We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Fast Regions-of-Interest Detection in Whole Slide Histopathology Images

Junzhou Huang and Ruoyu Li

Abstract

Detecting and localizing pathological region of interest (ROI) over whole slide pathological image (WSI) is a challenging problem. To reduce computational complexity, we introduced a two-stage superpixel-based ROI detection approach. To efficiently construct superpixels with fine details preserved, we utilized a novel superpixel clustering algorithm which cluster blocks of pixel in a hierarchical fashion. The major reduction of complexity is attributed to the combination of boundary update and coarse-to-fine refinement in superpixel clustering. The former maintains the accuracy of segmentation, meanwhile, avoids most of unnecessary revisit to the 'non-boundary' pixels. The latter reduces the complexity by faster localizing those boundary blocks. Detector of RoI was trained using handcrafted features extracted from super-pixels of labeled WSIs. Extensive experiments indicates that the introduced superpixel clustering algorithm showed lifted accuracy on lung cancer WSI detection at much less cost, compared to other classic superpixel clustering approaches. Moreover, the clustered superpixels do not only facilitate a fast detection, also deliver a boundary-preserving segmentation of ROI in whole slide images.

Keywords: region of interest, whole slide histopathology images, superpixel, segmentation, detection, unsupervised learning

1. Introduction

At our age, many hazardous infectious diseases, e.g. bird flu, and many different kind of cancers, e.g. lung cancer, are still the top threats to our personal health and the public sanitation as well. Automatic searching and localizing Regions of Interest (ROIs) on histopathological images is a crucial intermediate step between large-scale images acquisition and the computer-aided automated diagnosis that we pursue. As the fast development of deep learning techniques and the introduction of neural network models, e.g. convolutional neural networks (CNNs), to medical image understanding area, we are finally able to extend the boundary of modern medical image saliency detection, classification and segmentation [1–3]. Whole Slide Images (WSIs) are the digitized histopathology images taken over an entire slide of tissue, which retrains as much intact pathological information as possible. Therefore, a typical WSI, that usually has resolution at scale of $10^6 \times 10^6$, is $1.5 \sim 2.0$ Gigabyte large on disk, which is thousands times larger than those images from deep learning benchmark datasets, like MNIST [4] and CIFAR [5].

Therefore, traditional fully convolutional networks, used to work perfectly for medical image segmentation [6], are no longer applicable, because of the parameter scale that may explode and the rising risk of under-fitting along with lack of labeled WSIs for training. We need a brand-new cost-efficient solution designed especially for WSIs to handle such magnificent scale of data without losing too much performance. As far as we know, there are no existing convolutional neural networks who claim themselves to directly work on raw images at WSI scale without any downsampling or patching. The most popular walk-around for extracting features from WSIs is to first sample a bag of patches over WSIs and then train and execute inference on patches respectively. Then, aggregating the prediction from patch level to WSI level is to give final model output. Patch-based network [2] successfully handled classification task on WSIs, [7] enabled survival time inference purely based on tumor tissue WSIs. Although, these models applied to WSIs successfully saved most of computational cost by patching, they also dumped lots of task-relevant information hidden in those patches not being sampled. Besides, losing topological spatial information of patches after being sampled from WSI makes predictor treat patches equally, which is obviously not the optimal strategy.

Considering the practical clinic scenarios for image detection and segmentation techniques applied to CT [6] and MRI [8] and the associated pathophysiological procedures, we summarized some challenging but necessary technical requirements for any ROI detection and segmentation solutions for WSIs:

1. High time and energy efficiency. To make it scalable, the ROI detection and localization is supposed to be accomplished within short period of time with high recall and acceptable precision.

2. High fidelity and high trustworthiness on generated ROIs of WSI. We need to quickly and correctly classify if a proposed ROI belongs to, at least partly, ground truth ROIs. Because the ROI prediction may largely affect downstream tasks, e.g. disease diagnosis decisions.

Regions of interest (ROI) could have different definition according to particular scenarios. In this article, we name ROIs as the local regions filled with tumor cell cluster or other cancer-related cells such as lymphocyte. In past related works, ROI detection and segmentation are usually treated separately as two different tasks. The former is to quickly search and localize any suspicious regions on image according to predefined patterns. The output of this task may not have to be fine-detailed at pixel level, due to computational efficiency concern, and sometimes a bounding box that surrounds, at least partly, the ground-truth ROI is enough satisfactory. While, the latter task is to give a pixel-accurate contour of each detected ROIs, which is significantly more expensive. In fact, detection and segmentation are not strictly isolated, and on the opposite, the two tasks could be combined as one under some circumstances. Many CNNs based image segmentation models are indeed end-to-end solutions directly extract and learn hierarchical feature pyramid from raw channels of images to execute pixel-clustering at different level of granularity. Semantic segmentation network [9] is to obtain object detection and segmentation in single forward-pass of network. The advantages of applying deep neural network is from treating the feature design work as an optimization problem, and therefore CNNs are able to discover hidden representations that better serve prediction tasks than handcrafted descriptors, who are either over-localized or not robust. The requirement on high recall rules out patch-based WSI solutions. And patch-based methods obviously cannot handle segmentation of entire WSIs. However, in order to directly work on WSI input, the networks either

make the receptive field of convolutional operators large enough to cover any potential region of interest, or stack more layers with relatively small kernel to aggregate local features from entire ROI to form its high-level representations for classification. No matter what architecture is chosen, the total count of parameter in WSI segmentation network is going to be magnificent. Due to the expense of having quality annotation of all ROIs (i.e. tumor cells) on WSIs, the annotated ground-truths of segmentation for training models is quite constrained and very likely not enough to train a wide and deep network as described above that directly works on raw WSIs.

To work around this difficulty, in the method to be introduced, we first chose to still rely on handcrafted features as descriptors of patches to save massive feature aggregation calculations in CNNs, and in the meanwhile we also utilized the hierarchical pyramid structures appear between feature maps of consecutive convolutional layers of CNNs. While, different from what happened in CNNs, in the pyramid of introduced multi-level iterative method, feature vectors of descriptor are not changed along with level, because we did not have gradients back-propagated from loss to update feature formations, while the spatial segmentation did get updated at different granularity of patching. Without having ground-truth of segmentation of ROIs, we introduced superpixel clustering as an unsupervised way to learn spatial segmentation of image, since we do not have gradient to update the assignment of segmentations as well. At different level granularity, we divide the entire WSI into patches of different scale, then the introduced superpixel clustering method [10] is going to cluster patches based on several handcrafted local textual descriptors, preserving both topological consistency and appearance similarity. After superpixel constructed, we run a pre-trained classifier, e.g. SVM or CRF, to classify superpixels represented by the averaged descriptors of patches. Averaging of patch descriptors is to avoid additional difficulty of training a classifier for superpixels of different size and varying shape. This is also the biggest challenge for building an end-to-end fully convolutional network fed with clustered superpixels, since the shape of input tensor to any neural network cannot be undefined.

The main contributions of article is to decouple and reformulate ROI detection and semantic segmentation, that requires dense annotation, into an iterative execution of unsupervised superpixel clustering and classification at coarse-to-fine level of patching granularity. This semi-supervised approach largely replies on quality of superpixel clustering. To obtain better fine-detailed superpixels, we introduced a novel topology-preserved superpixel clustering algorithm to this problem. Besides, the approach introduced is also dependent on accurate classification of superpixels, especially at coarser levels, because any mistaken classification of coarse superpixel cannot be compensated in fine-grained superpixel refinement at next level of granularity. The recall of ROIs will benefit from the increased classification accuracy. Therefore, we trained compact but robust classifier, e.g. SVM, with minimal data requirement. On the other hand, without fine-tune, an improved segmentation of superpixels will automatically boost accuracy of a pre-trained classifier.

2. Related work

Superpixel is a common replacement of pixel with purposes more than saving computational cost. It clusters nearby pixels of similar attributes together as fundamental operational unit in downstream tasks, e.g. object detection, segmentation and even real-time tracking. In this session we introduced the state-of-the-art superpixel clustering algorithms and the combination of superpixel with deep neural networks (DNNs) in medical image understanding.

2.1 Superpixel clustering

One important feature of superpixel construction is that this is a pure unsupervised approach in which there are no annotated ground-truths in any format for guiding the label assignment on pixels. The pixels are clustered purely based on the attributes, such as appearance and physical location, etc. SLIC [11] is an iterative K-mean superpixel clustering that walk through all pixels. It is able to generate almost equally-sized superpixels with outstanding boundary adherence. And the time complexity could be further reduced by limiting search space to a small nearby area. While, iterating over entire pixels is still too expensive, stopping SLIC from being applied on large images like WSIs. If compromise part of accuracy, SEEDS [12], that started from randomly initialized superpixel partitions, focused on updating boundary pixel allocation only and proposed a fast energy function to evaluate each adaption of pixel label assignment by enforcing color homogeneity. Linear spectral clustering, a.k.a. LSC, combined normalized cut and K-mean clustering after discovering optimizing these two objective functions are in fact equivalent on the condition that defines similar function as inner product of feature vectors [13]. LSC also achieved satisfactory boundary adherence and color consistency within segmented superpixels with $\mathcal{O}(N)$ complexity, where N is the pixel number. Compared to SLIC, LSC saved computations from pre-allocation of pixel to large regions by eigenvector-based normalized cuts. And different from the two-stage Ncuts [14], LSC accomplished Ncuts and K-mean in one-stage. Similar to LSC, the computational complexity of SEEDS and SLIC is also approximated as $\mathcal{O}(N)$. Therefore, within visual comparison in **Figure 1**, we did not include expensive solutions such as ERS [15] with $\mathcal{O}(N^2 \log N)$ and EneOpt0 [16] with $\mathcal{O}(N^3)$ complexity. Because we only consider those approaches who are potentially feasible for segmenting whole slide images.
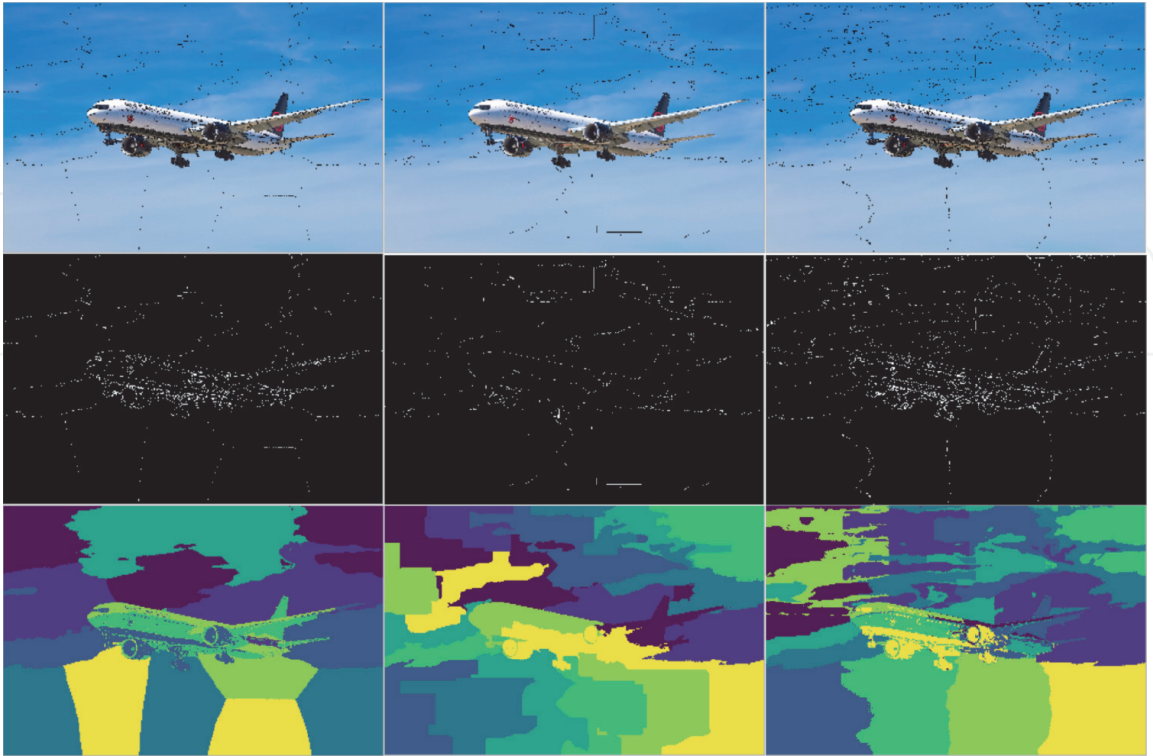

Figure 1.
Example of superpixel clustering on image with three classic solutions: LSC [13] (left), SEEDS [12] (middle) and SLIC [11] (right). The upper row is the edges of superpixels displayed on image. The middle row is the contours of superpixels. The bottom row is the segmentation mask filled with different color on different superpixel.

2.2 ROI and superpixel

Regions of interest (ROI) in histopathology whole slide images (WSIs) are usually those disease-related cells or the tissues of specific patterns, but they do not have descriptive definitions to form a category of objects. Due to the magnificent scale of WSI, the major challenge would be the scalability and the memory efficiency of algorithms. Bejnordi et al. [17] relied on cheap segmentation of superpixels on downsampled WSIs to filter out those regions irrelevant to ROIs. However, it did not correctly notice the inevitable influence of wrong classification of coarse superpixels, because the algorithm completely ruled out those regions from later more accurate segmentation and classification. Besides, the classifiers had to be trained multiple times with patches extracted from the superpixels of different magnification to work on different levels of granularity. Litjens et al. [18] reduced the workload of labeling and grading by two ways: by excluding the areas of definitely normal tissues within a single specimen or by excluding entire specimens which do not contain any tumor cells. Litjens et al. [18] presented a multi-resolution cancer detection algorithm to boost the latter. While it also suffered from the loss of recall as [17]. Another superpixel automated segmentation method is [19], which trained a classifier to predict where mitochondrial boundaries occurs using diverse cues from superpixel graph. While, because the selected superpixel clustering approach [11] did not offer satisfactory boundary adherence, the classifier encumber the overall detection performance. As summary, in order to accomplish a quick detection and segmentation of ROIs in WSI, a combination of superpixel clustering and pre-trained classifier seems a popular choice, while the performance bottleneck was the tradeoff between the efficiency and the quality of superpixel clustering, which directly determined classifier accuracy.

To reduce the intense computational cost in superpixel clustering, the algorithm to be introduced creatively combined the coarse-to-fine scheme [20] and the boundary-only update strategy proposed in SPSS [21]. In our method, clustering manipulated the rectangular blocks of pixel as basic unit and a coarse segmentation of superpixel would be constructed before a more fine-detailed refinement got executed. on each level of construction, only boundary blocks or their nearby neighbors got chance of label update. **Figure 2** illustrated the procedures of introduced superpixel clustering. Furthermore, the introduced boundary-only update strategy on next level would emphasize on differentiating foreground and background blocks, considering the boundaries between superpixels within ROIs are less important. The improvement brought by our algorithm on ROI detection accuracy has been proved and verified in [10, 22], where the method had quantitatively verify the improvement of the accuracy of ROI detection in histopathology images,
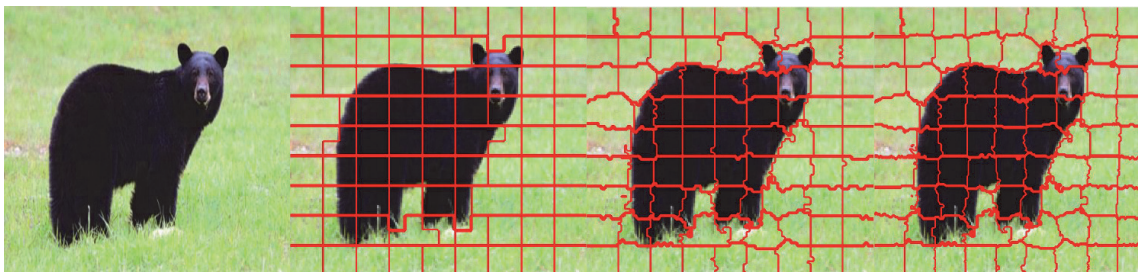

Figure 2.
An example of the coarse-to-fine/boundary-only update based superpixel segmentation algorithm first presented in [10]. The basic manipulation unit is the rectangular block instead of pixels during each stage. We start from a coarse segmentation and end with pixel-level refinement on superpixel boundary. The block size is respectively 10 × 10, 2 × 2, 1 × 1 (single pixel) from left to right.

e.g. lung cancer H&E-stained WSIs. **Figure 3** shows comparison of classic superpixel methods [11–13] on cancer patients WSI.

2.3 DNNs on superpixel

As success of deep neural networks in computer vision, many works have extended application of DNNs onto superpixel. Gadde et al. [23] introduced a bilateral inceptions module to accelerate convergence of CNNs with superpixel as network input for semantic segmentation. Kwak et al. [24] treated superpixels as "pooling" layer in neural network, but preserving low-level structures. Therefore, their framework trained semantic segmentation network without pixel-level ground-truth. To construct superpixels for small objects of complicated boundaries, [25] introduced a superpixel segmentation based on pixel features trained with affinity loss and segmentation error. In medical images domain, superpixels are also utilized as a topology-preserving simplification of data for deep network. The organ segmentation network in [26] worked on the descriptors extracted from superpixels clustered in CT images. And then CNN simply did a pixel-wise refinement based on the coarse segmentation given by superpixel. Different from previous works who simply utilized superpixels as reduction of image primitives, [27] proposed an end-to-end" Superpixel Sampling Network" (SSN) which contains differentiable superpixel construction together with learning a task-specific prediction.

The rest of article is organized as following: we first introduce the multi-resolution fast superpixel clustering with coarse-to-fine and boundary-only strategy to increase efficiency. Both mathematical explanation and illustrative examples will be given in Section 3. Then we elaborate the numerical results on classification accuracy and visual comparison of superpixels with classic methods on TCGA WSI dataset in Session 4. Lastly, conclusion and future work will be given in Session 5.
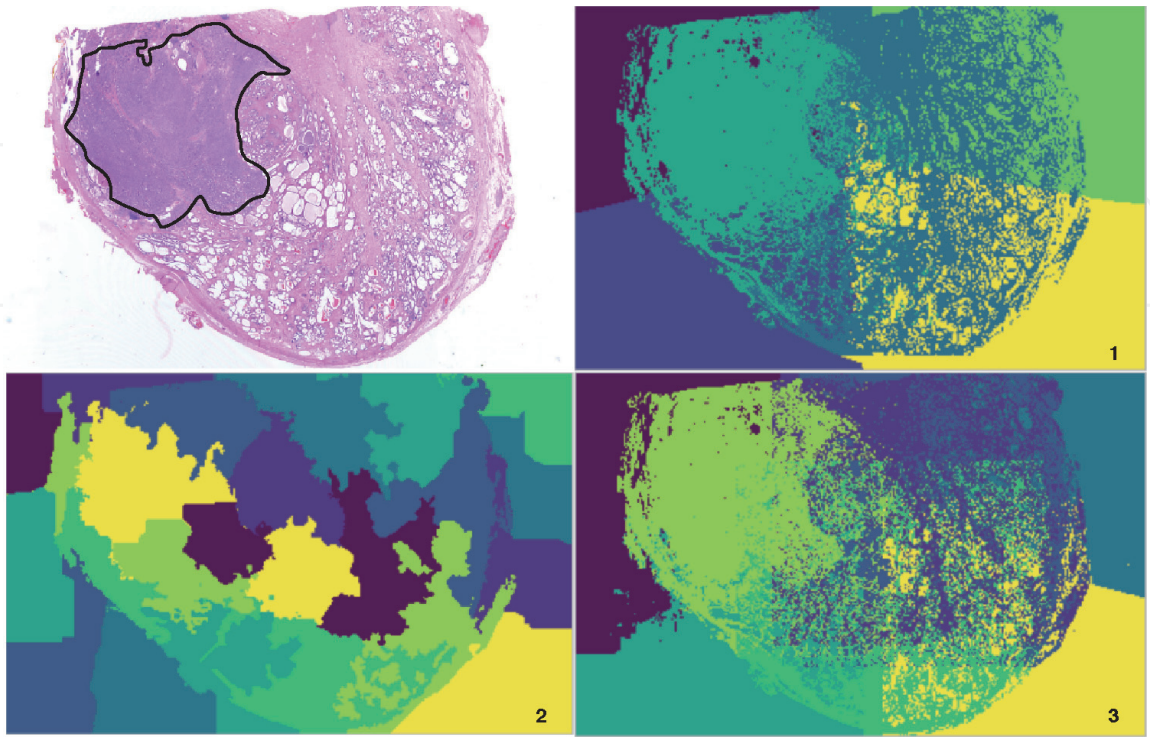

Figure 3.
Example of pathological whole slide image with ROI annotations and the superpixels generated by three classic solutions of linear complexity: (1) LSC [13], (2) SEEDS [12] and (3) SLIC [11].

3. Methodology

The detection framework introduced is not only going to propose bounding box to surround ROIs, but also is going to offer fine-detailed, boundary-adherent superpixel segmentation of them. On the other hand, an improved superpixel construction contributes the differentiation of ROI from background as well. Therefore, the proposed approach comprised two components: fine-detailed superpixel segmentation and superpixel classification. For reduction of computational expense, we chose not to accomplish superpixel segmentation at finest level in one shot. For instead, we first obtain a coarse superpixel segmentation from clustering big pixel blocks (e.g. 500 × 500). A pre-trained binary classifier then predicts label (ROI v.s. background) of superpixels. Afterwards, those superpixels labeled as ROI along with their neighbors will move to next round of segmentation at finer resolution. The process will be repeated until quality becomes satisfactory. Different from previous superpixel clustering methods [11, 21], the introduced algorithm gave topology-preserving superpixels. A better detection recall is expected as well, since our method did not completely rule out negatively labeled superpixel at coarse stage as [17, 18], and for instead we include negative neighbor superpixels to next level of segmentation.

3.1 Superpixel clustering and detection

3.1.1 Energy function

Think of superpixels of flexible number of blocks $\mathbf{S} = \{s_0, \cdots, s_{K-1}\}$, and the blocks belong to superpixel $\mathbf{S}_k$ as $\{b_0, \cdots, b_{M-1}\}$, we devised two representations of block: appearance and position. Appearance representation of block is the averaged RGB color over pixels in block as $\mathbf{C}$. Position representation of block is the relative position coordinates at center point of block as $\mathbf{P}$. At superpixel level, $\Theta = (\theta_0, \cdots, \theta_{K-1})$ and $\Xi = (\xi_0, \cdots, \xi_{K-1})$ are the center positions and the mean color vectors of superpixels. The objective function to be minimized consists of a series of energy functions and penalty terms. For appearance, total variance of three color channels are color energy function of superpixel $\mathbf{S}_k$ defined as:

$$E_{col}(\mathbf{S}_k) = \sum_{q=0}^{2} \frac{1}{\|s\|} \sum_{b \in \mathbf{S}_k} \left(c_b^q - \xi_k^q\right)^2, \tag{1}$$

also known as appearance coherence. For position, the averaged $l2$ distance from block position $\mathbf{P}_b$ to the center position of its superpixel is the position energy function, $E_{pos}(\mathbf{S}_k) = \frac{1}{\|S_k\|} \sum_{b \in \mathbf{S}_k} \|\mathbf{P}_b - \theta_k\|^2$. This is to ensure clustered blocks are geophysically close. Besides, to avoid seeing any superpixels with sophisticated boundary, we use the total boundary length as boundary penalty function. Furthermore, we constrain the minimal size of finalized superpixel to be at least 25% of initial size. If any update of block's belonging violates this constrain, we give infinity penalty to this update, therefore, the algorithm will reject such label assignment update.

$$P_{size}(\mathbf{S}_k) = \begin{cases} +inf, & size(\mathbf{S}_k) < 0.25 \times initialsize \\ 0, & otherwise. \end{cases} \tag{2}$$

Similar penalty would be applied, if the update causes any isolated blocks who are surrounded by blocks from other superpixels. This is to enforce all generated superpixels to be topologically connected.

3.1.2 Boundary-only update

To define boundary energy function, we need to define boundary block and length. If a block has any neighbor block from other superpixel, then it is a boundary block. The boundary length of block is the number of neighbor blocks that belong to other superpixel.

$$P_b(s) = \sum_{b \in \mathbf{S}_k} \sum_{b_n \in Neighbor(b)} S(\mathbf{S}_k, b_n), \qquad (3)$$

where $S(\mathbf{S}_k, b_n)$ is the indicator function of superpixel belonging for block, which return 0 if $b_n \in S_k$, otherwise 1. In our algorithm, we first stack entire initial boundary blocks into a queue, then the iterative superpixel clustering algorithm will work on boundary blocks only for consideration of updating label (i.e. superpixel assignment) of block. This is so-called 'boundary-only update'. In other words, the non-boundary blocks will not be considered for label change until they become boundary blocks. When the algorithm decides to update the label of a block, its neighbor will be considered to become new boundary blocks. When using the boundary-only update, there are two things to notice: 1) when update the label of block, it definitely change the list of boundary blocks; 2) we need to append the new boundary block to the end of the list because and follow the FIFO principle when deciding the order of blocks for consideration of changing label, in order to avoid the risk of divergence given by correlated dimensions in coordinate descent optimization. The candidate superpixel labels for a boundary block to swap are limited to its neighbor superpixels, otherwise it will trigger the topology connectivity penalty by having an isolated block. Given a trial of label update, the algorithm compares the objective function values before and after the change to see whether and how much the change is able to drive energy down.

We elaborate objective function each step of updating block-wise superpixel label assignment as below:

$$E(\mathbf{S}) = \sum_s \left(E_{col}(s) + \lambda_{pos} E_{pos}(s) + \lambda_b P_b(s) + P_{topo}(s) + P_{size}(s) \right), \qquad (4)$$

where λ_{pos}, λ_b are respectively the tradeoff coefficients for position energy function and boundary length penalty term. In practice, the regularization on superpixel size and topological connectivity will give infinite penalty on those superpixels of over-small size as $P_{size}(S_k) \approx inf$ and those of isolated blocks, i.e. $P_{topo}(S_k) \approx inf$. Therefore, the algorithm will always reject such label proposal that violates topology connectivity and size regularization. When superpixel assignment of a boundary block is updated, the algorithm will add its neighbor blocks to queue, because those non-boundary blocks are now next to other superpixels. The convergence will arrive when the queue is empty.

Algorithm 1 Multi-resolution ROI Detection (MROID).

---

superpixel number - K
for l = 1 to levelMax **do**
    if l = 1 **then**
        1. Initialize blocks **B** on level l size on entire image;
        2. Initialize K superpixels **S**; initialize Θ, Ξ
    else

---

1. Initialize blocks **B** on level l size within positive superpixels and their neighbor superpixels $\hat{\mathbf{S}}$. Initialize Θ, Ξ for $\hat{\mathbf{S}}$.

end if

Compute the mean color and position in each block;

Initialize L, the queue of boundary blocks on level l;

while length(L) $\neq 0$ **do**

    Pop out block b_i^l from the queue;

    $E_{before} = E(\mathbf{S})$;

    for $b_n \in \text{Neighbor}(b_i^l)$ **do**

        change label of b_i^l to neighbor b_n's label;

        $E_{after}(b_n) = E(\mathbf{S})$;

    end for

    find the $\hat{b}_n = arg \min_{b_n \in Neighbor(b_i^l)} E_{after}(b_n)$;

    if $E_{after}(b_n) < E_{before}$ then update label of b_i^l to that of $\hat{b}_n$.

    append $Neighbor\left(b_i^l\right)$ to L.

end while

run binary classifier on superpixels to predict ROI.

end for

3.1.3 Coarse-to-fine detection

Instead of processing WSI at different resolutions [17], we cluster superpixels at coarse-to-fine level of resolution. Yao et al. [10] adopted boundary-only update as well to save unnecessary revisit to non-boundary blocks, while the boundary blocks on WSI may still be too much for extensive iterations. To further reduce the amount of data brought to finer update with more intense computation, we utilized a pre-trained classifier, e.g. SVM, to predict whether the superpixel belongs part of ROI. For any superpixel moved to finer update, smaller blocks will be initialized within its region. For example, a 10×10 block will be divided into 25 block of size 2×2 arranged at 5×5 grid. Boundary block queue will be refilled with 2×2 blocks who sit on superpixel boundaries. The classifier was trained using features extracted from patches sampled from ROI and non-ROI regions over annotated WSIs. To deal with different cardinality of patch per superpixel, we use pooling patch features at inference time. Given that we did not downsample images, therefore, the classifier trained on raw WSIs is able to be reused with different level of superpixel. See **Figure 4** as illustration.

3.2 Complexity analysis

Pixel-wise superpixel constructions [11, 12] have $\mathcal{O}(N)$ complexity, where N is number of pixel, while it made them infeasible on WSIs of trillions of pixels. The introduced algorithm is able to reduce the complexity to scale of number of block i.e. $\mathcal{O}\left(\sum_{k=0}^{K-1} \|\mathcal{S}_k\|\right) \ll \mathcal{O}(N)$. The boundary-only update, first presented in [10], further constrains involved blocks to those boundary blocks. Considering the purpose of clustered superpixel, our algorithm combined detection and superpixel clustering together, and it only executes finer segmentation within those coarse superpixels who were classified as ROI. It saved the calculations wasted on updating the superpixels that do not contribute to ROI detection. Due to the reduced dimensionality, the convergence comes faster than pixel-wise clustering methods.
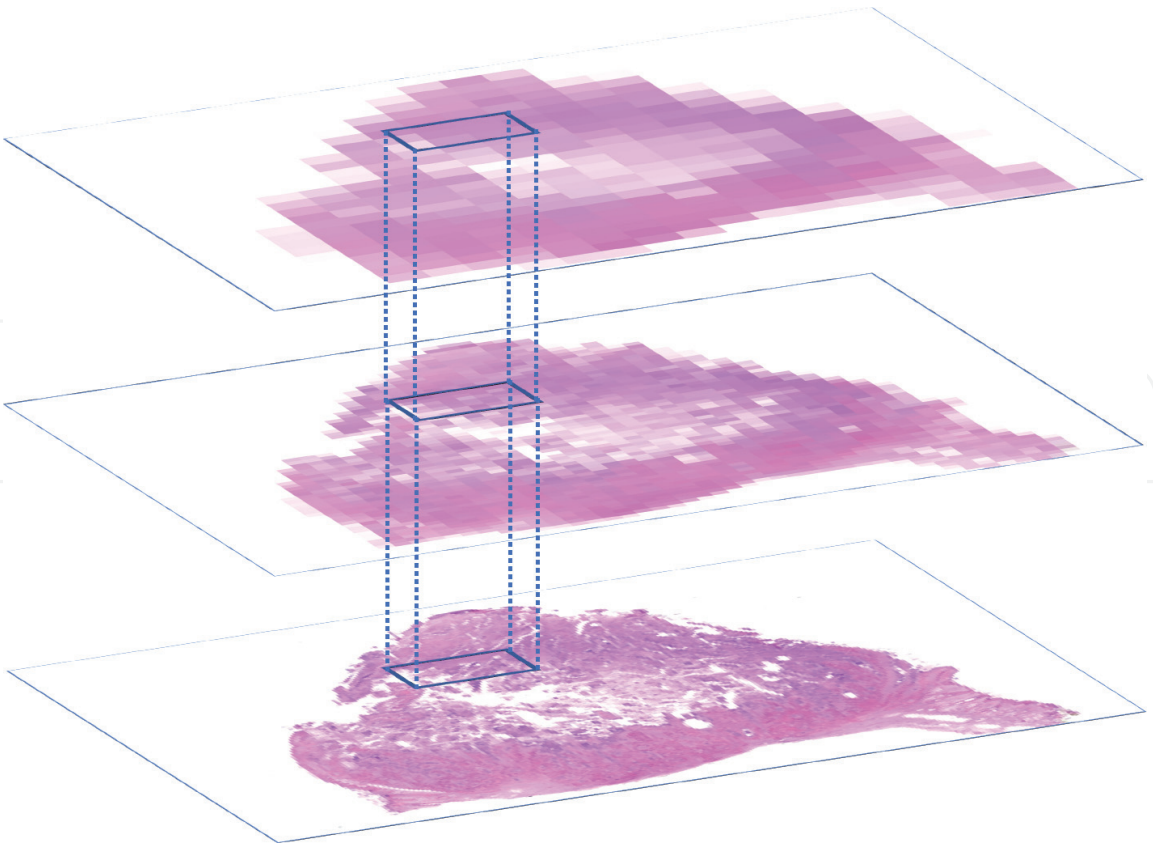
Figure 4.
An illustration of multi-resolution process of ROI detection on WSI. The example has 3 level of granularity in term of block size. Note that we did not downsample the WSI directly, which dump falsely many local details, and we still include neighbor superpixels close to positive ones at coarse classification to next level. If the bounding box is the ROI (a rough identifier), as resolution goes high, superpixels cover and surround the bounding box will get fine-detailed update.

4. Experiments

4.1 ROIs in lung cancer histopathology WSI

In histopathology images like lung cancer WSIs, the regions of interest are those areas consist of cancer cells or other tissues that may be related to tumor diagnosis. A fast detection approach of ROIs is to search and localize those regions on image at WSI scale, that usually have trillions of pixels. Traditional pixel-wise methods and neural network cannot directly work on WSI, due to the extraordinary data scale and image dimensionality. Downsampling of WSI reduces complexity but also loses local fine-detailed features. Superpixels first cluster those pixels of similar spacial, color and topological properties as whole, and then in downstream tasks e.g. detection and segmentation, the superpixels will act as minimal manipulatable unit, reducing image primitives and complexity. If superpixels were well constructed, the downstream will not be affected by the sparse representation of image. The tumor cells of lung cancer patients (not only for lung cancer, but also generally appear in other subtypes of cancer) infest as cell mass. If treat the regions where tumor cell mass appears as ROIs, we can easily see that the H&E stained histopathology images that those tumor cells are more deeply colored due to the massive reproduction of genetic materials inside tumor nuclei (See **Figure 5**).

4.2 Experimental setup

In the experimental stage, a random forest and a support-vector-machine (a.k.a. SVM) classifier were trained with local features extracted from regions defined by
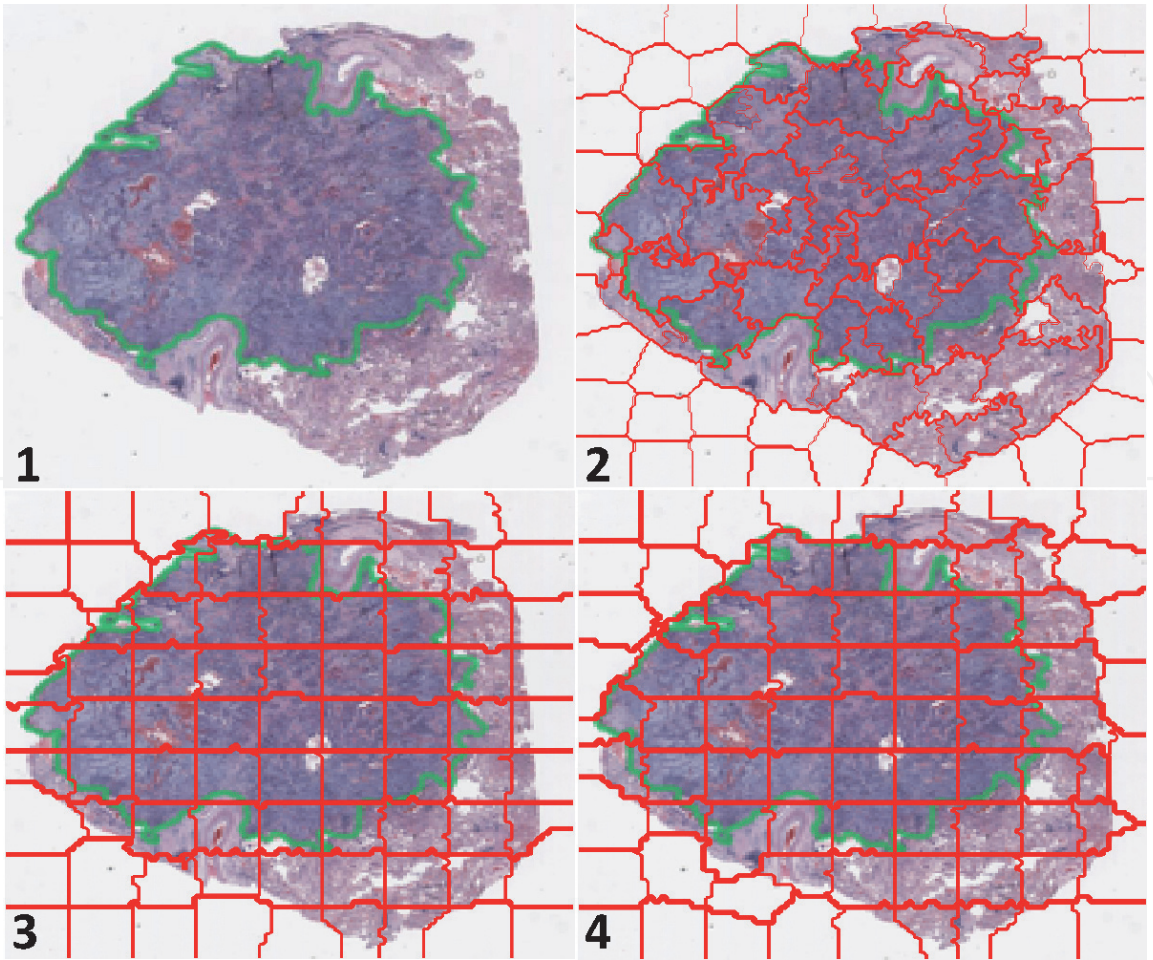
Figure 5.
The comparison of several superpixel clustering on lung cancer H&E stained WSI: 1) the origin (with ROI annotated), 2) SLIC [11], 3) SPSS [21], 4) our method. The ROI is contoured by green line.

the superpixels given by Algorithm 1. The total 384 dimensional features include local binary patterns and statistics derived from the histogram of the three-channel HSD color model as well as common texture features, e.g. color SIFT. The introduced method was compared against the superpixels generated by SLIC [11] and tetragonum (i.e. rectangular patches). The experiments used the adenocarcinoma and squamous cell carcinoma lung cancer WSIs from the NLST (National Lung Screening Trial) Data Portal1[1]. In superpixel classification, we executed feature extraction on the sampled patches (100×100) with 10% overlap with each other within each superpixel, we rule out patches sit across boundary avoiding noise. Lastly, we averaged the feature vectors of patches as representation of superpixel. When deciding ROI belonging for superpixel, if any part of ground-truth ROI fall into a superpixel, it will count as positive. The setup is rooted at the extremely high recall requirement for medical diagnosis. Given this setup, for better detection precision, superpixels should be better boundary adherent and clearly separated from background.

4.3 Experimental results

Due to the overwhelming fidelity of superpixels given by our algorithm, the classifier operated over the regions segmented by superpixels is able to deliver better classification results (See **Table 1**). Since the feature descriptors were built

————————

[1] https://biometry. nci.nih.gov/cdas/studies/nlst/

Classifier	Metric	MROID	SLIC [11]	Tetragonum
Random Forest	Error rate	**0.1326**	0.1933	0.2047
	Precision	**0.7127**	0.6835	0.6740
	Recall	**0.7333**	0.6108	0.6450
SVM	Error rate	**0.3011**	0.3343	0.3061
	Precision	**0.6754**	0.6672	0.6723
	Recall	**0.7450**	0.6604	0.6972

Table 1.
The table presents the comparison results of the proposed solution, MROID (numbers in bold), SLIC and tetragonum (non-superpixel) in term of classification statistics including: the rate of error classification, precision and recall. Tetragonum: Sliding rectangular windows.

on the patches segmented by contours of superpixels, the better the superpixel adhere to the boundaries, the better differentiability the features have for superpixel classification.

Figure 6 demonstrated the introduced multi-resolution coarse-to-fine superpixel segmentation in a lung cancer histopathology images. The algorithm first manipulated large block (180×180) to cluster superpixels, then move to finer segmentation with 10×10 blocks on the superpixels selected by the classifier. The recursive refinement continues until the block queue run out, which means energy loss converges. In **Table 1**, we compared the classification recall and precision using superpixels given by SLIC and our method as well as simply patches without any preprocessing like superpixel clustering. Our results showed that, compared to simple patching, utilizing superpixel may not always increase ROI recall but definitely lift precision. Compared to superpixel given by SLIC with sophisticated boundary, out method outperformed on both recall and precision. We also observed that, if superpixels do not adhere to boundary, a detection based on classification of superpixels of low segmentation accuracy leads to worse accuracy than a trivial patch based method. While, our method delivered best results at both recall and precision.
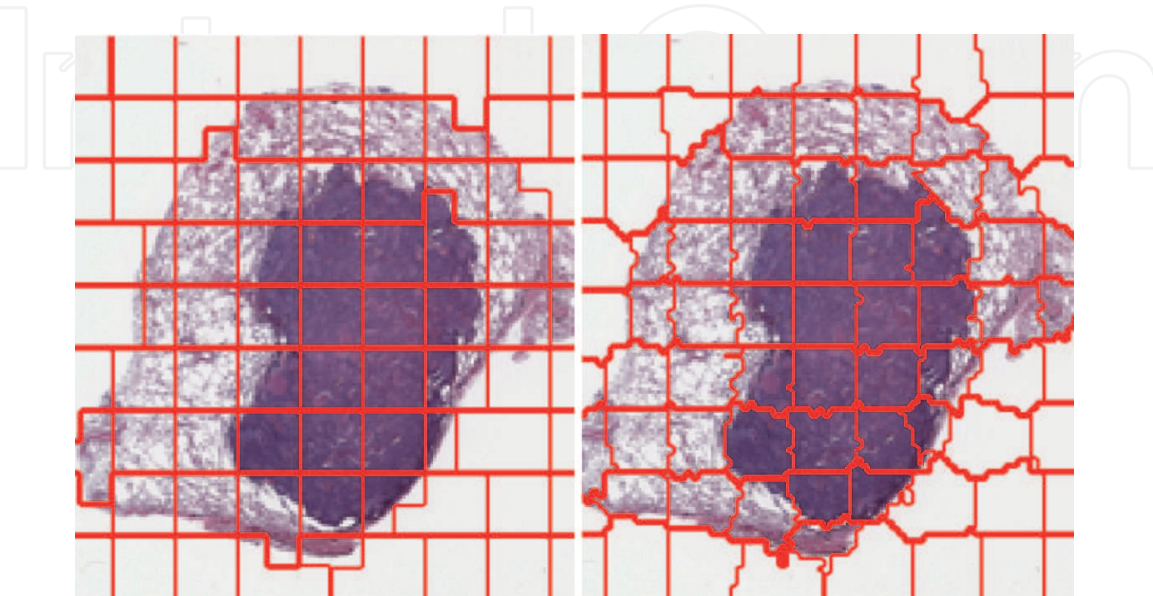

Figure 6.
A coarse-to-fine superpixel clustering on a lung cancer WSI from NLST. 1) coarse segmentation of superpixels using large blocks (180×180); 2) refined segmentation with small blocks within selected superpixels.

5. Conclusion

In the chapter, we presented a novel local feature based solution to fast search and detection of regions of interest (ROI) in whole slide lung cancer histopathology image. For superpixel clustering, we introduced coarse-to-fine multi-resolution segmentation of superpixel by manipulating blocks of different size. Besides, boundary-only update strategy also reduced the computational complexity to the scale of superpixel boundary length, irrelevant of image size.

We creatively embedded the ROI classification into superpixel clustering algorithm. Iteratively executing superpixel construction and ROI detection. A better superpixel will accelerate detection and lift accuracy, while on the other hand, a better classification of ROI on coarse superpixel guides superpixel segmentation at finer level. Our algorithm performed a faster and finer ROI detection and segmentation. The effectiveness and efficiency of our algorithm has been verified on large histopathology WSI database, e.g. NLST.

In future, as the development of neural network capable of flexible input size [28, 29], it is likely to merge superpixel construction and downstream tasks, e.g. semantic segmentation, classification together in neural network architecture, in which superpixels are clustered using hidden features, while superpixels boost feature learning as well.

Author details

Junzhou Huang* and Ruoyu Li
1 500 UTA Boulevard, The University of Texas at Arlington, Arlington, TX,
United States

*Address all correspondence to: jzhuang@uta.edu

IntechOpen

References

[1] Takács P, Manno-Kovacs A. MRI Brain Tumor Segmentation Combining Saliency and Convolutional Network Features. In2018 International Conference on Content-Based Multimedia Indexing (CBMI) 2018 Sep 4 (pp. 1–6). IEEE.

[2] Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based convolutional neural network for whole slide tissue image classification. InProceedings of the ieee conference on computer vision and pattern recognition 2016 (pp. 2424–2433).

[3] Bándi P, van de Loo R, Intezar M, Geijs D, Ciompi F, van Ginneken B, van der Laak J, Litjens G. Comparison of different methods for tissue segmentation in histopathological whole-slide images. In2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) 2017 Apr 18 (pp. 591–595). IEEE.

[4] LeCun Y, Cortes C, Burges CJ. MNIST handwritten digit database.

[5] Krizhevsky A, Hinton G. Convolutional deep belief networks on cifar-10. Unpublished manuscript. 2010 Aug;40(7):1–9.

[6] Zhou Z, Siddiquee MM, Tajbakhsh N, Liang J. Unet++: A nested u-net architecture for medical image segmentation. InDeep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support 2018 Sep 20 (pp. 3–11). Springer, Cham.

[7] Yao J, Zhu X, Jonnagaddala J, Hawkins N, Huang J. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. Medical Image Analysis. 2020 Jul 19:101789.

[8] Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In2016 fourth international conference on 3D vision (3DV) 2016 Oct 25 (pp. 565–571). IEEE.

[9] Chan L, Hosseini MS, Rowsell C, Plataniotis KN, Damaskinos S. Histosegnet: Semantic segmentation of histological tissue type in whole slide images. InProceedings of the IEEE International Conference on Computer Vision 2019 (pp. 10662–10671).

[10] Yao J, Boben M, Fidler S, Urtasun R. Real-time coarse-to-fine topologically preserving segmentation. InProceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015 (pp. 2947–2955).

[11] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence. 2012 May 29;34(11):2274–82.

[12] Van den Bergh M, Boix X, Roig G, de Capitani B, Van Gool L. Seeds: Superpixels extracted via energy-driven sampling. InEuropean conference on computer vision 2012 Oct 7 (pp. 13–26). Springer, Berlin, Heidelberg.

[13] Li Z, Chen J. Superpixel segmentation using linear spectral clustering. InProceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015 (pp. 1356–1363).

[14] Ren X, Malik J. Learning a classification model for segmentation. Innull 2003 Oct 13 (p. 10). IEEE.

[15] Liu MY, Tuzel O, Ramalingam S, Chellappa R. Entropy rate superpixel segmentation. InCVPR 2011 2011 Jun 20 (pp. 2097–2104). IEEE.

[16] Veksler O, Boykov Y, Mehrani P. Superpixels and supervoxels in an

energy optimization framework. InEuropean conference on Computer vision 2010 Sep 5 (pp. 211–224). Springer, Berlin, Heidelberg.

[17] Bejnordi BE, Litjens G, Hermsen M, Karssemeijer N, van der Laak JA. A multi-scale superpixel classification approach to the detection of regions of interest in whole slide histopathology images. InMedical Imaging 2015: Digital Pathology 2015 Mar 19 (Vol. 9420, p. 94200H). International Society for Optics and Photonics.

[18] Litjens G, Bejnordi BE, Timofeeva N, Swadi G, Kovacs I, Hulsbergen-van de Kaa C, van der Laak J. Automated detection of prostate cancer in digitized whole-slide images of H and E-stained biopsy specimens. InMedical Imaging 2015: Digital Pathology 2015 Mar 19 (Vol. 9420, p. 94200B). International Society for Optics and Photonics.

[19] Lucchi A, Smith K, Achanta R, Lepetit V, Fua P. A fully automated approach to segmentation of irregularly shaped cellular structures in EM images. InInternational Conference on Medical Image Computing and Computer-Assisted Intervention 2010 Sep 20 (pp. 463–471). Springer, Berlin, Heidelberg.

[20] Van den Bergh M, Roig G, Boix X, Manen S, Van Gool L. Online video seeds for temporal window objectness. InProceedings of the IEEE international conference on computer vision 2013 (pp. 377–384).

[21] Yamaguchi K, McAllester D, Urtasun R. Efficient joint segmentation, occlusion labeling, stereo and flow estimation. InEuropean Conference on Computer Vision 2014 Sep 6 (pp. 756–771). Springer, Cham.

[22] Li R, Huang J. Fast regions-of-interest detection in whole slide histopathology images. In: International Workshop on Patch-based Techniques in Medical Imaging 2015 Oct 9 (pp. 120-127). Springer, Cham.

[23] Gadde R, Jampani V, Kiefel M, Kappler D, Gehler PV. Superpixel convolutional networks using bilateral inceptions. InEuropean Conference on Computer Vision 2016 Oct 8 (pp. 597–613). Springer, Cham.

[24] Kwak S, Hong S, Han B. Weakly supervised semantic segmentation using superpixel pooling network. InAAAI 2017 Feb 4 (Vol. 1, p. 2).

[25] Tu WC, Liu MY, Jampani V, Sun D, Chien SY, Yang MH, Kautz J. Learning superpixels with segmentation-aware affinity loss. InProceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018 (pp. 568–576).

[26] Liu X, Guo S, Yang B, Ma S, Zhang H, Li J, Sun C, Jin L, Li X, Yang Q, Fu Y. Automatic organ segmentation for CT scans based on super-pixel and convolutional neural networks. Journal of digital imaging. 2018 Oct 1;31(5):748–60.

[27] Jampani V, Sun D, Liu MY, Yang MH, Kautz J. Superpixel sampling networks. InProceedings of the European Conference on Computer Vision (ECCV) 2018 (pp. 352–368).

[28] Li R, Wang S, Zhu F, Huang J. Adaptive graph convolutional neural networks. arXiv preprint arXiv: 1801.03226. 2018 Jan 10.

[29] Yang F, Sun Q, Jin H, Zhou Z. Superpixel Segmentation with Fully Convolutional Networks. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020 (pp. 13964–13973).