Medical Image Analysis 72 (2021) 102104

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Challenge Report

Spatially Constrained Context-Aware Hierarchical Deep Correlation Filters for Nucleus Detection in Histology Images

Sajid Javed [a,f], Arif Mahmood [b], Jorge Dias [a,f], Naoufel Werghi [a,f,*], Nasir Rajpoot [c,d,e]

[a] Khalifa University Center for Autonomous Robotic Systems (KUCARS), Khalifa University, Abu Dhabi, UAE.
[b] Department of Computer Science, Information Technology University, Lahore, Pakistan.
[c] Department of Computer Science, University of Warwick, Coventry, CV4 7AL, U.K.
[d] Department of Pathology, University Hospitals Coventry and Warwickshire, Walsgrave, Coventry, CV2 2DX, U.K.
[e] The Alan Turing Institute, London, NW1 2DB, U.K.
[f] Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, UAE.

ARTICLE INFO

Article history:
Received 4 June 2020
Revised 10 May 2021
Accepted 12 May 2021
Available online 24 May 2021

MSC:
41A05
41A10
65D05
65D17

Keywords:
Nucleus detection
Deep learning
Correlation filters
Computational pathology

ABSTRACT

Nucleus detection in histology images is a fundamental step for cellular-level analysis in computational pathology. In clinical practice, quantitative nuclear morphology can be used for diagnostic decision making, prognostic stratification, and treatment outcome prediction. Nucleus detection is a challenging task because of large variations in the shape of different types of nucleus such as nuclear clutter, heterogeneous chromatin distribution, and irregular and fuzzy boundaries. To address these challenges, we aim to accurately detect nuclei using spatially constrained context-aware correlation filters using hierarchical deep features extracted from multiple layers of a pre-trained network. During training, we extract contextual patches around each nucleus which are used as negative examples while the actual nucleus patch is used as a positive example. In order to spatially constrain the correlation filters, we propose to construct a spatial structural graph across different nucleus components encoding pairwise similarities. The correlation filters are constrained to act as eigenvectors of the Laplacian of the spatial graphs enforcing these to capture the nucleus structure. A novel objective function is proposed by embedding graph-based structural information as well as the contextual information within the discriminative correlation filter framework. The learned filters are constrained to be orthogonal to both the contextual patches and the spatial graph-Laplacian basis to improve the localization and discriminative performance. The proposed objective function trains a hierarchy of correlation filters on different deep feature layers to capture the heterogeneity in nuclear shape and texture. The proposed algorithm is evaluated on three publicly available datasets and compared with 15 current state-of-the-art methods demonstrating competitive performance in terms of accuracy, speed, and generalization.

1. Introduction

Nucleus detection in cancer Whole Slide Images (WSIs) stained with Hematoxylin and Eosin (H&E) dyes is a fundamental step in computational pathology (Javed et al. (2020c); Hu et al. (2009); Lucchinetti et al. (2011); Bui et al. (2019); Dunne and Going (2001); Javed et al. (2019, 2018a)). In clinical practice, manual analysis of individual nuclei is a laborious task, and it also suffers from inter-observer and intra-observer variability (Andrion et al. (1995); López et al. (2012)). Pathologists often use nuclear features to assess the degree of malignancy in the tumor microenvironment (Gurcan et al. (2009)). Automatic detection of nuclei can serve as a useful tool for downstream analysis to make better decisions in cancer diagnosis, prognostication, and therapy planning (Javed et al. (2020d); Demir and Yener (0000); Xing and Yang (2016); Veta et al. (2014); Gurcan et al. (2009)). However, nucleus detection is a challenging task because of the nuclear clutter and diverse nuclear morphology such as varying chromatic texture, shape, and size (Gurcan et al. (2009); Xing and Yang (2016)). Moreover, computational challenges also arise because WSIs are multi-gigapixel images or three-dimensional arrays stored in a multi-resolution format, and may contain billions of pixels at the highest resolution. For example, the WSI in Fig. 1(a) at the highest magnification level (40×) has the dimensions of 90K×44K pixels and

* Corresponding author at: Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, UAE.
E-mail address: Naoufel.Werghi@ku.ac.ae (N. Werghi).

https://doi.org/10.1016/j.media.2021.102104
1361-8415/

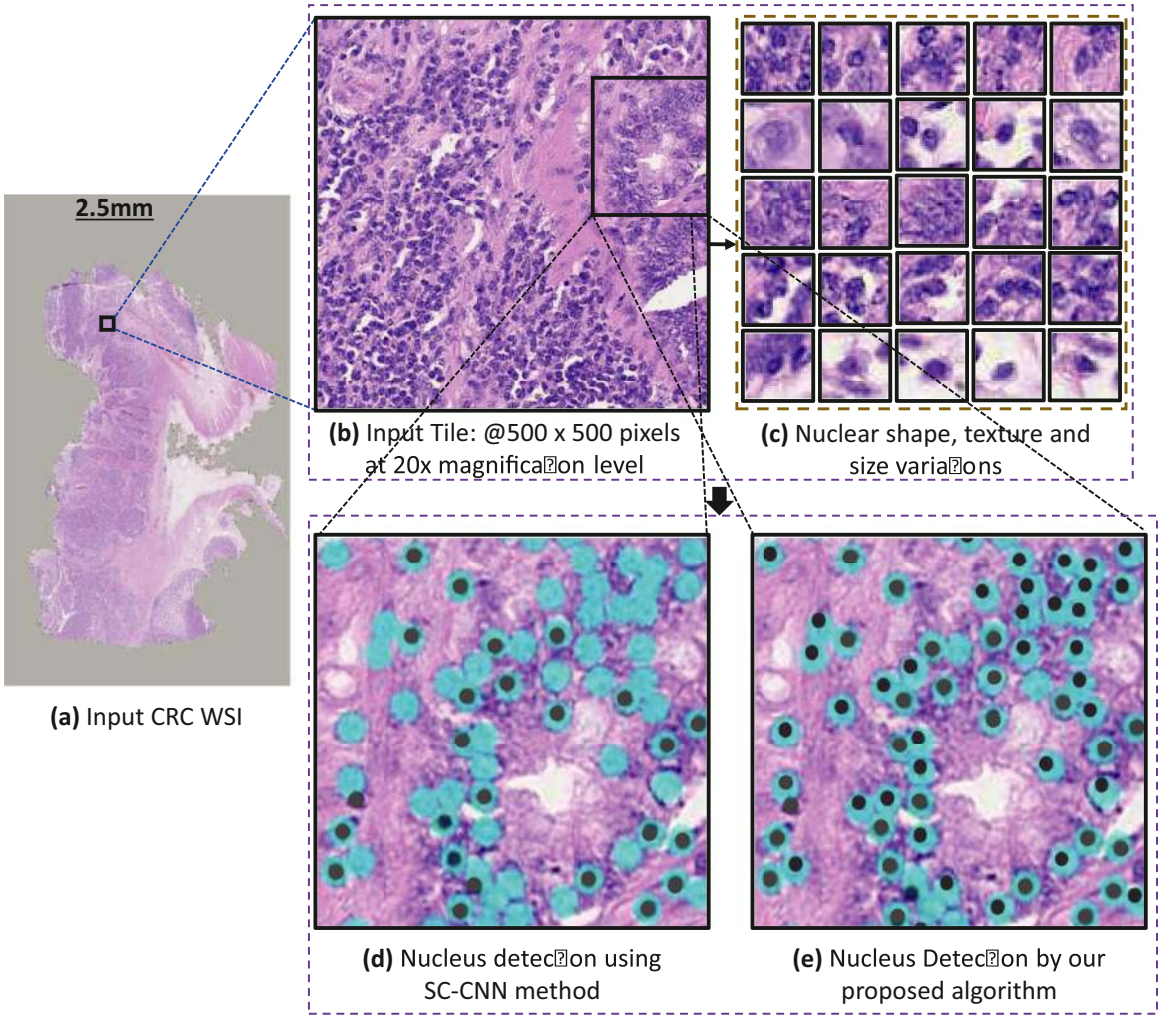


Fig. 1. An example of multi-gigapixel whole slide image of colorectal cancer and the results of the proposed algorithm for nucleus detection compared with current state-of-the-art SC-CNN method (Sirinukunwattana et al. (2016)) under varying nuclear shape, morphology, texture, and clutteredness. For better visualization of cluttered nuclei and complex tumor epithelial, regions of interest are enlarged in the black boxes with nuclei ground truth locations and detection results. The detected nuclei are marked by black dots, while the ground truth nuclei are represented as thicker cyan dots for visualization purposes.

contains tens of thousands of cell nuclei. Fig. 1 also shows the results of nucleus detection by our proposed algorithm and the state-of-the-art SC-CNN method (Sirinukunwattana et al. (2016)) under varying nuclear shape, size, texture, and clutter.

In the past few years, a number of potential methods have been reported for automatic detection of the nucleus in the literature such as (Höfener et al. (2018); Shi et al. (2018); Xing and Yang (2016); Gurcan et al. (2009); Veta et al. (2014); Demir and Yener (0000); Xue and Ray (2017); Su et al. (2016); Ahmad et al. (2018); Xing et al. (2019); Koohababni et al. (2018); Graham et al. (2019); Sirinukunwattana et al. (2016); Xie et al. (2018)). A few challenging datasets for different cancer tissue types have also been made available for nucleus detection and classification (Sirinukunwattana et al. (2016); Graham et al. (2019); Su et al. (2016)). Interested readers are encouraged to explore more details for nucleus detection in recent surveys (Su et al. (2016); Xing and Yang (2016); Gurcan et al. (2009); Veta et al. (2014); Demir and Yener (0000); Irshad et al. (2013)).

The existing nucleus detection methods (Xing and Yang (2016)) can be categorized into classical machine learning methods (Adiga et al. (2006); Yan et al. (2008); Su et al. (2014); Sommer et al. (2012)) and recent deep learning methods (Sirinukunwattana et al. (2016); Graham et al. (2019); Su et al. (2016)). In classical machine learning methods, hand-crafted features such as intensity and histogram are used with support vector machines and random decision forest classifiers to estimate the nuclear probability map. Adiga *et al.* used distance transform function for nucleus detection in breast cancer histology images (Adiga et al. (2006)). Yan *et al.* also used Euclidean distance transform to identify nucleus for watershed segmentation in fluorescence images (Yan et al. (2008)). Li *et al.* used intensity features in the distance map and then Gaussian filter for noise suppression (Li et al. (2009)). Despite the improvements, distance transform-based methods do not work well for densely clustered nuclei. Supervised learning methods have also been proposed to address the complex rich heterogeneous nature of histopathology images (Su et al. (2014); Sommer et al. (2012)). These methods usually train binary classifiers to reduce false detection. Some classical methods also assume individual nucleus as round or circular objects. This assumption does not hold in cases involving elongated and irregularly shaped nuclei (Xing and Yang (2016)).

The deep learning methods (Xie et al. (2018); Sirinukunwattana et al. (2016); Graham et al. (2019); Koohababni et al. (2018); Xing et al. (2019)) train an end-to-end convolutional neural network for generating nuclear probability

maps. These networks are invariant to nuclear size, texture, and morphology. Xing *et al.* proposed three different deep convolutional neural networks (CNNs) models for brain tumor, pancreatic, and breast cancer histology images for the purpose of nucleus detection (Xing et al. (2015)). Xie *et al.* proposed structured regression-based CNN which generates proximity patch with fast scanning strategy for nucleus detection (Xie et al. (2015)). Sirinukuwattana *et al.* proposed locality sensitive deep learning approach for nucleus detection (Sirinukunwattana et al. (2016)). In their approach, a spatially constrained CNN is employed to generate a probability map for a given input image using a sliding window strategy. Recently, Graham *et al.* proposed a deep CNN model for simultaneous segmentation and classification of the nucleus (Graham et al. (2019)). While deep learning-based methods perform well, they often struggle with the detection of cluttered nuclei which are quite common in the tumor microenvironment. Besides, when trained CNN models are applied to test data in a sliding window manner, processing large-scale images become computationally expensive. In general, deep learning methods require a large amount of training data which may not always be available. Moreover, deep learning approaches are resource-hungry requiring expensive computational systems (Xing and Yang (2016)).

In this work, we propose a nucleus detection algorithm using correlation filters which have been previously employed for visual object tracking application (Bolme et al. (2010); Henriques et al. (2014); Fiaz et al. (2019)). Compared to the end-to-end deep learning, the correlation filters are computationally effective and require a significantly less amount of training data (Henriques et al. (2014)). The correlation filters are also flexible and can detect complex and irregular-shaped nuclei without requiring handcrafted features.

In the current work, we formulate the nucleus detection problem as learning a set of robust and nucleus-specific correlation filters which have not been previously thoroughly investigated for such tasks. In the proposed algorithm, a set of training nuclei patches are cropped from histology images and 2-D Gaussian-shape response maps are considered as the ground truth nuclei regression targets. We aim to learn filters such that the maximum response is obtained when convolved with a nucleus region by solving a ridge regression problem. Nuclei in the test patches are then detected by a circular convolution operation which is implemented using Fast Fourier Transform (FFT) (Bolme et al. (2010); Henriques et al. (2014)). The performance of correlation filters may degrade because of the wide variations in the nuclear shape, texture, and morphology. To address these challenges, we propose to train correlation filters on multiple layers of deep features using a pre-trained VGG-19 model on a large-scale tissue images dataset (Kather et al. (2019)). We also propose the correlation filters to be learned in a structure-aware manner by constructing a graph across different nucleus components. The correlation filters are constrained to be eigenvectors of the Laplacian of the nuclei structure graph. Also, in order to improve the localization, we propose the correlation filter to be context-aware because the contextual information improves the discriminative ability of the filter.

Benefiting from the high generalizability of CNNs as feature extractors (Sharif Razavian et al. (2014)), correlation filters trained on features extracted from multiple layers of CNN are effective in capturing nuclear semantics as well as localization. Deep layers of CNNs are more effective to capture semantics; however, they are not ideal for capturing fine-grained spatial details required for nucleus localization. On the contrary, shallow layers are precise in localization but do not capture nuclear semantic information. Therefore, training correlation filters on both shallow and deep CNN layers is important to capture both nuclear semantics and localization.

In addition to directly utilizing deep features, we also propose to construct nuclear structure graphs using deep features at multiple levels. These graphs encode the pairwise similarity among different nuclear components thus enriching multiple levels of information to handle nuclear shape and texture variations. Using these graphs, we compute Laplacian matrices which have previously been used for graph partitioning, subspace-based data clustering, visual tracking, and moving object segmentation tasks (Yin et al. (2015); Javed et al. (2017, 2016, 2018b, 2020b, 2020a); Giraldo et al. (2020)). We constrain the correlation filters to be orthogonal to the Laplacian basis to better discriminate different nuclear components from the non-nuclear regions and also to discern each nucleus from the remaining nuclei in case of clutter. Such a synergy nurtures the proposed correlation filters with robustness to nuclear size, texture, and morphological variations while enhancing their localization capacity. Another limitation of correlation filters is the undesired boundary effects caused by the circular shifts of the training patches which may lead to degraded performance (Ahmad et al. (2018)). To address this problem, we propose to exploit local contextual information in the learning phase. The motivation here is that the surroundings of the nuclei can enhance the detection performance, especially if there is a non-nuclear region. The correlation filters are also constrained to be orthogonal to the contextual basis to better differentiate nuclear and non-nuclear regions.

Our proposed algorithm, dubbed as Spatially Constrained Context-aware hierarchical Deep Correlation Filters (SCC-DCF) for nucleus detection, is tested on three publicly available datasets including two colon cancer histology datasets CRCHistoPhenotypes (Sirinukunwattana et al. (2016)) and CoNSeP (Graham et al. (2019)) and a third multi-organ histology images PanNuke dataset (Gamper et al. (2020)). Our proposed algorithm is compared with 15 existing state-of-the-art methods and demonstrated superior performance. Following are the main contributions of this work:

- The correlation filters have not been thoroughly investigated for nucleus detection in computational pathology; we fill this research gap and demonstrate the strength of the correlation filters in terms of both accuracy and computational complexity. We also demonstrate the favorable performance of the proposed algorithm by using less training data.
- We propose to use deep features hierarchy and train a set of correlation filters which improves the nucleus detection performance as compared to classical color or stain features used in previous studies.
- In order to ensure the correlation filter to be structure-aware, we propose a spatial graph encoding pairwise similarities of nuclei different components. The filter is constrained to be orthogonal to the Laplacian basis of this spatial graph to address the unwanted boundary effects.
- We also incorporate local contextual information in the correlation filters which are encoded by selecting neighboring patches around each nucleus to improve nuclear localization.
- We proposed a novel objective function by incorporating structural constraints and the contextual information of each nucleus into the discriminative correlation filters framework. The objective function is solved using Alternating Direction Methods of Multiplier (ADMM) with a closed-form solution to each sub-problem.

The rest of the paper is organized as follows: Section 2 presents related work, Section 3 describes the main steps of our proposed algorithm, Section 4 presents experimental evaluations while the conclusion and future directions are provided in Section 5.

2. Literature review

Over the past few years, several methods have been proposed for nucleus detection in routine H&E histology images (Xing and

Yang (2016); Irshad et al. (2013)). In classical machine learning methods, Difference of Gaussian (DoG) (Cosatto et al. (2008)), Laplacian of Gaussian (LoG) (Byun et al. (2006)), Marker-based watershed approaches (Grau et al. (2004); Yang et al. (2006)), H-minima (Cheng et al. (2008)), Morphological operation-based methods (Park et al. (2012); Yang et al. (2006)), Maximally Stable Extremal Region (MSER) (Matas et al. (2004)), Radial Symmetry-based Voting Methods (Xing et al. (2013); Qi et al. (2011)), and supervised learning based on support vector machines and Random Decision Forest (RDF) classifiers (Mao et al. (2006)) have attracted significant attention.

Cosatto *et al.* proposed difference of Gaussian (DoG) followed by Hough transform to find radially symmetrical shapes (Cosatto et al. (2008)). However, this method demonstrated degraded performance in the case of clustered nuclei. To address this issue, Byun *et al.* employed Laplacian of Gaussian (LoG) filter in which nuclear size is known a priori on retinal images (Byun et al. (2006)). The scale-normalized LoG filter was used to detect cells on phase-contrast microscopy images (Lindeberg (1998)). However, these methods also failed in clutterred nuclei which exhibit weak boundaries. To tackle this issue, Al-Kofahi *et al.* have introduced a multi-scale LoG filter constrained by a Euclidean distance map (Al-Kofahi et al. (2009)). LoG-based methods are usually sensitive to large variations in nuclear size and the absence of clear nucleus boundaries (Xing and Yang (2016)). In order to handle nuclei clutteredness challenges, marker-based watershed approaches are widely utilized to locate and split close-by nucleus (Grau et al. (2004); Yang et al. (2006)). For instance, a variant of marker-controlled watershed segmentation using H-minima transform of nuclear shape is investigated (Jung and Kim (2010)). Although, H-minima based methods aim to reduce the false detection, they provide only limited robustness to a nucleus with heterogeneous texture.

Binary morphological filtering-based methods with a certain structuring element are also employed for nucleus detection (Park et al. (2012); Yang et al. (2006)). Park *et al.* have proposed an improved erosion operation that exploits a noise-robust measurement of convexity as the stopping criterion (Park et al. (2012)). Yuan *et al.* proposed to segment all nuclei with the help of thresholding, followed by morphological operation, distance transform, and watershed segmentation (Yang et al. (2006)). The centroids of the individual segmented nuclei were used as the detected points. However, these methods may not work well on a nucleus with heterogeneous intensity and cluttered background. To handle these issues, Maximally Stable Extremal Region (MSER) detector is also used to locate the nucleus (Matas et al. (2004)). By assuming that cells are approximately circular or elliptical, Parvin *et al.* introduced a kernel-based radial voting method to iteratively localize nucleus, which is relatively insensitive to image noise (Parvin et al. (2007)). Several other radial voting-based methods were also presented for automatic nucleus detection in pathology images (Xing et al. (2013); Qi et al. (2011)).

Supervised methods have also been proposed in the literature to solve the binary classification problem. Mao *et al.* have proposed a supervised learning approach for nucleus detection and segmentation in bladder cancer images (Mao et al. (2006)). A supervised learning method has also been proposed for mitotic cell detection in breast cancer images (Sommer et al. (2012)). A key limitation of these approaches is the lack of feature relevance for nucleus detection and the reduced capacity to accommodate a broad spectrum of nuclear shape, color, and texture variations remains limited.

Deep learning methods have recently become very popular to perform nucleus detection in histology images, because of their ability to perform feature discovery and generalization model (Xie et al. (2018); Sirinukunwattana et al. (2016); Graham et al. (2019); Koohababni et al. (2018); Xing et al. (2019)). Cruz *et al.* showed that a deep learning architecture outperforms methods based on handcrafted features (Cruz-Roa et al. (2013)). Ciregan *et al.* used Deep CNN for automatic detection of mitotic cells in breast cancer histology images (Cireşan et al. (2013)). Xie *et al.* proposed a structural regression model for CNN, where a nucleus center is detected if it has the maximum value in the proximity map (Xie et al. (2015)). Xu *et al.* proposed a stacked sparse autoencoder, where it learns high-level features of a nucleus centroid and then a softmax classifier is used to separate nucleus and non-nucleus image patches (Xu et al. (2014)). Sirinukunwattana *et al.* proposed SC-CNN method uses a regression approach to find the likelihood of a pixel being the center of a nucleus (Sirinukunwattana et al. (2016)). The probability values are topologically constrained in a way that in the vicinity of nucleus center the probability is higher. In (Xing et al. (2016)), a combination of CNNs for nucleus segmentation and a dictionary learning technique is proposed for refining segmentation. Hou *et al.* proposed a sparse convolutional autoencoder for simultaneous nucleus detection and feature extraction (Hou et al. (2019)). In their method, the detection is based on a fairly deep network of 15 layers comprising of multiple CNN branches that perform detection, segmentation, and image reconstruction. Koohababni *et al.* proposed mixture density networks in CRC WSIs (Koohababni et al. (2018)). Xie *et al.* proposed a multi-task learning method based on deep CNN encoder-decoder module (Xie et al. (2018)). Nucleus segmentation approaches are also proposed for nucleus centroid detection, but these methods require precisely annotated nucleus boundary (Graham et al. (2019)). Similarly, other segmentation methods such as NODE (Pinckaers and Litjens (2019)) can also be adapted to segment the nucleus instead of glandular structure segmentation. More recently, Tofighi *et al.* proposed shape prior deep networks for nucleus detection (Tofighi et al. (2019)).

Existing deep learning approaches for nucleus detection are promising; however, they require a significant amount of training data and expensive computational platforms to cater to the high computational requirements. In contrast, our goal here is to propose an algorithm that can be trained using much smaller training datasets that can be executed on typical desktop machines. The proposed SCC-DCF algorithm is a hybrid approach combining both classical correlation filters with deep features as well as graph-based structural constraints and nucleus contextual information to get the improved nucleus detection performance. To the best of our knowledge, such an algorithm has not been previously proposed for nucleus detection problem in histology images.

3. The proposed method

A block diagram of the proposed Spatially Constrained Context-aware Deep Correlation Filter (SCC-DCF) algorithm for nucleus detection is shown in Fig. 2. The main components of the proposed algorithm consist of pre-processing steps, deep features extraction, spatial graph construction, and training a set of robust correlation filters across different convolutional layers. In the first step, we crop a set of $m \times m$ nuclei patches from the training histology images where each nucleus is kept at the center of the patch. We also extract k contextual patches of the same size around each nucleus containing 50% overlap with a nucleus patch and 50% contextual information. These patches are selected in the same way for all nuclei covering k neighboring regions to reduce unwanted boundary effects during the correlation filter learning. In the second step, we input these training nuclei patches into the VGG-19 network for deep features extraction. We use a pre-trained model which is trained on a 100K tissue images dataset (Kather et al. (2019)) and we extract deep features from each of the five convolutional lay-

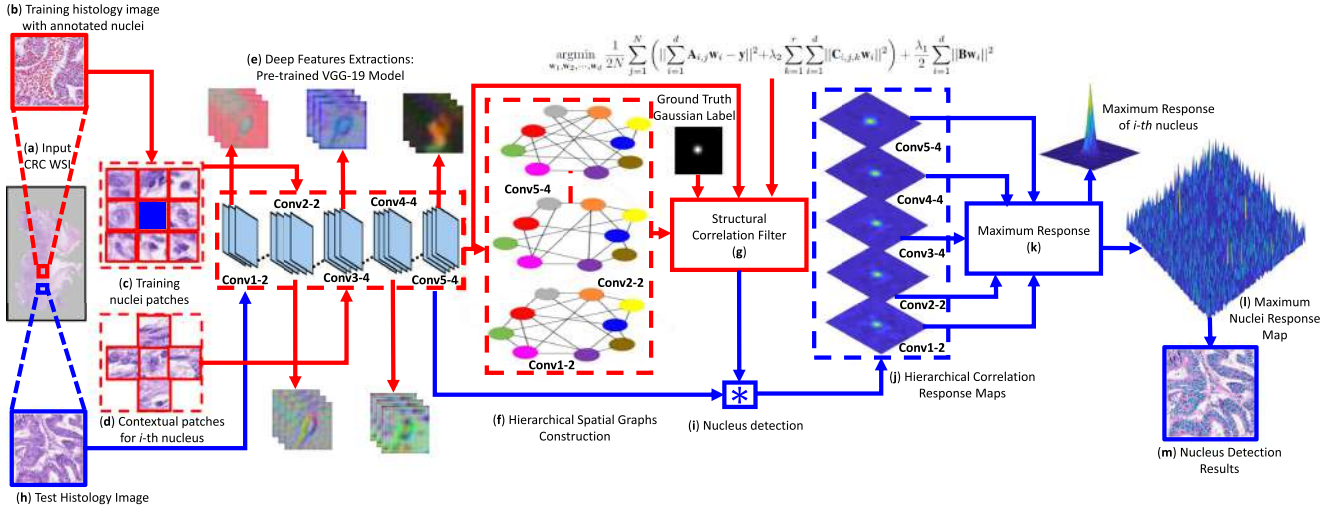


Fig. 2. Schematic illustration of our proposed SCC-DCF algorithm for nucleus detection in histology images. Steps (a)-(g) in red color show the training pipeline. Step (a) shows input whole slide image of colorectal cancer from CRCHistoPhenotypes dataset (Sirinukunwattana et al. (2016)). Step (b) shows training histology image with ground truth nuclei locations at 20x magnification level. Step (c) shows training nuclei patches of size 31×31 pixels where each nucleus is kept at the center. Step (d) shows contextual patches around j-th nucleus patch. Step (e) shows the deep features extraction of j-th nucleus using pre-trained VGG-19 model (Kather et al. (2019)). Step (f) shows the construction of spatial graph at each features hierarchy. Step (g) represents the objective function minimization and learning optimal correlation filters for each layer. Steps (h)-(m) in blue color show the testing pipeline where the learnt correlation filters are convolved on the test histology image and maximum response is computed for all layers for nucleus detection. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ers. In the third step, we compute the spatial structure graph by computing pairwise distances among different nuclei components using deep features. At each level of the hierarchy, a different spatial graph is constructed to capture the nucleus structure at that particular level. In the fourth step, we train a set of correlation filters on the deep features and enforce these filters to be orthogonal to the Laplacian basis of the spatial graph as well as to the contextual basis. A novel objective function is formulated incorporating multiple layers of deep features, and the information contained in the spatial graphs as well as contextual patches around each nucleus. This objective function is solved using the ADMM method with a closed-form solution. In the testing stage, we first extract the deep features from a test image followed by convolving the set of trained correlation filters. We compute the response of each filter on the corresponding deep feature of the test image and select the response with maximum amplitude where the peak of the filter shows the potential nucleus location. In the following subsections, we describe each component of the proposed SCC-DCF algorithm in more detail.

3.1. Pre-Processing steps

The hematoxylin dye is basic in nature and binds to the acidic DNA or RNA in the nucleus during the staining process. Therefore, most of the nucleus detection methods are designed by processing the hematoxylin channel only. We also extract the hematoxylin channel from the training and testing images using a color deconvolution method (Ruifrok et al. (2001)). In addition, the log transformation is applied to suppress the noise. The pixel values are then normalized to have a zero mean and unit variance for consistent intensity. These input images are zero-padded and multiplied by a cosine window to address the border effects of the convolution.

3.2. Deep feature extraction

We input the nuclei patches extracted from pre-processed training images and the contextual patches to the VGG-19 network for deep features extraction. We employ a VGG model which is trained for tissue classification (Kather et al. (2019)) and we extract deep features from each of the five convolutional layers. The earlier convolutional layers include features from *conv1-2, conv2-2*, and *conv3-4*, while later convolutional layers consist of *conv4-4* and *conv5-5*. The use of convolutional feature maps helps to encode the varying nuclear shapes and texture appearances (Fig. 2(e)). As the data propagates along with CNN, the semantic discrimination between different nuclei categories becomes strong, and the spatial resolution gradually reduces which helps in capturing coarse as well as fine information for nucleus localization. The pooling operations in the VGG-network result in the gradual reduction of spatial resolution with the increasing depth of the convolutional layers. However, low spatial resolution is insufficient for the accurate localization of the nucleus. We address this issue by resizing each feature map to the same size as the input image size of $n = m \times m$ pixels with bilinear interpolation and create input features matrix $\mathbf{X}_l \in \mathbb{R}^{n \times d}$, where d is the number of channels at the l-th level in the hierarchy. Since we have N number of nuclei, therefore, we concatenate $\mathbf{X}_l$ matrices into a global data matrix $\mathbf{M}_l \in \mathbb{R}^{n \times (N \times d)}$ such that a row in this matrix is a concatenation of the corresponding nucleus locations across all nuclei. The deep features visualization is shown in Fig. (2 (e)).

3.3. Proposed SCC-DCF formulation

Our aim is to learn a discriminative nucleus detection filter that can be applied to the test histology images to infer the locations of the nuclei. The correlation filter allows for dense sampling around the nucleus at a low computational cost. This is achieved by modeling all possible translations of the nucleus within a patch as cyclic shifts (Henriques et al. (2014)) as shown in Fig. 3 We concatenate all the cyclic shifts of $\mathbf{X}_l$ into a data matrix $\mathbf{A}^l \in \mathbb{R}^{n \times n \times d}$ resulting in circulant structure of this matrix. Such a structure facilitates a very efficient solution. If we multiply such a matrix with a filter $\mathbf{w}$ in vectorized form, it is equivalent to convolving the nucleus by that filter in the spatial domain. Let $\mathbf{y} \in \mathbb{R}^n$ be the regression target which is a vectorized image of a 2D Gaussian label.

Our aim is to learn a multi-channel filter $\mathbf{w}_i^l \in \mathbb{R}^n$ for the l-th layer by using the following ridge regression problem in the

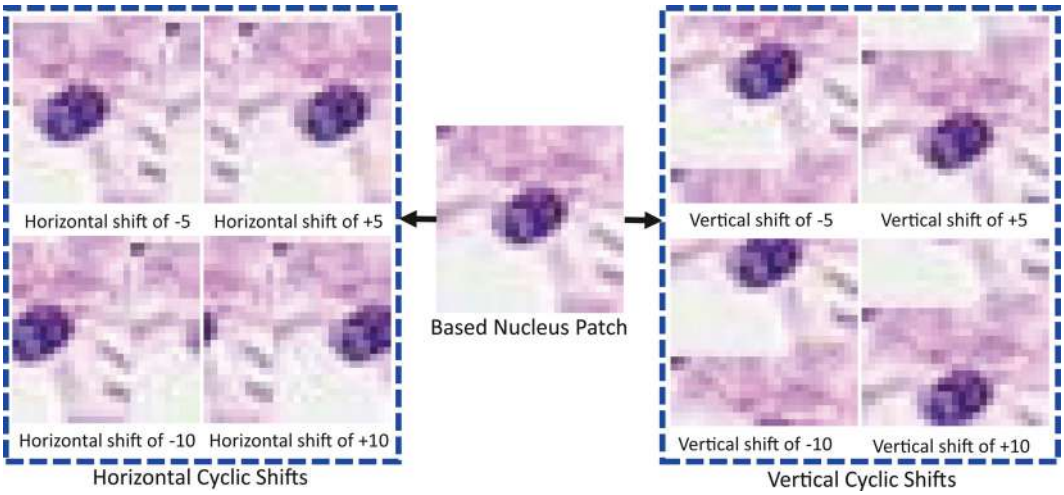


Fig. 3. An example of horizontal and vertical cyclic shift operation of a base nucleus patch. For illustration purpose only eight cyclic shifts are shown however in our algorithm pixel-wise cyclic shifts are performed.

Fourier domain (Hamed et al. (2013)) as:

$$\underset{\mathbf{w}_1^l,\mathbf{w}_2^l,\cdots,\mathbf{w}_d^l}{\operatorname{argmin}} \frac{1}{2N}\sum_{j=1}^{N}\left(||\sum_{i=1}^{d}\mathbf{A}_{i,j}^{l}\mathbf{w}_i^l-\mathbf{y}||^2+\lambda_1\sum_{i=1}^{d}||\mathbf{w}_i^l||^2\right), \tag{1}$$

where $\mathbf{A}_{i,j}^{l}\in\mathbb{R}^{n\times n}$ contains all circulant shifts of the i-th channel of the j-th nucleus feature map at l-th layer. λ_1 is a regularization parameter and N is the total number of training nuclei patches. The objective function in Eq. (1) is convex and has a unique global minimum. By taking gradient with respect to $\mathbf{w}^l$ and equating to zero leads to a closed-form solution for the filter. For the sake of simplicity, we ignore the superscript l and we consider $A_{i,j}$ and $\mathbf{w}$ for a particular layer.

In order to reduce the boundary effects of correlation filters, we introduce contextual information in the objective function defined by Eq. (1). For each vectorized nucleus patch $\mathbf{a}_j$, we extract r contextual patches $\mathbf{c}_{j,k}$, where $1\leq k\leq r$, covering 50% nucleus patch and 50% contextual information. The contextual information consists of the pixels outside the current nuclei patch. These pixels mainly contain non-nuclei regions however they may also contain parts of the other nuclei. The regression target labels are selected as zero for these contextual patches. This is because our aim is to enforce the correlation filter to generate a very low response on the contextual information of the nucleus and generate a high response on the nucleus itself. The same circulant shift operations are applied on each nucleus patch $\mathbf{a}_j$ as well as its contextual patches $\mathbf{c}_{j,k}$ to get $\mathbf{C}_{i,j,k}$, where $1\leq i\leq d$. The objective function defined by Eq. (1) incorporating contextual information can then be formulated as follows:

$$\underset{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_d}{\operatorname{argmin}} \frac{1}{2N}\sum_{j=1}^{N}\left(||\sum_{i=1}^{d}\mathbf{A}_{i,j}\mathbf{w}_i-\mathbf{y}||^2 +\lambda_2\sum_{k=1}^{r}\sum_{i=1}^{d}||\mathbf{C}_{i,j,k}\mathbf{w}_i||^2+\lambda_1\sum_{i=1}^{d}||\mathbf{w}_i||^2\right), \tag{2}$$

where in the second term minimizing $||\mathbf{C}_{i,j,k}\mathbf{w}_i||^2$ means that filter $\mathbf{w}_i$ is enforced to be orthogonal to the contextual patches $\mathbf{C}_{i,j,k}$. It is because while computing matrix-vector multiplication, inner products are computed between the rows of $\mathbf{C}_{i,j,k}$ and $\mathbf{w}_i$. Each row of $\mathbf{C}_{i,j,k}$ is a circulant shifted version of the contextual patch therefore minimization of $||\mathbf{C}_{i,j,k}\mathbf{w}_i||^2$ enforces $\mathbf{w}_i$ to be close to being orthogonal to each shifted version of the contextual patch. In order to make the correlation filter to be nucleus structure-aware, we propose to construct a spatial graph across different components of the nucleus. Since, the Laplacian matrix of this graph encodes the nucleus structure, therefore; we constrained the correlation filter to be orthogonal to these Laplacian basis. The proposed SCC-DCF objective function is formulated as follows:

$$\underset{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_d}{\operatorname{argmin}} \frac{1}{2N}\sum_{j=1}^{N}\left(||\sum_{i=1}^{d}\mathbf{A}_{i,j}\mathbf{w}_i-\mathbf{y}||^2 +\lambda_2\sum_{k=1}^{r}\sum_{i=1}^{d}||\mathbf{C}_{i,j,k}\mathbf{w}_i||^2\right)+\frac{\lambda_1}{2}\sum_{i=1}^{d}||\mathbf{B}\mathbf{w}_i||^2, \tag{3}$$

where $\mathbf{B}$ is the Laplacian basis of the spatial graph constructed across all deep feature channels and it encodes the spatial nucleus structure. In the following sub-sections, we provide more details of the graph construction and objective function.

3.4. Spatial graph regularization

For each level of the hierarchy, the deep features are computed using a deep neural network and a different spatial graph $\mathbf{G}_l$ is computed capturing the nuclear appearance variations at that level.

Let $\mathbf{G}_l=(\mathbf{V}_l,\mathbf{Z}_l)$ be an undirected weighted spatial graph at l-th level, where $\mathbf{V}_l$ and $\mathbf{Z}_l$ are the vertices and the edge weighted adjacency matrices. Vertex $\mathbf{V}_l(i)\in\mathbb{R}^{n\times d}$ contains feature values for the i-th nucleus across d channels which is represented as a row vector in the features matrix $\mathbf{M}_l$. Since, the correlation filter preserves the nucleus structure on the Riemann manifold (Yin et al. (2015)), therefore; the spatial nucleus structure ensures that if two nodes $\mathbf{V}_l(i)$ and $\mathbf{V}_l(j)$ are close in a data manifold then there corresponding filter coefficients should also be close. We consider spatial closeness among the nuclei feature maps, encoded, in the graph $\mathbf{G}_l$ using k-nearest neighbor method (Muja and Lowe (2014)). First, we search for the closest neighbors for all the rows in the features matrix $\mathbf{M}_l$ based on the Euclidean distance, where each vertex is connected to its k-nearest neighbors, so that if $\mathbf{V}_l(i)$ and $\mathbf{V}_l(j)$ are in the k-nearest neighbors of each other, we set:

$$Z_l(i,j)=\exp\left(-\frac{||\mathbf{V}_l(i)-\mathbf{V}_l(j)||_2^2}{2\sigma_s^2}\right), \tag{4}$$

where σ_s is a smoothing parameter and its value is the average distance among the vertices in $\mathbf{G}_l$. If two vertices $\mathbf{V}_l(i)$ and $\mathbf{V}_l(j)$ are within the k-nearest neighbors then they are connected:

$Z_l(i,j) > 0$, otherwise, $Z_l(i,j) = 0$. The normalized Laplacian matrix $\mathbf{L}_l$ is computed using the weighted adjacency matrix $\mathbf{Z}_l$ as:

$$\mathbf{L}_l = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{Z}_l\mathbf{D}^{-\frac{1}{2}}, \tag{5}$$

where $\mathbf{D} \in \mathbb{R}^{n\times n}$ is the degree matrix with its i-th diagonal element being equal to the sum of the i-th row of $\mathbf{Z}_l$ and $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the identity matrix. The eigenvectors of $\mathbf{L}_l$ act as Laplacian basis which encodes the nuclei structural information captured by the graph $\mathbf{G}_l$. To enforce the correlation filter to be structure-aware, we constrained $\mathbf{w}_i$ to act as the eigenvector of $\mathbf{L}_l$. For this purpose, we generalize the eigenvalue problem $\mathbf{w}_i^\top\mathbf{L}_l\mathbf{w}_i$ to be minimized independently for each channel as:

$$\Theta_l = \sum_{i=1}^{d}\mathbf{w}_i^\top\mathbf{L}_l\mathbf{w}_i. \tag{6}$$

The matrix $\mathbf{L}_l$ can also be symmetrically decomposed as:

$$\mathbf{L}_l = \mathbf{P}\Sigma\mathbf{P}^\top = (\Sigma^{\frac{1}{2}}\mathbf{P}^\top)^\top\Sigma^{\frac{1}{2}}\mathbf{P}^\top = \mathbf{B}_l^\top\mathbf{B}_l, \tag{7}$$

where $\mathbf{P} \in \mathbb{R}^{n\times n}$ is orthonormal matrix with each column being an eigenvector of $\mathbf{L}_l$, and $\Sigma \in \mathbb{R}^{n\times n}$ is a diagonal matrix with its diagonal element Σ_{ii} being a singular value of $\mathbf{L}_l$ (sorted as $0 \le \Sigma_{ii} \le \cdots \Sigma_{nn}$). The matrix $\mathbf{B}_l = \Sigma^{\frac{1}{2}}\mathbf{P}^\top$ is computed using all eigenvectors of $\mathbf{L}_l$. It may be considered as a combination of the scaled basis of the graph-Laplacian matrix $\mathbf{L}_l$ and thus defining a manifold structure of the nucleus. Substituting $\mathbf{L}_l = \mathbf{B}_l^\top\mathbf{B}_l$ in Eq. (6), we get:

$$\Theta_l = \sum_{i=1}^{d}\mathbf{w}_i^\top\mathbf{B}_l^\top\mathbf{B}_l\mathbf{w}_i = \sum_{i=1}^{d}||\mathbf{B}_l\mathbf{w}_i||^2. \tag{8}$$

The spatial structural constraint above can be interpreted as enforcing correlation filter $\mathbf{w}_i$ in each channel to be orthogonal to the eigenvectors of $\mathbf{L}_l$. Assuming that the manifold spanned by the non-nuclei patches is different from the manifold spanned by the nuclei patches, therefore, such a correlation filter will be able to better discriminate the nucleus from the non-nucleus region resulting in performance improvement.

3.5. SCC-DCF Optimization

We optimize SCC-DCF model (3) using ADMM method by solving one variable and fixing others (Boyd et al. (2011)). We first introduce two auxiliary variables as $\mathbf{p} = \mathbf{w}$ and a spatial filter $\mathbf{q} = \mathbf{w}$ to make the objective function separable. The constrained optimization problem in Lagrangian form is given by

$$\begin{aligned}\mathcal{L}(\mathbf{w},\mathbf{p},\mathbf{h},\mathbf{q},\mathbf{f}) = \frac{1}{2N}\sum_{j=1}^{N}\Bigg(||\sum_{i=1}^{d}\mathbf{A}_{i,j}\mathbf{w}_i - \mathbf{y}||^2 \\ +\lambda_2\sum_{k=1}^{r}\sum_{i=1}^{d}||\mathbf{C}_{i,j,k}\mathbf{p}_i||^2\Bigg) + \frac{\lambda_1}{2}\sum_{i=1}^{d}||\mathbf{B}\mathbf{q}_i||^2 \\ +\frac{\gamma}{2}\sum_{i=1}^{d}\|\mathbf{w}_i - \mathbf{p}_i + \mathbf{h}_i\|^2 + \frac{\gamma}{2}\sum_{i=1}^{d}\|\mathbf{w}_i - \mathbf{q}_i + \mathbf{f}_i\|^2,\end{aligned} \tag{9}$$

where $\mathbf{f}$ and $\mathbf{h}$ are auxilliary variables. Each sub-problem **w, p, q, f**, and **h** can then be solved efficiently using ADMM.

Solving Sub-problem w: By fixing other variables in (9) excluding $\mathbf{w}$, the sub-problem $\mathbf{w}^{(t+1)}$ at the $(t+1)$-th iteration can be written as:

$$\begin{aligned}\mathbf{w}^{(t+1)} = \min_{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_d}\frac{1}{N}\sum_{j=1}^{N}\Bigg(||\sum_{i=1}^{d}\mathbf{A}_{i,j}\mathbf{w}_i - \mathbf{y}||^2\Bigg) \\ +\gamma\sum_{i=1}^{d}\|\mathbf{w}_i - \mathbf{p}_i + \mathbf{h}_i\|^2 + \gamma\sum_{i=1}^{d}\|\mathbf{w}_i - \mathbf{q}_i + \mathbf{f}_i\|^2\end{aligned} \tag{10}$$

Using Parseval's theorem, the above equation can be rewritten in the Fourier domain as:

$$\begin{aligned}\operatorname*{argmin}_{\widehat{\mathbf{w}}}\frac{1}{N}\sum_{j=1}^{N}\Bigg(||\sum_{i=1}^{d}\widehat{\mathbf{x}}_{i,j}\odot\widehat{\mathbf{w}}_i - \widehat{\mathbf{y}}||^2\Bigg) \\ +\gamma\sum_{i=1}^{d}||\widehat{\mathbf{w}}_i - \widehat{\mathbf{p}}_i + \widehat{\mathbf{h}}_i||^2 + \gamma\sum_{i=1}^{d}||\widehat{\mathbf{w}}_i - \widehat{\mathbf{q}}_i + \widehat{\mathbf{f}}_i||^2\end{aligned} \tag{11}$$

where $\widehat{\mathbf{x}}_{i,j}$ is the i-th column vector of deep features matrix $\mathbf{X}_l$ for j-th nucleus in the Fourier domain and $\widehat{\mathbf{w}}_i$ denotes the DFT of the filter $\mathbf{w}_i$. From above equation, we can see that the k-th element of the label $\widehat{\mathbf{y}}_i$ only depends on the k-th element of the filter $\widehat{\mathbf{w}}_i$ and sample $\widehat{\mathbf{x}}_{i,j}$ across all d channels. Therefore, it can be further decomposed into n subproblems. Let $\mathcal{V}_k(\cdot)$ takes the k-th element for all d channels. Then, Eq. (11) becomes

$$\begin{aligned}\operatorname*{argmin}_{\mathcal{V}_k(\widehat{\mathbf{w}})}\frac{1}{N}\sum_{j=1}^{N}\left\|\mathcal{V}_k(\widehat{\mathbf{x}}_j)^\top.\mathcal{V}_k(\widehat{\mathbf{w}}) - \widehat{\mathbf{y}}(k)\right\|^2 \\ +\gamma\left\|\mathcal{V}_k(\widehat{\mathbf{w}}) - \mathcal{V}_k(\widehat{\mathbf{p}}) + \mathcal{V}_k(\widehat{\mathbf{h}})\right\|^2 + \gamma\left\|\mathcal{V}_k(\widehat{\mathbf{w}}) - \mathcal{V}_k(\widehat{\mathbf{q}}) + \mathcal{V}_k(\widehat{\mathbf{f}})\right\|^2\end{aligned} \tag{12}$$

By taking the derivative with respect to $\mathcal{V}_k(\widehat{\mathbf{w}})$ and setting it zero, we can get a closed-form solution

$$\begin{aligned}\mathcal{V}_k(\widehat{\mathbf{w}}) = \left(\tfrac{1}{N}\textstyle\sum_{j=1}^{N}\mathcal{V}_k(\widehat{\mathbf{x}}_j)\mathcal{V}_k(\widehat{\mathbf{x}}_j)^\top + 2\gamma\mathbf{I}\right)^{-1}\alpha, \\ \alpha = \tfrac{1}{N}\textstyle\sum_{j=1}^{N}\mathcal{V}_k(\widehat{\mathbf{x}}_j)\widehat{\mathbf{y}}(k) + \gamma\left(\mathcal{V}_k(\widehat{\mathbf{p}}) - \mathcal{V}_k(\widehat{\mathbf{h}}) + \mathcal{V}_k(\widehat{\mathbf{q}}) - \mathcal{V}_k(\widehat{\mathbf{f}})\right)\end{aligned} \tag{13}$$

Since, $\mathcal{V}_k(\widehat{\mathbf{x}}_j)\mathcal{V}_k(\widehat{\mathbf{x}}_j)^\top$ is a rank-1 matrix, Eq. (13) can be solved more efficiently using ShermanMorrison formula (Petersen et al. (2008)), we have

$$\mathcal{V}_k(\widehat{\mathbf{w}}) = \frac{1}{2\gamma}\left(\mathbf{I} - \frac{\frac{1}{N}\sum_{j=1}^{N}\mathcal{V}_k(\widehat{\mathbf{x}}_j)\mathcal{V}_k(\widehat{\mathbf{x}}_j)^\top}{2\gamma + \frac{1}{N}\sum_{j=1}^{N}\mathcal{V}_k(\widehat{\mathbf{x}}_j)^\top\mathcal{V}_k(\widehat{\mathbf{x}}_j)}\right)\alpha \tag{14}$$

Note that Eq. (14) only contains vector multiply-add operation and thus can be computed efficiently. The filter $\mathbf{w}$ can then be obtained by the inverse DFT of $\widehat{\mathbf{w}}$.

Solving Sub-problem q: In Eq. (9), fixing other variables excluding $\mathbf{q}$, the solution for sub-problem $\mathbf{q}^{t+1}$ at the (t + 1)-th iteration is given by

$$\mathbf{q}^{(t+1)} = (\tilde{\mathbf{B}}\tilde{\mathbf{B}}^\top + \lambda_1\mathbf{I})^{-1}(\gamma\tilde{\mathbf{w}} + \gamma\tilde{\mathbf{f}}), \tag{15}$$

where $\tilde{\mathbf{B}}$ represents the $dn \times dn$ diagonal matrix concatenated with d diagonal matrices Diag($\mathbf{B}$). The vectors $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{f}}$ denote the concatenated vectors of $\mathbf{w}_i$ and $\mathbf{f}_i$ accross d-channels.

Solving Sub-problem p: : In Eq. (9), fixing other variables excluding $\mathbf{p}$, the solution for sub-problem $\mathbf{p}^{t+1}$ at the (t + 1)-th iteration is given by:

$$\mathbf{p}^{(t+1)} = \frac{\lambda_2}{N}\sum_{j=1}^{N}(\sum_{k=1}^{r}\tilde{\mathbf{Q}}_{j,k}^\top\tilde{\mathbf{Q}}_{j,k} + \lambda_2\mathbf{I})^{-1}(\gamma\tilde{\mathbf{w}} + \gamma\tilde{\mathbf{h}}), \tag{16}$$

where $\tilde{\mathbf{Q}}$ represents the $dn \times dn$ diagonal matrix concatenated with d diagonal matrices Diag($\mathbf{C}$). The vector $\tilde{\mathbf{h}}$ denotes the concatenated vector of $\mathbf{h}_i$ accross d-channels.

Similarly, the variables **h, m**, and γ can be updated iteratively in (9) as:

$$\begin{aligned}\mathbf{h}^{(t+1)} &= \mathbf{w}^{(t+1)} - \mathbf{p}^{(t+1)} + \mathbf{h}^{(t)}, \\ \mathbf{f}^{(t+1)} &= \mathbf{w}^{(t+1)} - \mathbf{q}^{(t+1)} + \mathbf{f}^{(t)}, \\ \gamma^{(t+1)} &= \min(\gamma^{\max}, \rho\gamma^t),\end{aligned} \tag{17}$$

where ρ is a scalar term. Algorithm 1 summarizes the optimization procedure.

Algorithm 1: Proposed SCC-DCF algorithm for nucleus detection.

Input: N nuclei patches of size 31×31 pixels.
Initialization: Input feature matrix $\mathbf{A}_l \in \mathbb{R}^{n \times n \times d}$, λ_1, λ_2, $\gamma = 10$, $\gamma_{\max} = 100$, $\rho = 1.2$, $\mathbf{w}^0 = 0$, $\mathbf{h}^0 = 0$, and $\mathbf{f}^0 = 0$.
Compute $\mathbf{B} \in \mathbb{R}^{n \times n}$ using Eqs. (4)-(7).
Training:while *not converged* $(t = 0, 1, .)$ **do**
1. Compute $\mathbf{w}^{t+1}$ using (14).
2. Compute $\mathbf{q}^{(t+1)}$ using (15)
3. Compute $\mathbf{p}^{(t+1)}$ using (16).
5. Update $\mathbf{h}^{(t+1)}$, $\mathbf{f}^{(t+1)}$, and $\gamma^{(t+1)}$ using (17).

end
Output: $\mathbf{w}, \mathbf{q}, \mathbf{p}$
Testing Procedure:
1. Extract deep feaures from test image
2. Compute response $\mathbf{R}_l$ from each channel using Eq. (15)
3. Compute $\mathbf{R}_l$ using Eq. (18) and maximum response across l layers.

3.6. Nucleus detection

Given a test histology image with deep features $\mathbf{X}_l$ and learnt correlation filters $\mathbf{w}_i$ for each channel, the correlation response $\mathbf{R}_l$ at the l-th layer is then computed by

$$\mathbf{R}_l = \mathcal{F}^{-1}\left(\sum_{i=1}^{d} \hat{\mathbf{w}}_l^i \odot \hat{\mathbf{x}}_l^{i*}\right), \tag{18}$$

where $\hat{\mathbf{x}}_l^{i*}$ is the complex conjugate of Fourier representation of feature vector $\mathbf{x}_l^i$ at l-th layer and i-th channel. In order to integrate the response from multiple layers, we use the max rule across all layers. Such late fusion schemes have already been used in visual object tracking applications (Ma et al. (2015); Wang et al. (2018)). It provides computational efficiency as well as improved performance due to enabling the proposed algorithm for multi-resolution nucleus detection. The local maxima values represent the detected nuclei centers. Algorithm 1 recapitulates the different steps of training and testing stages in our approach.

4. Experimental evaluations

The proposed algorithm Spatially Constrained Context-aware hierarchical Deep Correlation Filters (SCC-DCF) is evaluated quantitatively as well as qualitatively on three different histology image datasets including two colon cancer datasets known as CRCHistoPhenotypes (Sirinukunwattana et al. (2016)) and Colorectal Nuclear Segmentation and Phenotypes (CoNSeP) (Graham et al. (2019)), and one multiple cancer types dataset known as PanNuke (Gamper et al. (2020)).

4.1. Compared methods

The results are compared with 15 existing state-of-the-art methods including:

- Structured Regression Convolutional Neural Network (SR-CNN)(Xie et al. (2015)): In this method, the last layer of the conventional CNN is replaced with the structured regression layer to encode the topological information for robust nucleus detection. Instead of labels, a proximity map is produced assigning higher values to the nucleus center.
- Spatially Constrained Convolutional Neural Network (SC-CNN) (Sirinukunwattana et al. (2016)): This method regresses the likelihood of each pixel is the center of a nucleus. During training, in the ground truth response map, high probability values are assigned in the vicinity of the nuclei centers.
- Shape Prior Convolutional Neural Network (SP-CNN) (Tofighi et al. (2018)): In the SP-CNN method, a set of canonical shapes is prepared with the help of a domain expert and used to perform nucleus detection using a deep neural network.
- Tunable Shape Prior Convolutional Neural Network (TSP-CNN) (Tofighi et al. (2019)): This method is the extended version of SP-CNN. The handcrafted canonical nuclei shapes are made learnable by introducing a new convolutional layer as well as regularization terms.
- Stacked Sparse Auto Encoder (SSAE) (Xu et al. (2015)): In this method, a sliding window is applied on each image patch in order to extract high-level features using an auto-encoder which are then classified as a nuclear or non-nuclear region.
- Crosswise SParse CNN (CSP-CNN) (Hou et al. (2019)): In this method, a convolutional autoencoder is used to decompose histology image patches into the foreground (nucleus) and background (non-nucleus).
- Local Isotropic Phase Symmetry measure (LIPSym) (Kuse et al. (2011)): In this method, a local symmetry measure is used for nucleus detection. This measure is designed to give high values at or near the nucleus center.
- CRImage (Yuan et al. (2012)): This method segments nuclei using morphological operation, followed by distance transform, and watershed method. The centroid of an individual segmented nucleus is then considered as detection.
- Vector Oriented Confidence Accumulation (VOCA) (Xie et al. (2018)): It is a deep convolutional encoder-decoder-based method that learns a confidence score, localization vector, and weight of contribution for each pixel.
- Mixture Density Networks (MDN) (Koohababni et al. (2018)): This method maps a single input image patch to a Gaussian probability density function of the nuclei center. If a local patch contains at least one nucleus, then a Gaussian mixture model is used to obtain the density functions.
- Correlation Filter (CF) (Ahmad et al. (2018)): This is the baseline method of the proposed SCC-DCF algorithm. It was initially proposed for nucleus detection by exploiting CFs.
- Horizontal and Vertical distances prediction Network (HoVer-Net) (Graham et al. (2019)): This method estimates vertical and horizontal distances of each pixel from its nearest nucleus center. These distances are then utilized to separate clustered nuclei, resulting in improved segmentation.
- Deep regression of the Distance map (DIST) (Naylor et al. (2018)):This method segments clustered nuclei by formulating the segmentation problem as a regression task of the distance map.
- Micro-Net (Raza et al. (2019)): It is trained at multiple resolutions of the input images and it connects the intermediate layers for better localization. It generates the output using multi-resolution deconvolution filters and it can handle variable object sizes and intensities.
- Mask-RCNN (He et al. (2017)): The original Mask R-CNN is re-trained for the task of nucleus segmentation by using smaller anchor boxes (Graham et al. (2019)).

For a fair comparison, the implementations were obtained from the original authors and their proposed parameters were used.

4.2. Variants of the proposed SCC-DCF algorithm

In addition, we also compared the performance of different variants and baselines of the proposed SCC-DCF algorithm including

- Context-Aware Correlation Filters (CACF): In this variant, no graph-based regularization is employed and we only utilized the hematoxylin channel of the histology images as used in the baseline CF setting (Ahmad et al. (2018)).
- Deep CACF (DCACF) using deep features: This variant is similar to CACF except we trained CFs by using the deep features extracted from the pre-trained VGG-19 network (Kather et al. (2019)).
- SCC-DCF-1: In this variant, we employed principal component analysis for dimensionality reduction to construct the spatial graph.

4.3. Experimental details

Following the protocol in (Sirinukunwattana et al. (2016)), we consider the detections within a radius of 6 pixels from the annotated ground truth center of the nucleus as true positive. In the case of multiple detections within the same 6 pixels region, only the detection closest to the ground truth center is considered as true positive.

We empirically observe that the two parameters $\lambda_1 = 0.3$ and $\lambda_2 = 0.5$, in model (3) are the best combination in all experiments. We crop the square region centered at the nucleus, in which the side length of the region is resized to $\sqrt{5n}$ (n represents the size of the nucleus). For the construction of spatial graphs, we used $k = 10$ nearest neighbors. The optimization hyper-parameters are set as $\gamma = 10$, $\gamma_{\max} = 100$, and $\rho = 1.3$ as suggested in (Danelljan et al. (2015)). The number of contextual patches is also set as $r = 4$.

4.4. Datasets

4.4.1. CRCHistoPhenotypes dataset (Sirinukunwattana et al. (2016))

This dataset[1] consists of 100 H&E stained CRC histology images each of size 500 × 500 pixels, cropped from non-overlapping regions of 10 WSIs from 9 different patients, at a pixel resolution of 0.55 m/pixel (20× optical magnification). The nuclei were manually annotated by an experienced pathologist. There is a total number of 29,756 nuclei marked at the center for detection purposes. The images consist of a variety of tissue appearances from both malignant and normal regions of WSIs. This is a challenging dataset with heterogeneous nuclear shapes and significant clutter, making the nucleus detection task difficult more challenging. Fig. 4 shows sample tissue images of CRC for nucleus detection from this dataset. We have employed 2-fold cross-validation by using 50 images for training and the remaining 50 images for testing in each fold as recommended by (Sirinukunwattana et al. (2016)).

4.4.2. Colorectal nuclear segmentation and phenotypes (consep) dataset (Graham et al. (2019))

This dataset[2] consists of 41 large patches each of size 1000 × 1000 pixels at 40× resolution level. These patches were extracted from H & E stained 16 CRC WSIs of different patients. This dataset consists of 24,319 manually annotated nuclei boundaries by experienced pathologists. The dataset was initially proposed for nucleus segmentation purposes; however, we employ it for nucleus detection by estimating the centroid of each nucleus segment and using it as ground truth. Similar to the CRCHistoPhenoypes dataset, this dataset also contains the challenges of heterogeneity and nuclei clutteredness. Fig. 5 shows a sample tissue image for nucleus detection from CoNSeP dataset. We have also employed the same training and testing splits as recommended by the original authors, consisting of 27 patches for training and the remaining 14 patches for testing (Graham et al. (2019)).

4.4.3. PanNuke Dataset (Gamper et al. (2020))

The PanNuke[3] is a large and diverse dataset for nucleus segmentation and classification that has been automatically annotated and validated by expert pathologists. This dataset contains 481 visual fields captured at the 40x resolution level from 19 distinct cancer types. Within these visual fields, 189,744 nuclei boundaries are exhaustively annotated. For a fair comparison of different models, patches of size 256×256 are extracted from these visual fields and randomly divided into three splits. Three folds evaluation is performed by selecting one of split as training, another one as validation, and the remaining one as testing split. In case no validation data is required, the validation and training splits are merged together and used as training splits. For quantitative comparisons, the results are the average across the three folds. The dataset was initially proposed for nucleus segmentation purposes; however, we employ it for nucleus detection by estimating the centroid of each nucleus segment and using it as ground truth. Similar to the other datasets, this dataset also contains the heterogeneity and nuclei clutteredness challenges. In addition to that, the PanNuke dataset contains a large number of distinct cancer types while other datasets contain only colon cancer histology images. Fig. 6 shows three sample images from this dataset.

4.5. Evaluation metrics

A detected nucleus is considered to be a true positive (TP) if it is within 6 pixel radius from the ground truth, otherwise; the detection is considered to be a false positive (FP). The ground truth not matching with any detections considered as false negatives (FN). For quantitative evaluations of SCC-DCF and its comparison with current state-of-the-art methods, we use Precision, Recall, and F_1 score, which is computed as:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{19}$$

where Precision is defined as $TP/(TP + FP)$ and Recall is $TP/(TP + FN)$. The aim is to maximize the F_1 measure so that its value is close to one. For a detailed comparison with current state-of-the-art methods, we have presented precision-recall (ROC) curves for different methods, as shown in Figs. 8 (a)-(c). The ROC curves are plotted by varying the threshold values applied to the predicted filter response before locating local maxima to avoid false-positive detections.

4.6. Evaluation on CRCHistophenotypes dataset

4.6.1. Visual assessment

The visual results of the proposed SCC-DCF are compared with current state-of-the-art methods including SC-CNN, SP-CNN, CF, TSP-CNN, and our variant CACF are shown in Fig. 4. The ground truth nuclei are marked as blue circles while the results of nucleus detection are marked as green circles. The results are shown for two different sample images containing cluttered tumor heterogeneity and lymphocytes. The results are displayed for the maximum F_1 score for each nucleus detection method.

Visual results demonstrate that the proposed SCC-DCF achieved excellent visual performance as it detected the majority of the nuclei in these images even in the presence of cluttered nuclei

[1] https://warwick.ac.uk/fac/sci/dcs/research/tia/data/crchistolabelednucleihe/.
[2] https://warwick.ac.uk/fac/sci/dcs/research/tia/data/hovernet/.
[3] https://warwick.ac.uk/fac/sci/dcs/research/tia/data/pannuke.

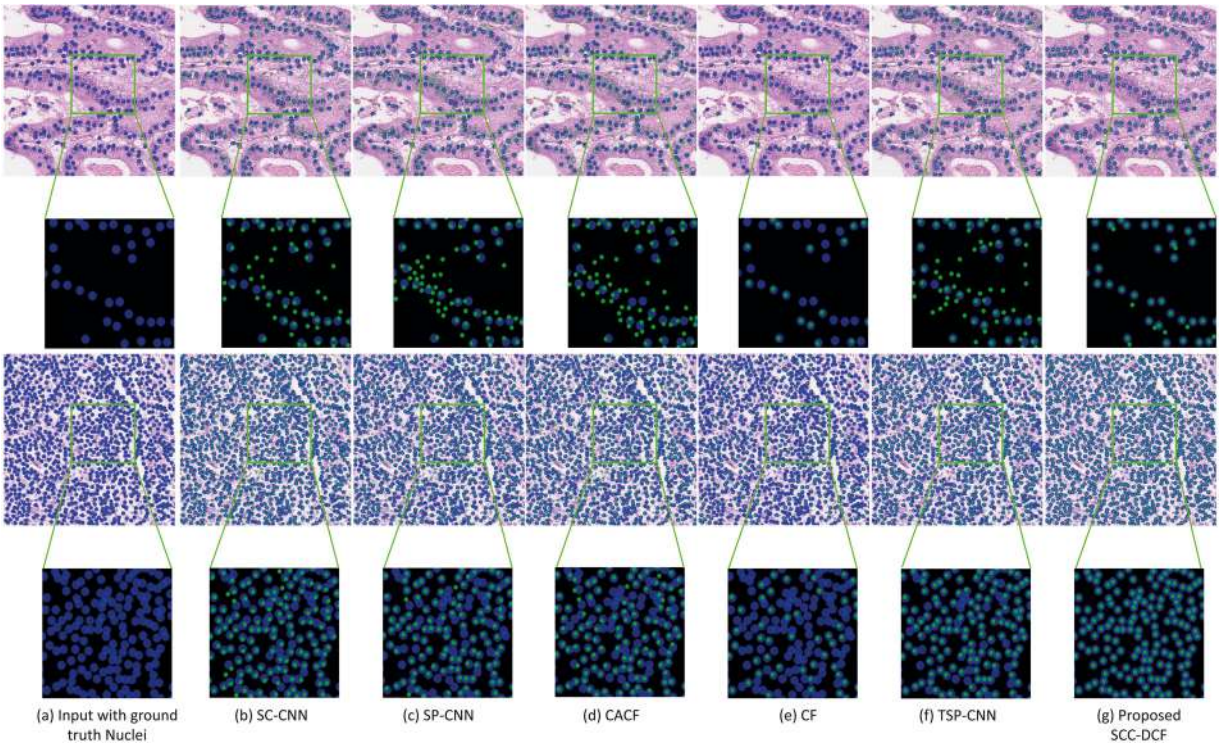


Fig. 4. Comparative visual assessment of the proposed SCC-DCF algorithm with current state-of-the-art methods on CRCHistoPhenotypes dataset (Sirinukunwattana et al. (2016)). The top row shows a test image with tumor heterogeneity while the test image in the bottom row shows the nuclei cluttered of lymphocytes. The ground truth nuclei locations are shown as blue circles, and nucleus detection results are shown as green circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

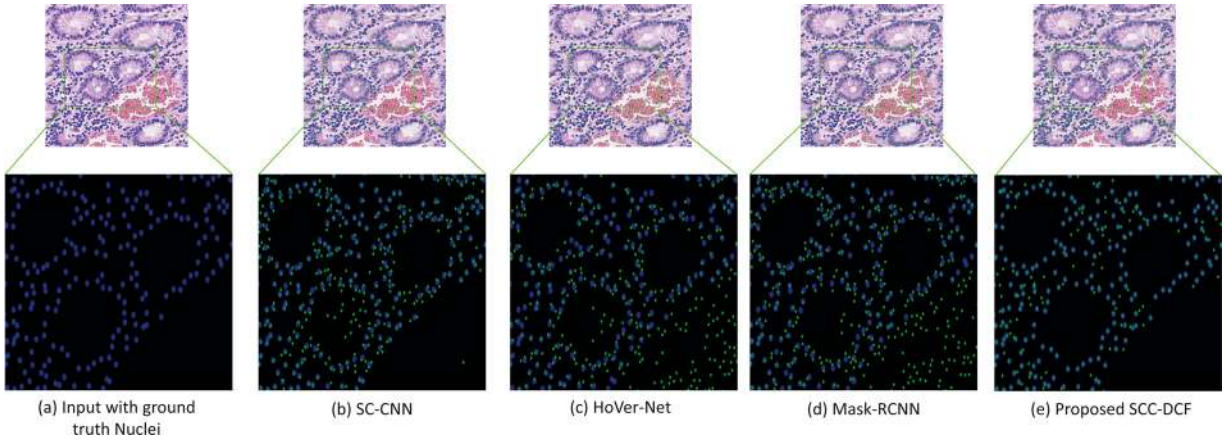


Fig. 5. Comparative visual assessment of the proposed SCC-DCF algorithm with current state-of-the-art methods on CoNSeP dataset (Graham et al. (2019)). The sample image shows benign epithelial nuclei heterogeneity and nuclei clutteredness. The ground truth nuclei locations are shown as blue circles and nucleus detection results are shown as green circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and varying nuclear shapes. CF has exhibited high false-negative rates while TSP-CNN, SP-CNN, and CACF methods have shown high false-positive rates in case of tumor heterogeneity (top row). The majority of the methods in the second row have performed well except CF and CACF which have relatively high false-negative rates, especially in the case of cluttered nuclei clusters. Both TSP-CNN and SP-CNN have also shown some miss-detections in the case of nuclei clutter. The proposed algorithm has also shown excellent results in the case of overlapping and clumped nuclei in the tumor epithelial region as shown in Fig. 7. This visual result demonstrates the localization capability of the proposed SCC-DCF algorithm leveraged by the spatial graph-based regularization.

4.6.2. Quantitative analysis

Table 1 compares the nucleus detection results of the proposed SCC-DCF algorithm with current state-of-the-art methods in terms of average precision, recall, and F_1 score. These results are reported by varying the threshold values to maximize the F_1 score. Overall, the proposed SCC-DCF has obtained the best F_1 score of 91.4%, which is 6.20% better than second-best performer TSP-CNN. Note that our variants, SCC-DCF-1 and DCACF, are the second and third best performers obtaining 90.6% and 88.5% F_1 score. Our variant CACF has also obtained an 85.0% F_1 score, which is comparable to TSP-CNN. The deep features in the DCACF method have resulted in a 3.5% performance boost over CACF while the structural graph-based regularization in the proposed SCC-DCF algorithm has

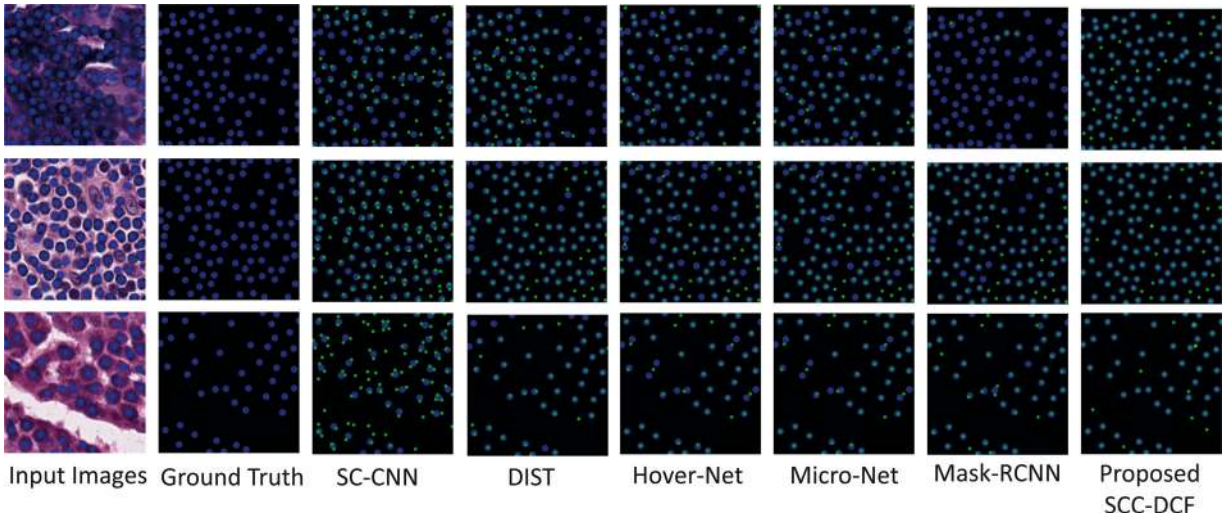


Fig. 6. Comparative visual assessment of the proposed SCC-DCF algorithm with current state-of-the-art methods on PanNuke dataset (Gamper et al. (2020)). From top to bottom: sample images of the breast, kidney, and adrenal gland cancers containing varying nuclear shape, size, and texture are shown in each row. From left to right, results of the nucleus detection are shown in each column for the compared methods including SC-CNN (Sirinukunwattana et al. (2016)), DIST (Naylor et al. (2018)), HoVet-Net (Graham et al. (2019)), Micro-Net (Raza et al. (2019)), Mask R-CNN (He et al. (2017)), and our proposed SCC-DCF algorithm. The ground truth nuclei locations are shown as blue circles while the nucleus detection results are shown as green circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

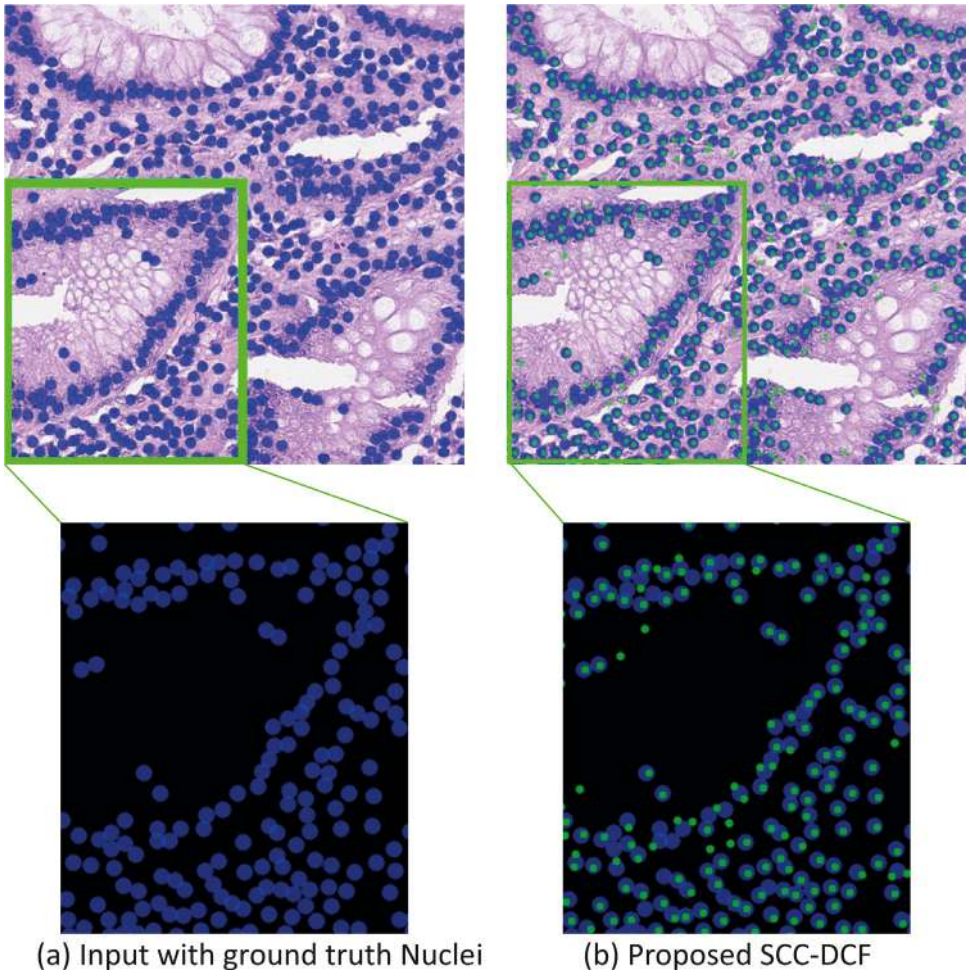
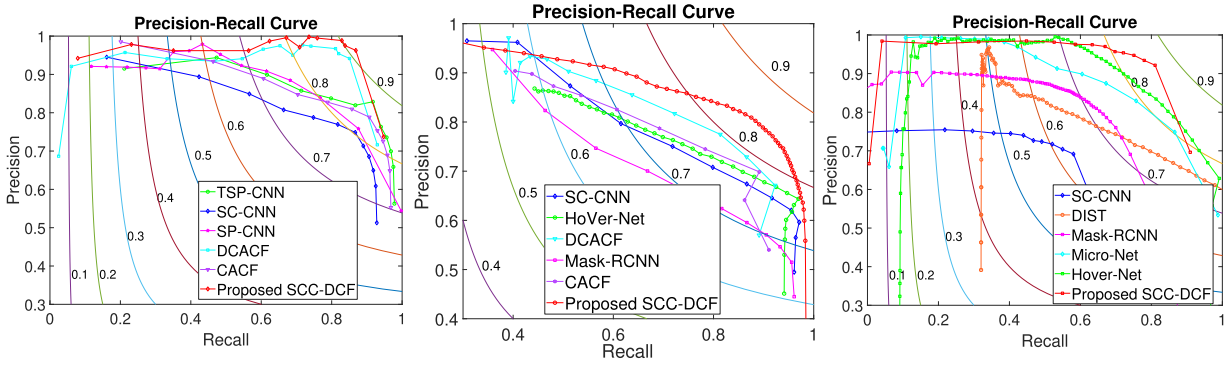


Fig. 7. Visual assessment of the proposed SCC-DCF algorithm in the epithelial region showing the case of overlapping and clumped nuclei. The test image is taken from the CRCHistoPhenotypes dataset (Sirinukunwattana et al. (2016)). The ground truth nuclei locations are shown as blue circles, and nucleus detection results are shown as green circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Comparative performance of the proposed algorithm with existing state-of-the-art methods in terms of the average precision, recall, and F_1 score on CRCHistoPhenotypes dataset (Sirinukunwattana et al. (2016)). The best and second-best performing methods are shown in red and blue, respectively.

Evaluation Measures	SC-CNN	TSP-CNN	SP-CNN	SR-CNN	SSAE	CSP-CNN	LIPSym	CRImage	VOCA	MDN	CF	CACF	DCACF	SCC-DCF-1	SCC-DCF
Precision	0.781	0.848	0.803	0.783	0.617	0.788	0.725	0.657	0.831	0.788	0.762	0.842	**0.925**	0.922	**0.981**
Recall	0.823	0.857	0.843	0.804	0.644	**0.886**	0.517	0.461	0.863	0.882	0.881	0.862	0.850	**0.892**	0.857
F_1 score	0.802	0.852	0.823	0.793	0.630	0.834	0.604	0.542	0.847	0.832	0.812	0.850	0.885	**0.906**	**0.914**



(a) CRCHistoPhenotypes dataset (b) CoNSeP dataset (c) PanNuke dataset

Fig. 8. Comparison of the proposed SCC-DCF, DCACF, and CACF algorithms with other state-of-the-art methods including TSP-CNN (Tofighi et al. (2019)), SP-CNN (Tofighi et al. (2018)), SC-CNN (Sirinukunwattana et al. (2016)), Hover-Net (Graham et al. (2019)), Mask-RCNN (He et al. (2017)), DIST (Naylor et al. (2018)), and Micro-Net (Raza et al. (2019)), in terms of precision-recall curves for nucleus detection on CRCHistoPhenotypes, CONSEP, and PanNuke datasets. Isolines in (a)-(c) indicate regions of the same F_1 scores.

further improved 2.90% performance over DCACF. The VOCA has obtained 84.7% performance and CSP-CNN has obtained 83.4% F_1 score.

Fig. 8 (a) shows the ROC curves of the best-performing methods, including SP-CNN, TSP-CNN, SC-CNN, CACF, DCACF, and our proposed SCC-DCF algorithm. The isolines corresponding to the regions of the same F_1 score are also shown in the background. It can be seen that the ROC curve for the proposed SCC-DCF algorithm approaches around 91.4% F_1 score, while the TSP-CNN remains the second-best method which obtained around 85.0% F_1 score.

4.7. Evaluation on consep dataset

4.7.1. Visual assessment

Fig. 5 shows the visual comparisons of the SCC-DCF algorithm with SC-CNN, HoVer-Net, and Mask-RCNN on a challenging test image of the CoNSeP dataset. The sample image contains tumor heterogeneity and the nuclei clutter with varying shapes. The ground truth nuclei are shown as blue circles, and nucleus detection results are shown as green circles.

SC-CNN has shown more false positives compared to the other methods, especially in the case of clustered nuclei. The localization of the proposed algorithm is significantly better compared to the other methods. The performance of HoVer-Net and Mask-RCNN has also remained low. Overall, the proposed algorithm has obtained the maximum true positive rate compared to the other methods.

4.7.2. Quantitative analysis

The proposed SCC-DCF algorithm is evaluated on the CoNSeP dataset quantitatively and compared with SC-CNN, CF, HoVer-Net, Mask-RCNN, Micro-Net, DIST, CACF, DCACF, and SCC-DCF-1 as shown in Table 2 by following the same experimental protocols defined in (Sirinukunwattana et al. (2016)). For the case of nucleus segmentation-based methods such as HoVer-Net, Mask-RCNN, Micro-Net, and DIST, the centroid of the segmented nucleus is considered as a nucleus detection and used for performance comparison. The other methods TSP-CNN and SP-CNN are only compared on the CRCHistoPhenotypes dataset because the training codes are not available, therefore these methods can not be trained on new datasets. The proposed SCC-DCF algorithm has obtained a maximum average F_1 score of 84.8% while the variants SCC-DCF-1, DCACF, and CACF obtained 82.2%, 79.4%, and 75.5% F_1 score. The nearest competitor is the nuclear segmentation-based method HoVer-Net which obtained 75.3% F_1 score. Since the CoNSeP dataset is more challenging than the CRCHistoPhenotypes dataset; therefore, the performance of all compared methods is reduced by 6.0% to 9.0%.

Fig. 8 (b) shows the ROC curves comparison of the proposed SCC-DCF algorithm with SC-CNN, HoVer-Net, DCACF, CACF, and Mask-RCNN on the CoNSeP dataset. The precision-recall curve for the proposed SCC-DCF algorithm approaches to around 84.0% F_1 score. Also, the recall range of 0.50 to 0.90 SCC-DCF algorithm ROC curve is significantly higher than the compared methods.

Table 2
Comparative performance of the proposed algorithm with existing state-of-the-art methods in terms of the average precision, recall, and F_1 score on CoNSeP dataset (Graham et al. (2019)). The best and second-best performing methods are shown in red and blue, respectively.

Methods	Precision	Recall	F_1 score
SC-CNN	0.751	0.803	0.737
CF	0.742	0.788	0.703
HoVer-Net	0.766	0.817	0.753
Mask-RCNN	0.743	0.719	0.713
Micro-Net	0.750	0.809	0.747
DIST	0.769	0.765	0.728
CACF	0.787	0.777	0.755
DCACF	**0.814**	**0.837**	0.794
SCC-DCF-1	0.837	0.808	**0.822**
SCC-DCF	**0.875**	**0.823**	**0.848**

Table 3
Comparative performance of the proposed algorithm with existing state-of-the-art methods in terms of the average precision, recall, and F_1 score on PanNuke dataset (Gamper et al. (2020)). The best and second-best performing methods are shown in red and blue, respectively.

Methods	Precision	Recall	F_1 score
SC-CNN	0.67	0.60	0.63
HoVer-Net	**0.82**	0.79	**0.80**
Mask-RCNN	0.76	0.68	0.72
Micro-Net	0.78	**0.82**	**0.80**
DIST	0.74	0.71	0.73
CACF	0.78	0.69	0.73
DCACF	0.77	0.75	0.76
SCC-DCF-1	0.79	0.80	0.79
SCC-DCF	**0.84**	**0.83**	**0.83**

4.8. Evaluation on pannuke dataset

4.8.1. Visual assessment

Fig. 6 shows the visual comparisons of the SCC-DCF algorithm with SC-CNN, DIST, HoVer-Net, Micro-Net, and Mask-RCNN on a challenging test images of the PanNuke dataset. These sample images belong to breast, kidney, and adrenal gland cancer types and contain tumor heterogeneity with varying nuclear shapes and sizes. The ground truth nuclei are shown as blue circles, and nucleus detection results are shown as green circles.

SC-CNN has shown more false positives compared to other methods in these test images. The nuclear segmentation methods DIST, Hover-Net, and Micro-Net show good nucleus detection results for the kidney cancer test image in the midlle. The breast cancer image on the top row poses a great challenge for these segmentation-based methods because they were not able to detect many nuclei. This effect is more pronounced for the Mask-RCNN where hardly 10 nuclei are detected out of more than 100 nuclei. Compared to these methods, our proposed algorithm has a very low false-negative rate. compared to SC-CNN and DIST methods, the proposed SCC-DCF has a low false-positive rate for test images.

4.8.2. Quantitative analysis

The proposed SCC-DCF algorithm is also evaluated on the PanNuke dataset quantitatively and compared with SC-CNN, CF, HoVer-Net, Mask-RCNN, Micro-Net, DIST, CACF, DCACF, and SCC-DCF-1 as shown in Table 3. The F_1 score reported in Table 3 is an average over three dataset folds. The proposed SCC-DCF algorithm has obtained 83.0% F_1 score while the nearest competitors are Hover-Net and Micro-Net both obtaining 80.0% F_1 score. The variant of the proposed algorithm SCC-DCF-1 has obtained 79.0% F_1 score that shows the importance of spatial graph into the correlation filters framework. On this dataset, these additional constraints has resulted in a 4.0% increase accuracy.

Fig. 8 (c) shows the ROC curves comparison of the proposed SCC-DCF algorithm with SC-CNN, DIST, Mask RCNN, Micro-Net, and Hover-Net on the PanNuke dataset. The precision-recall curve for the proposed SCC-DCF algorithm approaches to around 83.0% F_1 score.

4.9. Generalization evaluation

The generalization performance of the proposed SCC-DCF algorithm is also evaluated and compared with SC-CNN, TSP-CNN, and SP-CNN by training these methods on the CRCHistoPhenotypes dataset and testing on the CoNSeP dataset, as shown in Table 4. The proposed algorithm achieved best F_1 score of 62.9% among the compared methods which is 21.90% less than its own performance on CoNSeP dataset (Table 2) and 28.50% less than its performance on CRCHistoPhenotypes dataset (Table 1). Some reduction in the performance can also be attributed to the more challenging nature of the CoNSeP dataset. In contrast, the TSP-CNN has shown a 38.3% reduction as compared to its performance on the CRCHistoPhenotypes dataset (Table 1). The SC-CNN has shown a reduction of 34.9.% compared to CRCHistphenotypes dataset (Table 1) and 28.4% reduction compared to its performance on CoNSeP dataset (Table 2). Similarly, SP-CNN has also shown a reduction of 36.5% on the CRCHistophenotypes dataset (Table 1). Thus, compared to these existing methods, the proposed SCC-DCF algorithm has shown the best generalization performance.

Table 4
Evaluation of the proposed algorithm across datasets: training on CRCHistPhenotypes dataset (Sirinukunwattana et al. (2016)) and testing on CoNSeP dataset (Graham et al. (2019)). The best and second-best performing methods are shown in red and blue, respectively.

Methods	Precision	Recall	F_1 score
SC-CNN	0.346	**0.732**	0.453
SP-CNN	**0.457**	0.548	0.458
TSP-CNN	0.370	0.705	**0.469**
Proposed SCC-DCF	**0.618**	**0.641**	**0.629**

Table 5
Comparative performance of the proposed algorithm with SC-CNN methods using reduced training and testing splits on CRCHistoPhenotypes dataset (Sirinukunwattana et al. (2016)).

Data Splits	Methods	Precision	Recall	F_1 score
Training 30%	SC-CNN	0.716	0.703	0.657
Testing 70%	Proposed SCC-DCF	**0.876**	**0.777**	**0.823**
Training 20%	SC-CNN	0.380	0.879	0.521
Testing 80%	Proposed SCC-DCF	**0.871**	**0.743**	**0.801**
Training 10%	SC-CNN	0.341	0.580	0.414
Testing 90%	Proposed SCC-DCF	**0.791**	**0.651**	**0.714**

4.10. Evaluation on reduced training dataset

To further evaluate the strength of the proposed SCC-DCF algorithm, we reduced the training split on the CRCHistoPhenotypes dataset to 30%, 20%, and 10% and increased the testing split to 70%, 80%, and 90%, respectively. Because of the unavailability of the training codes of TSP-CNN and SP-CNN methods, we only compared our performance with SC-CNN as shown in Table 5. For the 30% training set, the performance of the SCC-DCF algorithm is 82.3% which is 9.10% reduced than its performance on 50% split used in Table 1. Compared to this, the performance of SC-CNN is reduced by 14.5% as observed in Table 1. For the case of 20% training data, the performance of the proposed SCC-DCF algorithm is 80.1%, which is 11.30% reduced than its performance on 50% training data. Compared to this, the performance of SC-CNN is reduced by 36.1%. At 10% training data, the proposed SCC-DCF algorithm observed the degradation of 20.0% while SC-CNN observed 38.8% degradation. Thus, the proposed algorithm has observed a graceful degradation with a reduced training dataset compared to the SC-CNN method. This superiority lies in the fact that our algorithm is based on correlation filters which can learn efficiently on relatively small training datasets as compared to deep neural network-based methods.

4.11. Computational complexity and execution time

We evaluate the computational complexity and execution time of the proposed algorithm which mainly depends on the optimization process and graph construction. We used the FLANN li-

braries for the graph construction using nearest neighbor strategy (Muja and Lowe (2014)). The spatial graph $\mathbf{G}_l$ complexity is $O(nd\log(n))$ where n is the number of pixels in each nucleus patch and d is the number of channels in each deep features hierarchy.

Since, Eq. (14) is separable in each pixel location, we solve this $\mathbf{w}$ element-wise and each element is a system of linear equations with d variables. Each sub-problem can be solved in $O(d)$ using Sherman-Morrison formula (Petersen et al. (2008)). Thus, the complexity of solving $\hat{\mathbf{w}}$ is $O(dn)$. Taking the DFT and inverse DFT into account, the complexity of solving $\mathbf{w}$ is $O(dn\log(n))$. The computational cost for solving both $\mathbf{p}$ and $\mathbf{q}$ sub-problems is $O(dn)$. Hence, the complexity of our SCC-DCF model is $O(n_1 dn\log(n))$, where n_1 represents the maximum number of iterations.

The execution time of the proposed algorithm is measured on a PC with an Intel Core i7 4.0 GHz, Titan Xp GPU, and 128 GB RAM. Compared with current state-of-the-art methods including SC-CNN, TSP-CNN, and HoVer-Net, the computational time of the proposed SCC-DCF algorithm is significantly reduced. On the CRCHistoPhenotypes dataset, the SCC-DCF algorithm takes 15.31 minutes for training, while testing time is 4.10 minutes on 50 test images each of size 500×500 pixels using MATLAB implementation. The time on each image varies depending upon the number of nuclei. Our variants SCC-DCF-1 (graph construction with compressed features using PCA), DCACF, and CACF takes 13.92, 13.43, and 3.30 minutes for training, and 4.10, 4.10, and 2.3 minutes for testing. In contrast, the SC-CNN method takes more than 8 hours of training time and 15.20 minutes for testing the same 50 images using Python implementation. The testing time of TSP-CNN and HoVer-Net using Python implementations are comparable with the testing time of the proposed algorithm in Matlab.

On the CoNSeP dataset, our proposed SCC-DCF algorithm took 16.71 minutes for training and 3.93 minutes for testing 14 images. The HoVer-Net method takes 380 minutes to train and 8.61 minutes for testing the same 14 images (Graham et al. (2019)). Thus, the proposed SCC-DCF algorithm is computationally attractive and its python implementation would be even more efficient.

5. Conclusions

In this work, novel constraints based on the spatial structure of the nucleus and its local contextual information are proposed in the discriminative correlation filter framework to handle varying nuclei shapes, texture, and clutter. The first constraint incorporates the spatial structure of the nuclei by constructing a dense graph across different nucleus components based on hierarchical deep features. The spatial graphs are computed at each resolution level of the deep features. The second constraint assists the correlation filter to discriminate between the nucleus and non-nucleus region. Both these constraints reduce the boundary effects during training and enhance nuclear localization. The proposed objective function containing spatial graph-based constraints and contextual information in the DCF framework is solved using the ADMM optimization method where the closed-form solution of each sub-problem is derived in a more efficient manner. At each level, we independently compute the constrained correlation filter response and maxima is seeked across all levels. The proposed algorithm dubbed as SCC-DCF has shown significant performance improvement on large-scale nucleus detection challenging datasets compared to 15 existing state-of-the-art methods. In the future, we will explore the strength of correlation filters for nucleus classification and tissue phenotyping problems.

Declaration of Competing Interest

- All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.
- This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.
- The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript
- The following authors have a following academic affiliations with organizations

CRediT authorship contribution statement

Sajid Javed: Conceptualization, Methodology, Software, Investigation, Project administration, Writing - original draft, Writing - review & editing. **Arif Mahmood:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing. **Jorge Dias:** Conceptualization, Methodology, Project administration, Writing - review & editing, Supervision. **Naoufel Werghi:** Conceptualization, Methodology, Project administration, Writing - review & editing, Supervision. **Nasir Rajpoot:** Conceptualization, Methodology, Project administration, Writing - review & editing, Supervision.

Acknowledgements

This publication acknowledges the support provided by the Khalifa University of Science and Technology under Award No. RC1-2018-KUCARS and electrical engineering and computer science departmnt at KU.The last author of this work, Nasir Rajpoot (NR), is supported by the UK Medical Research Council (No. MR/P015476/1). NR is also supported by the PathLAKE digital pathology consortium, which is funded from the Data to Early Diagnosis and Precision Medicine strand of the government's Industrial Strategy Challenge Fund, managed and delivered by UK Research and Innovation (UKRI).

References

Adiga, U., Malladi, R., Fernandez-Gonzalez, R., de Solorzano, C.O., 2006. High-throughput analysis of multispectral images of breast cancer tissue. IEEE T-IP 15 (8), 2259–2268.

Ahmad, A., Asif, A., Rajpoot, N., Arif, M., et al., 2018. Correlation filters for detection of cellular nuclei in histopathology images. J of M. S. 42 (1), 7.

Al-Kofahi, Y., Lassoued, W., Lee, W., Roysam, B., 2009. Improved automatic detection and segmentation of cell nuclei in histopathology images. IEEE T-BME 57 (4), 841–852.

Andrion, A., Magnani, C., Betta, P., Donna, A., Mollo, F., Scelsi, M., Bernardi, P., Botta, M., Terracini, B., 1995. Malignant mesothelioma of the pleura: interobserver variability.. J. of Cli. Path. 48 (9), 856–860.

Bolme, D.S., Beveridge, J.R., Draper, B.A., Lui, Y.M., 2010. Visual object tracking using adaptive correlation filters. IEEE CVPR.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. and Trends in ML 3 (1), 1–122.

Bui, M.M., Asa, S.L., Pantanowitz, L., Parwani, A., van der Laak, J., Ung, C., Balis, U., Isaacs, M., Glassy, E., Manning, L., 2019. Digital and computational pathology: bring the future into focus. J. of Path. Info. 10.

Byun, J., Verardo, M.R., Sumengen, B., Lewis, G.P., Manjunath, B., Fisher, S.K., 2006. Automated tool for the detection of cell nuclei in digital microscopic images: Application to retinal images.

Cheng, J., Rajapakse, J.C., et al., 2008. Segmentation of clustered nuclei with shape markers and marking function. IEEE T-BME 56 (3), 741–748.

Cireşan, D.C., Giusti, A., Gambardella, L.M., Schmidhuber, J., 2013. Mitosis detection in breast cancer histology images with deep neural networks. In: MICCAI, pp. 411–418.

Cosatto, E., Miller, M., Graf, H.P., Meyer, J.S., 2008. Grading nuclear pleomorphism on histological micrographs. IEEE ICPR.

Cruz-Roa, A.A., Ovalle, J.E.A., Madabhushi, A., Osorio, F.A.G., 2013. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. MICCAI.

Danelljan, M., Hager, G., Khan, F.S., Felsberg, M., 2015. Learning spatially regularized correlation filters for visual tracking. IEEE ICCV.

Demir, C., Yener, B.,. Automated cancer diagnosis based on histopathological images: a systematic survey.

Dunne, B., Going, J., 2001. Scoring nuclear pleomorphism in breast cancer. Histopathology 39 (3), 259–265.
Fiaz, M., Mahmood, A., Javed, S., Jung, S.K., 2019. Handcrafted and deep trackers: recent visual object tracking approaches and trends. ACM CSUR 52 (2), 43.
Gamper, J., Koohbanani, N.A., Graham, S., Jahanifar, M., Khurram, S.A., Azam, A., Hewitt, K., Rajpoot, N., 2020. Pannuke dataset extension, insights and baselines. arXiv preprint arXiv:2003.10778.
Giraldo, J.H., Javed, S., Bouwmans, T., 2020. Graph moving object segmentation. IEEE T-PAMI. 1–1
Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N., 2019. Hover-net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. MIA 58, 101563.
Grau, V., Mewes, A., Alcaniz, M., Kikinis, R., Warfield, S.K., 2004. Improved watershed transform for medical image segmentation using prior information. IEEE T-MI 23 (4), 447–458.
Gurcan, M.N., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., Yener, B., 2009. Histopathological image analysis: a review. IEEE Rev. in Bio. Eng. 2, 147.
Hamed, G., Sim, T., Lucey, S., 2013. Multi-channel correlation filters. IEEE ICCV.
He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. IEEE CVPR.
Henriques, J.F., Caseiro, R., Martins, P., Batista, J., 2014. High-speed tracking with kernelized correlation filters. IEEE Trans. Pattern Anal. Mach. Intell. 37 (3), 583–596.
Höfener, H., Homeyer, A., Weiss, N., Molin, J., Lundström, C.F., Hahn, H.K., 2018. Deep learning nuclei detection: a simple approach can deliver state-of-the-art results. Comp. Med. Imag. and Grap. 70, 43–52.
Hou, L., Nguyen, V., Kanevsky, A.B., Samaras, D., Kurc, T.M., Zhao, T., Gupta, R.R., Gao, Y., Chen, W., Foran, D., et al., 2019. Sparse autoencoder for unsupervised nucleus detection and representation in histopathology images. PR 86, 188–200.
Hu, L.S., Baxter, L., Smith, K., Feuerstein, B., Karis, J., Eschbacher, J., Coons, S., Nakaji, P., Yeh, R., Debbins, J., et al., 2009. Relative cerebral blood volume values to differentiate high-grade glioma recurrence from posttreatment radiation effect: direct correlation between image-guided tissue histopathology and localized dynamic susceptibility-weighted contrast-enhanced perfusion mr imaging measurements. A. J. of Nuer. Rad. 30 (3), 552–558.
Irshad, H., Veillard, A., Roux, L., Racoceanu, D., 2013. Methods for nuclei detection, segmentation, and classification in digital histopathology: a reviewcurrent status and future potential. IEEE Rev. in Bio. Eng. 7, 97–114.
Javed, S., Fraz, M.M., Epstein, D., Snead, D., Rajpoot, N.M., 2018. Cellular Community Detection for Tissue Phenotyping in Histology Images. In: Computational Pathology and Ophthalmic Medical Image Analysis, pp. 120–129.
Javed, S., Mahmood, A., Al-Maadeed, S., Bouwmans, T., Jung, S.K., 2018. Moving object detection in complex scene using spatiotemporal structured-sparse rpca. IEEE T-IP 28 (2), 1007–1022.
Javed, S., Mahmood, A., Bouwmans, T., Jung, S.K., 2016. Spatiotemporal low-rank modeling for complex scene background initialization. IEEE T-CSVT 28 (6), 1315–1329.
Javed, S., Mahmood, A., Bouwmans, T., Jung, S.K., 2017. Background–foreground modeling based on spatiotemporal sparse subspace clustering. IEEE T-IP 26 (12), 5840–5854.
Javed, S., Mahmood, A., Dias, J., Werghi, N., 2020. Cs-rpca: Clustered sparse rpca for moving object detection. In: IEEE ICIP, pp. 3209–3213.
Javed, S., Mahmood, A., Dias, J., Werghi, N., 2020. Robust structural low-rank tracking. IEEE T-IP 29, 4390–4405.
Javed, S., Mahmood, A., Fraz, M.M., Koohbanani, N.A., Benes, K., Tsang, Y.-W., Hewitt, K., Epstein, D., Snead, D., Rajpoot, N., 2020. Cellular community detection for tissue phenotyping in colorectal cancer histology images. MIA 63, 101696.
Javed, S., Mahmood, A., Werghi, N., Benes, K., Rajpoot, N., 2020. Multiplex cellular communities in multi-gigapixel colorectal cancer histology images for tissue phenotyping. IEEE T-IP 29, 9204–9219.
Javed, S., Mahmood, A., Werghi, N., Rajpoot, N., 2019. Deep multiresolution cellular communities for semantic segmentation of multi-gigapixel histology images. ICCV-W.
Jung, C., Kim, C., 2010. Segmenting clustered nuclei using h-minima transform-based marker extraction and contour parameterization. IEEE T-BME 57 (10), 2600–2604.
Kather, J.N., Krisam, J., Charoentong, P., Luedde, T., Herpel, E., Weis, C.-A., Gaiser, T., Marx, A., Valous, N.A., Ferber, D., et al., 2019. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. PLoS Med. 16 (1).
Koohababni, N.A., Jahanifar, M., Gooya, A., Rajpoot, N., 2018. Nuclei detection using mixture density networks. International Workshop on Machine Learning in Medical Imaging.
Kuse, M., Wang, Y.-F., Kalasannavar, V., Khan, M., Rajpoot, N., 2011. Local isotropic phase symmetry measure for detection of beta cells and lymphocytes. JPI 2.
Li, F., Zhou, X., Ma, J., Wong, S.T., 2009. Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. IEEE T-MI 29 (1), 96–105.
Lindeberg, T., 1998. Feature detection with automatic scale selection. IJCV 30 (2), 79–116.
López, C., Lejeune, M., Bosch, R., Korzynska, A., García-Rojo, M., Salvadó, M.-T., Álvaro, T., Callau, C., Roso, A., Jaén, J., 2012. Digital image analysis in breast cancer: an example of an automated methodology and the effects of image compression.. Stud. in H. Tech. and Inf. 179, 155–171.
Lucchinetti, C.F., Popescu, B.F., Bunyan, R.F., Moll, N.M., Roemer, S.F., Lassmann, H., Brück, W., Parisi, J.E., Scheithauer, B.W., Giannini, C., et al., 2011. Inflammatory cortical demyelination in early multiple sclerosis. N. E. J. of Med. 365 (23), 2188–2197.
Ma, C., Huang, J.-B., Yang, X., Yang, M.-H., 2015. Hierarchical convolutional features for visual tracking. IEEE ICCV.
Mao, K.Z., Zhao, P., Tan, P.-H., 2006. Supervised learning-based cell image segmentation for p53 immunohistochemistry. IEEE T-BME 53 (6), 1153–1163.
Matas, J., Chum, O., Urban, M., Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. Im. A. Vis. Comp. 22 (10), 761–767.
Muja, M., Lowe, G.D., 2014. Scalable nearest neighbor algorithms for high dimensional data. IEEE T-PAMI 36 (11), 2227–2240.
Naylor, P., Laé, M., Reyal, F., Walter, T., 2018. Segmentation of nuclei in histopathology images by deep regression of the distance map. IEEE T-MI 38 (2), 448–459.
Park, C., Huang, J.Z., Ji, J.X., Ding, Y., 2012. Segmentation, inference and classification of partially overlapping nanoparticles. IEEE T-PAMI 35 (3). 1–1
Parvin, B., Yang, Q., Han, J., Chang, H., Rydberg, B., Barcellos-Hoff, M.H., 2007. Iterative voting for inference of structural saliency and characterization of subcellular events. IEEE T-IP 16 (3), 615–623.
Petersen, K.B., Pedersen, M.S., et al., 2008. The matrix cookbook, vol. 7. TUD 15.
Pinckaers, H., Litjens, G., 2019. Neural ordinary differential equations for semantic segmentation of individual colon glands. arXiv preprint arXiv:1910.10470.
Qi, X., Xing, F., Foran, D.J., Yang, L., 2011. Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. IEEE T-BME 59 (3), 754–765.
Raza, S.E.A., Cheung, L., Shaban, M., Graham, S., Epstein, D., Pelengaris, S., Khan, M., Rajpoot, N.M., 2019. Micro-net: a unified model for segmentation of various objects in microscopy images. MIA 52, 160–173.
Ruifrok, A.C., Johnston, D.A., et al., 2001. Quantification of histochemical staining by color deconvolution. Ana. and Quan. Cyt. and His. 23 (4), 291–299.
Sharif Razavian, A., Azizpour, H., Sullivan, J., Carlsson, S., 2014. Cnn features off-the-shelf: an astounding baseline for recognition. IEEE CVPRW.
Shi, X., Sapkota, M., Xing, F., Liu, F., Cui, L., Yang, L., 2018. Pairwise based deep ranking hashing for histopathology image classification and retrieval. PR 81, 14–22.
Sirinukunwattana, K., e Ahmed Raza, S., Tsang, Y.-W., Snead, D.R., Cree, I.A., Rajpoot, N.M., 2016. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images.. IEEE T-MI 35 (5), 1196–1206.
Sommer, C., Fiaschi, L., Hamprecht, F.A., Gerlich, D.W., 2012. Learning-based mitotic cell detection in histopathological images. In: ICPR, pp. 2306–2309.
Su, H., Xing, F., Lee, J., Peterson, C., Yang, L., 2014. Automatic myonuclear detection in isolated single muscle fibers using robust ellipse fitting and sparse representation. IEEE/ACM T-CBB 11 (4), 714–726.
Su, H., Xing, F., Yang, L., 2016. Robust cell detection of histopathological brain tumor images using sparse reconstruction and adaptive dictionary selection. IEEE T-MI 35 (6), 1575–1586.
Tofighi, M., Guo, T., Vanamala, J.K., Monga, V., 2019. Prior information guided regularized deep learning for cell nucleus detection. IEEE T-MI 38 (9), 2047–2058.
Tofighi, M., Guo, T., Vanamala, J.K.P., Monga, V., 2018. Deep networks with shape priors for nucleus detection. IEEE ICIP.
Veta, M., Pluim, J.P., Van Diest, P.J., Viergever, M.A., 2014. Breast cancer histopathology image analysis: areview. IEEE Rev. in Bio. Eng. 61 (5), 1400–1411.
Wang, N., Zhou, W., Tian, Q., Hong, R., Wang, M., Li, H., 2018. Multi-cue correlation filters for robust visual tracking. IEEE CVPR.
Xie, C., Vanderbilt, C.M., Grabenstetter, A., Fuchs, T.J., 2018. Voca: Cell nuclei detection in histopathology images by vector oriented confidence accumulation.
Xie, Y., Xing, F., Kong, X., Su, H., Yang, L., 2015. Beyond classification: structured regression for robust cell detection using convolutional neural network. MICCAI.
Xing, F., Su, H., Yang, L., 2013. An integrated framework for automatic ki-67 scoring in pancreatic neuroendocrine tumor. MICCAI.
Xing, F., Xie, Y., Shi, X., Chen, P., Zhang, Z., Yang, L., 2019. Towards pixel-to-pixel deep nucleus detection in microscopy images. BMC Bio. 20 (1), 1–16.
Xing, F., Xie, Y., Yang, L., 2015. An automatic learning-based framework for robust nucleus segmentation. IEEE T-MI 35 (2), 550–566.
Xing, F., Xie, Y., Yang, L., 2016. An automatic learning-based framework for robust nucleus segmentation. IEEE T-MI 35 (2), 550–566.
Xing, F., Yang, L., 2016. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. IEEE Rev. in Bio. Eng. 9, 234–263.
Xu, J., Xiang, L., Hang, R., Wu, J., 2014. Stacked sparse autoencoder (ssae) based framework for nuclei patch classification on breast cancer histopathology. IEEE ISBI.
Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., Madabhushi, A., 2015. Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. IEEE T-MI 35 (1), 119–130.
Xue, Y., Ray, N., 2017. Cell detection with deep convolutional neural network and compressed sensing. arXiv preprint arXiv:1708.03307.
Yan, P., Zhou, X., Shah, M., Wong, S.T., 2008. Automatic segmentation of high-throughput rnai fluorescent cellular images. IEEE T-ITB 12 (1), 109–117.
Yang, X., Li, H., Zhou, X., 2006. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy. IEEE T-CS-I 53 (11), 2405–2414.
Yin, M., Gao, J., Lin, Z., 2015. Laplacian regularized low-rank representation and its applications. IEEE T-PAMI 38 (3), 504–517.
Yuan, Y., Failmezger, H., Rueda, O.M., Ali, H.R., Gräf, S., Chin, S.-F., Schwarz, R.F., Curtis, C., Dunning, M.J., Bardwell, H., et al., 2012. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. STM 4 (157). 157ra143–157ra143