Deep Multiresolution Cellular Communities for Semantic Segmentation of Multi-Gigapixel Histology Images

4 authors:

Sajid Javed
Khalifa University of Science and Technology
127 PUBLICATIONS **3,156** CITATIONS

SEE PROFILE

Arif Mahmood
Information Technology University (ITU)
169 PUBLICATIONS **3,716** CITATIONS

SEE PROFILE

Naoufel Werghi
Khalifa University
320 PUBLICATIONS **3,671** CITATIONS

SEE PROFILE

Nasir Rajpoot
The University of Warwick
290 PUBLICATIONS **13,721** CITATIONS

SEE PROFILE

Deep Multiresolution Cellular Communities for Semantic Segmentation of Multi-Gigapixel Histology Images

Sajid Javed[1], Arif Mahmood[2], Naoufel Werghi[1], and Nasir Rajpoot[3]

[1] Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates.

[2] Department of Computer Science, Information Technology University, Lahore, Pakistan.

[3] Department of Computer Science, University of Warwick, Coventry, CV4 7AL, UK

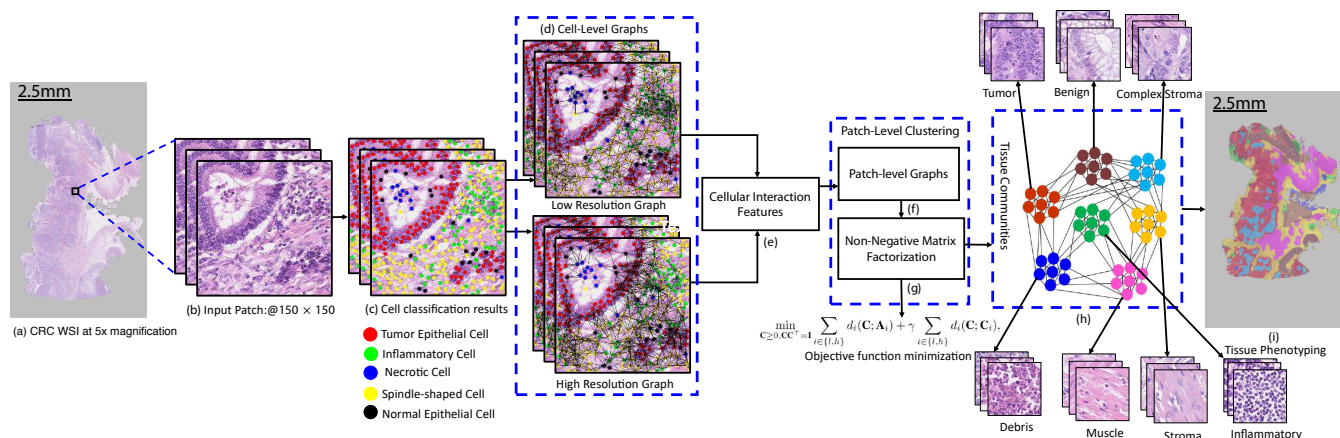{sajid.javed,naoufel.werghi}@ku.ac.ae, arif.mahmood@itu.edu.pk, N.M.Rajpoot@warwick.ac.uk

Figure 1: Schematic illustration of the semantic segmentation (or *tissue phenotyping*) problem in computational pathology and an overview of our proposed deep multiresolution cellular community detection for tissue phenotyping. (a) the input multi-gigapixel whole slide image (WSI) of ColoRectal Cancer (CRC) from our local university hospital; (b) Shows an input patch of size 150×150 captured at $20\times$; (c) results of cell detection and classification using spatially constrained deep neural network [26], where red, green, blue, yellow, and black colors represent tumor epithelial cell, inflammatory cell, debris or necrotic cell, spindle-shaped cell, and normal epithelial cell, respectively; (d) low- and high-resolution cell-level graphs – these graphs are constructed using cell-level coarse and fine =convolutional features of the deep neural network; (e)-(f) computation of cellular interaction features and the construction of low- and high-resolution patch-level graphs; (g) our proposed novel objective function minimized using a non-negative matrix factorization algorithm for clustering patch-level graphs into meaningful tissue components including tumor, stroma, muscle, inflammatory, debris or necrotic, benign, and complex stroma shown in (h); (i) results of the different tissue components overlaid on input CRC WSI (a). This problem is similar to the semantic segmentation problem for natural scene analysis in the computer vision community.

Abstract

Tissue phenotyping in cancer histology images is a fundamental step in computational pathology. Automatic tools for tissue phenotyping assist pathologists for digital profiling of the tumor microenvironment. Recently, deep learning and classical machine learning methods have been proposed for tissue phenotyping. However, these methods do not integrate the cellular community interaction features which present biological significance in tissue phenotyping context. In this paper, we propose to exploit deep multiresolution cellular communities for tissue phenotyping from multi-level cell graphs and show that such communities offer better performance compared to the deep learning and texture-based methods. We propose to use deep features extracted from two distinct layers of a deep neural network at the cell-level, in order to construct cellular graphs encoding cellular interactions at multiple scales. From these graphs, we extract cellular interaction-based features, which are then employed to construct patch-level graphs. Multiresolution communities are detected by considering the patch-level graphs as layers of multi-level graphs, and also by proposing novel objective function based on non-negative matrix factorization. We report results of our experiments on two datasets for colon cancer tissue phenotyping and demonstrate excellent performance of the proposed algorithm as compared to current state-of-the-art methods.

1. Introduction

Computational pathology is a fast-growing research area in medical imaging research [5, 20, 28, 31]. Cell detection, classification, and tissue phenotyping are considered as some of the main objectives of computational pathology which operates on digitized Whole Slide Images (WSIs) of tissue slide stained with routine Hematoxylin & Eosin (H&E) dyes [14, 16, 26, 27]. These cancer WSIs contain tens of billions of pixels (multi-giga pixels image data) at the highest resolution level of 40× shown in Fig. 1 (where we show a 150×150 pixel patch of the WSI at 20× magnification level), posing computational challenges as they are much larger in terms of image resolution than natural images. The application of tissue phenotyping can serve as building blocks for the development of computational pathology tools for systematic digital profiling of the spatial Tumor MicroEnvironment (TME) [16, 21, 22, 24, 27]. In clinical practice, such tools can be employed for better cancer grading and prognostication [12, 20, 22, 27].

Because of the wide applications and significance of automatic tissue phenotyping, several methods have been proposed in the literature [3, 7, 14–16, 19, 25, 30, 34]. Among these methods, patch based texture analysis is a popular approach for tissue phenotyping in which classifiers are trained on texture features of the histology image patches computed using local binary patterns, Gabor filters, or histograms features [3, 16, 17, 19, 25, 30]. Bionci *et al.* proposed to employ the SVM classifier on perception-like features for tissue classification [3]. Linder *et al.* proposed a simple SVM classifier trained on a set of local binary patterns and contrast measure features [19]. These studies were mainly limited to only stroma and tumor tissue phenotypes. However, ColoRectal Cancer (CRC) tissue consists of a rich mixture of several different tissue phenotypes including smooth muscle, inflammatory, necrotic, complex stroma, and benign tissue, as shown in Figs. 1 (h)-(i). To address the challenge of detailed tissue phenotyping, Sarkar *et al.* have proposed a saliency-guided dictionary learning method where Gabor features were computed and trained for binary and multi-class tissue classification [25]. Kather *et al.* also proposed a set of six textures features for multi-class tissue image classification [16]. Although texture-based feature analysis methods may be attractive due to their simplicity, texture features do not fully capture the biological significance of tissue components resulting in performance degradation [16].

Deep learning methods have also been proposed to classify WSIs into distinct meaningful tissue regions [11, 13, 22, 33, 34]. These methods use deep neural networks to extract convolutional features from histology image patches which are then used to train the classifiers. For instance, Xu *et al.* proposed an SVM classifier trained on CNN features extracted from a pre-trained AlexNet model for multi-class

tissue classification and segmentation in CRC [34]. Kather *et al.* recently used the VGG-19 network for end-to-end classification of tissues into eight distinct classes and predicted the survival analysis of CRC patients [15].

Recently, Javed *et al.* have introduced the notion of cellular community detection for tissue phenotyping [14]. In their work, they used a network for cellular interaction-based features which are then used for the detection of patch-level tissue communities. This technique enhanced the modeling of cellular community interaction. However, a limitation of that work is that the interaction between nearby cells extend beyond a single cellular network, and therefore, a more holistic model would best represent interaction between cellular communities. To that end, we propose to use a deep multi-resolution cellular networks to capture cellular interactions at different levels. To the best of our knowledge, such a multi-level graph network has not been used before for tissue phenotyping in histology images.

In our approach, we consider two types of cellular networks capturing cellular interactions encoded at two different resolutions by a deep neural network. At high resolution, fine cell details are encoded while at low-resolution semantic cell information is encoded as cellular features. These two types of features are then used for the construction of two distinct cellular networks which are then exploited for the computation of cellular interaction features at fine and coarse levels. The cellular interaction features are further used for the construction of two patch-level graph networks which we consider as two layers of a multi-layer graph. Each layer consists of the same number of nodes but varying number of edges. Multiresolution community detection is performed using a non-negative matrix factorization method on the adjacency matrices of the multi-layer graph [8]. Here, we propose a novel objective function that learns a low-dimensional subspace from each graph layer that are then fused into a common subspace on the Grassman manifold [1, 6, 23]. We dub our proposed algorithm as DEep MUltiresolution Cellular-communities (DeMuC) for tissue phenotyping and the main notion of DeMuC algorithm is presented in Fig. 1.

The rest of this paper is organized as follows. A review of recent literature on tissue phenotyping is given in Section 2. Section 3 describes the proposed algorithm in detail. Experiments and results are discussed in Section 4, and finally, conclusions and future directions are given in Section 5.

2. Related Work

In the past few years, the computational pathology research community has presented a number of tissue phenotyping studies [2, 3, 7, 11, 14–17, 19, 24, 25, 30, 32–34]. Broadly, we group tissue image classification methods into two categories including texture feature-based methods [3, 16, 17, 19, 25, 30] and deep learning-based methods
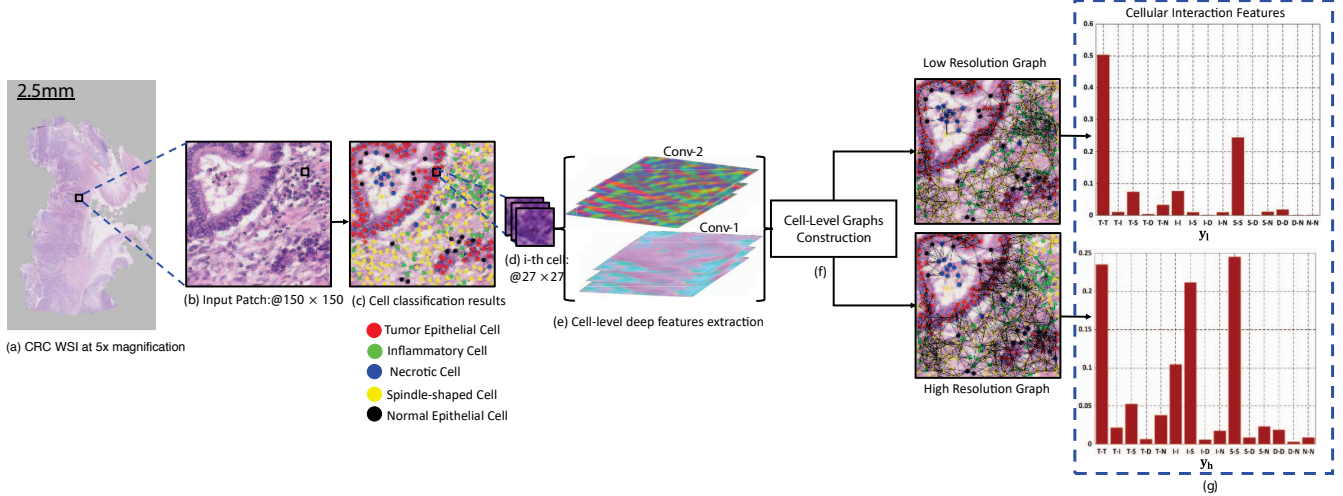
Figure 2: An illustration of the low-resolution and high-resolution cellular interaction features computation which are used in the patch-level graphs construction. Steps (a)-(c) show the input WSI of CRC, the input patch, and the results of the cell detection and classification. Step (d) shows the i-th cell cropped from (c) containing 27×27 pixels. Step (e) shows the low-resolution and high-resolution deep features extracted from the cell patch shown in (d). Step (f) shows the construction of two cell-level graphs at a coarse and fine level, and step (g) shows the 15 dimensional cellular interaction features vectors computed from cell-level graphs

[2, 7, 11, 15, 24, 32, 33].

The texture feature-based methods compute the local tissue features and train classifiers to predict different tissue phenotypes [16, 25]. For example, Tamura *et al.* proposed five different perception-like features including coarseness, contrast, directionality, line-likeness, and roughness texture measures [30]. Bionci *et al.* used the SVM classifier on these features for binary tissue classification [3]. Linder *et al.* proposed to use LBP with contrast measure to train a linear SVM classifier for tumor epithelium and stromal tissue components [19]. These studies presented encouraging results however, they were limited for the discrimination of either one or two tissue classes such as tumor and stroma. Kather *et al.* recently proposed a multi-texture features analysis method for eight class tissue classification in colon cancer histology images [16]. They used a set of six different texture features including, LBP, lower order and higher order histogram features, Gabor features, gray level co-occurrence matrix, and perception-like features. We observe that texture features do not fully capture the biological significance of the tissue components and these features are not robust in discriminating tumor from the complex stromal component.

A number of deep learning methodologies have also been proposed in the literature for tissue phenotyping [11, 15, 32, 33]. These methods train end-to-end CNNs for direct patch-based tissue classification. The training process learns a rich hierarchy of convolutional features of each tissue class and then predicts the tissue type based on the soft-max classifier. Xu *et al.* proposed a deep network for breast cancer histology images [33]. Huang *et al.* incorpo-

rated the idea of domain adaptation in the classical Alex-Net and Google Net [11]. Yu *et al.* extracted deep features from CRC WSIs and then used a linear SVM classifier for classification and segmentation [34]. Bejnordi *et al.* also proposed deep networks for the classification of breast cancer WSIs [2]. More recently, Kather *et al.* proposed a large-scale study to predict survival from CRC WSIs [15]. They used a pre-trained VGG-19 model and fine-tuned this network on nine large-scale tissue classes to estimate deep tumor-stroma score. The aforementioned deep learning methods produced good results for multi-class tissue classification. However, these methods require a very large amount of histology data for training, which is not always available since WSIs require exhaustive annotations by experienced pathologists which incur a high cost.

In contrast, in our approach, we propose a semi-supervised algorithm for seven class tissue classification. We use automatic cell classification for estimating cellular interaction among different cellular components. Our cellular interaction features encode more biological significance as compared to texture features and deep convolutional features and hence, has more potential for improving the performance of tissue phenotyping.

3. Proposed Algorithm

The pipeline of the proposed algorithm is shown in Fig. 1 and the computation of cellular interaction features is presented in Fig. 2. The main components of the proposed algorithm include a cell detection and classification network which is also used as a deep multi-resolution fea-

tures extractor, the construction of multi-resolution cell-level graphs, computation of cellular interaction features, construction of multi-layer patch-level graph, and estimation of cellular communities using non-negative matrix factorization method. In the later subsections, we describe each step of the proposed algorithm in detail.

3.1. Cell Detection, Classification, and Feature Extraction

We first extract non-overlapping patches from each WSI of CRC. In this work, we use a patch of size 150×150 pixels captured at 20x magnification level. For cell detection, we use a spatially constrained deep neural network proposed by Sirinukunwattana *et al.* [26]. For cell classification, a deep neural network is trained to classify five different cell types including Tumor epithelial (T), Debris (D), Inflammatory (I), Spindle-shaped (S), and Normal epithelial (N). This network consists of two convolutional layers each followed by a max-pooling layer which is followed by three fully connected layers. The output of the network is a set of five different types of cell nuclei's as shown in Fig. 2 (c). For each cellular component, the same network is also used for multi-resolution deep features extraction. After each convolutional layer, the coefficients are extracted and used as features for the input cell patch. The first convolutional layer provides fine image details while the second convolutional layer provides the semantic information regarding the input cell as shown in Fig. 2 (e).

3.2. Cell-Level Graphs Construction

The features extracted from each convolutional layer are used for the construction of the cell-level graph among the cells in a particular patch. The strength of the connection between two cells is assumed to be inversely proportional to the Euclidean distance between their respective features as given below.

$$A(i,j) = \begin{cases} 1, & \text{if } \exp(-\frac{||f_i - f_j||_2}{2\sigma^2}) \leq \tau_1, \\ 0, & \text{otherwise} . \end{cases} \quad (1)$$

where $A_{i,j}$ is the edge weight between i-th and j-th cells, f_i and f_j are the corresponding cell deep features and σ is the bandwidth parameter controlling edge weigh decay with increasing distance between two cells in the feature space. The edges between cells which are relatively at a larger distance are removed based on a threshold, τ_1. The edges between cells which are relatively closer, having distance less than the threshold, are assigned a unit weight. This will result in a binary adjacency matrix representing an unweighted graph. The two cell-level graphs corresponding to coarse and fine resolutions are shown in Fig. 2. It should be noted that the number of cellular components as a nodes

are same in both cell-level graphs but varying edges which show different kind of similarities among cell types.

3.3. Patch-Level Cellular Interaction Features

Using the graphs of cells contained in a particular patch, we compute the patch-level cellular interaction features, $\mathbf{y}_p \in \mathbb{R}^m$, where p is the patch index and m is the dimension of cellular interaction features ($m = 15$ in this work). If there is an edge between two cells, we assume those cells are interacting with each other. For five cellular types, we obtain fifteen different cellular interactions including T to T, T to I, T to S, T to D, T to N, I to I, I to S, I to D, I to N, S to S, S to D, S to N, D to D, D to N, and N to N as shown in Fig. 2 (g). The probability distribution of these cellular interactions is used as a patch-level feature. The probability distribution is estimated as a histogram of different types of cellular interactions. All patch features are concatenated into a matrix $\mathbf{Y} = \{\mathbf{y}_p\}_{p=1}^n$, where n is the total number of patches in a CRC WSI. We get two such input matrices $\mathbf{Y}_l$ and $\mathbf{Y}_h$ corresponding to low and high-resolution features.

3.4. Multi-layer Patch-Level Graphs Construction

Using cellular interaction features, we construct two undirected graphs $\mathbf{G}_l = (\mathbf{V}_l, \mathbf{A}_l)$ and $\mathbf{G}_h = (\mathbf{V}_h, \mathbf{A}_h)$ corresponding to low resolution and high-resolution features. A vertex $\mathbf{v}_l$ in $\mathbf{G}_l$ corresponds to $\mathbf{y}_l$ in the low resolution features matrix $\mathbf{Y}_l$. $\mathbf{A}_l \in \mathbb{R}^{n \times n}$ is the adjacency matrix of $\mathbf{G}_l$, which is computed by employing chi-squared distance as:

$$A_l(i,j) = \begin{cases} 1, & \text{if } \exp\left(-\frac{1}{\sigma}\sum_{k=1}^m \frac{(\mathbf{y}_l(i,k) - \mathbf{y}_l(j,k))^2}{\mathbf{y}_l(i,k) + \mathbf{y}_l(j,k)}\right) \leq \tau_2, \\ 0, & \text{otherwise} . \end{cases}$$

$$(2)$$

similarly, the adjacency matrix $\mathbf{A}_h \in \mathbb{R}^{n \times n}$ of graph $\mathbf{G}_h$ is computed using Eq. (2). τ_2 is a threshold on the chi-square similarity shown by Eq. (2). If the similarity between two patches (nodes) is less than a τ_2, then it is considered as zero or unconnected, otherwise, it is considered as 1.

3.5. DeMuC Mathematical Formulation

A network community may be defined as a set of nodes more densely connected to each other compared to the other nodes on that network. The problem of community detection is to find the community assignment of all nodes. Let $\mathbf{C} \in \{0,1\}^{n \times k}$, be the community membership matrix for n nodes and k communities, $\mathbf{C}_{ij} = 1$, if node i belongs to community j, and $\mathbf{C}_{ij} = 0$, otherwise. Since the communities considered in the tissue phenotyping problem are non-overlapping therefore, only one element in each row of $\mathbf{C}$ can be 1 and all others will be zero, $\sum_{j=1}^k \mathbf{C}_{ij} = 1$.

For a given adjacency matrix $\mathbf{A}_i$, for $i \in \{l, h\}$, the community matrix $\mathbf{C}_i$, for $i \in \{l, h\}$, is found using a

projective non-negative matrix factorization method [35]. We factorize each adjacency matrix as $\mathbf{A}_l \approx \mathbf{C}_l \mathbf{C}_l^\top \mathbf{A}_l$ and $\mathbf{A}_h \approx \mathbf{C}_h \mathbf{C}_h^\top \mathbf{A}_h$. The community matrices $\mathbf{C}_l$ and $\mathbf{C}_h$ are a low-rank representation of the corresponding adjacency matrices $\mathbf{A}_l$ and $\mathbf{A}_h$. These matrices are further constrained to be non negative as $\{\mathbf{C}_l, \mathbf{C}_h\} \geq 0$, and orthogonal as $\mathbf{C}_h^\top \mathbf{C}_h = \mathbf{C}_l^\top \mathbf{C}_l = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. To enforce these properties, the objective function can be written as:

$$\min_{\mathbf{C}_i \geq 0, \mathbf{C}_i \mathbf{C}_i^\top = \mathbf{I}} ||\mathbf{A}_i - \mathbf{C}_i \mathbf{C}_i^\top \mathbf{A}_i||_F^2, \text{ for } i \in \{l, h\}, \quad (3)$$

where $||\cdot||_F$ is the Frobenius norm and it is equal to $||\mathbf{C}||_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{k} |c(i,j)|^2}$.

For the multi-layer network having low resolution and high resolution networks as layers, we intend to compute a common community matrix $\mathbf{C}$ across both layers such that $\mathbf{C}$ is close to both $\mathbf{C}_l$ and $\mathbf{C}_h$. Combining the individual and combined community matrices computation, we formulate our novel objective function as follows:

$$\min_{\mathbf{C} \geq 0, \mathbf{C}\mathbf{C}^\top = \mathbf{I}} \sum_{i \in \{l,h\}} d_i(\mathbf{C}; \mathbf{A}_i) + \gamma \sum_{i \in \{l,h\}} d_i(\mathbf{C}; \mathbf{C}_i), \quad (4)$$

where $\sum_{i \in \{l,h\}} d(\mathbf{C}; \mathbf{A}_i)$ is the objective function for clustering individual layers, and $\sum_{i \in \{l,h\}} d(\mathbf{C}; \mathbf{C}_i)$ is the loss function which minimizes the distance between the consensus low-rank community representation matrix $\mathbf{C}$ and each individual low-rank community representation matrix $\mathbf{C}_i$ ($\mathbf{C}_l$ and $\mathbf{C}_h$ already computed using Eq. (3)), and $\gamma > 0$ is the parameter that assigns relative importance to both terms while minimizing Eq. (4). In the following subsections, we provide the solutions of the proposed objective functions (3) and (4).

3.6. DeMuC Optimization

In this section, we derive the multiplicative update rules to solve the objective functions (3) and (4). Similar to [18], these multiplicative update rules are used for finding the local minimum of the optimization problems (3)-(4).

3.6.1 Update Rules for Solving (3)

We provide the multiplicative update rule using projective non-negative matrix factorization method [35] for each adjacency matrix $\mathbf{A}_i$, to estimate each low-dimensional subspace known as low-rank community representation matrix $\mathbf{C}_i$ as:

$$\forall i \in \{l, h\}, \mathbf{C}_i(jk) \leftarrow \mathbf{C}_i(jk) \frac{[\mathbf{A}_i \mathbf{A}_i^\top \mathbf{C}_i]_{jk}}{[\mathbf{C}_i \mathbf{C}_i^\top \mathbf{A}_i \mathbf{A}_i^\top \mathbf{C}_i]_{jk}}, \quad (5)$$

where $\mathbf{C}_i(jk)$ is the j-th element in the k-th community for i-th low-rank community matrix $\mathbf{C}_i$. (5) converges to the optimal solution if the difference between the matrices $\mathbf{C}_i(t)$ and $\mathbf{C}_i(t-1)$ is less than a tolerance factor ζ, where t is the iteration index.

3.6.2 Update Rules for Solving (4)

The multiplicative update rule for the consensus low-rank community representation matrix $\mathbf{C}$ in the objective function (4) can be derived similarly to (5) using projective non-negative matrix factorization method as follows:

$$\mathbf{C}(jk) \leftarrow \mathbf{C}(jk) \frac{[\mathbf{A}_{avg}\mathbf{C}]_{jk}}{[\mathbf{C}\mathbf{C}^\top \mathbf{A}_{avg}\mathbf{C}]_{jk}}, \quad (6)$$

$$\text{where } \mathbf{A}_{avg} = \sum_{i \in \{l,h\}} \mathbf{A}_i \mathbf{A}_i^\top + \gamma \mathbf{C}_i \mathbf{C}_i^\top \quad (7)$$

while we derive the solution for our collective objective function (4) and then formulate its multiplicative update rules. The first term $d_i(\mathbf{C}; \mathbf{A}_i)$, $\forall i \in \{l, h\}$, in the objective function (4) is equivalent to (3).

For the second term $d(\mathbf{C}; \mathbf{C}_i)$, $\forall i \in \{l, h\}$, in (4), we utilize the orthonormal property of non-negative, low-rank community representation matrices, $\mathbf{C}_i$, and propose a distance measure based on this property. Dong *et al.* proposed to estimate subspace on Grassman manifold [6]. A Grassman manifold $\mathbb{G}(k, n)$ is a set of k-dimensional linear subspaces in $\mathbb{R}^n$. Given that, each orthonormal low-rank community reperesentation matrix, $\mathbf{C}_i \in \mathbb{R}^{n \times k}$, spaning the corresponding k-dimensional non-negative subspace, $span(\mathbf{C}_i)$ in $\in \mathbb{R}^n$, is mapped to a unique point on the Grassman manifold $\mathbb{G}(k, n)$. The geodesic distance between two subspaces can be computed by projection distance [8]. For instance, the squared distance between two subspaces, $\mathbf{C}_i$ and $\mathbf{C}_j$, is computed as follows:

$$d_{proj}^2 = \sum_{i=1}^{k} sin^2 \theta_i = k - \sum_{i=1}^{k} cos^2 \theta_i = k - tr(\mathbf{C}_i \mathbf{C}_i^\top \mathbf{C}_j \mathbf{C}_j^\top), \quad (8)$$

where $\{\theta_i\}_{i=1}^{k}$ are principal angles between k-dimensional subspaces, $span(\mathbf{C}_i)$ and $span(\mathbf{C}_j)$. Following this approach, we can write the second part $d(\mathbf{C}; \mathbf{C}_i)$, $\forall i \in \{l, h\}$, in our collective objective function (4) as:

$$\begin{aligned} \forall i \in \{l, h\}, d_i(\mathbf{C}; \mathbf{C}_i) &= k - tr(\mathbf{C}\mathbf{C}^\top \mathbf{C}_i \mathbf{C}_i^\top) \\ &= ||\mathbf{C}\mathbf{C}^\top - \mathbf{C}_i \mathbf{C}_i^\top||_F^2. \end{aligned} \quad (9)$$

where the consensus low-rank community representation matrix $\mathbf{C}$ is computed using the multiplicative update rule defined by (6).

We can now formulate the multiplicative update rules for our objective function (4). Following the constrained optimization theory [4] and non-negative matrix factorization [18], we first substitute Eq. (3) and Eq. (9) into Eq. (4) as follows:

$$\min_{\mathbf{C}\geq 0}\Psi = \min_{\mathbf{C}\geq 0,\mathbf{C}\mathbf{C}^\top=\mathbf{I}} \sum_{i\in\{l,h\}} ||\mathbf{A}_i - \mathbf{C}\mathbf{C}^\top\mathbf{A}_i||_F^2 + \\ \gamma \sum_{i\in\{l,h\}}(k - tr(\mathbf{C}\mathbf{C}^\top\mathbf{C}_i\mathbf{C}_i^\top)), \quad (10)$$

by taking the derivative of (10), we get

$$\nabla_C\Psi = -\sum_{i\in\{l,h\}} 4\mathbf{A}_i\mathbf{A}_i^\top\mathbf{C} + \sum_{i\in\{l,h\}} 2\mathbf{A}_i\mathbf{A}_i^\top\mathbf{C}\mathbf{C}^\top\mathbf{C} + \\ 2\mathbf{C}\mathbf{C}^\top\mathbf{A}_i\mathbf{A}_i^\top\mathbf{C} - 2\gamma\sum_{i\in\{l,h\}}\mathbf{C}_i\mathbf{C}_i^\top\mathbf{C} \quad (11)$$

where the first two terms under summation can be decomposed into two non-negative terms, namely:

$$\nabla_C\Psi_i(\mathbf{C};\mathbf{A}_i) = [\nabla_C\Psi_i(\mathbf{C};\mathbf{A}_i)]^+ - [\nabla_C\Psi_i(\mathbf{C};\mathbf{A}_i)]^-, \quad (12)$$

where $[\nabla_C\Psi_i(\mathbf{C};\mathbf{A}_i)]^- = 4\mathbf{A}_i\mathbf{A}_i^\top\mathbf{C} \geq 0$, $[\nabla_C\Psi_i(\mathbf{C};\mathbf{A}_i)]^+ = 2\mathbf{A}_i\mathbf{A}_i^\top\mathbf{C}\mathbf{C}^\top\mathbf{C} + 2\mathbf{C}\mathbf{C}^\top\mathbf{A}_i\mathbf{A}_i^\top\mathbf{C} \geq 0$ are non-negative terms, respectively. To incorporate the orthonormality constraint into the update rule, we employ the notion of natural gradient [23]. Since, the columns of matrix $\mathbf{C}$ span a vector subspace known as Grassman manifold $\mathbb{G}(k,n)$, i.e., $span(\mathbf{C}) \in \mathbb{G}(k,n)$, therefore, the ordinary gradient of the optimization problem (11) does not represent its steepest direction, but it represents a natural gradient [1]. We define a natural gradient to optimize our objective function (4) under the orthonormality constraint.

The natural gradient of Ψ on Grassman manifold at $\mathbf{C}$ can be written in terms of the ordinary gradient as follows [23]:

$$\tilde{\nabla}_{\mathbf{C}}\Psi = \nabla_{\mathbf{C}}\Psi - \mathbf{C}\mathbf{C}^\top\nabla_{\mathbf{C}}\Psi, \quad (13)$$

where $\nabla_{\mathbf{C}}\Psi$ is the ordinary gradient given by (11). Following the KKT condition and preserving the non-negativity of $\mathbf{C}$, the multiplicative update rules for matrix $\mathbf{C}$ using the natural gradient is as follows:

$$\mathbf{C}(jk) \leftarrow \mathbf{C}(jk)\frac{[\tilde{\nabla}_{\mathbf{C}}\Psi]^-}{[\tilde{\nabla}_{\mathbf{C}}\Psi]^+}, \quad (14)$$

where the non-negative parts of the normal gradient are written as follows

$$[\tilde{\nabla}_{\mathbf{C}}\Psi]^- = \Big(\sum_{i\in\{l,h\}}\mathbf{A}_i\mathbf{A}_i^\top + \gamma\mathbf{C}_i\mathbf{C}_i^\top\Big)\mathbf{C}$$
$$[\tilde{\nabla}_{\mathbf{C}}\Psi]^+ = \mathbf{C}\mathbf{C}^\top\Big(\sum_{i\in\{l,h\}}\mathbf{A}_i\mathbf{A}_i^\top + \gamma\mathbf{C}_i\mathbf{C}_i^\top\Big)\mathbf{C} \quad (15)$$

the consensus low-rank representation matrix $\mathbf{C}$ computed from (4) is used for node assignment to the k-th community. The entries in the i-th row of matrix $\mathbf{C}$ after row normalization are interpreted as a posterior probability that a node i belongs to each of the k composite communities. In our experiments, we apply a hard clustering procedure that is where a node is assigned to the k-th cluster that has the largest probability value.

4. Experimental Evaluations

We evaluated the performance of the proposed tissue phenotyping DeMuC algorithm both qualitatively and quantitatively on two publicly available datasets proposed by Kather *et al.*, including Colon Cancer Histology Images (CCHI-1) [16] and CCHI-2 [15]. A total of 5,000 histology images in the CCHI-1 dataset are divided into eight different tissue classes including tumor epithelium, simple stroma, complex stroma, lymphocytes, debris, mucosal glands, adipose, and background with 625 images in each class. In CCHI-2 dataset, a total of 7,180 images are divided into nine difference tissue classes including tumor colorectal adenocarcinoma epithelium (1,233 images), cancer-associated stroma or complex stroma (421 images), normal colon mucosa (741 images), smooth muscle (592 images), mucus (1,035 images), lymphocytes (634 images), debris (339 images), adipose (1,338 images), and background (847 images) tissue. We did not consider adipose and background tissue classes in both datasets as these tissue images do not contain any cellular components. So, we tested a total of 3,750 tissue images of 6 classes in the CCHI-1 dataset and a total of 4,995 tissue images of 7 classes in the CCHI-2 dataset. In both datasets, the tissue images are manually annotated by experienced pathologists and non-overlapping patches of sizes 150×150 and 224×224 are extracted from CRC WSIs in CCHI-1 and CCHI-2. Sample images of six and seven tissue classes from both datasets are shown in Fig. 3.

We used two main parameters γ and k to optimize the proposed objective function (4). γ is set according to $1/\sqrt{n}$, while k denotes the number of communities which is set as $k = 6$ for the CCHI-1 dataset and $k = 7$ for the CCHI-2 dataset. We compared the performance of the proposed DeMuC algorithm with 9 state-of-the-art methods including KM-CD [27], SDLs [25], B6F-SVM [16], DFOD [32], SHIRC [29], TPCD [14], DenseNet [10], ResNet101 [9],
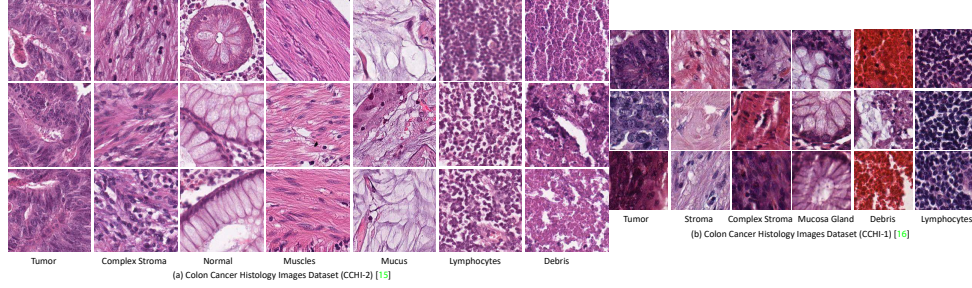
Figure 3: Samples images from both CCHI-1 [16] and CCHI-2 [15] dataset. (a) Shows the exemplar images of seven different tissue types including Tumor, Complex Stroma, Normal, Muscles, Mucus, Lymphocytes, and Debris from the CCHI-1 dataset [15]. (b) Shows the exemplar images of six distinct tissue types including Tumor, Stroma, Complex Stroma, Mucosa Gland, Debris, and Lymphocytes [16].

and SVM-CNN [34]. The methods, KM-CD and TPCD, exploited cellular interaction features and performed k-medoid and mean-shift clustering methods. The method, B6F-SVM, used a set of six different textural features and trained SVM classifier for multi-class tissue classification. SDLs, DFOD, and SHIRC, trained dictionary on textural features of the histology images. While, deep learning methods, DenseNet and ResNet101, are end-to-end trained for direct patch-based classification. CNN-SVM exploited the deep features from the pre-trained AlexNet model and trained the SVM classifier for tissue classification. The original author's implementation was used for the TPCD, KM-CD, SDLs, B6F-SVM, DFOD, and SHIRC, methods, while, we implemented deep learning methods, DenseNet, ResNet101, and SVM-CNN, which were pre-trained on the ImageNet database. We replaced the classification layer and fine-tuned these networks with stochastic gradient descent with a momentum of 0.8. We randomly divided both datasets into a 70% training set and a 30% testing set for multi-class histology images classification. We trained all networks on a desktop workstation with two Nvidia Titan Xp GPUs with a mini-batch size of 256 and a learning rate of 0.0003 for 130 epochs. In all these networks, the data augmentation techniques such as rotational invariance with random horizontal and vertical flips were used for training.

All experiments are carried out on a machine with an Intel Core i7 4.0 GHz CPU and 64 GB RAM on which our proposed objective function takes 9 iterations in 24.2 seconds for the CCHI-1 dataset and 15 iterations and 30.8 secs for the CCHI-2 dataset for the classification of histology images. The quantitative results are compared in terms of the True Positive (TP) rate and F_1 score as performance measures. The aim is to maximize TP and F_1 score for a more accurate classification of tissue components.

4.1. Evaluation on CCHI-1 Dataset

Table 1 shows the classification performance in terms of TP score in % and F_1 score as compared to existing state-of-the-art methods. On average, the proposed algorithm has performed favorably better than the compared methods in terms of both Average True Positive (AvTP) rate and av-

Table 1: Comparative performance of multi-class tissue classification on the Colon Cancer Histology Images-1 (CCHI-1) dataset [16]. The true positive rate in % and average F-score of the six tissue phenotypes are shown for the 10 methods. The true positive rate gives the % of test images classified correctly. The two best results are shown in red and blue fonts respectively.

Methods	Tumor	Stroma	Complex	Mucosa	Debris	Lympho	AvTP	F-score
KM-CD [27]	80.6	91.4	77.2	80.5	96.0	93.5	86.5	0.85
B6F-SVM[16]	86.3	86.4	72.4	92.4	88.1	89.3	85.8	0.89
DFOD [32]	79.3	89.3	81.5	78.8	92.4	92.6	86.6	0.81
SHIRC [29]	72.8	85.8	72.0	77.1	92.1	92.8	82.1	0.80
SDLs [25]	69.3	85.7	66.7	72.6	92.7	93.4	80.0	0.79
Deep Learning Methods	Tumor	Stroma	Complex	Mucosa	Debris	Lympho	AvTP	F-score
DenseNet [10]	85.8	90.7	87.5	86.8	93.1	97.4	90.2	0.89
SVM-CNN [34]	85.8	90.7	87.5	86.8	93.1	97.4	90.2	0.84
ResNet101 [9]	89.3	91.9	90.2	90.0	96.0	97.3	92.4	0.89
TPCD [14]	90.0	92.8	89.4	89.5	96.3	97.2	92.5	0.92
DeMuC	90.1	93.2	92.5	90.5	96.4	97.5	93.4	0.93

erage F_1 score. For instance, DeMuC has achieved 93.4% AvTP, which is 6.8% greater than the classical handcrafted features method such as DFOD in terms of AvTP. In terms of the F_1 score, DeMuC obtained 0.93, which is 4% better than the classical method, B6F-SVM, and 1.0% greater than TPCD. The excellent performance of the proposed algorithm is because of the encoding of multi-layer graph information into the composite community objective function.

For the tumor tissue component, our proposed algorithm DeMuC produced the best results of 90.1%. While the remaining methods could not achieve a true positive rate of more than 86% excluding ResNet101 (89.3%) and TPCD (90.0%) methods. It shows that the tumor tissue community is one of the difficult components for almost all of the compared methods. In the stroma tissue community, only the proposed algorithm, DeMuC, achieved the best true positive rate of 93.2%, which is significantly better than others. All of the remaining methods attained a true positive rate of more than 85.0%, which demonstrates that the stroma tissue images did not pose a great challenge for the compared methods.

The complex stroma is more complicated tissue component than other components for all of the compared methods. Only two methods ResNet101 and our proposed algorithm, DeMuC, achieved the true positive score of more than 90.0%. Our proposed algorithm obtained the best true positive score of 92.5% which is 2.3% better than

Table 2: Comparative performance of multi-class tissue classification on the Colon Cancer Histology Images-2 (CCHI-2) dataset [15]. The true positive rate in % and average F-score of the six tissue phenotypes are shown for the 10 methods. The true positive rate gives the % of test images classified correctly. The two best results are shown in red and blue fonts respectively.

Methods	Tumor	Muscle	Complex	Mucus	Debris	Lympho	Norm	AvTP	F-score
KM-CD [27]	79.1	80.1	76.1	79.4	95.5	92.4	78.1	82.9	0.84
B6F-SVM[16]	85.4	80.2	71.4	81.6	84.2	85.6	82.3	81.5	0.87
DFOD [32]	78.1	77.3	79.1	80.1	90.1	89.1	79.3	81.8	0.79
SHIRC [29]	70.1	69.3	71.5	76.6	89.1	87.3	71.5	76.4	0.81
SDLs [25]	67.2	65.6	64.1	70.2	90.5	90.1	67.2	73.5	0.80
Deep Learning Methods	Tumor	Stroma	Complex	Mucosa	Debris	Lympho	Norm	AvTP	F-score
DenseNet [10]	84.2	91.5	88.1	94.5	80.2	92.2	93.6	89.1	0.92
SVM-CNN [34]	84.3	89.4	84.5	90.1	79.8	90.2	85.4	86.2	0.82
ResNet101 [9]	90.1	87.4	89.1	88.1	84.3	84.1	95.4	88.3	0.90
TPCD [14]	92.0	90.5	79.8	94.1	87.6	94.0	90.6	89.8	0.89
DeMuC	95.1	91.5	90.1	95.8	90.0	90.7	92.5	92.2	0.93

ResNet101 deep features method. Most compared methods attained a true positive score of less than 80.0%. In terms of Mucosa gland tissue images, our proposed algorithm DeMuC produced the best results 90.5%, in terms of the true positive rate. The remaining methods attained good performance as compared to other tissue components for the mucosa gland class. Debris and Lymphocytes did not pose a great challenge for all of the compared methods. Our proposed algorithm, DeMuC, achieved the best results.

4.2. Evaluation on CCHI-2 Dataset

Table 2 shows the multi-class tissue classification performance in terms of TP score in % and F_1 score on the CCHI-2 dataset as compared to existing state-of-the-art methods. On the average, the proposed algorithm has achieved significantly better results in terms of both AvTP and F_1 score than the compared methods. For instance, DeMuC has achieved 92.2% AvTP and 0.93 F_1 score, which is 9.0% and 6.0% greater than the classical handcrafted features methods such as KM-CD and B6F-SVM. In comparison with deep learning methods, the proposed algorithm obtained 2.4% and 1.0% better results in terms of AvTP and F_1 score. This boost in the performance confirms the advantages of enforcing the composite community structure in the objective function in our proposed algorithm.

In terms of tumor tissue community, our proposed algorithm, DeMuC, produced the best results of 95.1%, while ResNet101 and TPCD methods achieved a significantly better true positive rate of 90.1% and 92.0%, respectively. The remaining texture analysis-based and dictionary learning-based methods, could not obtain a true positive rate of more than 80.0% excluding B6F-SVM (true positive rate of 85.4%). It shows that the tumor tissue community posed a great challenge for the classical handcrafted features-based methods. Since these methods could not fully capture the heterogeneous nature of the tumor component.

In the muscle tissue community, only the proposed algorithm, DeMuC, achieved the best true positive rate of 91.5%, which is comparable with DenseNet method. All of the remaining deep learning methods attained a true positive

rate of more than 80.0%. Only two handcrafted features-based methods, KM-CD and B6F-SVM, produced a true positive rate of more than 80.0% while, muscle community was the major burden for the remaining methods. The complex stroma was the most difficult tissue community for the majority of the compared methods since none of the methods could achieve a true positive rate of more than 90.0%. Only the proposed algorithm, DeMuC, obtained a 90.1% true positive rate which was 1.0% greater than the ResNet101 method.

For the mucus tissue component, our proposed algorithm, DeMuC, performed favorably better with 95.8% true positive scores. The remaining compared methods show some discrepancy in a true positive score for the mucus tissue images. Majority of the compared methods obtained good results for Debris and Lympho tissue components. The proposed algorithm, DeMuC, achieved 90.0% and 90.7%, true positive rate which is favorably better compared to B6F-SVM, DFOD, SHIRC, ResNet101 and SVM-CNN. In terms of normal colon mucosa (Norm) tissue type, DeMuC achieved 92.5% while, ResNet101 method obtained a 95.4% true positive rate. Many methods were not able to handle Norm tissue images accurately as most of these methods could not attain true positive rate more than 80%.

5. Conclusions

In this work, a novel semi-supervised cellular community detection algorithm has been proposed for tissue phenotyping based on a deep neural network for cell detection, classification, and clustering of image patches into biologically meaningful groups or communities. First, deep neural networks are used for cell detection and classification and then two cell-level graphs are constructed. Based on cellular interaction features at the patch level, two patch-level graphs are constructed using a chi-squared distance measure. A novel objective function is proposed which enforces non-negativity matrices constraints for estimating composite communities for tissue phenotyping. The proposed algorithm has shown better performance than end-to-end deep learning methods as well as several existing algorithms based on handcrafted features. In future, we aim to exploit more graph layers into the objective function and investigate their biological significance and clinical relevance.

Acknowledgment

References

[1] S.-I. Amari. Natural gradient works efficiently in learning. *Neu. Comp.*, 10(2):251–276, 1998. 2, 6

[2] B. E. Bejnordi, M. Mullooly, R. M. Pfeiffer, S. Fan, P. M. Vacek, D. L. Weaver, S. Herschorn, L. A. Brinton, B. van Ginneken, N. Karssemeijer, et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Modern Pathology*, page 1, 2018. 2, 3

[3] F. Bianconi, A. Álvarez-Larrán, and A. Fernández. Discrimination between tumour epithelium and stroma via perception-based features. *Neurocomputing*, 154:119–126, 2015. 2, 3

[4] S. Boyd and L. Vandenberghe. *Convex optimization*. CUP, 2004. 6

[5] M. M. Bui, S. L. Asa, L. Pantanowitz, A. Parwani, J. van der Laak, C. Ung, U. Balis, M. Isaacs, E. Glassy, and L. Manning. Digital and computational pathology: Bring the future into focus. *JPI*, 10, 2019. 2

[6] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov. Clustering on multi-layer graphs via subspace analysis on grassmann manifolds. *IEEE T-SP*, 62(4):905–918, 2013. 2, 5

[7] Y. Du, R. Zhang, A. Zargari, T. C. Thai, C. C. Gunderson, K. M. Moxley, H. Liu, B. Zheng, and Y. Qiu. Classification of tumor epithelium and stroma by exploiting image features learned by deep convolutional neural networks. *An. of BE*, pages 1–12, 2018. 2, 3

[8] V. Gligorijević, Y. Panagakis, and S. Zafeiriou. Non-negative matrix factorizations for multiplex network analysis. *IEEE T-PAMI*, 41(4):928–940, 2018. 2, 5

[9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *IEEE CVPR*, 2016. 6, 7, 8

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *IEEE CVPR*, 2017. 6, 7, 8

[11] Y. Huang, H. Zheng, C. Liu, X. Ding, and G. K. Rohde. Epithelium-stroma classification via convolutional neural networks and unsupervised domain adaptation in histopathological images. *IEEE J-BHI*, 21(6):1625–1632, Nov 2017. 2, 3

[12] A. Huijbers, R. Tollenaar, G. v Pelt, E. Zeestraten, S. Dutton, C. McConkey, E. Domingo, V. Smit, R. Midgley, B. Warren, et al. The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Annals of Oncology*, 24(1):179–185, 2012. 2

[13] A. Janowczyk and A. Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *JPI*, 7, 2016. 2

[14] S. Javed, M. M. Fraz, D. Epstein, D. Snead, and N. M. Rajpoot. Cellular community detection for tissue phenotyping in histology images. In *Computational Pathology and Ophthalmic Medical Image Analysis*, pages 120–129. Springer, 2018. 2, 6, 7, 8

[15] J. N. Kather, J. Krisam, P. Charoentong, T. Luedde, E. Herpel, C.-A. Weis, T. Gaiser, A. Marx, N. A. Valous, D. Ferber, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine*, 16(1):e1002730, 2019. 2, 3, 6, 7, 8

[16] J. N. Kather, C.-A. Weis, F. Bianconi, S. M. Melchers, L. R. Schad, T. Gaiser, A. Marx, and F. G. Zöllner. Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, 6:27988, 2016. 2, 3, 6, 7, 8

[17] S. Kothari, J. H. Phan, A. N. Young, and M. D. Wang. Histological image classification using biologically interpretable shape-based features. *BMC MI*, 13(1):9, 2013. 2

[18] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Advan. NIPS*, 2001. 5, 6

[19] N. Linder, J. Konsti, R. Turkki, E. Rahtu, M. Lundin, S. Nordling, C. Haglund, T. Ahonen, M. Pietikäinen, and J. Lundin. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. *Diagnostic pathology*, 7(1):22, 2012. 2, 3

[20] D. N. Louis, M. Feldman, A. B. Carter, A. S. Dighe, J. D. Pfeifer, L. Bry, J. S. Almeida, J. Saltz, J. Braun, J. E. Tomaszewski, et al. Computational pathology: a path ahead. *Archives of pathology & laboratory medicine*, 140(1):41–50, 2015. 2

[21] A. Madabhushi and G. Lee. Image analysis and machine learning in digital pathology: Challenges and opportunities. *MIA*, 33:170–175, 2016. 2

[22] M. Nalisnik, M. Amgad, S. Lee, S. H. Halani, J. E. V. Vega, D. J. Brat, D. A. Gutman, and L. A. Cooper. Interactive phenotyping of large-scale histology imaging data with histomicsml. *Scien. Rep.*, 7(1):14588, 2017. 2

[23] Y. Panagakis, C. Kotropoulos, and G. R. Arce. Non-negative multilinear principal component analysis of auditory temporal modulations for music genre classification. *IEEE T-ASLP*, 18(3):576–588, 2009. 2, 6

[24] Sari and C. Gunduz-Demir. Unsupervised feature extraction via deep learning for histopathological classification of colon tissue images. *IEEE T-MI*, page 1, 2018. 2, 3

[25] R. Sarkar and S. T. Acton. Sdl: Saliency-based dictionary learning framework for image similarity. *IEEE T-IP*, 27(2):749–763, 2018. 2, 3, 6, 7, 8

[26] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. Snead, I. A. Cree, and N. M. Rajpoot. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE T-MI*, 35(5):1196–1206, 2016. 1, 2, 4

[27] K. Sirinukunwattana, D. Snead, D. Epstein, Z. Aftab, I. Mujeeb, Y. W. Tsang, I. Cree, and N. Rajpoot. Novel digital signatures of tissue phenotypes for predicting distant metastasis in colorectal cancer. *Scien. Rep.*, 8(1):13692, 2018. 2, 6, 7, 8

[28] D. R. Snead, Y.-W. Tsang, A. Meskiri, P. K. Kimani, R. Crossman, N. M. Rajpoot, E. Blessing, K. Chen, K. Gopalakrishnan, P. Matthews, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology*, 68(7):1063–1072, 2016. 2

[29] U. Srinivas, H. S. Mousavi, V. Monga, A. Hattel, and B. Jayarao. Simultaneous sparsity model for histopathological image representation and classification. *IEEE T-MI*, 33(5):1163–1179, 2014. 6, 7, 8

[30] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE T-SMC*, 8(6):460–473, 1978. 2, 3

[31] J. van der Laak, N. Rajpoot, and D. Vossen. The promise of computational pathology. *The Pathologist*, (38):16–26, Jan 2018. 2

[32] T. H. Vu, H. S. Mousavi, V. Monga, G. Rao, and U. A. Rao. Histopathological image classification using discriminative feature-oriented dictionary learning. *IEEE T-MI*, 35(3):738–751, 2016. 2, 3, 6, 7, 8

[33] J. Xu, X. Luo, G. Wang, H. Gilmore, and A. Madabhushi. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing*, 191:214–223, 2016. 2, 3

[34] Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang, M. Lai, I. Eric, and C. Chang. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC BI*, 18(1):281, 2017. 2, 3, 7, 8

[35] Z. Yang and E. Oja. Linear and nonlinear projective nonnegative matrix factorization. *IEEE T-NN*, 21(5):734–749, 2010. 5