

A TV-based Image Processing Framework for Blind Color Deconvolution and Classification of Histological Images

Fernando Pérez-Bueno^{a,1,*}, Miguel López-Pérez^a, Miguel Vega^b, Javier Mateos^a, Valery Naranjo^c, Rafael Molina^a, Aggelos K. Katsaggelos^d

^a*Dpto. Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, Spain*

^b*Dpto. de Lenguajes y Sistemas Informáticos, Universidad de Granada, Spain*

^c*Dpto. de Comunicaciones, Universidad Politécnica de Valencia, Spain*

^d*Dept. of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA*

Abstract

In digital histopathological image analysis, two conflicting objectives are often pursued: closeness to the original tissue and high classification performance. The former objective tries to recover images (stains) that are as close as possible to the ones obtained by staining the tissue with a single dye. The latter objective requires images that allow the extraction of better features for an improved classification, even if their appearance is not close to single stained tissues. In this paper we propose a framework that achieves both objectives depending on the number of stains used to mathematically decompose the scanned image. The proposed framework uses a total variation prior for each stain together with the similarity to a given reference color-vector matrix. Variational inference and an evidence lower bound are utilized to automatically estimate all the latent variables and model parameters. The proposed methodology is tested on real images and compared to classical and state-of-the-art methods for histopathological blind image color deconvolution and prostate cancer classification.

Keywords: Blind Color Deconvolution, histopathological images, Variational Bayes, Prostate Cancer

1. Introduction

Histopathological tissues are usually stained with a combination of stains that bind to specific proteins on the tissue. Hematoxylin and Eosin (H&E) is one of the most commonly

*Corresponding author

Email addresses: fpb@ugr.es (Fernando Pérez-Bueno), mlopez@decsai.ugr.es (Miguel López-Pérez), mvega@ugr.es (Miguel Vega), jmd@decsai.ugr.es (Javier Mateos), vnaranjo@dcom.upv.es (Valery Naranjo), rms@decsai.ugr.es (Rafael Molina), aggk@eecs.northwestern.edu (Aggelos K. Katsaggelos)

¹This work was sponsored in part by Ministerio de Ciencia e Innovación under Contract BES-2017-081584 and project DPI2016-77869-C2-2-R.

used combination of stains. Hematoxylin stains cell nuclei while eosin stains cytoplasm and extracellular matrix components [1]. In digital brightfield microscopy, stained slides are then scanned to obtain high resolution whole-slide images (WSI). WSI analysis requires a lot of time and effort and computer-aided diagnosis (CAD) systems have become a valuable ally for pathologists. These systems frequently make use of the information provided by the different stains separately [2]. The separation of the stains in a WSI is known as Color Deconvolution (CD) and aims at estimating each stain concentration at each pixel location. Usually, the color spectral properties of each stain are also unknown since they vary from image to image. Color variations have a wide range of origins: different scanners, stain manufactures, or staining procedures, among others that create inter- and intra-laboratory differences. A study on color variation sources can be found in [3]. Blind CD techniques estimate image specific stain color-vectors together with stain concentrations.

CD is usually considered as a branch of color normalization. Tosta *et al.* [3] classified normalization methods into histogram matching, color transfer, and spectral matching. Normalization does not always require CD. Histogram matching methods do not use it, which leads to information loss as stains are assumed to be equally distributed. Color transfer usually separates histological regions identified by a segmentation step or between dyes. Although they usually involve deconvolution steps, it is not their main objective but a way to apply an statistical based color correction. Spectral matching techniques require to identify image specific spectral properties through CD. One of the first CD methods was proposed by Ruifrok *et al.* [4]. They obtained a set of globally standard color-vectors for hematoxylin, eosin and 3,3'-Diaminobenzidine (DAB), by measuring the relative absorption of each stain in single-stained images. The proposed set of stain color-vectors was calibrated for processing and digitization at the authors' laboratory. While these color vectors have been widely used, they do not take into account inter-slide variability. Empirical determination of the color-vector using single-stained tissue was used in [5, 6]. Aside from techniques that require the user to select pixels corresponding to each stain [7], several methods have been proposed to tackle inter-slide variability. In [2] Non-negative Matrix Factorization (NMF) is used to solve the problem formulated as a blind source separation one. This line of research was further developed in [8] and [9] by adding regularization and sparsity terms to take into account that a type of stain is only bounds to certain biological structures. Singular Value Decomposition (SVD), proposed in [10] to separate H&E images, was extended by McCann *et al.* [11] by taking into account the interaction between eosin and hematoxylin. The use of Non-Negative Least Squares (NNLS) to improve the performance of NMF is proposed in [12] to obtain a faster and less memory demanding method. Clustering techniques were explored in [13] where the stain vectors are estimated by projecting the input color image onto the Maxwellian chromaticity plane to form clusters, each one corresponding to one stained tissue type. In [14], to estimate the stain color-vector matrix, the image is segmented into background and pixels belonging to each stain using supervised relevant vector machines. The mean color of the pixels in each class is utilized as the stain color vector. Alsubaie *et al.* [15, 16], following [17], applied Independent Component Analysis (ICA) in the wavelet domain where the independence condition among sources is relaxed. Astola *et al.* [18] states that the method in [10] obtains better results applied in the linearly inverted RGB-space

and not in the (logarithmically inverted) absorbency space. In [19] a loss function based on the authors’ experience is optimized to obtain the image stain color-vectors. For further information on classical and state-of-the-art methods, the interested reader might check the reviews published in [20, 3].

In this paper, we present a framework for blind color deconvolution and classification of histological images. Depending on the number of stains used to mathematically model the observed image, the framework can be utilized to either recover the original H&E stains or to produce an H&E separation that boosts the performance of image classifiers. Within the framework, the proposed Bayesian blind CD problem algorithm, extends our previous work in [21] and [22]. In [21], a prior on the color-vectors, favouring similarity to a reference stain color-vectors, as well as a smoothness Simultaneous Autoregressive (SAR) prior model on each stain concentrations was used. As the SAR prior tends to oversmooth the edges of the image structures, in [22], we proposed the use of a Total Variation (TV) prior on each stain. The TV prior reduces the noise in the images while preserving sharp edges [23]. All model parameters were experimentally determined. In this paper, we extend the work in [22] by applying the Variational Bayes inference [24] and an evidence lower bound to automatically estimate all the latent variables and model parameters for blind color deconvolution and classification purposes. The proposed framework has been tested on additional real images for blind color deconvolution, where the fidelity to a ground-truth stain separation is assessed, and, for the first time, on classification tasks.

The rest of the paper is organized as follows: in section 2 the problem of color deconvolution is mathematically formulated. Following the Bayesian modelling and inference, in section 3 we propose a fully Bayesian algorithm for the estimation of the concentrations and the color-vector matrix as well as all the model parameters. In section 4, the proposed framework is evaluated on H&E stained images and its performance is compared with other classical and state-of-the-art CD methods in two different scenarios: color deconvolution and prostate cancer classification. Finally, section 5 concludes the paper.

2. Problem Formulation

Digital brightfield microscopes usually store a stained histological specimen’s WSI as an RGB color image of size $M \times N$, represented by the $MN \times 3$ matrix, $\mathbf{I} = [\mathbf{i}_R \ \mathbf{i}_G \ \mathbf{i}_B]$. Each color plane is stacked into a $MN \times 1$ column vector $\mathbf{i}_c = (i_{1c}, \dots, i_{MNc})^T, c \in \{R, G, B\}$. Each value i_{ic} represents the transmitted light on color band $c \in \{R, G, B\}$ for the pixel i of the slide.

CAD systems, on the other side, usually work with images in the *Optical Density* (OD) space. In this space, the intensity is linear with the amount of each stain absorbed by a sample. The OD of an image channel, $\mathbf{y}_c \in \mathbb{R}^{MN \times 1}$, is defined as $\mathbf{y}_c = -\log_{10}(\mathbf{i}_c/\mathbf{i}_c^0)$, where \mathbf{i}_c^0 denotes the incident light, and the division operation and $\log_{10}(\cdot)$ function are computed

element-wise. The observed OD image $\mathbf{Y} \in \mathbb{R}^{MN \times 3}$ has three OD channels, i.e.,

$$\mathbf{Y} = [\mathbf{y}_R \ \mathbf{y}_G \ \mathbf{y}_B] = \begin{bmatrix} \mathbf{y}_{1,:}^T \\ \vdots \\ \mathbf{y}_{MN,:}^T \end{bmatrix} = \begin{bmatrix} y_{1R} & y_{1G} & y_{1B} \\ \vdots & \vdots & \vdots \\ y_{MNR} & y_{MNG} & y_{MNB} \end{bmatrix}. \quad (1)$$

The Beer-Lambert law, for a slide stained with n_s stains, establishes that

$$\mathbf{Y}^T = \mathbf{M}\mathbf{C}^T + \mathbf{N}^T, \quad (2)$$

where $\mathbf{M} \in \mathbb{R}^{3 \times n_s}$ is the color-vector matrix,

$$\mathbf{M} = [\mathbf{m}_1 \ \dots \ \mathbf{m}_{n_s}] = \begin{bmatrix} \mathbf{m}_R^T \\ \mathbf{m}_G^T \\ \mathbf{m}_B^T \end{bmatrix} = \begin{bmatrix} m_{R1} & \dots & m_{Rn_s} \\ m_{G1} & \dots & m_{Gn_s} \\ m_{B1} & \dots & m_{Bn_s} \end{bmatrix} \in \mathbb{R}^{3 \times n_s}, \quad (3)$$

with each column \mathbf{m}_s in matrix \mathbf{M} being a unit ℓ_2 -norm stain color-vector containing the relative RGB color composition of the corresponding stain in the OD space, $\mathbf{C} \in \mathbb{R}^{MN \times n_s}$ is the stain concentration matrix,

$$\mathbf{C} = \begin{bmatrix} c_{11} & \dots & c_{1n_s} \\ \vdots & \ddots & \vdots \\ c_{MN1} & \dots & c_{MNn_s} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_{1,:}^T \\ \vdots \\ \mathbf{c}_{MN,:}^T \end{bmatrix} = [\mathbf{c}_1 \ \dots \ \mathbf{c}_{n_s}], \quad (4)$$

with the s -th column $\mathbf{c}_s = (c_{1s}, \dots, c_{MN_s})^T$, $s \in \{1, \dots, n_s\}$, representing the concentrations of the s -th stain, and the i -th row $\mathbf{c}_{i,:}^T = (c_{i1}, \dots, c_{in_s})$, $i = 1, \dots, MN$, representing the contribution of each stain to the i -th \mathbf{Y} pixel value, $\mathbf{y}_{i,:}$, and \mathbf{N} is a random matrix of size $MN \times 3$ with i.i.d. zero mean Gaussian components with variance β^{-1} , representing the noise introduced by the image capture system.

In the following section we use Bayesian modeling and inference to estimate both \mathbf{C} and \mathbf{M} , from \mathbf{Y}

3. Bayesian Modelling and Inference

Bayesian methods start with a prior distribution on the unknowns. In this paper we adopt the TV prior, which smooths the image noise while preserving its edges, for each one of the independent stain concentration vectors \mathbf{c}_s , that is,

$$p(\mathbf{C}|\boldsymbol{\alpha}) = \prod_{s=1}^{n_s} p(\mathbf{c}_s|\alpha_s) \propto \prod_{s=1}^{n_s} \exp[-\alpha_s \text{TV}(\mathbf{c}_s)], \quad (5)$$

with $\alpha_s > 0$ controlling the image smoothness. The TV function is defined for any \mathbf{c}_s , $s \in \{1, \dots, n_s\}$, as

$$\text{TV}(\mathbf{c}_s) = \sum_{i=1}^{MN} \sqrt{(\Delta_i^h(\mathbf{c}_s))^2 + (\Delta_i^v(\mathbf{c}_s))^2}, \quad (6)$$

where the operators $\Delta_i^h(\mathbf{c}_s)$ and $\Delta_i^v(\mathbf{c}_s)$ correspond to the horizontal and vertical first order differences of \mathbf{c}_s at pixel i , respectively.

The color-vector matrix $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_{n_s}]$ varies, as previously discussed, from image to image. However, images from the same laboratory usually have similar colors and we can benefit from this prior knowledge. Ruifrok *et al.* [4] proposed a procedure to obtain a laboratory dependant standard color-vectors. Although those vectors are not exact for every single image, they are representative and widely used. To take into account these considerations, we incorporate similarity to a reference color-vector matrix $\underline{\mathbf{M}} = [\underline{\mathbf{m}}_1, \dots, \underline{\mathbf{m}}_{n_s}]$ into the color-vector matrix prior model as

$$p(\mathbf{M}|\boldsymbol{\gamma}) = \prod_{s=1}^{n_s} p(\mathbf{m}_s|\gamma_s) \propto \prod_{s=1}^{n_s} \gamma_s^{\frac{3}{2}} \exp\left(-\frac{1}{2}\gamma_s\|\mathbf{m}_s - \underline{\mathbf{m}}_s\|^2\right), \quad (7)$$

where γ_s , $s = 1, \dots, n_s$, controls our confidence on the accuracy of $\underline{\mathbf{m}}_s$.

Finally, from the degradation model in (2), we have

$$p(\mathbf{Y}|\mathbf{M}, \mathbf{C}, \beta) = \prod_{i=1}^{MN} p(\mathbf{y}_{i,:}|\mathbf{M}, \mathbf{c}_{i,:}, \beta) = \prod_{i=1}^{MN} \mathcal{N}(\mathbf{y}_{i,:}|\mathbf{M}\mathbf{c}_{i,:}, \beta^{-1}\mathbf{I}_{3\times 3}). \quad (8)$$

With all these ingredients, we define the joint probability distribution as

$$p(\mathbf{Y}, \mathbf{C}, \mathbf{M}, \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) = p(\mathbf{Y}|\mathbf{C}, \mathbf{M}, \beta) p(\mathbf{C}|\boldsymbol{\alpha}) p(\mathbf{M}|\boldsymbol{\gamma}) p(\beta) p(\boldsymbol{\alpha}) p(\boldsymbol{\gamma}), \quad (9)$$

where $p(\boldsymbol{\gamma})$, $p(\boldsymbol{\alpha})$ and $p(\beta)$ are improper distributions of the form $p(w) \propto \text{const}$.

Following the Bayesian paradigm, inference will be based on the posterior distribution $p(\Theta|\mathbf{Y})$ with $\Theta = \{\mathbf{C}, \mathbf{M}, \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}\} = \{\mathbf{c}_1, \dots, \mathbf{c}_{n_s}, \mathbf{m}_1, \dots, \mathbf{m}_{n_s}, \beta, \alpha_1, \dots, \alpha_{n_s}, \gamma_1, \dots, \gamma_{n_s}\}$, the set of all unknowns.

Since the above posterior cannot be obtained in closed form, several approaches have been proposed to approximate it. In this paper we use the mean-field variational Bayesian model [25] to approximate $p(\Theta|\mathbf{Y})$ by the distribution $q(\Theta)$ of the form

$$q(\Theta) = q(\beta) \prod_{s=1}^{n_s} q(\mathbf{m}_s)q(\mathbf{c}_s)q(\alpha_s)q(\gamma_s), \quad (10)$$

where $q(\beta)$, $q(\alpha_s)$, $q(\gamma_s)$, $s = 1, \dots, n_s$, are assumed to be degenerate distributions. The optimal $q(\Theta)$ minimizes the Kullback-Leibler divergence [26] defined as

$$\mathbf{KL}(q(\Theta) \parallel p(\Theta|\mathbf{Y})) = \int q(\Theta) \log \frac{q(\Theta)}{p(\Theta|\mathbf{Y})} d\Theta \quad (11)$$

$$= \int q(\Theta) \log \frac{q(\Theta)}{p(\Theta, \mathbf{Y})} d\Theta + \log p(\mathbf{Y}). \quad (12)$$

The Kullback-Leibler divergence is always non negative and equal to zero if and only if $q(\Theta) = p(\Theta|\mathbf{Y})$.

Even with this factorization, the TV prior for \mathbf{C} hampers the evaluation of this divergence. To solve this problem, we define for α_s , \mathbf{c}_s , and any N -dimensions vector $\mathbf{u}_s \in (R^+)^{MN}$, $s = 1, \dots, n_s$, the functional

$$\mathcal{M}_s(\mathbf{c}_s, \mathbf{u}_s | \alpha_s) = \exp \left[-\frac{\alpha_s}{2} \sum_{i=1}^{MN} \frac{(\Delta_i^h(\mathbf{c}_s))^2 + (\Delta_i^v(\mathbf{c}_s))^2 + u_{is}}{\sqrt{u_{is}}} \right]. \quad (13)$$

Now, using the inequality for $w \geq 0$ and $z > 0$, $\sqrt{wz} \leq \frac{w+z}{2} \Rightarrow \sqrt{w} \leq \frac{w+z}{2\sqrt{z}}$, we can write

$$\exp[-\alpha_s \text{TV}(\mathbf{c}_s)] \geq \mathcal{M}_s(\mathbf{c}_s, \mathbf{u}_s | \alpha_s), \quad s = 1, \dots, n_s. \quad (14)$$

We, then, define

$$\mathcal{M}(\mathbf{C}, \mathbf{U} | \boldsymbol{\alpha}) = \prod_s \mathcal{M}_s(\mathbf{c}_s, \mathbf{u}_s | \alpha_s), \quad (15)$$

where $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_{n_s}]$ and $F(\Theta, \mathbf{U}, \mathbf{Y}) = p(\mathbf{Y} | \mathbf{M}, \mathbf{C}, \beta) \mathcal{M}(\mathbf{C}, \mathbf{U}, \boldsymbol{\alpha}) p(\mathbf{M}, \boldsymbol{\gamma}) p(\beta) p(\boldsymbol{\alpha}) p(\boldsymbol{\gamma})$ to obtain the inequality

$$\log p(\Theta, \mathbf{Y}) \geq \log F(\Theta, \mathbf{U}, \mathbf{Y}). \quad (16)$$

We have then found a lower bound, $F(\Theta, \mathbf{U}, \mathbf{Y})$, for the joint probability $p(\Theta, \mathbf{Y})$ defined in (9). Utilizing this lower bound in (12), we minimize $\mathbf{KL}(q(\Theta) || F(\Theta, \mathbf{U}, \mathbf{Y}))$ instead of $\mathbf{KL}(q(\Theta) || p(\Theta | \mathbf{Y}))$.

As shown in [25], the mean field variational distribution approximation establishes that for each unknown $\theta \in \Theta$, $q(\theta)$ will have the form

$$q(\theta) \propto \exp \langle \log F(\mathbf{Y}, \mathbf{C}, \mathbf{M}, \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \theta)}, \quad (17)$$

where $\Theta \setminus \theta$ represents all the variables in Θ except θ and $\langle \cdot \rangle_{q(\Theta \setminus \theta)}$ denotes the expected value calculated using the distribution $q(\Theta \setminus \theta)$. For variables with a degenerate posterior approximation, that is, for $\theta \in \{\beta, \alpha_1, \dots, \alpha_{n_s}, \gamma_1, \dots, \gamma_{n_s}\}$, the value where the posterior degenerates is

$$\hat{\theta} = \arg \max_{\theta} \langle \log F(\mathbf{Y}, \mathbf{C}, \mathbf{M}, \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \theta)}. \quad (18)$$

For the rest of the variables, that is, for $\theta \in \{\mathbf{m}_1, \dots, \mathbf{m}_{n_s}, \mathbf{c}_1, \dots, \mathbf{c}_{n_s}\}$, when point estimates are required, the expected value, that is, $\hat{\theta} = \langle \theta \rangle_{q(\theta)}$ is used.

Let us now explicitly obtain analytical expressions for these estimates.

3.1. Concentration Update

According to (17), the estimation of the distributions on the concentrations $q(\mathbf{c}_s)$ is obtained as

$$q(\mathbf{c}_s) \propto \exp \langle \log F(\mathbf{Y}, \mathbf{C}, \mathbf{M}, \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \mathbf{c}_s)}, \quad (19)$$

where

$$\langle \log F(\mathbf{Y}, \mathbf{C}, \mathbf{M}, \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \mathbf{c}_s)} = \langle \log p(\mathbf{Y} | \mathbf{C}, \mathbf{M}, \beta) \rangle_{q(\Theta \setminus \mathbf{c}_s)} + \langle \log \mathcal{M}(\mathbf{C}, \mathbf{U}, \boldsymbol{\alpha}) \rangle_{q(\Theta \setminus \mathbf{c}_s)}. \quad (20)$$

To calculate the first term of the sum, we rewrite the distribution probability in (8) as

$$\begin{aligned}
p(\mathbf{Y}|\mathbf{M}, \mathbf{C}, \beta) &\propto \beta^{\frac{1}{2}} \prod_{i=1}^{MN} \exp\left(-\frac{1}{2}\beta\|\mathbf{y}_{i,:} - \sum_{s=1}^{n_s} c_{is}\mathbf{m}_s\|^2\right) \\
&= \beta^{\frac{1}{2}} \prod_{i=1}^{MN} \exp\left(-\frac{1}{2}\beta\|\mathbf{y}_{i,:} - c_{is}\mathbf{m}_s - \sum_{k \neq s} c_{ik}\mathbf{m}_k\|^2\right) \\
&= \beta^{\frac{1}{2}} \prod_{i=1}^{MN} \exp\left(-\frac{1}{2}\beta \sum_{s=1}^{n_s} \left[-2c_{is}\mathbf{m}_s^T \left(\mathbf{y}_{i,:} - \sum_{k \neq s} c_{ik}\mathbf{m}_k\right) + c_{is}^2\|\mathbf{m}_s\|^2\right]\right. \\
&\quad \left. + \text{const}\right), \tag{21}
\end{aligned}$$

where we have separated the contribution of the s -th stain to each observed image pixel from the rest of stains and const indicates the term which does not depend on \mathbf{c}_s .

Then, we calculate $\langle \log p(\mathbf{Y}|\mathbf{C}, \mathbf{M}|\beta) \rangle_{q(\Theta \setminus \mathbf{c}_s)}$ as

$$\begin{aligned}
\langle \log p(\mathbf{Y}|\mathbf{C}, \mathbf{M}|\beta) \rangle_{q(\Theta \setminus \mathbf{c}_s)} &= \left\langle -\frac{\beta}{2} \sum_{i=1}^{MN} \sum_{s=1}^{n_s} \left[-2c_{is}\mathbf{m}_s^T \left(\mathbf{y}_{i,:} - \sum_{k \neq s} c_{ik}\mathbf{m}_k\right) + c_{is}^2\|\mathbf{m}_s\|^2\right] \right\rangle \\
&= -\frac{\beta}{2} (-2\mathbf{c}_s^T \mathbf{z}^{-s} + \|\mathbf{c}_s\|^2 \langle \|\mathbf{m}_s\|^2 \rangle), \tag{22}
\end{aligned}$$

where \mathbf{z}^{-s} is a column vector with components

$$z_i^{-s} = \langle \mathbf{m}_s \rangle^T \mathbf{e}_{i,:}^{-s} \quad \text{with} \quad \mathbf{e}_{i,:}^{-s} = \mathbf{y}_{i,:} - \sum_{k \neq s} \langle c_{ik} \rangle \langle \mathbf{m}_k \rangle, \quad i = 1, \dots, MN. \tag{23}$$

From (13), we can calculate

$$\begin{aligned}
\langle \log \mathcal{M}(\mathbf{C}, \mathbf{U}, \boldsymbol{\alpha}) \rangle_{q(\Theta \setminus \mathbf{c}_s)} &= \left\langle -\frac{\boldsymbol{\alpha}_s}{2} \sum_{i=1}^{MN} \frac{(\Delta_i^h(\mathbf{c}_s))^2 + (\Delta_i^v(\mathbf{c}_s))^2 + \mathbf{u}_{is}}{\mathbf{u}_{is}} \right\rangle \\
&= -\frac{\boldsymbol{\alpha}_s}{2} (\mathbf{c}_s)^T [(\Delta^h)^T \mathbf{W}(\mathbf{u}) \Delta^h + (\Delta^v)^T \mathbf{W}(\mathbf{u}) \Delta^v] \mathbf{c}_s + \text{const}, \tag{24}
\end{aligned}$$

where $\mathbf{W}(\mathbf{u}_s)$ is a diagonal matrix of the form $\mathbf{W}(\mathbf{u}_s) = \text{diag}(u_{is}^{-1/2})$, for $i = 1, \dots, MN$.

Hence,

$$\begin{aligned}
\langle \log F(\mathbf{Y}, \mathbf{C}, \mathbf{M}|\beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \mathbf{c}_s)} &= -\frac{\beta}{2} (-2\mathbf{c}_s^T \mathbf{z}^{-s} + \|\mathbf{c}_s\|^2 \langle \|\mathbf{m}_s\|^2 \rangle) \\
&\quad -\frac{\boldsymbol{\alpha}_s}{2} (\mathbf{c}_s)^T [(\Delta^h)^T \mathbf{W}(\mathbf{u}) \Delta^h + (\Delta^v)^T \mathbf{W}(\mathbf{u}) \Delta^v] \mathbf{c}_s + \text{const}, \tag{25}
\end{aligned}$$

which, from (17), produces $q(\mathbf{c}_s) = \mathcal{N}(\mathbf{c}_s | \langle \mathbf{c}_s \rangle, \Sigma_{\mathbf{c}_s})$, where

$$\Sigma_{\mathbf{c}_s}^{-1} = \beta \langle \|\mathbf{m}_s\|^2 \rangle \mathbf{I}_{MN \times MN} + (\Delta^h)^T \mathbf{W}(\mathbf{u}_s) \Delta^h + (\Delta^v)^T \mathbf{W}(\mathbf{u}_s) \Delta^v \quad (26)$$

$$\langle \mathbf{c}_s \rangle = \beta \Sigma_{\mathbf{c}_s} \mathbf{z}^{-s}, \quad (27)$$

where Δ^h and Δ^v represent the convolution matrices associated with the first order horizontal and vertical differences, respectively. Note that the matrix $\mathbf{W}(\mathbf{u}_s)$ can be interpreted as a spatial adaptivity matrix since it controls the amount of smoothing at each pixel location depending on the strength of the intensity variation at that pixel, as expressed by the horizontal and vertical intensity gradient.

3.2. Color-Vector Update

In a similar way, using (23), we calculate the distribution of \mathbf{m}_s ,

$$\begin{aligned} \langle \log F(\mathbf{Y}, \mathbf{C}, \mathbf{M} | \beta, \boldsymbol{\alpha}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \mathbf{m}_s)} &= \langle \log p(\mathbf{Y} | \mathbf{C}, \mathbf{M}, \beta) \rangle_{q(\Theta \setminus \mathbf{m}_s)} + \langle \log p(\mathbf{M}, \boldsymbol{\gamma}) \rangle_{q(\Theta \setminus \mathbf{m}_s)} \\ &= -\frac{\beta}{2} \left(\|\mathbf{m}_s\|^2 \sum_{i=1}^{MN} \langle c_{is}^2 \rangle - 2\mathbf{m}_s^T \sum_{i=1}^{MN} \langle c_{is} \rangle \mathbf{e}_{i,:}^{-s} \right) - \frac{1}{2} \gamma_s \|\mathbf{m}_s - \underline{\mathbf{m}}_s\|^2 + \text{const}, \end{aligned} \quad (28)$$

which, from (17), produces $q(\mathbf{m}_s) = \mathcal{N}(\mathbf{m}_s | \langle \mathbf{m}_s \rangle, \Sigma_{\mathbf{m}_s})$,

where

$$\Sigma_{\mathbf{m}_s}^{-1} = \left(\beta \sum_{i=1}^{MN} \langle c_{is}^2 \rangle + \gamma_s \right) \mathbf{I}_{3 \times 3}, \quad (29)$$

$$\langle \mathbf{m}_s \rangle = \Sigma_{\mathbf{m}_s} \left(\beta \sum_{i=1}^{MN} \langle c_{is} \rangle \mathbf{e}_{i,:}^{-s} + \gamma_s \underline{\mathbf{m}}_s \right). \quad (30)$$

Notice that $\langle \mathbf{m}_s \rangle$ may not be a unitary vector even if $\underline{\mathbf{m}}_s$ is. To obtain unitary vectors, we can always replace $\langle \mathbf{m}_s \rangle$ by $\langle \mathbf{m}_s \rangle / \|\langle \mathbf{m}_s \rangle\|$ and $\Sigma_{\mathbf{m}_s}$ by $\Sigma_{\mathbf{m}_s} / \|\langle \mathbf{m}_s \rangle\|^2$.

3.3. U Update

The use of the majorization with the functional in (13) introduces a new set of parameters, \mathbf{U} , that need to be estimated along with the concentrations and the color-vectors matrix. To estimate the \mathbf{U} matrix, we need to solve, for each $s \in \{1, \dots, n_s\}$,

$$\hat{\mathbf{u}}_s = \arg \min_{\mathbf{u}_s} - \langle \log \mathcal{M}_s(\alpha_s, \mathbf{c}_s, \mathbf{u}_s) \rangle_{q(\mathbf{c}_s)}, \quad (31)$$

whose solution is given by

$$\hat{u}_{is} = \arg \min_{u_{is}} \frac{\langle (\Delta_i^h(\mathbf{c}_s))^2 + (\Delta_i^v(\mathbf{c}_s))^2 \rangle + u_{is}}{\sqrt{u_{is}}} = \langle \Delta_i^h(\mathbf{c}_s)^2 \rangle + \langle \Delta_i^v(\mathbf{c}_s)^2 \rangle. \quad (32)$$

3.4. Parameter Update

Finally, the estimates of the noise, concentration, and color-vectors parameters are obtained according (18) as

$$\hat{\beta}^{-1} = \frac{\text{tr}(\langle (\mathbf{Y}^T - \mathbf{M}\mathbf{C}^T)(\mathbf{Y}^T - \mathbf{M}\mathbf{C}^T)^T \rangle_{\mathbf{q}(\Theta)})}{3MN}, \quad (33)$$

$$\hat{\alpha}_s^{-1} = \frac{\text{tr} \left(\left((\Delta^h)^T (\Delta^h) + (\Delta^v)^T (\Delta^v) \right) \langle \mathbf{c}_s \mathbf{c}_s^T \rangle \right)}{MN}, \quad (34)$$

$$\hat{\gamma}_s^{-1} = \frac{\text{tr}(\langle (\mathbf{m}_s - \underline{\mathbf{m}}_s)(\mathbf{m}_s - \underline{\mathbf{m}}_s)^T \rangle)}{3}. \quad (35)$$

3.5. Calculating the expectations and concentration covariance matrices

To estimate the concentrations and color-vectors, the expectations $\langle c_{is}^2 \rangle$ in (30) and $\langle \|\mathbf{m}_s\|^2 \rangle$ in (26) need to be calculated. Also, the computation of the matrix $\Sigma_{\mathbf{c}_s}$, defined in (26), is an issue due to the size of WSI images. In this section we explicitly calculate the mentioned expected values and address the concentrations covariance matrix calculation issue.

Notice that $\langle c_{is}^2 \rangle$ can be calculated using (27) and $\langle \|\mathbf{m}_s\|^2 \rangle$ can be easily calculated from (29) resulting in

$$\sum_{i=1}^{MN} \langle c_{is}^2 \rangle = \sum_{i=1}^{MN} \langle c_{is} \rangle^2 + \text{tr}(\Sigma_{\mathbf{c}_s}), \quad \langle \|\mathbf{m}_s\|^2 \rangle = \|\langle \mathbf{m}_s \rangle\|^2 + \text{tr}(\Sigma_{\mathbf{m}_s}). \quad (36)$$

The matrix $\Sigma_{\mathbf{c}_s}$ must be explicitly calculated to find its trace and also to calculate \hat{u}_{is} . However, since its calculation is very intense, following [27], we approximate the covariance matrix as follows. We first approximate $\mathbf{W}(\mathbf{u}_s)$ using $\mathbf{W}(\mathbf{u}_s) \approx z(\mathbf{u}_s)\mathbf{I}$, where $z(\mathbf{u}_s)$ is calculated as the mean value of the diagonal values in $\mathbf{W}(\mathbf{u}_s)$, that is, $z(\mathbf{u}_s) = \frac{1}{MN} \sum_i \frac{1}{\sqrt{u_{is}}}$. We then use the approximation

$$\Sigma_{\mathbf{c}_s}^{-1} \approx \beta \langle \|\mathbf{m}_s\|^2 \rangle \mathbf{I}_{MN \times MN} + \alpha_s z(\mathbf{u}_s) (\Delta^h)^T (\Delta^h) + \alpha_s z(\mathbf{u}_s) (\Delta^v)^T (\Delta^v) = \mathbf{B}. \quad (37)$$

Note that the matrix \mathbf{B} is a block circulant matrix with circulant blocks (BCCB), thus, computing its inverse can be very efficiently performed in the discrete Fourier domain. Finally, we have

$$\begin{aligned} \langle \Delta_i^h(\mathbf{c}_s)^2 \rangle + \langle \Delta_i^v(\mathbf{c}_s)^2 \rangle &\approx (\Delta_i^h(\langle \mathbf{c}_s \rangle))^2 + (\Delta_i^v(\langle \mathbf{c}_s \rangle))^2 \\ &+ \frac{1}{MN} \text{tr} \left[\mathbf{B}^{-1} \times \left((\Delta^h)^T (\Delta^h) + (\Delta^v)^T (\Delta^v) \right) \right]. \end{aligned} \quad (38)$$

Algorithm 1 Variational Bayesian TV Blind Color Deconvolution

Require: Observed image \mathbf{I} and reference (prior) color-vector matrix $\underline{\mathbf{M}}$.

Obtain the observed OD image \mathbf{Y} from \mathbf{I} and set $\langle \mathbf{m}_s \rangle^{(0)} = \mathbf{m}_s$, $\Sigma_{\mathbf{m}_s}^{(0)} = \mathbf{0}$, $\Sigma_{\mathbf{c}_s}^{(0)} = \mathbf{0}$, $\langle \mathbf{c}_s \rangle^{(0)}$, $\forall s = 1, \dots, n_s$, from the matrix \mathbf{C} obtained as $\mathbf{C}^T = \underline{\mathbf{M}}^+ \mathbf{Y}^T$, with $\underline{\mathbf{M}}^+$ the Moore-Penrose pseudo-inverse of $\underline{\mathbf{M}}$, and $n = 0$.

while convergence criterion is not met **do**

1. Set $n = n + 1$.
2. Obtain $\beta^{(n)}$, $\alpha^{(n)}$ and $\gamma^{(n)}$ from (33), (34) and (35), respectively.
3. Using $\langle \mathbf{c}_s \rangle^{(n-1)}$ and $\Sigma_{\mathbf{c}_s}^{(n-1)}$, $\forall s \in \{1, \dots, n_s\}$, update the new variational parameters $\hat{\mathbf{u}}_s^{(n)}$ from (32).
4. Using $\langle \mathbf{c}_s \rangle^{(n-1)}$, $\Sigma_{\mathbf{c}_s}^{(n-1)}$ and $\langle \mathbf{m}_s \rangle^{(n-1)}$, update the color-vectors $\langle \mathbf{m}_s \rangle^{(n)}$ and $\Sigma_{\mathbf{m}_s}^{(n)}$ from (30) and (29), $\forall s$.
5. Using $\langle \mathbf{m}_s \rangle^{(n)}$, $\Sigma_{\mathbf{m}_s}^{(n)}$ and $\hat{\mathbf{u}}_s^{(n)}$, update the concentrations $\Sigma_{\mathbf{c}_s}^{(n)}$ and $\langle \mathbf{c}_s \rangle^{(n)}$ from (26) and (27), $\forall s$.

end while

Output the color-vector $\hat{\mathbf{m}}_s = \langle \mathbf{m}_s \rangle^{(n)}$ and the concentrations $\hat{\mathbf{c}}_s = \langle \mathbf{c}_s \rangle^{(n)}$.

3.6. Proposed Algorithm

Based on the previous derivations, we propose the Variational Bayesian TV Blind Color Deconvolution in Algorithm 1. The algorithm starts from the observed RGB image and a reference (prior) color vector matrix. Using this reference color-vector matrix as an starting point, the algorithm estimates in an iterative way, the model and variational parameters value, the distribution on the concentrations and distribution on the color-vectors.

The linear equations problem in (27), used in step 5 of Alg. 1, has been solved using the conjugate gradient approach while the color-vectors update in step 4 of the algorithm has been directly calculated from the equations due to the small size of the problem. On convergence, the algorithm returns point estimates of the color-vectors and concentrations as the mean value of the estimated distributions. Finally, from Alg. 1, an RGB image of each separated stain, $\hat{\mathbf{I}}_s^{\text{sep}}$, can be obtained as

$$(\hat{\mathbf{I}}_s^{\text{sep}})^T = \exp_{10}(-\hat{\mathbf{m}}_s \hat{\mathbf{c}}_s^T). \quad (39)$$

4. Experimental results

As previously indicated, blind color deconvolution algorithms are used for visual inspection and automatic classification of images. These may be conflicting goals since the most accurate color deconvolved images, in the sense of closeness to each single dye, are not usually the ones that lead to the highest performance in classification.

In this section, we will show that, depending on the number of components used in the deconvolution process, the proposed methodology can obtain either the most accurate color images or produce stains that lead to the highest classification performance. To do so, we have designed two set of experiments. In the first one, the proposed method is

applied on the *Warwick Stain Separation Benchmark* (WSSB) dataset [16] (a dataset where the ground-truth color-vectors are known) and its results are compared to classical and state-of-the-art deconvolution methods both visually and numerically. We will show that the proposed method outperforms the competing methods when two components are used. We also presents results on prostate cancer detection using the histopathological SICAPv1 database [28]. On this carefully annotated dataset, color deconvolution is used to separate H&E stains from which a set of features are extracted. Following [28], those features are then used to train a group of state-of-the-art supervised classification methods to distinguish between benign and pathological images. In this classification scenario we will show that the proposed method outperforms its competitors when three components are used.

The experiments carried out will then indicate that the introduced framework can be used for accurate reconstruction of original stains and to obtain better classification results depending on the number of stains used to decompose the image.

4.1. Color Deconvolution Experiments

In this first experiment, we assess the quality of the color deconvolution methods for accurate H&E separation. For this purpose, we used the *Warwick Stain Separation Benchmark* (WSSB) [16] dataset as a test-bed. WSSB contains 24 H&E stained images of different tissues (breast, colon and lung) from different laboratories which have been captured with different microscopes. For each image, its ground truth stain color-vector matrix, \mathbf{M}_{GT} , was manually obtained by medical experts. The median value of a set of image pixels with a single stain was used. The pixels were selected based on biological structures: nuclei for hematoxylin and cytoplasm for eosin. The ground truth concentrations were obtained in [16] from the ground-truth color-vector matrix as $\mathbf{C}_{GT}^T = \mathbf{M}_{GT}^+ \mathbf{Y}^T$. From those ground-truth concentrations and color-vectors, a RGB image for each stain separately is obtained by applying (39). A sample breast image from the WSSB dataset is shown in Fig. 1a and its ground truth RGB separation is depicted in Fig. 1b.

The proposed framework was compared with the classical non-blind method by Ruifrok *et al.*[4], the classical blind color deconvolution by Macenko *et al.*[10], and the state-of-the-art methods by Vahadane *et al.*[8], Alsubaie *et al.*[16], and Hidalgo-Gavira *et al.*[21]. The proposed Algorithm 1 was run on this dataset until the criterion $\| \langle \mathbf{c}_s \rangle^{(n)} - \langle \mathbf{c}_s \rangle^{(n-1)} \|^2 / \| \langle \mathbf{c}_s \rangle^{(n)} \|^2 < 10^{-5}$ was met by all stains, that is, $s = 1, 2, \dots, n_s$. Since different tissues may have different color characteristics, the reference (prior) color-vector matrix \mathbf{M} was obtained by selecting, by non-medical experts, a single pixel from each type of tissue, breast, colon and lung, containing mainly hematoxylin and another pixel containing mainly eosin. When a third component is utilized, following the most commonly used implementation of Ruifrok’s method [29], the color representing the third component of each reference color-vector was calculated as the complementary of the first two colors. For all the competing algorithms, parameters were selected following the recommendations in the original paper or the reference software freely available.

The resulting H-only and E-only images were compared both visually and numerically by means of the *Peak Signal to Noise Ratio* (PSNR) and *Structural Similarity* (SSIM) metrics. Numerical results, presented in Table 1, show that using two stains, i. e., $n_s = 2$, the

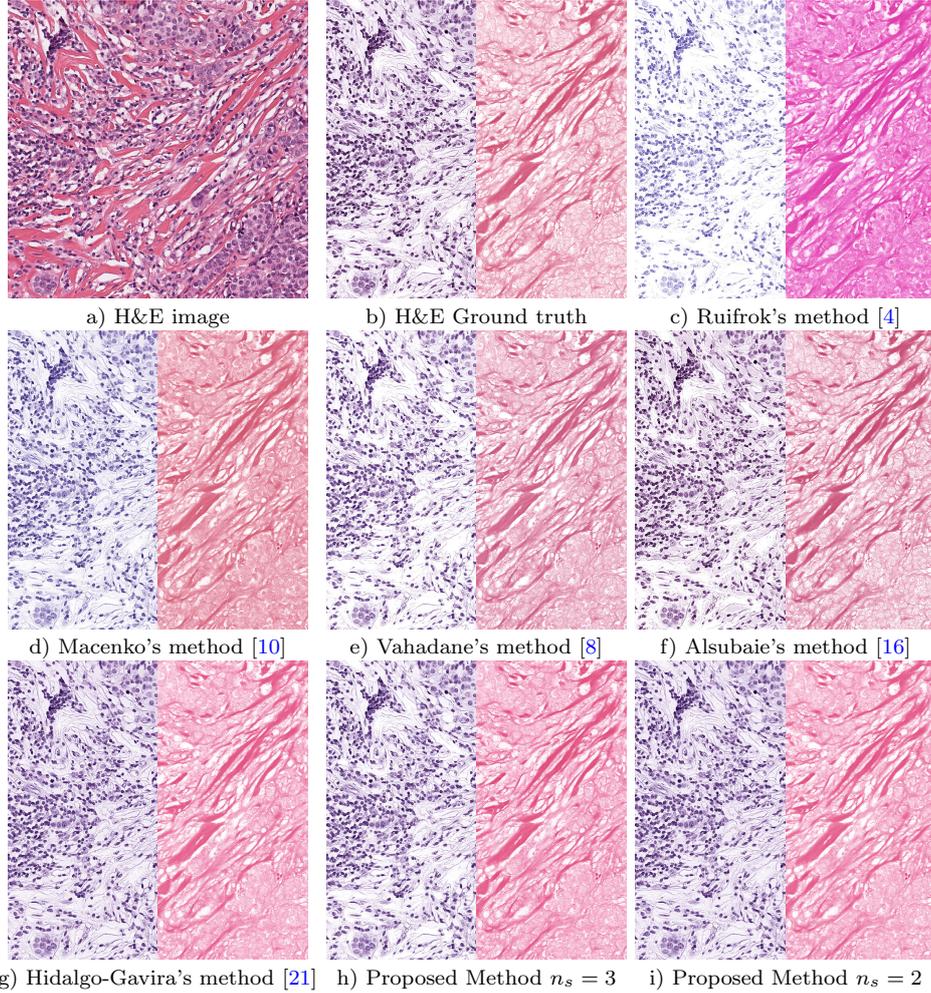


Figure 1: A breast H&E stained image from the WSSB dataset in [16], its ground truth separated H-only and E-only images, and its separation results by the competing and proposed methods. Hematoxylin and eosin separations are presented on the left and right hand sides of each image, respectively.

proposed method produces higher PSNR and SSIM values than the competing models except for SSIM in lung images where a slightly higher value is obtained by the Hidalgo-Gavira's method.

The separated H- and E-only images from the observed image in Fig. 1a are shown in Fig. 1(c-i). The proposed method and the methods by Vahadane and Hidalgo-Gavira produce H&E images very similar to the ground truth separation in Fig. 1b. Note also that the images obtained by Hidalgo-Gavira's method and the proposed one with two and three components are very similar. Notice, however, that the H-only images produced by the proposed method (Fig. 1h-i) are sharper and nuclei are clearer which will be useful, as we will later see, for classification. Both methods use the same prior model on the color-vectors, but they differ on the prior on the concentrations. While Hidalgo-Gavira's method uses a SAR model, ours uses a TV-based one. This model produces sharper images than those

Table 1: PSNR and SSIM for the different methods on the WSSB dataset [16].

Image	Stain	Ruifrok's method [4]		Macenko's method [10]		Vahadane's method [8]		Alsubaie's method [16]		Hidalgo-Gavira's method [21]		Proposed method $n_s = 3$		Proposed method $n_s = 2$	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Colon	H	22.27	0.8141	23.91	0.8095	25.83	0.8851	21.11	0.7241	28.57	0.9542	24.83	0.9005	28.62	0.9544
	E	20.70	0.7456	21.55	0.6365	26.29	0.8904	21.94	0.8540	27.58	0.9139	25.97	0.8695	27.60	0.9161
Breast	H	15.27	0.6215	26.24	0.9552	25.46	0.9239	24.60	0.8068	28.81	0.9528	27.71	0.9538	29.14	0.9560
	E	17.66	0.7644	23.62	0.9336	27.68	0.9550	25.92	0.9380	26.60	0.9464	26.84	0.9510	26.76	0.9492
Lung	H	22.47	0.7987	19.52	0.7389	25.87	0.8912	20.62	0.5551	32.91	0.9763	25.00	0.8374	33.10	0.9757
	E	22.05	0.7734	18.09	0.5088	25.53	0.8195	23.95	0.8939	30.77	0.9306	25.81	0.8426	31.02	0.9353
Mean	H	20.00	0.7448	23.22	0.8345	25.72	0.9100	22.11	0.6953	30.10	0.9611	25.85	0.8972	30.29	0.9621
	E	20.14	0.7611	21.08	0.6930	26.50	0.8883	23.94	0.8953	28.32	0.9303	26.21	0.8877	28.46	0.9336

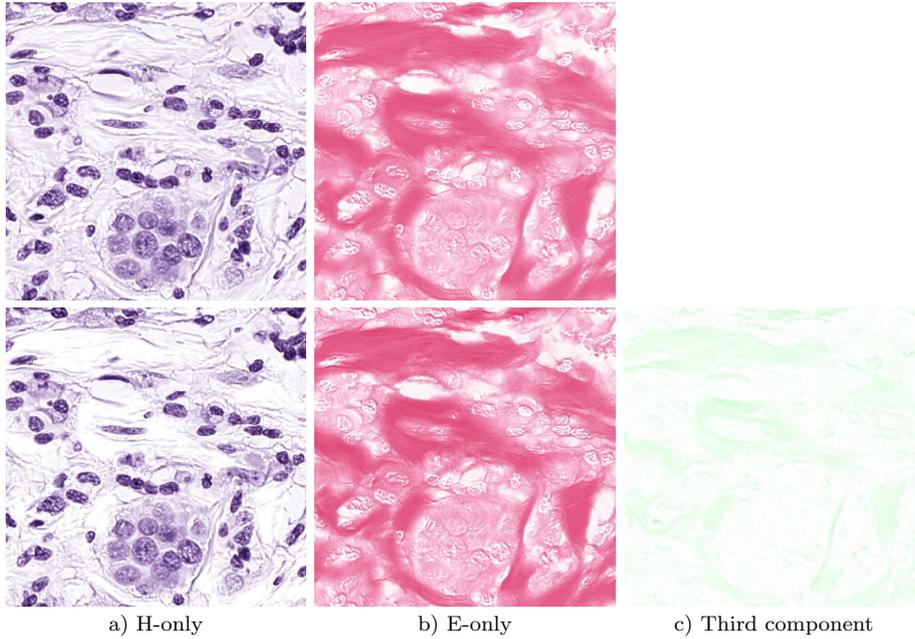


Figure 2: Detail of the H-only, E-only and third component separations of the bottom left corner of Fig. 1a obtained with the proposed method using two components (top) and three components (bottom).

obtained by Hidalgo-Gavira's method and is richer in details than Vahadane's method all the above is reflected in higher PSNR and SSIM values, see Table 1.

When a third component is used, the separation obtained by the proposed method, see Fig. 1h, is not so close to the ground-truth. Zoomed in areas of the bottom left corner of Fig. 1(h-i) are shown in Figure 2 for a better visual inspection. Colors are visually similar to the ones obtained when using two components, but some pixel information, specially from the background in the hematoxylin band, has been displaced to the third component. It can be observed that the third component has bright values, that is, only a small fraction of the information originally in the other bands is captured by this one, and nuclei in the H-only image appear brighter and are more clearly separated when three components are used, which will be extremely useful for classification. However, this implies a separation from the ground-truth images and, hence, lower values of PSNR and SSIM. In spite of the lower objective quality measure values, the separation in three components leads, as we will

Table 2: Computational time in seconds for the different methods on the WSSB dataset [16].

Method	Ruifrok [4]	Macenko [10]	Vahadane [8]	Alsubaie [16]	Hidalgo-Gavira [21]	Proposed $n_s = 3$	Proposed $n_s = 2$
Whole Dataset	147.68	141.47	375.10	210.13	357.03	877.67	507.28
Mean per image	6.15	5.89	15.62	8.75	14.87	36.56	21.13

see in the next section, to a better classification.

To conclude this section, Table 2 contains a computational time comparison between the competing methods. The method by Ruifrok is the fastest one. As complexity increases, the methods require higher computational time. Method by Vahadane requires as much time as the method by Hidalgo-Gavira but achieves lower PSNR and SSIM values. The proposed method takes longer than the competing ones but the higher computational burden is accompanied by higher figures-of-merit as already shown in Table 1. Note, also, that the proposed method estimates the model parameters together with the color-vector matrix and the concentrations, increasing the running time but making the method parameter free.

4.2. Prostate Cancer Classification Experiments

In this section we study how the use of different stain deconvolution methods affects the performance of classifiers. We use the SICAPv1 database, a prostate cancer histopathological database recently presented in [28]. The database contains 79 H&E WSIs from 48 patients scanned at 40x magnification, 19 correspond to benign prostate tissue biopsies (negative class) and 60 to pathological prostate tissue biopsies (positive class). In each pathological WSI, malignant regions were annotated by expert pathologists. The whole dataset was divided into a training set of 60 WSI (17 benign and 43 pathological), and a test set of 19 WSI (2 benign and 17 pathological). The images were downsampled to 10x scale and those in the training set were divided into patches of size 512×512 pixels and 1024×1024 pixels with a 50% overlap. Using this scale and patch size it is possible to capture complete glands in the 512×512 patches. Patches containing more than a 75% of background were discarded. Benign patches were extracted from benign WSI. Malignant patches were considered only if they contain at least a 25% of malignant tissue. Following [28] we will use cross-validation on this subset of the training set to assess the performance of the classifiers. Figure 3 shows an example of malignant patches with the areas annotated by the pathologist. In this experiment we only consider the dataset with patches of size 1024×1024 since it produced the best results in the patch classification experiments carried out in [28]. This training dataset contains 1909 patches from benign WSIs and 344 from pathological ones.

The dataset was color deconvolved using the proposed and competing methods. The H&E concentration image in the OD space was used to extract features to be utilized as input to the classifiers. Following [28], we used the concatenation of Local Binary Patterns Variance (LBPV) [30] and Geodesic granulometries (GeoGran) features [28]. LBPV features capture the texture and contrast information from the hematoxylin. GeoGran is an H&E granulometry based descriptor in the OD space recently proposed in [28] for prostate cancer classification. It encodes the structure of the glands by recovering, from the hematoxylin, the structure of the nuclei which formed the gland frontiers (those that enclosed their lumen

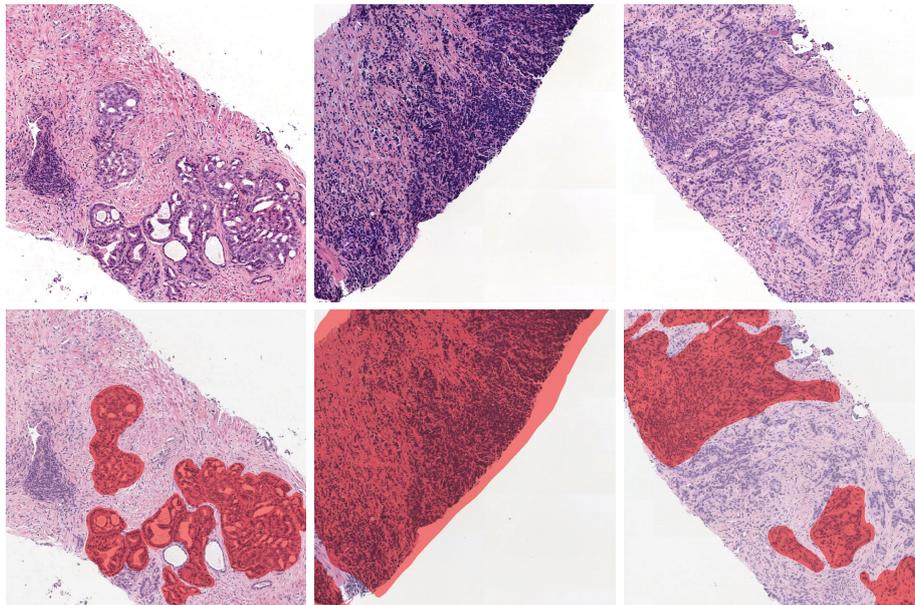


Figure 3: Top row: Patches extracted from SICAPv1 database. Bottom row: The same patches with the pathological areas annotated by the pathologist colored in red.

Table 3: AUC of the proposed and competing deconvolution methods with different classifiers.

Method	RF	GP	XgBoost	DGP
Ruifrok [4]	0.9789±0.0187	0.9855±0.0089	0.9764±0.0218	0.9737±0.0239
Macenko [10]	0.9315±0.0273	0.9535±0.0276	0.9425±0.0209	0.8802±0.0792
Vahadane [8]	0.9222±0.0318	0.9479±0.0321	0.9295±0.0325	0.9420±0.0436
Alsubaie [16]	0.9262±0.0586	0.9442±0.0294	0.9246±0.0612	0.9344±0.0581
Hidalgo-Gavira [21]	0.9157±0.0528	0.9542±0.0332	0.9228±0.0540	0.8997±0.0810
Proposed $n_s = 2$	0.9242±0.0579	0.9498±0.0332	0.9294±0.0824	0.9249±0.0638
Proposed $n_s = 3$	0.9798±0.0174	0.9856±0.0082	0.9797±0.0160	0.9718±0.0208

and cytoplasm). It also utilizes how distinguishable is, in the eosin, the lumen and nuclei structure from the rest of the stroma. This information is relevant to discriminate between pathological and benign tissues. The combination of LBPV and GeoGran features, which obtained the best classification results in the mentioned paper, allows to collect texture and structural information in the image, creating a descriptor able to accurately classify histopathological images.

A set of shallow and deep classifiers were trained with those descriptors and their results were compared. We used Random Forest (RF) [31], Extreme Gradient Boosting (XgBoost) [32], Gaussian Processes (GP)[33] and Deep Gaussian Processes (DGP)[34]. The tree-based ensemble models and the shallow and Deep GP can capture complex patterns in data and they are state-of-art classifiers. RF and XgBoost are configured with 1000 estimators and maximum depth of 20 and 30, respectively. A learning rate of 0.01 is chosen for XgBoost. Following the same approach as in [28] we use variational inference on a single-layer GP classifier with a RBF kernel [35]. For DGP, doubly stochastic variational inference [36] in a

Table 4: Accuracy of the proposed and competing deconvolution methods with different classifiers.

Method	RF	GP	XgBoost	DGP
Ruifrok [4]	0.9408±0.0301	0.9512±0.0272	0.9324±0.0505	0.9349±0.0337
Macenko [10]	0.8656±0.0277	0.8883±0.0561	0.8904±0.0205	0.8043±0.0399
Vahadane [8]	0.8870±0.0284	0.8826±0.0531	0.8830±0.0299	0.8996±0.0317
Alsubaie [16]	0.8825±0.0557	0.8793±0.0438	0.8730±0.0769	0.8885±0.05985
Hidalgo-Gavira [21]	0.8799±0.0105	0.8706±0.0445	0.8881±0.0673	0.8693±0.0810
Proposed $n_s = 2$	0.8914±0.0579	0.9029±0.0426	0.8910±0.0824	0.8797±0.0649
Proposed $n_s = 3$	0.9422±0.0375	0.9519±0.0319	0.9420±0.0339	0.9349±0.0257

Table 5: Accuracy of the proposed methods with different classifiers in train and test.

Method	RF	GP	XgBoost	DGP
$n_s = 2$ train	0.9789±0.0036	0.9794±0.0030	0.9697±0.0033	0.9401±0.0166
$n_s = 2$ test	0.8914±0.0579	0.9029±0.0426	0.8910±0.0824	0.8797±0.0649
$n_s = 3$ train	0.9774±0.0026	0.9878±0.0041	0.9796±0.0001	0.9605±0.0059
$n_s = 3$ test	0.9422±0.0375	0.9519±0.0319	0.9420±0.0339	0.9349±0.0257

three-layer classifier with RBF kernel and 100 inducing points per layer was used.

For each classifier, a five-fold cross-validation was applied to compare its performance with each deconvolution method. To avoid correlation between training and test sets, patches from the same image and patient were assigned to the same fold. Since the training set has more benign than pathological patches, an usual scenario on medical applications, balanced classifiers were built with all the pathological instances and a subset of the benign ones. The final prediction will be the average of the predictions of each classifier. The area under the ROC curve (AUC) obtained by the different deconvolution methods and classifiers is presented in Figure 4 and Table 3. Accuracy is shown in Table 4.

From Table 3, the best results are obtained using the proposed method with $n_s = 3$ and GP, with an AUC of 0.9856. The proposed method with $n_s = 3$ also obtains the best results among the shallow classifiers being the Ruifrok’s method the one obtaining the best result with the DGP classifier. When the proposed method is run with only two components results are also competitive but not as good as the ones obtained with three components. The curves in Figure 4 clearly show the advantage of the proposed $n_s = 3$ method and Ruifrok’s over the others. Average results of the method with $n_s = 2$ are also visible. From Table 4, the proposed method with $n_s = 3$ reaches the highest accuracy for all the classification methods. Notice that Ruifrok’s method was used for color deconvolution in [28] and so the figures of merits reported in the first line of Table 3 coincide with those reported in Table 5 in [28]. Finally, we would like to mention that in [28] a comparison with the deep learning methods VGG19, Inception v3, and Xception was carried out. The deep learning methods use as input the original RGB images, so the values reported for them in [28] are valid here. GPs and DGPs perform similarly and are competitive to VGG19, the best performing deep learning method in [28].

To assess the generalization capability of our model, we show in Table 5 the accuracy of the proposed method obtained for the train and test sets when performing cross validation.

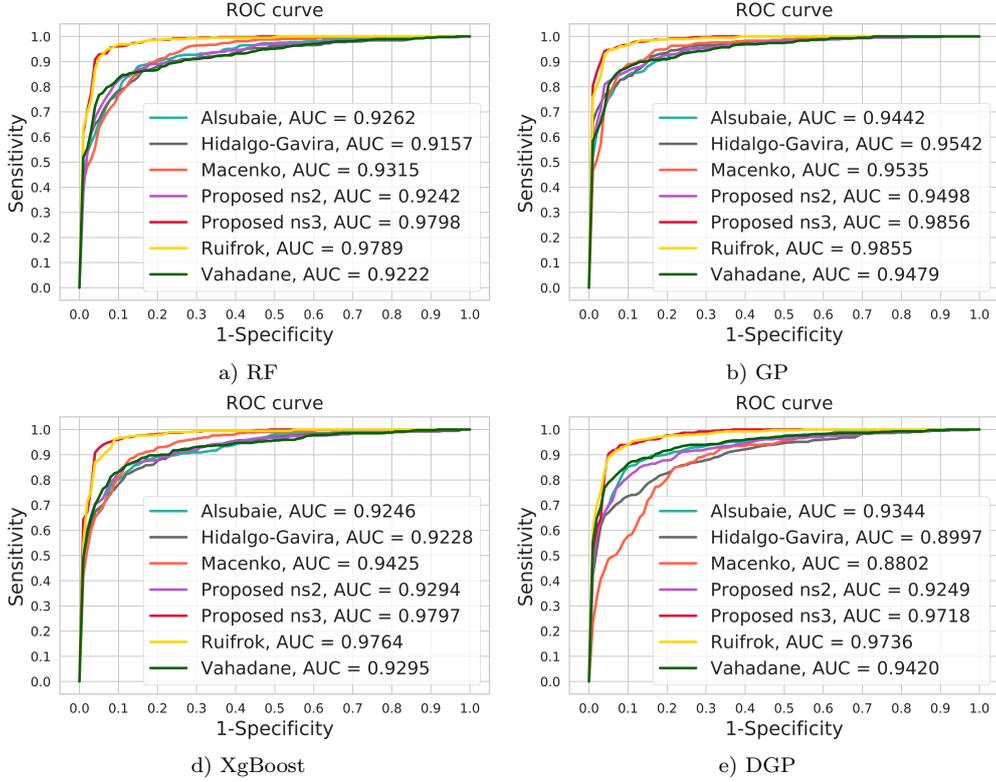


Figure 4: ROC curves for the competing methods and classifiers. Each sub-image contains all deconvolution methods AUC for a single classifier.

The use of $n_s = 2$ induced a higher overfitting to the train data in all the classification methods, reducing their generalization capability. For the GP and DGP models, Figure 5 includes the evolution of accuracy in train and test during the training procedure. Both GP and DGP models obtain a high accuracy from the beginning of the training and quickly converge. The overfitting when using $n_s = 2$ is visible in both models. The values obtained in training data using $n_s = 2$ and $n_s = 3$ are similar while the results obtained in testing data with $n_s = 2$ are much lower than the ones obtained with $n_s = 3$.

For classification, the use of a third component capturing residual information is clearly an advantage although the obtained images are not as close to the ground-truth separations as those obtained using $n_s = 2$. As seen in section 4.1, the third component is mainly capturing background information from the hematoxylin channel. An example of component concentration values in the OD space, which are used to extract the features, is shown in Fig. 6. The hematoxylin is used to extract LBPV and GeoGran features, that is, textures and nuclei structure. Due to the prostate tissue characteristics, the cytoplasm captures eosin and, partly, hematoxylin, so it appears also on the background of the hematoxylin band (see Fig. 6a). When three components are used, this background information is displaced to the third component. This also leads to a clearer hematoxylin (Fig. 6c) where nucleus information, belonging to the gland frontiers, is enhanced while the nucleus information

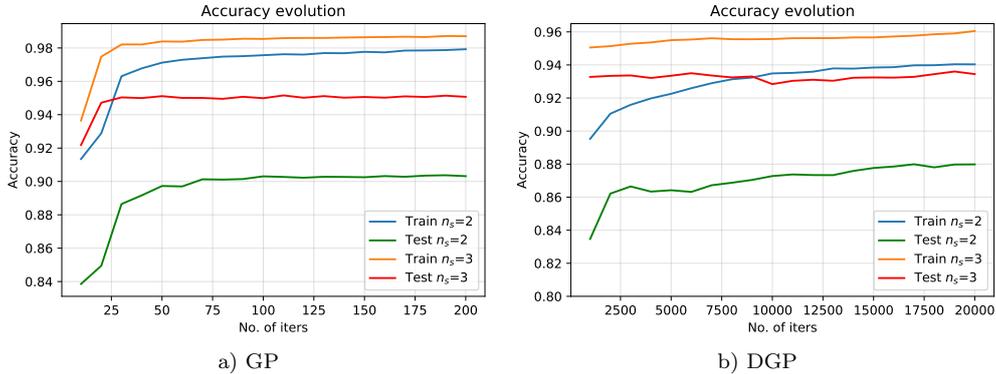


Figure 5: Train and test accuracy during the training procedure. a) GP model. b) DGP model.

belonging to the stroma (non-discriminative) appears in the third component. This allows to obtain less noisy features.

The eosin band is used to obtain GeoGran features to capture stroma information and identifies whether is invaded by nuclei or not. The use of three components makes the eosin band slightly more contrasted, which allows to obtain better descriptors. The joint use of descriptors extracted from hematoxylin and eosin bands by the proposed method using three component leads to an increased classification performance. The use of the TV prior, which produces sharper edges, also helps the feature extractors and, hence, the classifier.

4.2.1. Whole slide image evaluation

Our ultimate goal is to analyse full WSI images. To extend patch-wise classification to WSI classification, each WSI was split into overlapping patches. For each pixel, the probability of being cancerous was estimated by bilinearly interpolating the predicted probabilities of its four closest patches (using Euclidean distance to the center of the patches). A pixel-wise probability map was then obtained for each WSI. To assess the proposed method performance on this task, we deconvolved the train and test sets using $n_s = 3$. The GP classifier was then trained with the 60 images of the training set and used to predict the 19 WSIs in the test set. To obtain a better map resolution, 512×512 patches were used with 75% overlap. Figure 7 illustrates the result on a WSI of the test set. Probability maps are represented as heat maps. Red and blue colors are assigned to highest and lowest probabilities of being cancerous, respectively. The obtained probabilities correctly identify the annotated areas. Figure 8 shows zoomed in regions of interest.

5. Conclusions

In this work we have presented a framework for blind color deconvolution and classification of histological images. In this framework, we have developed a novel variational Bayesian blind color deconvolution algorithm which automatically estimates the concentration of stains, the color-vector matrix, and all the model parameters. It takes into account the spatial relations between pixels by means of a TV prior model, as well as the similarity

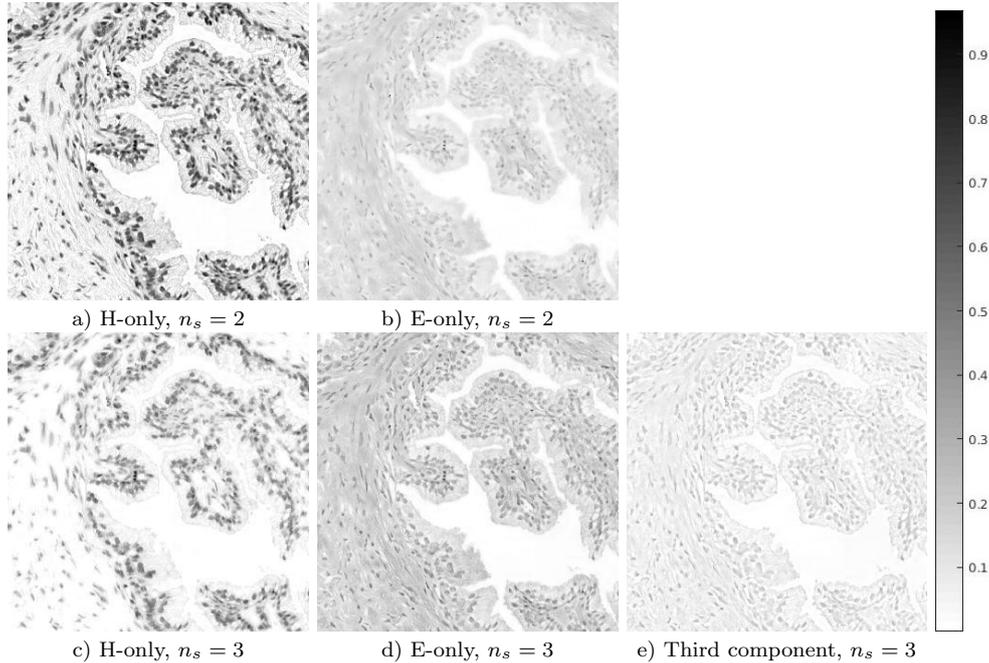


Figure 6: Detail of the H-only, E-only and third component concentration values in the OD space for a patch of SICAPv1 database obtained by the proposed method using two (top) and three components (bottom). The color-map of the images is inverted for a better visualization.

to a reference color-vector matrix. The use of the non-quadratic TV energy helps to reduce the noise in the images while preserving sharp edges.

For H&E stained images, color deconvolution with two components can be used in order to capture all stain details when visual inspection is needed. Classification algorithms, however, benefit from a clearer separation between classes. The use of a third, residual, component helps that separation by capturing information that is not completely explained by only one of the two stains. We found that, when using a third component, we obtain a clearer hematoxylin background, while nucleus information is enhanced and nuclei appear more clearly. The eosin is not severely modified, but the contrast of the image is increased which meliorates the discrimination power of this band. The use of a third component reduces the SSIM and PSNR values, but it helps the geodesic and LBPV descriptors to extract the relevant information and leads to better classification results.

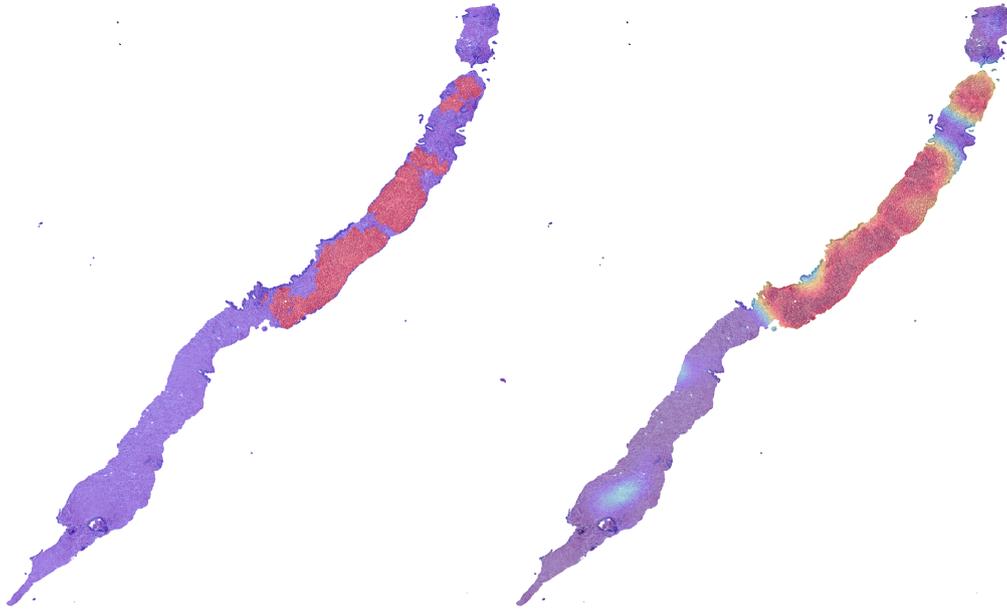


Figure 7: Full WSI comparison. Left: Areas annotated by the pathologists. Right: Probability maps (heat maps) obtained by the proposed method with $n_s = 3$ and GP classifier with 512×512 patches.

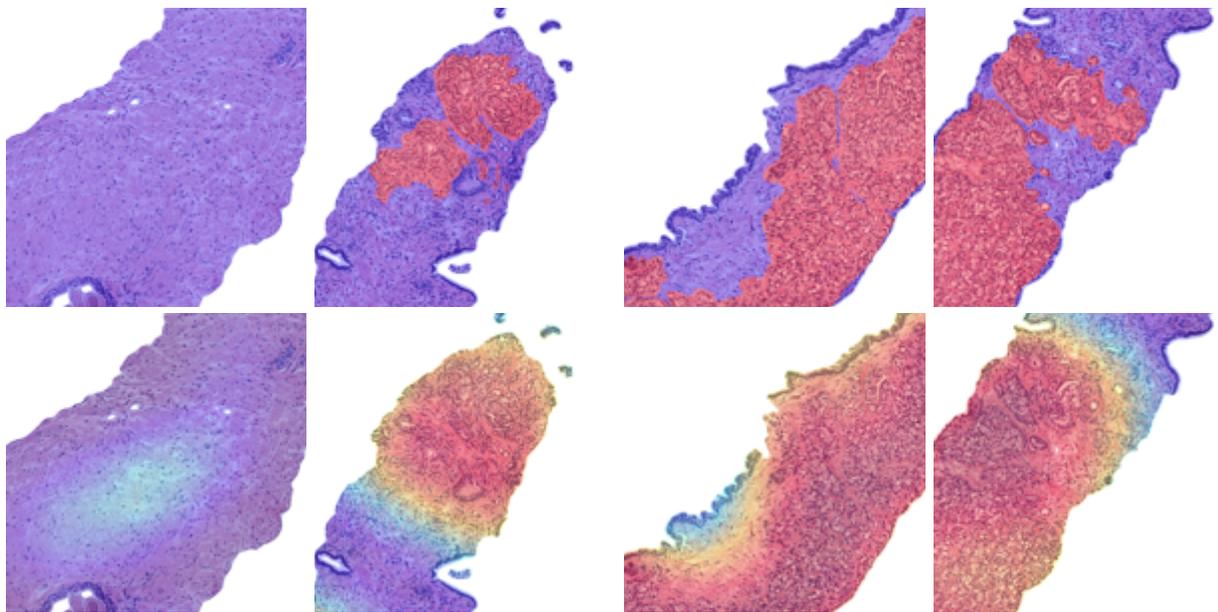


Figure 8: Regions of interest from Figure 7. Top row: annotations by a pathologist. Bottom row: Probability maps (heat maps) obtained.

6. References

References

- [1] A. H. Fischer, K. A. Jacobson, J. Rose, R. Zeller, Hematoxylin and Eosin Staining of Tissue and Cell Sections, Cold Spring Harbor Protocols (2008).
- [2] A. Rabinovich, S. Agarwal, C. Laris, J. H. Price, S. J. Belongie, Unsupervised color decomposition of histologically stained tissue samples, in: *Advances in Neural Information Processing Systems*, 2004, pp. 667–674.
- [3] T. A. A. Tosta, P. R. de Faria, L. A. Neves, M. Z. do Nascimento, Computational normalization of H&E-stained histological images: Progress, challenges and future potential, *Artificial Intelligence in Medicine* 95 (2019) 118 – 132.
- [4] A. C. Ruifrok, D. A. Johnston, Quantification of histochemical staining by color deconvolution, *Analytical and quantitative cytology and histology* 23 (2001) 291–299.
- [5] P. A. Bautista, Y. Yagi, Staining correction in digital pathology by utilizing a dye amount table, *Journal of digital imaging* 28 (3) (2015) 283–294.
- [6] T. Abe, H. Haneishi, P. A. Bautista, Y. Murakami, M. Yamaguchi, N. Ohyama, Y. Yagi, Color correction of red blood cell area in H&E stained images by using multispectral imaging, in: *4th European Conference on Colour in Graphics, Imaging, and Vision and 10th International Symposium on Multi-spectral Colour Science, CGIV*, 2008, pp. 432–436.
- [7] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images, *IEEE Computer Graphics and Applications* 21 (5) (2001) 34–41.
- [8] A. Vahadane, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A. M. Schlitter, I. Esposito, N. Navab, Structure-preserving color normalization and sparse stain separation for histological images, *IEEE Transactions on Medical Imaging* 35 (2016) 1962–1971.
- [9] J. Xu, L. Xiang, G. Wang, S. Ganesan, M. Feldman, N. N. Shih, H. Gilmore, A. Madabhushi, Sparse non-negative matrix factorization (SNMF) based color unmixing for breast histopathological image analysis, *Computerized Medical Imaging and Graphics* 46 (2015) 20–29.
- [10] M. Macenko, M. Niethammer, et al., A method for normalizing histology slides for quantitative analysis, in: *International Symposium on Biomedical Imaging (ISBI)*, 2009, pp. 1107–1110.
- [11] M. T. McCann, J. Majumdar, et al., Algorithm and benchmark dataset for stain separation in histology images, in: *International Conference on Image Processing (ICIP)*, 2014, pp. 3953–3957.
- [12] D. Carey, V. Wijayathunga, A. Bulpitt, D. Treanor, A novel approach for the colour deconvolution of multiple histological stains, in: *Proceedings of the 19th Conference of Medical Image Understanding and Analysis*, 2015, pp. 156–162.
- [13] M. Gavrilovic, J. C. Azar, et al., Blind color decomposition of histological images, *IEEE Transactions on Medical Imaging* 32 (2013) 983–994.
- [14] A. M. Khan, N. Rajpoot, D. Treanor, D. Magee, A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution, *IEEE Transactions on Biomedical Eng.* 61 (6) (2014) 1729–1738.
- [15] N. Alsubaie, S. E. A. Raza, N. Rajpoot, Stain deconvolution of histology images via independent component analysis in the wavelet domain, in: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 2016, pp. 803–806.
- [16] N. Alsubaie, N. Trahearn, S. E. A. Raza, D. Snead, N. Rajpoot, Stain deconvolution using statistical analysis of multi-resolution stain colour representation, *PLOS ONE* 12 (2017) e0169875.
- [17] N. Trahearn, D. Snead, I. Cree, N. Rajpoot, Multi-class stain separation using independent component analysis, in: *Medical Imaging 2015: Digital Pathology*, 2015, p. 94200J.
- [18] L. Astola, Stain separation in digital bright field histopathology, in: *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 2016, pp. 1–6.
- [19] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, J. Shi, C. Xue, Adaptive color deconvolution for histological WSI normalization, *Computer Methods and Programs in Biomedicine* 170 (2019) 107–120.

- [20] S. Roy, A. K. Jain, S. Lal, J. Kini, A study about color normalization methods for histopathology images, *Micron* 114 (2018) 42–61.
- [21] N. Hidalgo-Gavira, J. Mateos, M. Vega, R. Molina, A. K. Katsaggelos, Variational Bayesian blind color deconvolution of histopathological images, *IEEE Transactions on Image Processing* accepted for publication.
- [22] M. Vega, J. Mateos, R. Molina, A. K. Katsaggelos, Variational Bayes color deconvolution with a total variation prior, in: *27th European Signal Processing Conference, EUSIPCO 2019*, 2019, p. TuEP3.7.
- [23] S. Villena, M. Vega, R. Molina, A. Katsaggelos, A non-stationary image prior combination in super-resolution, *Digital Signal Processing* 32 (2014) 1–10.
- [24] P. Ruiz, X. Zhou, J. Mateos, R. Molina, A. Katsaggelos, Variational Bayesian blind image deconvolution: A review, *Digital Signal Processing* 47 (2015) 116–127.
- [25] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006, pp. 454–455.
- [26] S. Kullback, *Information Theory and Statistics*, Dover Pub., 1959.
- [27] S. D. Babacan, R. Molina, A. K. Katsaggelos, Parameter estimation in TV image restoration using variational distribution approximation, *IEEE Transactions Image Processing* (2008) 326–339.
- [28] A. E. Esteban, M. Lopez-Perez, A. Colomer, M. A. Sales, R. Molina, V. Naranjo, A new optical density granulometry-based descriptor for the classification of prostate histological images using shallow and deep Gaussian processes, *Computer Methods and Programs in Biomedicine* 178 (2019) 303–317.
- [29] G. Landini, Colour deconvolution, <https://blog.bham.ac.uk/intellimic/g-landini-software/colour-deconvolution/>, accessed: 2019-10-30.
- [30] Z. Guo, L. Zhang, D. Zhang, Rotation invariant texture classification using LBP variance (LBPV) with global matching, *Pattern Recognition* 43 (3) (2010) 706–719.
- [31] M. Valkonen, K. Kartasalo, K. Liimatainen, M. Nykter, L. Latonen, P. Ruusuvuori, Metastasis detection from whole slide images using local features and random forests, *Cytometry Part A* 91 (6) (2017) 555–565.
- [32] A. Pimkin, G. Makarchuk, V. Kondratenko, M. Pisov, E. Krivov, M. Belyaev, Ensembling neural networks for digital pathology images classification and segmentation, *Lecture Notes in Computer Science* 10882 LNCS (2018) 877–886.
- [33] C. Rasmussen, C. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2006.
- [34] A. Damianou, N. Lawrence, Deep Gaussian processes, *Journal of Machine Learning Research* 31 (2013) 207–215.
- [35] M. Opper, C. Archambeau, The variational Gaussian approximation revisited, *Neural Comput.* 21 (3) (2009) 786–792.
- [36] H. Salimbeni, M. Deisenroth, Doubly stochastic variational inference for deep Gaussian processes, in: *Advances in Neural Information Processing Systems*, 2017, pp. 4591–4602.



Fernando Pérez Bueno received the degree in Telecommunications Engineering from the Universidad de Granada in 2015 and the M.S. degree in Data Science and Computer Engineering from the University of Granada in 2019. He started the Ph.D. degree in 2018 at the University of Granada under the supervision of Prof. Molina, being a member of the Visual Information Processing Group in the Department of Computer Science and Artificial Intelligence. His research interests focus on the use of Bayesian modeling and inference to solve different problems related to image restoration and machine learning. Centered in cancer histopathological images improvement and classification. During his research, he has addressed several problems, such as blind image deconvolution, image denoising and image classification.



Miguel López Pérez received the degree in Mathematics from the Universidad de Granada in 2017 and the M.S. degree in Data Science and Computer Engineering from the University of Granada in 2018. He started the Ph.D. degree in 2018 at the University of Granada under the supervision of Prof. Molina and Prof. Katsaggelos, being a member of the Visual Information Processing Group in the Department of Computer Science and Artificial Intelligence. His research interests focus on the use of Bayesian modeling, specially Gaussian processes and their application to image processing, Computer Vision and classification problems, working usually with medical imaging problems.



Miguel Vega was born in 1956, in Spain. He received the Bachelor Physics degree from the Universidad de Granada, Granada, Spain, in 1979, and the Ph.D. degree from the Universidad de Granada, in 1984. He is with the E.T.S. Ing. Informática of the Universidad de Granada (Departamento de Lenguajes y Sistemas Informáticos), where he is an Associate Professor and teaches software engineering. His research focuses on image processing (multichannel and super-resolution image reconstruction). He has collaborated at several projects from the Spanish Research Council.



Javier Mateos received the degree in computer science in 1991 and the Ph.D. degree in computer science in 1998, both from the University of Granada. He joined the Department of Computer Science and Artificial Intelligence, University of Granada, in 1992 where he is currently a Professor. He was the deputy head of the Department of Computer Science and Artificial Intelligence from 2014 to 2018. He is coauthor of “Superresolution of Images and Video” (Claypool, 2006) and “Multispectral Image fusion Using Multiscale and Super-resolution Methods” (VDM Verlag, 2011). He received the IEEE International Conference on Signal Processing Algorithms, Architectures, Arrangements, and Applications best paper award (2013) and was finalist for the IEEE International Conference on Image Processing Best Student Paper Award (2010). He is conducting research on image and video processing, including image restoration, image and video recovery, super-resolution from (compressed) stills and video sequences, pansharpening and image classification. Dr. Mateos is a senior member of the IEEE. He serves as an editor of “Digital Signal Processing” (2011–present) and as an Associate Editor for the “IEEE Transactions on Image Processing” (2014–2017).



Valery Naranjo received the Ph.D. degree in telecommunications in 2002. She is currently a Professor at the Universitat Politècnica de València, Spain. She teaches Digital Signal Processing and Digital Image Processing in the telecommunication, biomedical degree and in the master's degree in technology, systems and communication networks of the Communications Department of the Universitat Politècnica de València. She has worked in the field of signal, image and video processing for different applications such as video surveillance, coding, and restoration. Her research interests also include medical image processing, hyperspectral image analysis, and physiological signal processing.



Rafael Molina was born in 1957. He received the degree in mathematics (statistics) and the Ph.D. degree in optimal design in linear models from the University of Granada, Granada, Spain, in 1979 and 1983, respectively. He became Professor of Computer Science and Artificial Intelligence at the University of Granada, Granada, Spain, in 2000. He is the former Dean of the Computer Engineering School at the University of Granada (1992–2002) and Head of the Computer Science and Artificial Intelligence department of the University of Granada (2005–2007). His research interest focuses mainly on using Bayesian modeling and inference in problems like image restoration (applications to astronomy and medicine), superresolution of images and video, blind deconvolution, computational photography, source recovery in medicine, compressive sensing, low-rank matrix decomposition, active learning, fusion, and classification. Prof. Molina serves as an Associate Editor of Applied Signal Processing (2005–2007); the IEEE Transactions on Image Processing (2010–present); and Progress in Artificial Intelligence (2011–present); and an Area Editor of Digital Signal Processing (2011–present). He is the recipient of an IEEE International Conference on Image Processing Paper Award (2007) and an ISPA Best Paper Award (2009). He is a coauthor of a paper awarded the runner-up prize at Reception for early-stage researchers at the House of Commons.



Aggelos K. Katsaggelos (F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Greece, in 1979, and the M.S. and Ph.D. degrees in Electrical Engineering from the Georgia Institute of Technology, in 1981 and 1985, respectively. In 1985, he joined the Department of Electrical Engineering and Computer Science at Northwestern University, where he is currently a Professor, holder of the Joseph Cummings Chair. He is a member of the Academic Staff, NorthShore University Health System, an affiliated faculty at the Department of Linguistics and he has an appointment with the Argonne National Laboratory. He was previously the holder of the Ameritech Chair of Information Technology and the AT&T Chair and the co-Founder and Director of the Motorola Center for Seamless Communications. He has published extensively in the areas of multimedia processing and communications (over 250 journal papers, 500 conference papers and 40 book chapters) and he is the holder of 26 international patents. He is the co-author of Rate-Distortion Based Video Compression (Kluwer, 1997), Super-Resolution for Images and Video (Claypool, 2007), Joint Source-Channel Video Transmission (Claypool, 2007), Machine Learning Refined (Cambridge University Press, 2016) and The Essentials of Sparse Modeling and Optimization (Springer, 2017, forthcoming). He has supervised 56 PhD dissertations so far. Among his many professional activities Prof. Katsaggelos was Editor-in-Chief of the IEEE Signal Processing Magazine (1997-2002), a BOG Member of the IEEE Signal Processing Society (1999-2001), a member of the Publication Board of the IEEE Proceedings (2003-2007), and he is currently a member of the Awards Board of the Signal Processing Society. He is a Fellow of the IEEE (1998) and SPIE (2009) and the recipient of the IEEE

Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), the IEEE Signal Processing Society Technical Achievement Award (2010), an IEEE Signal Processing Society Best Paper Award (2001), an IEEE ICME Paper Award (2006), an IEEE ICIP Paper Award (2007), an ISPA Paper Award (2009), and a EUSIPCO paper award (2013). He was a Distinguished Lecturer of the IEEE Signal Processing Society (2007-2008).