remove blood and background, and then use k-means in the L*a*b color space. This same RBG tresholding protocol is also used in [28] to remove blood. Recently, Adamo *et al.* [29] choose between red or green channels, before using the Otsu's threshold to detect blood vessels. The initial mask is then processed using pixel connectivity. A Maximum Likelihood Estimation is implemented by Mosaliganti *et al.* [30] to classify the pixels into four classes (blood, cytoplasm, nuclei, and background) in the RGB color space. In [31], Swiderska *et*

al. use the magenta channel of the CMYK space to detect blood areas that are later classified into hemorrhages or vessels using mathematical morphology and a decision tree. Chen *et al.* [32] use Hue and saturation to detect intracerebral hemorrhage. HSV was used by Fernandez-Carrobles *et al.* [33] in combination with logical operators and morphology and by Aldasoro *et al.* [34] combined with region-growing.

More complex approaches combine the choice of color space with clustering, mathematical morphology, classification, or DL. Mercan *et al.* [35] use a combination of LBP and L*a*b histograms to feed a clustering algorithm. Blood is identified as one of the clusters, which is used later to represent WSIs as bag-of-words. In [36], Swidereska *et al.* use a U-Net model to detect damaged tissue. Although the technique is only tested on IHC-stained brain tissues, the authors consider it possible to use it for WSI analysis of other tissues and stain protocols. Clymer *et al.* [37] use a pretrained RetinaNet model to detect blood vessels at low resolution. In other cases, blood is identified as a region to discard. Wetteland *et al.* [38] use a segmentation CNN to find several tissue classes, including blood and damaged tissue. This work was extended to multi-scale in [39], and combined with clustering to include low-probability patches in [40]. Finally, Kanwal *et al.* [41] evaluate the detection of blood and damaged tissue from diagnostically relevant tissue using five popular CNN architectures.

2. Material and Methods

In this section we first introduce the BKSVD method [4] for blind color deconvolution of histological images, then provide the details for the proposed blood robust BKSVD, and finally present the data that has been used in this work.

2.1. Bayesian k-SVD for Blind Color Deconvolution

Although the color in the slide is a combination of both H&E stains, each biological structure presents structure-specific color properties [6, 4]. This allows pathologists to distinguish structures based on their color. In [4] a Bayesian framework is used to approach the problem as a dictionary learning problem, with sparse concentrations [6], which encourages the framework to find the color-vector matrix $\mathbf{M}$ that better represents the differential staining of the structures in the image.

Following (1), the OD observed image $\mathbf{Y}$ is modeled with a Gaussian distribution $\mathrm{p}(\mathbf{Y}|\mathbf{C},\mathbf{M},\beta)$, where β controls the noise precision. That is,

$$\mathrm{p}(\mathbf{Y}|\mathbf{C},\mathbf{M},\beta) = \prod_{q=1}^{Q} \mathcal{N}(\mathbf{y}_q|\mathbf{M}\mathbf{c}_q, \beta^{-1}\mathbf{I}_{3\times 3}). \tag{2}$$

The sparsity of the solution is promoted by a zero-mean Laplace prior on the concentrations, that is,

$$\mathrm{p}(\mathbf{c}_q) \propto \exp\big(-\sqrt{\lambda_q}\|\mathbf{c}_q\|_1\big), \tag{3}$$

with $\lambda_q > 0$ the scale parameter. Unfortunately, the non-conjugacy of this distribution with the likelihood in (2) makes inference intractable. To circumvent this problem, the prior is replaced with a two-tiered hierarchical prior where

$$\mathbf{c}_q \sim \mathcal{N}(\mathbf{c}_q|\mathbf{0}, \mathbf{\Gamma}_q), \tag{4}$$

is a zero-mean normal distribution with diagonal covariance $\mathbf{\Gamma}_q = \text{diag}(\boldsymbol{\gamma}_q)$, and the positive-valued γ_{sq} is sampled from a Gamma hyperprior given by $\gamma_{sq} \sim \Gamma(1, \lambda_q/2)$. Independence is assumed yet again, so that $\text{p}(\boldsymbol{\gamma}_q) = \prod_s \text{p}(\gamma_{sq})$. The samples produced using this scheme follow a Laplace distribution [4] while keeping the inference tractable. Finally, a flat prior $\text{p}(\mathbf{M})$ on the color-vector matrix is used, and unit norm for each column $\mathbf{m}_s$ is assumed on their posterior estimation.

The true posterior $\text{p}(\mathbf{\Theta}|\mathbf{Y}) = \text{p}(\mathbf{Y}, \mathbf{\Theta})/\text{p}(\mathbf{Y})$, where $\mathbf{\Theta} = \{\beta, \mathbf{M}, \mathbf{C}\}$ is the set of unknowns, is approximated with variational inference [4]. WSIs are usually very large, rendering their processing computationally expensive. The number of pixels necessary to estimate the color-vector matrix $\hat{\mathbf{M}}$ is reduced by using a uniform random sampling of the WSI. The estimations $\hat{\mathbf{M}}$ and $\hat{\mathbf{C}}$ are then initialized to $\underline{\mathbf{M}}$ and $\hat{\mathbf{C}} = \underline{\mathbf{M}}^{+}\mathbf{Y}$, where $\underline{\mathbf{M}}$ is the Ruifrok's standard matrix [12] and $\underline{\mathbf{M}}^{+}$ the Moore-Penrose pseudo-inverse of $\underline{\mathbf{M}}$. After iteratively optimizing the model parameters and estimations for $\hat{\mathbf{M}}$ and $\hat{\mathbf{C}}$ are provided. Notice that the method in [4] did not consider the presence of blood. When the image contains blood artifacts $\mathbf{M}$ will be wrongly estimated because it will be forced to have two color vectors (H&E).

2.2. Robust Blind Color Deconvolution and Blood Detection

Blood and other artifacts hamper the accurate estimation of the color-vector matrix and the successful separation of the stains. Developing a robust and accurate BCD method is closely linked to the correct identification of these elements, which leads to the combined BCD and blood detection approach. In this work, we extend the BKSVD method introduced above and propose its use for blood detection and robust BCD estimation. Figure 1 illustrates the proposed method.

As discussed above, blood gets stained in a different color, commonly used for its detection [9]. Therefore, it can be seen as an additional effective stain, requiring a third stain channel to represent it. A third channel is often used by BCD [12, 13, 16], commonly referred to as the "background" channel. The content of the third channel is considered to be residual when only two stains are present in the image. However, it is important to keep the staining procedure in mind. While pixels in the image are a combination of H&E, only the blood suffers from metachromasia and should use the third channel. Therefore, including a third channel when not necessary, hampers the quality of the H&E estimated color-vectors $\hat{\mathbf{m}}_h$, $\hat{\mathbf{m}}_e$, and the corresponding concentrations [16].

BKSVD [4] uses two channels to separate H&E images, but it can be extended to include more channels. To account for the potential presence of blood or other elements in the image, we divide the BCD process into two stages. The first stage

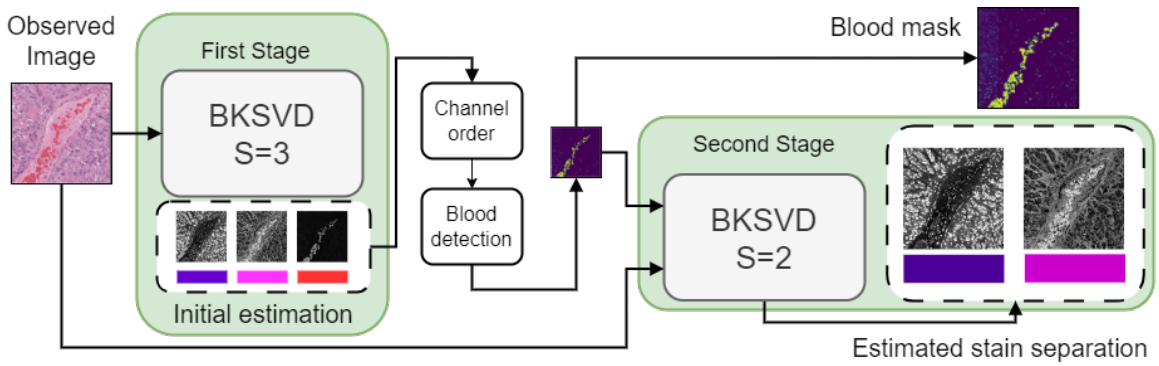

Figure 1: Overview of the proposed method. In the first stage, the image is separated into three stain channels. After checking the right order of the channels, see the text, the method identifies the blood pixels. The blood mask is used to exclude those pixels from the second stage, that provides a robust estimation of the H&E stains in the image. The binary blood mask is displayed using python viridis color map for better visualization. Purple and yellow represent 0 and 1, respectively.

employs three channels to represent all elements in the image, while the second stage uses only two channels and focuses on the H&E stains. Blood detection is performed as an intermediate step between both stages.

The first stage starts from the H&E Ruifrok's standard matrix $\underline{\mathbf{M}}$ [12] and introduce a third color-vector $\underline{\mathbf{m}}_b$ which is orthogonal to both $\underline{\mathbf{m}}_h$ and $\underline{\mathbf{m}}_e$ [16]. Despite this initialization, the method does not guarantee that the estimated $\hat{\mathbf{M}}$ and $\hat{\mathbf{C}}$ preserve the channel order. Maintaining the correct order of the stain channels becomes crucial, as we require to know which ones correspond to H&E. Thus, we calculate the correlation between the columns of $\hat{\mathbf{M}}$ and the reference $\underline{\mathbf{m}}_h$, $\underline{\mathbf{m}}_e$. We then select $\hat{\mathbf{m}}_h$, $\hat{\mathbf{m}}_e$ as those with the highest correlation to their respective references. The remaining channel is designated as the third channel, representing blood, and the order of $\hat{\mathbf{C}}$ is adjusted accordingly. The output of this first stage is the initial staining estimation.

At this point, the goal is to identify blood and refine the quality of the stain separation. The dictionary learning approach of BKSVD is intended to find a representation of the image where each pixel is assigned to a single channel whenever possible, pixels assigned to or exhibiting higher values in the third channel are indicative of an inaccurate representation by the H&E channels. This was used in our prior work [11], where the values in the blood channel $\hat{\mathbf{C}}_{:,3}$ are thresholded to detect anomalies. However, this approach entails manually establishing a threshold on the blood channel, which can vary between images from different sources. In this paper, we propose to take advantage of the Gaussian posterior distribution $q(\mathbf{c}_q)$ that derives from the BKSVD model using Bayesian inference. Note that, rather than a single estimation for the concentration values, BKSVD estimates a whole probability distribution we can use, see [4] for details. This approach is at the basis of anomaly detection using generative models [42]. The detection framework is based on the negative logarithm of $q(\mathbf{c}_q)$, which can be built using the outputs of the first BKSVD stage.

$$-\log q(\mathbf{c}_q) = \frac{\hat{\beta}}{2} \underbrace{\|\mathbf{y}_q - \sum_s \hat{\mathbf{m}}_s c_{q,s}\|_2^2}_{\text{Error term}} + \frac{1}{2} \underbrace{\mathbf{c}_q^T \mathbf{\Gamma}_q^{-1} \mathbf{c}_q}_{\text{Prior term}} + const \tag{5}$$

which correspond to a Gaussian distribution with mean and covariance matrix given by

$$\hat{\mathbf{c}}_q = \hat{\beta} \mathbf{\Sigma}_{\mathbf{c}_q} \hat{\mathbf{M}}^{\mathrm{T}} \mathbf{y}_q, \tag{6}$$

$$\mathbf{\Sigma}_{\mathbf{c}_q} = \left(\hat{\beta} \hat{\mathbf{M}}^{\mathrm{T}} \hat{\mathbf{M}} + \hat{\mathbf{\Gamma}}_q^{-1} \right)^{-1}. \tag{7}$$

Leaving aside constant terms, Eq. (5) comprises two key components, with both playing a relevant role in the detection process. The first evaluates the difference between a sample $\mathbf{y}_q$ and its reconstruction using the estimated stains and concentrations. Thus, it plays the role of a reconstruction error. The latter measures how close the estimated $\mathbf{c}_q$ is to the prior $p(\mathbf{c}_q)$ (4), serving as a regularization that penalizes non-sparse pixels. The parameters $\hat{\beta}$ and $\mathbf{\Gamma}_q^{-1}$, represent the variance of the prior distributions in (2) and (4), respectively. They are automatically calculated within the BKSVD model and serve to balance both terms according to their contribution to the marginal likelihood [42].

It is important to notice that we are focused on anomalies mostly expressed in the third (blood) channel, but this is not considered in (5). Therefore, we include some changes to create an anomaly score function. In the error term, we exclude the blood channel for the reconstruction process. This induces a higher error in pixels that require the use of that channel. In contrast, the prior term maintains all three channels.

Then, we need to look at the weighting parameters $\hat{\beta}$ and $\mathbf{\Gamma}_q^{-1}$. Unfortunately, the use of random sampling in the BKSVD method [4] renders $\hat{\beta}$ and $\mathbf{\Gamma}_q^{-1}$ unavailable for the entire image, as they are solely calculated for a representative sample of pixels. $\hat{\beta}$ is a global measure of the noise and can be considered to represent the entire image. In contrast, $\mathbf{\Gamma}_q^{-1}$ is a diagonal covariance matrix iteratively calculated at each pixel to maximize the posterior. To maintain computational efficiency, we choose a simplified model and approximate it with the same diagonal matrix for the entire image $\hat{\gamma}\mathbf{I}$. The scalar value $\hat{\gamma}$, is computed as the median value of the estimated γ_{sq} in $\mathbf{\Gamma}_q^{-1}$. Consequently, the proposed anomaly score is as follows:

$$AS_q = \frac{\hat{\beta}}{2} \underbrace{\|\mathbf{y}_q - \sum_{s=h,e} \hat{\mathbf{m}}_s c_{q,s}\|_2^2}_{\text{Error term}} + \frac{1}{2\hat{\gamma}} \underbrace{\mathbf{c}_q^T \mathbf{c}_q}_{\text{Prior term}}. \tag{8}$$

In this framework, (8) delivers $\mathbf{AS} \in \mathbb{R}^{1\times Q}$, where pixels conforming to the model

exhibit values close to zero, while anomalies yield significantly higher values. The value of Eq. (8) is used to mark and discard pixels that deviate from the model and cannot be adequately represented using hematoxylin and eosin alone. Given that this value approximates a Laplacian probability distribution, we can set a threshold according to the percentile P of data that we want to capture.

$$thr = \mu - b \times \log(2 - 2P) \tag{9}$$

where P is typically set to 0.95, and μ and b are the location and scale parameters of a Laplacian distribution calculated from the samples of (8) using the maximum likelihood estimator [43]. Then, the binary blood mask $\mathbf{B}$ is computed as:

$$B_q = \begin{cases} 1 & \text{if } AS_q > thr, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Finally, in the second BCD stage, we use only the blood-free pixels to re-estimate the H&E color-vector matrix and concentrations, starting from Ruifrok's reference $\underline{\mathbf{M}}$. This ensures a better representation of those pixels in terms of H&E, reducing the deviation caused by blood pixels. The complete procedure is summarized in Algorithm 1. [1]

Algorithm 1 Robust BKSVD for blood detection

Require: Observed image $\mathbf{I}$, initial normalized $\underline{\mathbf{M}}$, percentile P.
Ensure: Estimated stain color-vector matrix, $\hat{\mathbf{M}}$, concentrations, $\hat{\mathbf{C}}$, blood mask $\mathbf{B}$.

1: Obtain the OD image $\mathbf{Y}$ from $\mathbf{I}$.
2: First stage: set $S = 3$. Estimate $\hat{\mathbf{M}}$ and $\hat{\mathbf{C}}$ using BKSVD [4].
3: Sort $\hat{\mathbf{M}}$ and $\hat{\mathbf{C}}$ using the correlation of the columns of $\hat{\mathbf{M}}$ and $\underline{\mathbf{M}}$.
4: Calculate thr using (9)
5: Create blood mask $\mathbf{B}$ using (10).
6: Get blood-free pixels $\mathbf{Y}_{bf}$ by removing blood-positive pixels from $\mathbf{Y}$.
7: Second stage: set $S = 2$. Re-estimate $\hat{\mathbf{M}}$ for H&E on blood-free pixels $\mathbf{Y}_{bf}$ using BKSVD.
8: **return** $\hat{\mathbf{M}}$, $\hat{\mathbf{C}} = \hat{\mathbf{M}}^{+}\mathbf{Y}$, and $\mathbf{B}$.

2.3. Databases

In order to thoroughly validate the proposed method, a diverse range of images from different centers, both with and without the presence of blood, is essential. This section describes the four databases used in this paper.

- **Synthetic Blood Dataset (SBD)** We constructed the SBD using the Warwick Stain Separation Benchmark (WSSB) [14]. WSSB comprises 24 H&E images

[1] The code will be made available at https://github.com/vipgugr/ upon acceptance of the paper.

of different tissues from different laboratories, each with a known ground truth color-vector matrix $\mathbf{M}_{GT}$. These images are free from artifacts.

To introduce controlled levels of blood, we synthetically combine the WSSB images with blood images acquired from real WSIs. We incrementally add blood pixels as additional columns to each image. This allows us to measure the effect of different amounts of blood on the estimation of the color-vector matrix and the segmentation of blood pixels. Which is important to address the adaptability of methods to different scenarios. The amounts of blood considered are $\{0, 0.1, 0.2, \ldots, 0.9\}$ times the size of the original images, which creates a dataset of 240 images. All synthetically added pixels are labeled as blood.

- **TCGA Blood Dataset** To test the method on real images containing blood, we used 8 breast biopsies from The Cancer Genome Atlas (TCGA). Breast biopsies often contain blood due to the biopsy procedure [3] and it is also possible to find blood vessels in the tissue. We selected 16 2000×2000 H&E image patches where blood was manually labeled at pixel level. The labeling was performed by a trained technician with the Qpath annotation tool, using a combination of manual drawing and the color-sensitive wand.

- **Camelyon-17** We use the training set of Camelyon-17 breast cancer challenge [44]. It includes 500 WSIs from 5 medical centers and binary labels for cancer classification. It is intended to assess the performance of classification methods, and it is of special interest given the color variation between the centers. We use the subset of patches that have been used for classification experiments in [5, 4]. It contains 224×224 non-overlapping patches from the 500 WSIs in the Camelyon-17 training set.

- **Camelyon Blood Dataset** Some of the images from the Camelyon-17 dataset include patches with a significant amount of blood pixels. To increase the number and variety of real images on which our method is tested, we have manually identified and labeled these patches in 25 images. A patch was labeled as blood if it contained a minimum of 25% of blood pixels. This results in a dataset of 527 blood patches and 7786 tissue patches. Tissue patches contain no blood or less than 25% of blood. This dataset will be made publicly available in Zenodo [45].

3. Experiments

We have designed a set of experiments to assess our method. They are organized into three subsections: (i) Blood detection, (ii) Color vector estimation in the presence of blood, and (iii) cancer classification. In the first subsection, we examine the blood detection capability of the proposed method on synthetic and real images. Then, we assess the robustness to estimate the stains when blood is present in the image. Finally, we evaluate the impact of blood on a breast cancer classification task.

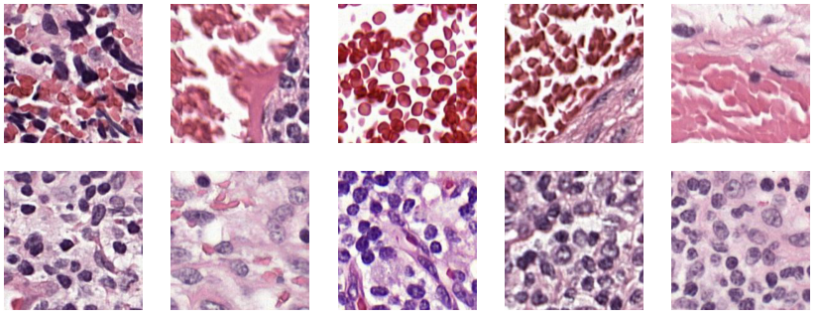

Figure 2: Example patches from the Camelyon Blood Dataset. Each column depicts two patches from the same image, one labeled as blood (top) and one tissue (bottom).

3.1. Blood Detection

In this experiment, we evaluate the performance of the proposed method in identifying blood within the image. We compare the results with other approaches for blood detection. We use the Area Under the ROC Curve (AUC) to provide an overall assessment across various thresholds, the Jaccard index (intersection over union), and the F1-score (or Dice coefficient). We compare the proposed approach with our previous work in [11], which corresponds to thresholding the third concentration channel $\mathbf{C}_{:,3}$, and the blood detection approaches in [29, 32, 31]. These methods involve thresholding the red RGB, the saturation HSV, and the magenta CMYK channels, respectively. They are denoted as R-RGB, S-HSV, and M-CMYK, respectively. For a better understanding of the model, we also include the results using the error and prior terms in (8).

Table 1: ROC AUC for blood detection on the SBD images

Amount	Method						
of blood	$\mathbf{C}_{:,3}$	Error	Prior	Proposed	R-RGB	M-CMYK	S-HSV
0.1	0.9778	0.9729	0.9624	**0.9864**	0.7825	0.9489	0.9758
0.2	0.9820	0.9780	0.9572	**0.9853**	0.7779	0.9447	0.9726
0.3	0.9814	0.9776	0.9516	**0.9833**	0.7734	0.9397	0.9684
0.4	0.9720	0.9669	0.9502	**0.9836**	0.7745	0.9402	0.9686
0.5	**0.9791**	0.9745	0.9494	0.9764	0.7744	0.9398	0.9685
0.6	0.9587	0.9518	0.9475	**0.9813**	0.7740	0.9392	0.9679
0.7	0.9705	0.9653	0.9482	**0.9815**	0.7746	0.9392	0.9680
0.8	0.9750	0.9689	0.9468	**0.9807**	0.7748	0.9391	0.9679
0.9	0.9644	0.9586	0.9426	**0.9799**	0.7749	0.9386	0.9674
Mean	0.9734	0.9683	0.9507	**0.9820**	0.7757	0.9410	0.9695

Table 1 summarizes the performance in terms of AUC for the synthetic SBD images. Notably, the RGB approach exhibits the lower discriminative power. In contrast, $\mathbf{C}_{:,3}$ outperforms the CMYK and HSV approaches. It is worth noting that using the error and prior terms from Eq. (8) separately does not improve the use of $\mathbf{C}_{:,3}$. However, when both terms are combined in the proposed approach, we achieve a notably enhanced detection capability.

To calculate the Jaccard index and F1-score, we need to determine an appropriate threshold for each method. In [11] we conducted an explorative analysis to determine the threshold for the blood channel. In this work, the proposed method is thresholded as detailed in section 2.2. We set the threshold to capture the 95 percentile for the Laplace distribution. For all the other approaches we use Otsu's method, which provides a dynamic (per image) threshold to separate pixels into two classes and it is a common approach for histological images [29, 3, 44]. Pixels with values above the selected threshold are identified as blood and the resulting mask is then compared with the ground truth label.

Tables 2 and 3 summarize the results for the above-discussed thresholds. The RGB approach is not close to the ground truth. CMYK, HSV, and $\mathbf{C}_{:,3}$ provide a relatively fair estimation, but their performance depends on the amount of blood in the image. Despite the high AUC value in table 1, Otsu's threshold is not appropriate for the error term. In contrast, it shows a good performance in the prior term. The proposed approach outperforms other methods, achieving the highest Jaccard index and F1-score in most cases, and also obtaining the best mean value. The controlled amount of blood in this dataset allows us to evaluate how the methods behave under different conditions. While the effect of larger amounts of blood in table 1 is subtle, their influence on the choice of a threshold clearly affects the performance of the methods in tables 2 and 3. It is worth noticing that the proposed approach is the most stable and the best choice when small amounts of blood are present in the image, which is often the case for real images.

Table 2: Jaccard index for different approaches to blood detection on the SBD images.

Amount	Method						
of blood	$\mathbf{C}_{:,3}$	Error	Prior	Proposed	R-RGB	M-CMYK	S-HSV
0.1	0.3567	0.5522	0.5027	**0.8354**	0.0320	0.2208	0.3187
0.2	0.5071	0.5183	0.6441	**0.8277**	0.0602	0.3821	0.4931
0.3	0.5882	0.4877	0.6911	**0.7985**	0.0857	0.4822	0.6049
0.4	0.6479	0.4863	0.7254	**0.8071**	0.1050	0.5560	0.6781
0.5	0.6874	0.4488	0.7371	**0.7907**	0.1217	0.6103	0.7248
0.6	0.7213	0.4470	0.7492	**0.7899**	0.1367	0.6513	0.7577
0.7	0.7446	0.4360	0.7541	0.7805	0.1488	0.6849	**0.7833**
0.8	0.7641	0.4177	0.7574	0.7764	0.1604	0.7119	**0.8027**
0.9	0.7790	0.4167	0.7586	0.7715	0.1707	0.7340	**0.8176**
Mean	0.6441	0.4679	0.7022	**0.7975**	0.1135	0.5593	0.6645

We now evaluate the methods on the real images from the TCGA dataset. The results are presented in table table 4. The proposed approach also performs well on these images. Although HSV achieves a higher AUC, its detection capability is reduced when a threshold is required. In contrast, the proposed approach achieves a fair AUC value and outperforms all other methods in terms of the Jaccard index and F1-score.

Figure 3 provides visual insight. It depicts some of the TCGA images along with their binary masks and plotted together with the resulting blood-mask using the methods with the thresholds as in table 4. The RGB, CMYK, and HSV approaches include

Table 3: F1-score for different approaches to blood detection on the SBD images.

Amount	Method						
of blood	$\mathbf{C}_{:,3}$	Error	Prior	Proposed	R-RGB	M-CMYK	S-HSV
0.1	0.5113	0.6952	0.6476	**0.9014**	0.0619	0.3485	0.4443
0.2	0.6608	0.6719	0.7715	**0.9001**	0.1133	0.5361	0.6312
0.3	0.7311	0.6472	0.8096	**0.8821**	0.1574	0.6365	0.7337
0.4	0.7782	0.6468	0.8354	**0.8893**	0.1894	0.7032	0.7933
0.5	0.8082	0.6089	0.8446	**0.8797**	0.2164	0.7488	0.8290
0.6	0.8328	0.6129	0.8533	**0.8804**	0.2399	0.7815	0.8530
0.7	0.8491	0.6025	0.8570	**0.8736**	0.2584	0.8069	0.8711
0.8	0.8624	0.5824	0.8595	0.8722	0.2758	0.8266	**0.8845**
0.9	0.8725	0.5839	0.8606	0.8689	0.2910	0.8423	**0.8946**
Mean	0.7674	0.6280	0.8155	**0.8831**	0.2004	0.6923	0.7705

Table 4: Detection metrics for different approaches to blood detection on TCGA Images

Method	$\mathbf{C}_{:,3}$	Error	Prior	Proposed	R-RGB	M-CMYK	S-HSV
ROC AUC	0.8080	0.8594	0.8519	0.9010	0.6106	0.8000	**0.9371**
Jaccard	0.3248	0.2553	0.1783	**0.3256**	0.0923	0.1254	0.1722
F1 score	0.4358	0.3735	0.2891	**0.4480**	0.1584	0.2105	0.2665

a significant amount of false positives. Among these, HSV is the most accurate, producing precise masks for some of the images. RGB is less effective, particularly when images feature a white background, as it results in pixels with elevated values in all RGB channels. This figure is particularly interesting to illustrate the behavior of the error and prior terms. In the case of the error term, it can be seen that most blood pixels are detected. However, there is also a notable presence of false positives, indicating that some pixels intended to be represented solely with H&E are erroneously using the blood channel. The prior may not be as effective at detecting blood, but it detects non-sparse pixels that either use multiple channels or exhibit unusually high values. The combination of both terms in the proposed approach leads to the most accurate masks, closely followed by $\mathbf{C}_{:,3}$, which is more unstable across images, therefore leading to lower figures of merit in table 4.

Finally, the proposed method is evaluated on the patches from the Camelyon Blood dataset. For this dataset, the processing is made on a per-image basis. A single color-vector matrix is calculated for all patches in the image, and the same threshold for **AS** (8) is used to calculate **B** for each patch. In accordance with the labeling protocol, a patch is predicted as blood if it contains more than 25% of predicted blood pixels.

Within this framework, the proposed method achieves a mean AUC value of 0.89, with a sensitivity of 0.91 for the blood patches, and a specificity of 0.86. A visual inspection of the false positives revealed that some of them contain small amounts of blood or other artifacts such as blur, or cauterized or folded tissues. This observation suggests that the proposed method might be valuable in detecting such elements. The prior term assigns a high score to excessively dark pixels, which are typically not sparse as they use all channels. This probably explains the detection of cauterized or folded tissues. Similarly, blurred pixels blend with their neighbors, which probably requires

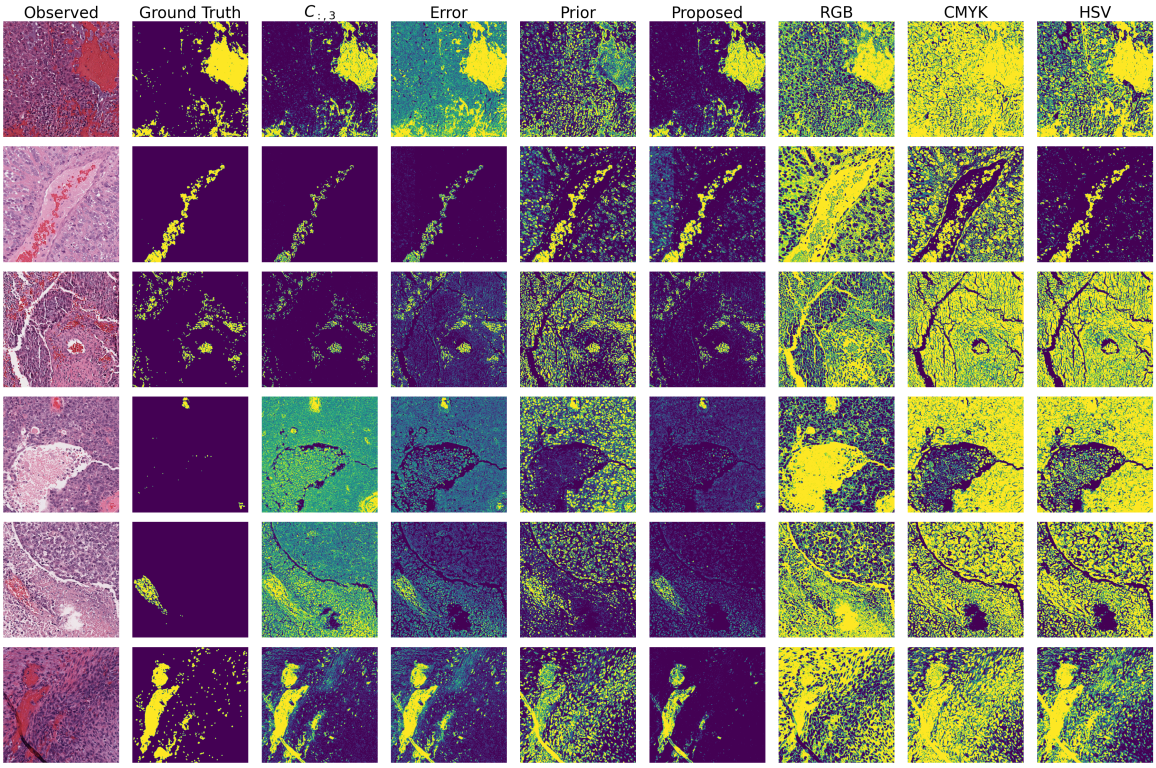

Figure 3: Blood detection qualitative comparison on images from the TCGA dataset. From left to right: Observed image, ground truth mask, and predicted masks for the different methods. The binary masks are displayed using python viridis color map for better visualization. Purple and yellow represent 0 and 1, respectively.

the use of multiple channels to represent their observed value. Figure 4 provides visual examples of these cases.

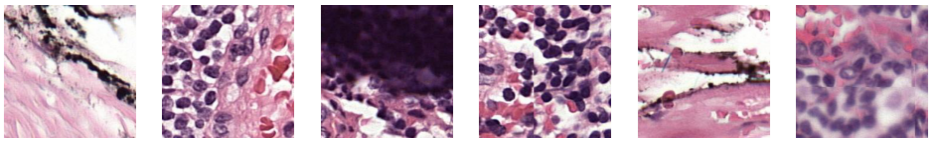

Figure 4: Examples of false positives from the Camelyon Blood dataset. They present cauterized tissue, small amounts of blood, folded tissue, or blur. These patches were detected to have more than 25% of blood pixels by the proposed method.

As a conclusion for this section, the proposed approach effectively identifies blood in the images. The results from the experiments presented above highlight the superior performance of the proposed approach compared to commonly used domains (RGB, CMYK, and HSV) in synthetic and real images. We have also shown a promising performance on blood detection at a patch level. Blood detection methods often involve a combination of thresholding with logical operators, morphology, or region-growing [31, 33, 34]. In this work, however, we focused on assessing the detection

capability of the different color channels. While RGB, CMYK, and HSV channels might potentially benefit from more complex segmentation techniques, the same holds true for the proposed approach. Exploring further filtering and enhancements for the detection mask is a task that should be addressed in future research.

3.2. Color vector estimation in the presence of blood

As discussed earlier, the presence of blood and other artifacts hinders the estimation of the color-vector matrix and its associated concentrations. In this experiment, we study this effect and assess the resilience of various BCD methods when blood is present in the image. First, we evaluate how blood affects the estimation of the color vector-matrix, then how it affects the concentration structure. We compare our proposed method with the methods by Macenko *et al.*[13] (MAC), the robust method by Vahadane *et al.*[6] (VAH), Alsubaie *et al.*[14] (ALS), the reference-prior based Bayesian methods using Total Variation (TV) [16], and Super Gaussian (SG) [5], the recent Deep Bayesian approach in [23] (BCDNet), and BKSVD [4]. In this section, we use the SBD images, as they include the stain-separation ground truth.

Table 5: Mean Euclidean distance between ground truth $\mathbf{M}_{GT}$ and obtained $\mathbf{M}$ color-vector matrix for different BCD methods.

Amount of blood	Method MAC	VAH	ALS	TV	SG	BCDNet	BKSVD	Proposed
0	0.2428	0.1019	0.3059	0.3328	0.3119	0.2748	0.0838	**0.0719**
0.1	0.3004	0.0972	0.3291	0.3783	0.3119	0.2749	0.2427	**0.0727**
0.2	0.3074	0.1116	0.3983	0.3889	0.3119	0.2766	0.2484	**0.0595**
0.3	0.3090	0.1227	0.3824	0.3971	0.3119	0.2778	0.2487	**0.0879**
0.4	0.3099	0.1340	0.4116	0.4061	0.3119	0.2784	0.2515	**0.1023**
0.5	0.3101	0.1398	0.4305	0.4141	0.3119	0.2788	0.2510	**0.1146**
0.6	0.3096	0.1893	0.4272	0.4244	0.3119	0.2791	0.2525	**0.1328**
0.7	0.3091	0.2494	0.4012	0.4321	0.3119	0.2794	0.2489	**0.1233**
0.8	0.3089	0.2710	0.5008	0.4417	0.3119	0.2798	0.2481	**0.1449**
0.9	0.3090	0.3002	0.5228	0.4503	0.3119	0.2799	0.2492	**0.1612**
Mean	0.3016	0.1717	0.4110	0.4066	0.3119	0.2780	0.2325	**0.1071**

To evaluate the estimated H&E color-vector matrix $\hat{\mathbf{M}}$ we compare it with the expected $\mathbf{M}_{GT}$ using the Euclidean distance. The comparison of the competing methods is presented in Table 5. The results demonstrate that the current state-of-the-art stain separation methods struggle when blood is present in the image. MAC, ALS, and BKSVD methods show a noticeable degradation in their estimations as soon as blood appears. The similarity prior in TV, SG, and BCDNet helps mitigate the impact of blood, but their reliance on the standard matrix from [12] as a reference, keeps them distant from the ground truth. Only VAH is capable of dealing with blood up to a certain extent. VAH redundantly estimates the color-vector matrix and chooses the median value for the whole image. However, this would not be as effective if blood pixels were distributed throughout the images. In contrast, the proposed method not only shows an excellent performance even with the increasing amount of blood present in the images, but it also achieves the most accurate estimation when blood is absent.

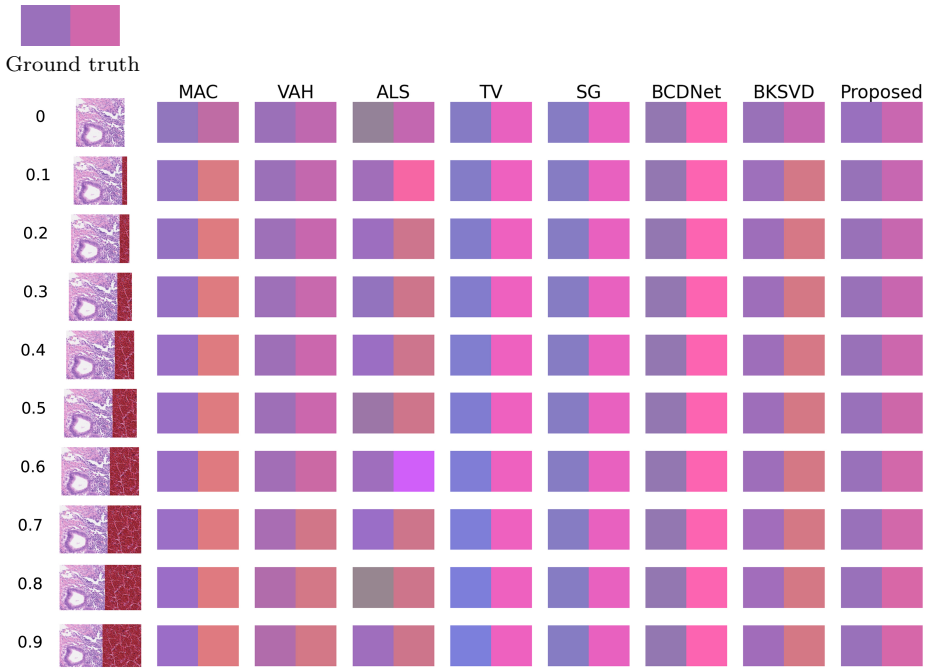

Figure 5: Qualitative comparison of color-matrices on an SBD image. The first row corresponds to the clean image and the amount of blood increases with the row number. Each column depicts the obtained $\mathbf{M}$ for different state-of-the-art methods.

A qualitative comparison of the methods for an example image, adding increasing amounts of blood is shown in figure 5. The color-vector matrix obtained by the different methods is depicted as a tuple where each square shows the normalized color vector for each stain, H and E, respectively. Small amounts of blood affect most of the methods. The effect of blood is similar for non-robust methods and most appreciated in the E channel, which becomes reddish as the amount of blood increases. The E channel changes to represent blood rather than eosin, and the H channel shifts to represent the mix of H&E which are closer in color between them than with blood.

Finally, we evaluate the quality of the structure for the H&E stain separation in terms of Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM). Tables 6 and 7 include the comparison of the H&E stain separation obtained for the images in the WSSB dataset. In this case, the estimation and separation are calculated with the synthetic image including blood, but only the original pixels are considered when calculating the metrics, as the ground truth is only available for those pixels.

The superior color-vector estimation achieved by the proposed method results in the highest fidelity in terms of PSNR and SSIM for the reconstructed H-only and E-only images. For VAH and BKSVD, the values significantly decrease when blood is present. For MAC and ALS the impact is less pronounced, possibly due to their initial estimation not being too close to the ground truth. The effect of the similarity prior, as discussed earlier for TV, SG, and BCDNet, is also evident in these results.

Figure 6 depicts the H-only and E-only images for the ground truth, the BKSVD, and the proposed method for an example image. To illustrate the effect of blood, we have selected a patch with 0.5 of synthetically added blood. In comparison to the ground truth, the shift of the estimated colors for the BKSVD can be clearly appreciated. The eosin channel is more reddish and does not correctly capture the connective tissue. This information is moved to the hematoxylin channel, which captures more information than expected. On the contrary, the proposed method accurately separates the tissue into their respective channels. The blood mask presented in figure 6.d.2 shows that very few tissue pixels are marked as blood.

Table 6: PSNR for the H-only and E-only images obtained using the different methods on the synthetic WSSB dataset.

Amount of blood	Method							
	MAC	VAH	ALS	TV	SG	BCDNet	BKSVD	Proposed
0	18.16	19.90	19.57	20.10	20.11	24.53	30.70	**31.06**
0.1	18.47	20.18	21.57	19.98	20.11	24.54	23.68	**29.18**
0.2	18.42	19.75	21.53	19.74	20.11	24.52	23.51	**31.37**
0.3	18.39	19.10	21.38	19.56	20.11	24.51	23.09	**31.04**
0.4	18.37	18.86	21.60	19.32	20.11	24.51	23.17	**29.95**
0.5	18.35	18.82	21.03	19.14	20.11	24.49	23.10	**29.52**
0.6	18.33	18.57	21.15	18.91	20.11	24.49	22.84	**27.20**
0.7	18.32	18.44	21.83	18.75	20.11	24.49	22.81	**27.33**
0.8	18.32	18.28	21.17	18.55	20.11	24.49	22.80	**25.85**
0.9	18.32	18.16	20.03	18.39	20.11	24.48	22.61	**25.75**
Mean	18.35	19.01	21.09	19.24	20.11	24.50	23.83	**28.83**

Table 7: SSIM for the H-only and E-only images obtained using the different methods on the synthetic WSSB dataset.

Amount of blood	Method							
	MAC	VAH	ALS	TV	SG	BCDNet	BKSVD	Proposed
0	0.6522	0.8085	0.6717	0.6842	0.6922	0.8652	0.9296	**0.9584**
0.1	0.7084	0.8130	0.6894	0.6791	0.6922	0.8651	0.8385	**0.9399**
0.2	0.7049	0.7544	0.6943	0.6686	0.6922	0.8649	0.8346	**0.9617**
0.3	0.7028	0.6987	0.7019	0.6604	0.6922	0.8647	0.8410	**0.9601**
0.4	0.7011	0.6791	0.6930	0.6508	0.6922	0.8647	0.8350	**0.9596**
0.5	0.7002	0.6752	0.7209	0.6432	0.6922	0.8648	0.8329	**0.9424**
0.6	0.6987	0.6620	0.7001	0.6336	0.6922	0.8648	0.8208	**0.9339**
0.7	0.6983	0.7005	0.7441	0.6268	0.6922	0.8648	0.8329	**0.9371**
0.8	0.6992	0.6978	0.7143	0.6182	0.6922	0.8648	0.8294	**0.9233**
0.9	0.7000	0.7041	0.6770	0.6112	0.6922	0.8648	0.8289	**0.9186**
Mean	0.6966	0.7193	0.7007	0.6476	0.6922	0.8648	0.8424	**0.9435**

3.3. Cancer classification

The ultimate objective of histological image preprocessing is to facilitate the automated diagnosis task. In this final experiment, we evaluate the impact of the proposed method on breast cancer classification using the Camelyon-17 dataset. Following [4] all images are deconvolved and the resulting concentrations are employed to feed a

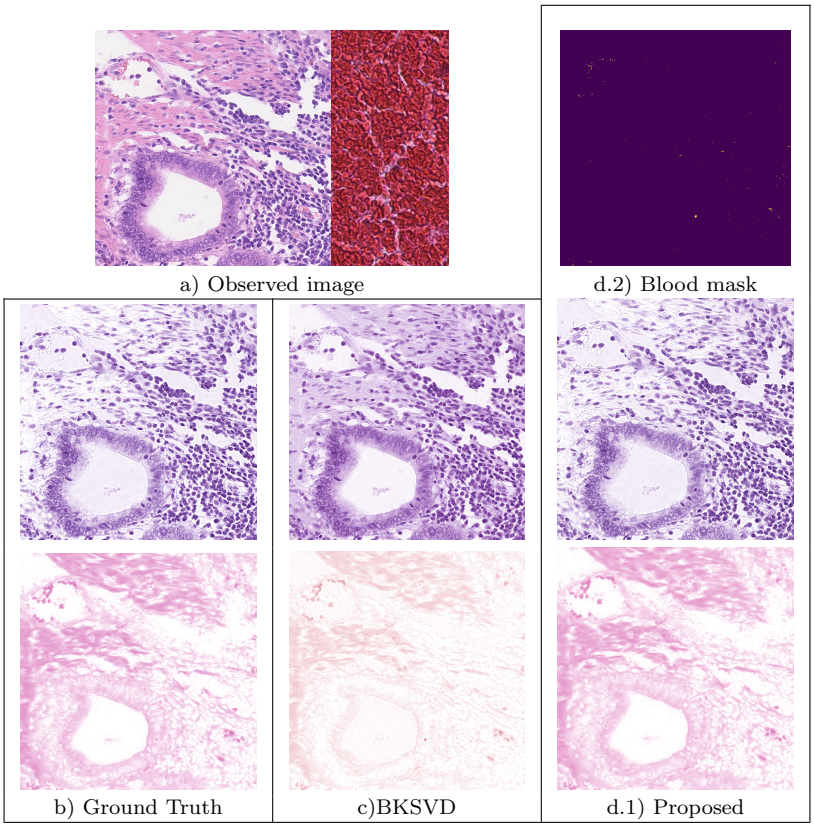

Figure 6: Effect of blood in the estimated stain separation. a) Observed image with synthetically added blood. b) H-only and E-only ground truth for the tissue pixels of the observed image in a. c) Stain separation by the BKSVD method. d.1) Stain separation by the proposed method. d.2) Blood mask obtained by the proposed method for those pixels.

VGG19 classifier. For training, we use data from four centers, while the last center is kept for testing. The choice of the testing center was based on the higher color deviation observed in [4]. This configuration allows the classifier to serve as a benchmark for assessing the stain separation quality. Using the results from the proposed method, we explore two different classification approaches: (i) feeding the classifier with the H&E concentrations as a 2-channel image, and (ii) normalizing the images, following [4], before feeding them to the classifier. We compare these approaches to BKSVD, which was the best-performing method in [4], and the use of original images.

Table 8: Classification performance of the VGG19 classifier in Camelyon-17 test center.

	Original Images	BKSVD H&E	BKSVD Normalized	Proposed H&E	Proposed Normalized
AUC	0.9491	0.9835	0.9818	**0.9857**	**0.9857**
F1-score	0.8852	0.9349	0.9103	0.9175	**0.9405**
Accuracy	0.8876	0.9364	0.9163	0.9229	**0.9405**

The results are presented in table 8. Interestingly, even though blood is not explicitly

mentioned as a relevant feature of the Camelyon-17 dataset [44], taking it into account leads to a better classification performance. The proposed method not only enhances classification when compared to the use of the original images but also outperforms the non-robust BKSVD. Both approaches, using H&E concentrations and the normalized images, achieve the highest AUC scores, with the latter obtaining better F1-score and accuracy figures.

4. Limitations and future work

This study introduces, for the first time, the use of BCD techniques for blood detection in histological images. Therefore, it is not exempt from limitations. In this section we highlight areas where additional research is needed.

- We have focused this study on blood detection. As previously discussed, this common artifact has a severe impact on the estimation of the color-vector matrix. Our results hint that the method could be suitable for the detection of other types of artifacts such as blur, folds, or cauterized tissue. Further evaluation and research to address these artifacts is of interest to the field of computational pathology.

- Our results are tested with H&E staining only. This is the most common protocol and it is certainly the most used in computational pathology. However, it is interesting to extend this research to other staining protocols such as H&E and Saffron or immunohistopathology. The BKSVD method can be easily extended to other staining protocols if the Beer-Lambert law is satisfied, just by changing the initial normalized matrix. If the proposed approach is suitable for detecting blood and artifacts on other staining protocol is yet to be investigated.

- This paper presents evaluations on blood detection, stain separation and cancer classification. Our experiments cover three different tissue types and up to 5 different laboratories. However, we have only evaluated the classification performance on the Camelyon-17 dataset for breast cancer. It would be interesting to further assess the effect of blood detection on other pathologies and CAD-related tasks such as segmentation. This requires datasets with classification labels where the presence of artifacts is acknowledged at least, if not labeled.

- The computational cost of the BKSVD method is significantly lower than other similar approaches [4] and it is designed to scalate when working with large images. However, the proposed method requires two stages to refine the quality of the stain separation, which raises the computational cost of the proposed approach. In the future, we plan to integrate both stages achieving blood detection and accurate stain separation simultaneously rather than sequentially.

5. Conclusions

In this work, we have extended the BKSVD method for stain separation to account for the presence of blood in histological images. Through this novel Bayesian framework, it is possible to detect blood and achieve a robust estimation of the stains. The proposed approach leverages the estimated posterior distribution for the concentration, which incorporates the reconstruction error and a prior term. The estimation procedure is split into two stages. The initial stage uses three channels to correctly represent all elements in the image. After identifying the blood pixels using the posterior, the subsequent stage uses only two channels and focuses on accurately estimating the color and concentrations of the H&E stains.

This work introduces the use of BCD techniques for blood detection, establishing a connection between the fields of color-preprocessing and artifact detection. This intersection benefits both domains. On the one hand, the presence of blood hampers the estimation of the stains in the images and achieving correct stain separation. On the other hand, BCD stain separation provides a valuable framework for the effective detection of blood.

Our method has been tested on synthetic and real images containing blood, showing excellent performance. It exhibits the capability to identify blood pixels, accurately estimate the color-vector matrix, and separate the stains. This ensures tissue fidelity, even with the presence of artifacts, while also providing a mask for potentially anomalous pixels.

We have also evaluated the impact of considering blood on breast cancer classification, which resulted in a substantial enhancement in the performance of a VGG19 classifier when applied to the Camelyon-17 dataset.

We believe that our work remarks on an important issue of stain separation techniques that are the basis for the color-preprocessing of histopathological images. They are severely affected by blood in the images, leading to unexpected results. In future research, we plan to extend these results and address various other types of artifacts in histological images, such as blur, cauterized, or folded tissue.

Acknowledgements

This work was supported by project PID2022-140189OB-C22 funded by MCIN / AEI / 10.13039 / 501100011033, Spain. The work by Fernando Pérez-Bueno was supported in part by JDC2022-048784-I, funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/PRTR.

References

[1] S. Morales, K. Engan, V. Naranjo, Artificial intelligence in computational pathology – challenges and future directions, Digit. Signal Process. (2021) 103196.

[2] T. S. Perry, Andrew ng x-rays the ai hype, IEEE Spectrum, (2021).

[3] N. Kanwal, F. Pérez-Bueno, A. Schmidt, R. Molina, K. Engan, The devil is in the details: Whole slide image acquisition and processing for artifacts detection, color variation, and data augmentation. a review., IEEE Access 10 (2022) 58821–58844.

[4] F. Pérez-Bueno, J. G. Serra, M. Vega, J. Mateos, R. Molina, A. K. Katsaggelos, Bayesian k-SVD for h and e blind color deconvolution. applications to stain normalization, data augmentation and cancer classification, Comput. Med. Imaging Graph. 97 (2022) 102048.

[5] F. Pérez-Bueno, M. Vega, M. A. Sales, J. Aneiros-Fernández, V. Naranjo, R. Molina, A. K. Katsaggelos, Blind color deconvolution, normalization, and classification of histological images using general super gaussian priors and bayesian inference, Comput. Meth. Prog. Bio. 211 (2021) 106453.

[6] A. Vahadane, et al., Structure-preserving color normalization and sparse stain separation for histological images, IEEE Trans. Med. Imag. 35 (2016) 1962–1971.

[7] D. Tellez, et al., Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology, Med. Image Anal. 58 (2019) 101544.

[8] A. H. M. Linkon, M. M. Labib, T. Hasan, M. Hossain, Marium-E-Jannat, Deep learning in prostate cancer diagnosis and gleason grading in histopathology images: An extensive study, Informatics in Medicine Unlocked 24 (2021) 100582.

[9] F. Bukenya, et al., An automated method for segmentation and quantification of blood vessels in histology images, Microvas. Res. 128 (2020) 103928.

[10] A. Anghel, et al., A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology, Front. Med. 6 (2019).

[11] F. Pérez-Bueno, K. Engan, R. Molina, A robust BKSVD method for blind color deconvolution and blood detection on H&E histological images, in: Artificial Intelligence in Medicine. AIME 2023., 2023, pp. 207–217.

[12] A. C. Ruifrok, D. A. Johnston, Quantification of histochemical staining by color deconvolution, Anal. Quant. Cytol. Histol. 23 (2001) 291–299.

[13] M. Macenko, et al., A method for normalizing histology slides for quantitative analysis, in: Int. Symp on biomed Imaging (ISBI), 2009, pp. 1107–1110.

[14] N. Alsubaie, et al., Stain deconvolution using statistical analysis of multi-resolution stain colour representation, PLOS ONE 12 (2017) e0169875.

[15] N. Hidalgo-Gavira, J. Mateos, M. Vega, R. Molina, A. K. Katsaggelos, Variational Bayesian blind color deconvolution of histopathological images, IEEE Trans. Image Process. 29 (1) (2020) 2026–2036.

[16] F. Pérez-Bueno, M. López-Pérez, M. Vega, J. Mateos, V. Naranjo, R. Molina, et al., A TV-based image processing framework for blind color deconvolution and classification of histological images, Digit. Signal Process. 101 (2020) 102727.

[17] A. Bentaieb, G. Hamarneh, Adversarial stain transfer for histopathology image analysis, IEEE Trans. Med. Imag. 37 (3) (2018) 792–802.

[18] F. G. Zanjani, S. Zinger, B. E. Bejnordi, J. A. W. M. 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, pp. 573–577.

[19] M. T. Shaban, C. Baur, N. Navab, S. Albarqouni, Staingan: Stain style transfer for digital histological images, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 953–956.

[20] R. Duggal, et al., SD-Layer: Stain Deconvolutional Layer for CNNs in Medical Microscopic Imaging, in: Medical Image Computing and Computer Assisted Intervention - MICCAI 2017. Lecture Notes in Computer Science. Springer, Cham, Vol. 10435, 2017, pp. 435–443.

[21] Y. Zheng, et al., Stain standardization capsule for application-driven histopathological image normalization, IEEE J. biomed and Health Inform. 25 (2) (2021) 337–347.

[22] S. Abousamra, D. J. Fassler, L. Hou, Y. Zhang, R. R. Gupta, T. M. Kurç, L. F. Escobar-Hoyos, D. Samaras, B. Knudson, K. R. Shroyer, J. Saltz, C. Chen, Weakly-supervised deep stain decomposition for multiplex ihc images, 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (2020) 481–485.

[23] S. Yang, F. Pérez-Bueno, F. M. Castro-Macías, R. Molina, A. K. Katsaggelos, Deep bayesian blind color deconvolution of histological images, in: 2023 IEEE International Conference on Image Processing (ICIP), 2023, pp. 710–714.

[24] N. Trahearn, et al., Multi-class stain separation using independent component analysis, in: Medical Imaging 2015: Digital Pathology, 2015, p. 94200J.

[25] M. Niethammer, et al., Appearance Normalization of Histology Slides, in: Machine Learning in Medical Imaging, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 58–66.

[26] N. T. Kim, et al., An Original Approach for Quantification of Blood Vessels on the Whole Tumour Section, Anal. Cell. Pathol. 25 (2) (2003) 63–75.

[27] O. Sertel, et al., Texture classification using nonlinear color quantization: Application to histopathological image analysis, in: 2008 IEEE Int. Conf. on Acoust. Speech Signal Process., 2008, pp. 597–600.

[28] O. Sertel, et al., Histopathological Image Analysis Using Model-Based Intermediate Representations and Color Texture: Follicular Lymphoma Grading, J. Signal Processing Systems 55 (1) (2009) 169.

[29] A. Adamo, A. Bruno, G. Menallo, M. G. Francipane, M. Fazzari, R. Pirrone, E. Ardizzone, W. R. Wagner, A. D'Amore, Blood vessel detection algorithm for tissue engineering and quantitative histology, Annals of Biomedical Engineering 50 (4) (2022) 387–400.

[30] K. Mosaliganti, et al., An imaging workflow for characterizing phenotypical change in large histological mouse model datasets, J. biomed Inform. 41 (6) (2008) 863–873.

[31] Z. Swiderska-Chadaj, et al., Automatic quantification of vessels in hemorrhoids whole slide images, in: IEEE Int. Conf. Comput. Probl. Elec. Eng., 2016, pp. 1–4.

[32] Z. Chen, et al., Histological quantitation of brain injury using whole slide imaging: A pilot validation study in mice, PLOS ONE 9 (3) (2014) 1–10.

[33] M.-M. Fernández-Carrobles, et al., A morphometric tool applied to angiogenesis research based on vessel segmentation, Diagnostic Pathology 8 (Suppl 1) (2013) S20.

[34] C. Reyes-Aldasoro, et al., An automatic algorithm for the segmentation and morphological analysis of microvessels in immunostained histological tumour sections, J. Microscopy 242 (3) (2011) 262–278.

[35] E. Mercan, et al., Localization of diagnostically relevant regions of interest in whole slide images, in: 2014 22nd Int. Conf. on Pattern Recognition, IEEE, 2014, pp. 1179–1184.

[36] Z. Swiderska-Chadaj, et al., Deep learning for damaged tissue detection and segmentation in ki-67 brain tumor specimens based on the u-net model, Bulletin of the Polish Academy of Sciences: Technical Sciences (2018) 849–856.

[37] D. Clymer, et al., Decidual vasculopathy identification in whole slide images using multiresolution hierarchical convolutional neural networks, The American J. Pathology 190 (10) (2020) 2111–2122.

[38] R. Wetteland, et al., Multiclass tissue classification of whole-slide histological images using convolutional neural networks, ICPRAM 2019 - Proceedings of the 8th Int. Conf. on Pattern Recognition Applications and Methods (2019) 320–327.

[39] R. Wetteland, K. Engan, et al., A multiscale approach for whole-slide image segmentation of five tissue classes in urothelial carcinoma slides, Technol. Cancer Res. Treat. 19 (2020).

[40] R. Wetteland, et al., Semi-supervised tissue segmentation of histological images, Colour and visual computing Symp/CEUR Workshop Proceedings, 2020.

[41] N. Kanwal, S. Fuster, F. Khoraminia, T. C. Zuiverloon, C. Rong, K. Engan, Quantifying the effect of color processing on blood and damaged tissue detection in whole slide images, in: 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), IEEE, 2022, pp. 1–5.

[42] F. Pérez-Bueno, L. García, G. Maciá-Fernández, R. Molina, Leveraging a probabilistic pca model to understand the multivariate statistical network monitoring framework for network security anomaly detection, IEEE/ACM Transactions on Networking 30 (3) (2022) 1217–1229.

[43] R. M. Norton, The double exponential distribution: Using calculus to find a maximum likelihood estimator, The American Statistician 38 (2) (1984) 135–136.

[44] P. Bándi, O. Geessink, Q. Manson, M. Van Dijk, M. Balkenhol, M. Hermsen, B. Ehteshami Bejnordi, B. Lee, K. Paeng, A. Zhong, Q. Li, F. G. Zanjani, S. Zinger, K. Fukuta, D. Komura, V. Ovtcharov, S. Cheng, S. Zeng, J. Thagaard, A. B. Dahl, H. Lin, H. Chen, L. Jacobsson, M. Hedlund, M. Çetin, E. Halıcı, H. Jackson, R. Chen, F. Both, J. Franke, H. Küsters-Vandevelde, W. Vreuls, P. Bult, B. van Ginneken, J. van der Laak, G. Litjens, From detection of individual metastases to classification of lymph node status at the patient level: The CAMELYON17 challenge, IEEE Transactions on Medical Imaging 38 (2) (2019) 550–560.

[45] F. Pérez-Bueno, K. Engan, R. Molina, A Blood Dataset from Camelyon 17, [Data set] Zenodo (2024). doi:10.5281/zenodo.11268269.