Published in final edited form as:

Cytometry A. 2016 October ; 89(10): 893–902. doi:10.1002/cyto.a.22929.

A Supervised Learning Framework for Pancreatic Islet Segmentation With Multi-Scale Color-Texture Features and Rolling Guidance Filters

Yue Huang^{1,2,†}, Chi Liu^{2,†}, John F. Eisses³, Sohail Z. Husain³, and Gustavo K. Rohde^{2,4,*}

¹School of Information Science and Engineering, Xiamen University, Xiamen, China

²Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, 15213, Pennsylvania

³Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, 15224, Pennsylvania

⁴Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, 15213, Pennsylvania

Abstract

Islet cell quantification and function is important for developing novel therapeutic interventions for diabetes. Existing methods of pancreatic islet segmentation in histopathological images depend strongly on cell/nuclei detection, and thus are limited due to a wide variance in the appearance of pancreatic islets. In this paper, we propose a supervised learning pipeline to segment pancreatic islets in histopathological images, which does not require cell detection. The proposed framework firstly partitions images into superpixels, and then extracts multi-scale color-texture features from each superpixel and processes these features using rolling guidance filters, in order to simultaneously reduce inter-class ambiguity and intra-class variation. Finally, a linear support vector machine (SVM) is trained and applied to segment the testing images. A total of 23 hematoxylin-and-eosin-stained histopathological images with pancreatic islets are used for verifying the framework. With an average accuracy of 95%, training time of 20 min and testing time of 1 min per image, the proposed framework outperforms existing approaches with better segmentation performance and lower computational cost.

Key terms

pancreatic islet; histopathological image segmentation; supervised learning; rolling guidance filter; multi-scale features

Islet cell function is important for glucose homeostasis (1). Loss of islet cell function can result in diabetes, a chronic pathologic process that, if left unchecked, results in the multi-

*Correspondence to: Gustavo K. Rohde, Department of Biomedical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213. gustavor@cmu.edu.

[†]Yue Huang and Chi Liu contributed equally to this work.

Additional Supporting Information may be found in the online version of this article.

organ dysfunction, severe morbidity and death (2). Diabetes can result from islet destruction (Type I diabetes) or insulin resistance that progresses to islet dysfunction and islet loss (Type II diabetes). Both types of diabetes lead to endorgan injury unless treated with neurologic, cardiovascular, renal and ocular complications (2). During the early stages of the disease islet numbers are thought to increase to compensate for insulin need (3). Currently, analysis of islet health relies on glucose monitoring as a marker of islet function, rather than actual evaluation of islet quantity and health. In experimental models where islet quantification is used to assess islet health and function, manual evaluation by trained pathologists requires tedious and time intensive analysis of hematoxylin and eosin (H&E) sections. Automated quantification of islet cell density would allow objective analysis of novel therapeutic strategies for regeneration of islet cells including insulin producing β -cells for the treatment of diabetes.

According to the literature, there have been few works contributing to pancreatic islet segmentation in histopathological images, and most of the existing works strongly depend on cell detection and segmentation, under the assumption that pancreatic islets have a high cell density. Floros et al. proposed a novel graph-based pancreatic islet segmentation approach based on cell nuclei detection with randomized tree ensembles (4). After detection of each nucleus in the histopathological images, a classifier is applied to recognize the different categories of nuclei, such as α -cell, β -cell and so on. The pancreatic islet can be segmented by constructing a graph to represent the high cell density. This approach has been applied to automated assessment of β -cell areas and density per islet based on immunofluorescence staining (5). However, this method is still limited by its assumption that there is only one islet in each image, and that the islet usually has high density cells (4). An automated system, called pancreas++, has also been developed to detect and quantify pancreatic islet cells in microscopic images (6). Similar to Ref. 4, this algorithm is also based on cell detection. The system includes thresholding, nearest-neighbor interpolation, and active contour models for detecting large contiguous islet regions. However, the system works well for fluorescent images only, where all the objects in the images are cells with different colors that can be detected after simple thresholding. A similar study has been reported by Brissova et al. (7), who developed an image analysis pipeline for pancreatic islet analysis by identifying both α -cells and β -cells, but the approach is also designed for fluorescent images only.

Therefore, it is still a challenge to segment pancreatic islets in histopathological images, due to variance in sizes, shapes, and contexts of the pancreatic islets. Additionally, unknown objects may also exist in the background with a similar pattern (cell density) to the islets. As shown in Figure 1, the islet in (a) is in a regular shape containing dense cells, the islet in (b) is small with only few internal cells (on the bottom right), and the islet in (c) has an irregular shape and has a similar texture pattern with the objects in the background. Therefore, cell detection based pancreatic islet segmentation is not an optimal option for histopathological image analysis.

There have been some reports contributing to histopathological images segmentation with unsupervised learning methods, such as N-point correlation functions (N-pcfs), level sets, and graph run-length (8–10). Besides, machine learning methods have also been used for

histopathological image analysis (11, 12). These approaches treat the segmentation as a pattern recognition problem, where each pixel or small region is treated as a sample (13). A classifier is trained with the extracted features from the training samples and the ground-truth labels are provided by pathologists. The testing samples, which have the same extracted features as the training samples, will be classified by the trained classifier into two categories: objects of interest or background (14). Finally, the regions of interest (ROIs) in the testing images are detected and segmented once all samples have been classified (15). For example, Basavanahally et al. proposed an image-based detection and grading system for lymphocytic infiltration in HER2+ breast cancer histopathological images, where each lymphocytic infiltration candidate is described with three graphs as its feature in the supervised classification (16). The segmentation method in Ref. 17 is based on multiple instance learning. The basic idea of the work is that it considers each histopathological image as a bag, and then considers the image patches in the image as instances. The work aims to identify the image patches from the images with labels, that is, it predicts the unknown instance (image patch) labels from the known bag (image) labels. A supervised learning segmentation method using pixel-wise classification has also been proposed and applied for detection of pancreatic acinar cells in histopathological images (18–20). The features in the classification are defined with the multi-scale intensity neighborhoods in each pixel.

Since the method in Ref. 18 is used for a similar application to the proposed problem, it can potentially be applied to pancreatic islet segmentation. However, the framework still has some limitations. First, pixel-wise classification is not an optimal option for segmenting larger ROIs, such as islets or glands. This is because examination of local details using pixel-wise features cannot sufficiently represent regional characteristics available in a global view, and it is also computationally inefficient. For example, Ozdemir et al. proposed a tissue classification model with the features combining pixel-wise structural and graph-based patterns for describing each gland in the histopathological images (21). Segmentation via region-based classification firstly partitions the image into many small regions, such as image patches or superpixels, and then classifies each small region into either object or background (22). The features extracted from the sub-regions tend to capture global characteristics rather than pixel-wise features. In addition, these methods are more efficient since the classification is only applied to a limited number of subregions. Secondly, multi-scale feature extraction has taken inspiration from the real-world diagnosis, where pathologists zoom in and zoom out of particular regions by switching lenses (18). The most widely used method to generate multi-scale images is convolving the image with well-known filters, such as a Gaussian filter (23–26). However, since pancreatic islets can have a variety of appearances, this can result in intra-class ambiguity in classification-based segmentation. It is desirable to suppress the intra-class ambiguity while simultaneously preserving the large inter-class variances in images at different scales. To the best of our knowledge, no previous study has focused on the filter design for multi-scale feature extraction in histopathological image processing.

Therefore, in our work, a supervised framework based on superpixel-wise classification is proposed for pancreatic islet segmentation in histopathological images. There are two main steps in the proposed framework, superpixel generation and superpixel classification. In the

first step, images are partitioned into superpixels using color deconvolution and clustering methods. In the second step, a multi-scale feature extraction method is proposed for the supervised framework. Multi-scale texture-color features are extracted from images at different scales, which are generated by convolving the original image with a new filter, called a rolling guidance filter (27). The proposed work is validated using 23 hematoxylin-and-eosin (H&E) stained histopathological images with pancreatic islets.

Our contributions can be summarized as follows: (a) a superpixel-classification based pancreatic islet segmentation method is proposed, which is more efficient than segmentation via pixel-wise classification; (b) pancreatic islets are segmented without any contextual features, thus there is no longer any requirement for a cell detection procedure; (c) a new filter, called a rolling guidance filter, is used to generate images at different scales for multi-scale feature extraction, in order to enhance the discriminative power of the extracted features.

The remainder of the article is organized as follows. In the next section, the sample preparation and the image processing and image analysis framework are described in detail. After acquisition of the H&E histopathological images from an image scanner, each image in the training dataset is partitioned into many superpixels, and their labels are generated using the ground-truth obtained from pathologists. A multi-scale texture-color feature extractor is applied to all superpixels in the training images, using a new scale definition known as a rolling guidance filter. A linear support vector machine is trained, and the testing images are segmented when all of the superpixels are classified as either background or object. The segmentation results from the proposed framework are provided and compared with other pancreatic islet segmentation methods in the results section. The performances are further analyzed in the discussion section, followed by a final conclusion.

Materials

Images of pancreatic islet within the pancreatic parenchyma were provided by researchers at Children's Hospital Pittsburgh of UPMC. Six-week-old male Swiss Webster mice (Charles River, Wilmington, MA) weighing 20–25 g were fed standard laboratory chow with free access to water. All animal experiments were performed using a protocol approved by the University of Pittsburgh Institutional Animal Care and Use Committee. Isolated pancreas tissue from five separate mice were fixed in 4% PFA overnight. The fixed tissue was embedded in paraffin, sectioned, and stained with H&E. Whole sections were scanned using a 60× objective in an Aperio scanner (Aperio ePathology Solutions, Leica Microsystems, Buffalo Grove, IL) to generate high resolution images. The spatial resolution of the image is 0.14 $\mu\text{m}/\text{pixel}$. Partitioned images were generated from whole images and used in the development and analysis. Twenty-three images including pancreatic islet and other pancreatic cell types were used in the analysis. The width range of 23 images is [650, 698], while the length range of 23 images is [900, 953]. The average image size was 937×670 pixel².

Methods

The proposed supervised framework is summarized in Figure 2. Both the training procedure in A and the testing procedure in B contain superpixel segmentation and multi-scale feature extraction, which are the main components of the proposed framework. The techniques used for superpixel segmentation and multi-scale feature extraction will be described in detail within this section.

Superpixel Generation

In this step, all of the images, including both training and testing, are partitioned into a large number of superpixels. An ideal superpixel segmentation method should accurately adhere to object boundaries, and should have fast computation and be memory-efficient (28). Simple linear iterative clustering (SLIC) is widely used methods in superpixel generation, which adopts a k-means clustering approach to efficiently grouping pixels (29). The study in Ref. (30) proposed SLIC-spectrum, an improved SLIC method, that simultaneously considers both global image information and local features based on grouping pixels into superpixels with weighted k-means clustering in a high-dimensional space. In histopathological image analysis, H&E staining is well-defined, and the hematoxylin channel and the eosin channel can be separated with color deconvolution (31). Therefore, SLIC-spectrum-HE is proposed by introducing the definition of H&E staining to SLIC-spectrum. First, a four-dimensional vector $[h, e, x, y]$ is defined for each pixel, where h and e denote the hematoxylin channel and the eosin channel respectively, and x, y denotes the coordinates. The range of each component in the vector is linearly normalized into $[0, 1]$ (30). Second, the feature vector of each pixel is then mapped to a higher dimensional space using a nonlinear kernel. The mapping procedure for a pixel p in the input image space is defined as

$$\phi(p) = \frac{1}{w(p)} (c_c \cos \frac{\pi}{2} h_p, c_c \sin \frac{\pi}{2} h_p, c_c \cos \frac{\pi}{2} e_p, c_c \sin \frac{\pi}{2} e_p, c_s \cos \frac{\pi}{2} x_p, c_s \sin \frac{\pi}{2} x_p, c_s \cos \frac{\pi}{2} y_p, c_s \sin \frac{\pi}{2} y_p), \quad (1)$$

where p is the linear pixel index in the image, $\phi(\cdot)$ is the mapping procedure ($\mathbb{R}^4 \rightarrow \mathbb{R}^8$), h_p and e_p denote the haematoxylin and eosin components generated by color deconvolution, respectively. x_p, y_p are the vertical and horizontal coordinates in the image plane. c_c and c_s are two parameters used to control the relative significance of color and spatial information when calculating similarity between data points (30). It has been investigated and concluded in Ref. 30 that objective functions of weighted k-means and normalized cuts share the same optimum points. After mapping, the weighted k-means is then applied to group pixels in the high dimensional feature space, by minimizing the object function as follows:

$$L_{k-m} = \sum_{k=1}^K \sum_{p \in CL_k} w(p) \|\phi(p) - C_k\|^2, \quad (2)$$

$$C_k = \frac{\sum_{q \in CL_k} w(q) \phi(q)}{\sum_{q \in CL_k} w(q)}, \quad (3)$$

where q and p represent the linear index in the image space, C_k is the center of the k th ($k = 1, 2, \dots, K$) cluster CL_k , K is the number of clusters, and $w(p)$ and $w(q)$ are the weights assigned to pixels p and q , respectively.

Multi-Scale Feature Extraction

After superpixel segmentation, the multi-scale features are extracted for each superpixel in order to train the classifier and obtain prediction labels for the testing data. As mentioned in the introduction, the pancreatic islets may have various sizes, shapes, and contexts, and unknown objects may also exist in the background with a similar pattern (cell density) to the islets. To be considered as a classification problem, the segmentation performance is limited to large intra-class variances, and relatively small inter-class variances. Therefore, the ideal extracted features should have discriminative power across different classes. Rather than scale definition using a well-known filter, e.g., a Gaussian filter, the proposed method defines the image scales and the corresponding image pyramid with a new type of filter, which is called a rolling guidance filter (27).

The rolling guidance filter is designed to iteratively remove small structures and preserve larger structures, e.g. edges between objects. It is also considered to be an effective and efficient tool to enhance inter-class variance and suppress intra-class variance at different levels. This method provides a new type of image pyramid, with each scale representing information in the scale only (27). The basic equation for the rolling guidance filter is defined as

$$J^{t+1}(p) = \frac{1}{K_p} \sum_{q \in N(p)} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{\|J^t(p) - J^t(q)\|^2}{2\sigma_r^2}\right) I(q), \quad (4)$$

where I is the input image, p and q are the coordinates of two pixels in I , $N(p)$ denotes the neighboring pixels of p , $K_p = \sum_{q \in N(p)} \exp(-\|p - q\|^2 / 2\sigma_s^2) \exp(-\|J^t(p) - J^t(q)\|^2 / 2\sigma_r^2)$ is used for normalization purposes, σ_s and σ_r denote the spatial and range weights, respectively, J is the guidance image at the t th iteration, with J^t generated by convolving I with a Gaussian filter. The equation above can be considered to represent a filter that smooths the input I guided by the structure of J^t . The term $\exp(-\|p - q\|^2 / 2\sigma_s^2)$ enables thorough suppression of small textures at different levels during iterations, thus intra-class

variances are minimized. Meanwhile, global color variances and large edges are maintained by the term $\exp(-\|J^t(p) - J^t(q)\|^2 / 2\sigma_r^2)$, thus preserving inter-class variances (23). Thus the proposed scale definition enhances the discriminative power of the extracted features.

Texture-color features have been reported to separate cells from H&E stained histopathological images in Ref. (32). However, the reported texture-color feature extraction method is more suitable to pixel-wise classification in segmentation of small objects, e.g., cell or nuclei. After convolving with rolling filters, texture-color features are extracted in image at each scale. In the proposed framework, a method called segment based fractal texture analysis (SFTA) is used to describe the texture pattern due to the observation that histopathological images with more jagged edges and prominent color differences tend to have higher fractal dimension (islet), images with more uniform texture properties (background) will have closer fractal dimension. In addition, SFTA has been reported to outperform many well-known feature extractors with higher classification accuracy and less computational cost, such as histogram, Gabor, Haralick, co-occurrence gray level matrices (GLCM) and even combinations of some of them (33). The algorithm of SFTA is briefly summarized as follows. First, it decomposes images into various binary thresholded images using many pairs of lower and upper threshold values; second, a widely used measurement, called Hausdorff dimension, is used to calculate the fractional dimensions in each thresholded image; finally, the output of the SFTA feature vector is constructed as the size, mean gray level, and fractal dimension of resulting binary images. Therefore, the texture-color feature vector in each superpixel at each scale is defined as:

$$F = [F_t, F_c], \quad (5)$$

$$F_c = [m_h, m_e, st_h, st_e], \quad (6)$$

where F_t is texture feature vector from SFTA and m_h , m_e , st_h , and st_e denote the mean and standard deviation values of haematoxylin and eosin channels respectively for each superpixel. The feature vector is also normalized by the mean and variance of the training data.

After feature extraction, a linear support vector machine (SVM), which is one of the most widely used classifiers, is used in the proposed framework for binary classification (34). Details of the SVM implementation are provided in the next section.

Results

Framework Implementations

A machine with an Intel(R) Core (TM) i5-3210M CPU @ 2.5 GHz Processor, 4G RAM and a 64-bit operating system is used to implement and evaluate the proposed methods. All code is implemented by Matlab 2015a. Only a few parameters are required to set up the implementation of the framework for optimal performance. The width range of 23 images is

[650, 698], while the length range of 23 images is [900, 953]. During the superpixel segmentation set up, the number of superpixels is manually set as 200 for all of the images.

For the feature extraction step, the multi-scale color-texture features are defined using three scales, two of which are generated by a rolling guidance filter, and the other is the original image. The number of iterations used to generate the two images for the two scales as well as the original image is 5 and 10, respectively. A 50-dimensional vector F is defined for each scale, so each superpixel is denoted by a 150-dimensional vector containing the three scales. Since a feature-label pair is required for each superpixel in the classifier training, more details on the label assignment of each superpixel after feature extraction will be provided. In the proposed work, the ground-truth is labeled by a pathologist at a pixel-level. Although it is assumed that superpixels will group pixels from the same class, the algorithm cannot guarantee that the generated superpixels will perfectly adhere to the pancreatic islet boundaries, so superpixel label assignment is required. Since there are variations in size of the pancreatic islets, it is difficult to determine the number of superpixels for each islet. So superpixel generation is implemented in the original image, and then the label of each superpixel is defined by a simple majority voting strategy. The strategy used is that if more than half of the pixels in one superpixel belong to the islet, then the superpixel is given an islet label, and if more than half of the pixels in the superpixel belong to background, then the superpixel is given a background label.

Linear SVM is implemented using the academic package LIBSVM (35). The penalty constant ϕ in SVM is optimized using a double cross-validation strategy. That is, the dataset is divided into training data and testing data by a leave-one-image-out method; and then the training dataset is further divided into training data and validation data using this leave-one-image-out method. So ϕ is optimized in the training data and validation data, and then is applied to the individual testing data. The values of the tunable parameter ϕ and corresponding validation performances are provided in the supplemental files. In the training procedure, the superpixel labels, along with the extracted features from the training data, are used to train the classifier, and then the classifier is used to identify superpixels in the testing images as either background or islet. In the testing procedure, if a superpixel is assigned to one class, then all of its pixels will be categorized as the same class. Therefore, a testing image can be segmented after all of its superpixels have been classified.

Results and Comparisons

In this section, both qualitative and quantitative results are shown and compared with existing methods. As mentioned in the introduction, most of the computer-aided methods for pancreatic islet segmentation require cell detection. The method reported in Refs. 19 and 20 is also a supervised learning segmentation framework for pancreatic histopathological images, but is based on pixel-wise classification. Thus, the method in (10, 11) is included for comparison. This method is reproduced using their public source code in Matlab and the same parameter settings as given in their article. The segmentation results of the four images are provided in Figure 3 as examples, and it can be seen that the boundaries of the identified islets overlap on the original images. It can be clearly observed that the proposed method (left column) outperforms the pixel-wise segmentation method in (10, 11) with more

accurate boundary adherence to the pancreatic islets. A quantitative segmentation comparison is additionally provided in Table 1. The segmentation performances are validated using the binary ground-truth image that originated from manual labeling, where background is denoted by 0 and islets are denoted by 1. Metric accuracy is a global measurement equal to the ratio of the total number of correctly classified pixels to the number of pixels in the images. In addition, another measurement called dice metric (dm) is also added to the evaluation, which was also used in other studies (18, 36), and is given by the following equation:

$$dm = \frac{2 |T_p \cap T_g|}{|T_p| + |T_g|}, \quad (7)$$

where T_p denotes the segmented region of pixels for a given islet, T_g denotes the ground-truth set of pixels for the same islet, and $|\cdot|$ represents the set size. A value of dm that is closer to 1 denotes better performance. The value of dm is calculated for all islets in each image, and the value of the mean and standard deviation are also provided. It can also be clearly observed that in the quantitative validation, the proposed method outperforms the method in Refs. 10 and 11, with higher values of both accuracy and dice metrics. The statistical analysis of the t test also demonstrates statistically significant differences between the two methods for both measurements. As well as segmentation performance, a comparison of the computational cost for each method is also shown in Table 2, where the values are all averaged across all of the images. It can also be observed that the proposed framework is more efficient than the method in Refs. 10 and 11 with less computation time for both the training and testing procedure. This can be explained by the fact that the other method is pixel-wise, and thus each pixel in the image will be examined in the classification, while the proposed method is based on superpixel-wise classification, so only a limited number of superpixels are analyzed during testing. As mentioned in the Materials section, the average image size is 937 pixels $\times$ 670 pixels.

Discussion

In this section, we provide further discussion on the reasons that the proposed pancreatic islet segmentation framework provides better performance. Each step of the framework is validated, including superpixel generation and superpixel classification (including feature extraction and classifier design) individually using the following three steps.

Step 1. It should be mentioned that some pixels have the wrong labels assigned due to under-segmentation errors from the superpixel generation, rather than the superpixel classification. For example, if a generated superpixel contains an overlap of both background and islet, and the label of the superpixel is assigned to be background, then the pixels from the islet in this superpixel will be misclassified. Therefore, superpixel generation plays an important role in the final segmentation. Examples of the superpixels generated using the proposed SLIC-spectrum-HE method and SLIC are shown in Figure 4, where it can be observed from the magnified areas of Figure 4b and 4c that the proposed superpixel generation method has better performance in adhering to the pancreatic islet boundary, thus there are fewer

superpixels with segmentation errors using the proposed method than with SLIC. Additionally, we also set up a framework called ‘SLIC baseline’ which changes the SLIC-spectrum-HE method of the proposed framework with SLIC, while keeping the remaining parts unchanged. A quantitative comparison between ‘SLIC baseline’ and the proposed framework with the same measurements and statistical analysis as in the results section is provided, as shown in Figure 5 and Table 3. It can also be concluded that the proposed superpixel generation method contributes to a better segmentation performance of the framework.

Step 2. We set up a new framework called “Gaussian filter baseline,” which defines a multi-scale texture-color feature with a different scale definition for each superpixel. In other words, the image for each scale is generated by convolving the image with a Gaussian filter instead of the proposed rolling guidance filter. The remaining parts of ‘Gaussian filter baseline’ are the same as the proposed framework. There are two rounds of quantitative validation. The first round evaluates the performances of the superpixel-wise classification, with the well-known measurements of sensitivity, specificity, accuracy and area under the receiver-operator characteristic curve (AUC). A larger AUC value denotes a better classification performance. Comparisons of the superpixel-wise classification with different scale definitions are provided in Table 4. Both mean and standard deviation values are provided (mean $\pm$ standard deviation), and the related statistical analysis with the t test demonstrates the statistically significant difference between the two methods. After superpixel-wise classification, all pixels in the same superpixel will be categorized with the same label, thus the testing image is segmented. Therefore, the second round is similar to the results section and step 1, where the islet segmentation performance comparisons between Gaussian filter baseline and the proposed framework is quantitatively validated in terms of accuracy, dice metrics and related statistical analysis, as shown in Table 5. The segmentation results from Gaussian filter baseline and the proposed framework are also provided for two sample images and compared in Figure 6. The experiments of this step support our assertion that the rolling guidance filter also contributes to better performance of the proposed framework.

Step 3. Similar to step 2, we set up a framework called “linear SVM baseline” which substitutes the linear kernel in SVM with a Gaussian kernel. There is no statistically significant difference between SVM and kernel SVM, which can also be explained. The difference in objective functions between linear SVM and kernel SVM is the nonlinear mapping procedure for the data, which aims to find a linear hyper-plane to separate the data by projecting the data to a higher dimensional space. In the multi-scale feature extraction step of the proposed framework, each superpixel is denoted by a 150-dimensional vector aggregated from three scales, so there are limited performance improvements to be obtained from another mapping method. A similar conclusion can be found in reference (37).

We will also analyze the framework to identify the error sources of the proposed segmentation framework. The sources of error arise from both superpixel generation and superpixel classification. With the ground-truth image, it is possible to calculate the number of under-segmented superpixels, which have an overlap of both islets and non-islets in each image, and then get the mean value across the whole dataset. With the SLIC method, the

mean value of under-segmented superpixels per image is 18; while this number decreases to 10 per image with the proposed superpixel generation method. As we mentioned in the results section, all pixels in one superpixel will be assigned to the same label after superpixel classification, so there will be misclassified pixels in under-segmented superpixels, resulting in segmentation errors. As shown in step 2 for the step validation, it can be concluded that even with the proposed method, the superpixel classification accuracy is not 100%, as there are still some segmentation errors due to the superpixel classification.

The segmentation results are further validated using a new metric on a per-islet basis. All of the islets that were segmented by the proposed framework have been examined, and the number of true-positives, false-positives, and false-negatives were calculated. A true-positive islet is defined as a detected islet that is located inside a true islet, a false-negative islet is a missed islet, and a false positive islet is defined as the detected islet locates in the background. Examples of true-positive islets, false-negative islets and false-positive islets are shown in Figure 7, indicated using different color rectangles. In total, there are 28 islets in the ground-truth images in the dataset, and the proposed framework detects 37 islets. The number of true-positive islets, false-negative islets and false-positive islets are 24, 4, and 13, respectively.

Conclusions

This work has proposed a supervised learning based framework for the pancreatic islet segmentation task in histopathological images, in order to remove the requirement for cell detection in islet segmentation. The proposed framework includes superpixel generation and superpixel classification. Superpixels are generated with a weight k-means clustering method on a high-dimensional space. Each superpixel is then classified as either an islet or a non-islet with multi-scale texture-color features. Rather than convolving each image with a widely used Gaussian filter, the image is generated for each scale using a rolling guidance filter, in order to suppress intra-class variances and enhance inter-class variances in the pattern recognition based segmentation framework. Experiments on pancreatic islet segmentation in histopathological images and corresponding comparisons with other existing works have demonstrated the effectiveness and efficiency of the proposed works. The proposed framework can also be considered as a general tool to segment regions of interest in other histopathological image processing tasks. As for future works, the hand-crafted feature extraction step in proposed framework can be replaced by convolutional neural networks (CNNs), which has the ability of extracting features driven from the data automatically and adaptively.

Acknowledgments

Grant sponsor: National Institutes of Health; Grant number: R01 (DK103002); Grant sponsor: The National Natural Science Foundation of China; Grant numbers: 61571382, 61571005, 81301278, 61172179 and 61103121; Grant sponsor: The CCF-Tencent Open Fund.

Literature Cited

1. Ashcroft FM, Rorsman P. Diabetes mellitus and the beta cell: The last ten years. *Cell*. 2012; 148:1160–1171. [PubMed: 22424227]

2. WHO. Part 1: Diagnosis and classification of diabetes mellitus. Geneva: World Health Organization; 1999. Definition, diagnosis and classification of diabetes mellitus and its complications.
3. Weir GC, Bonner-Weir S. Islet B cell mass in diabetes and how it relates to function, birth, and death. *Ann N Y Acad Sci.* 2013; 1281:92–105.
4. Floros, X., Fuchs, TJ., Rechsteiner, MP., Spinas, G., Moch, H., Buhmann, JM. Proceedings of the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part II. London: Springer; 2009. Graph-based pancreatic islet segmentation for early type 2 diabetes mellitus on histopathological tissue; p. 633–640.
5. Rechsteiner MP, Floros X, Boehm BO, Marselli L, Marchetti P, Stoffel M, Moch H, Spinas GA. Automated assessment of beta-cell area and density per islet and patient using TMEM27 and BACE2 immunofluorescence staining in human pancreatic beta-cells. *PLoS One.* 2014; 9:e98932. [PubMed: 24905913]
6. Chen H, Martin B, Cai H, Fiori JL, Egan JM, Siddiqui S, Maudsley S. Pancreas++: Automated quantification of pancreatic islet cells in microscopy images. *Front Physiol.* 2012; 3:482. [PubMed: 23293605]
7. Brissova M, Fowler MJ, Nicholson WE, Chu A, Hirshberg B, Harlan DM, Powers AC. Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy. *J Histochem Cytochem.* 2005; 53:1087–1097. [PubMed: 15923354]
8. Bunyak, F., Hafiane, A., Palaniappan, K. Histopathology tissue segmentation by combining fuzzy clustering with multiphase vector level sets. In: Arabnia, RH., Tran, Q-N., editors. *Software Tools and Algorithms for Biological Systems.* New York, NY: Springer New York; 2011. p. 413–424.
9. Mosaliganti K, Janoos F, Irfanoglu O, Ridgway R, Machiraju R, Huang K, Saltz J, Leone G, Ostrowski M. Tensor classification of N-point correlation function features for histology tissue segmentation. *Med Image Anal.* 2009; 13:156–166. [PubMed: 18762444]
10. Tosun AB, Gunduz-Demir C. Graph run-length matrices for histopathological image segmentation. *IEEE Trans Med Imaging.* 2011; 30:721–732. [PubMed: 21097378]
11. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. *IEEE Rev Biomed Eng.* 2009; 2:147–171. [PubMed: 20671804]
12. Kothari S, Phan JH, Stokes TH, Wang MD. Pathology imaging informatics for quantitative analysis of whole-slide images. *J Am Med Inform Assoc.* 2013; 20:1099–1108. [PubMed: 23959844]
13. Gultekin T, Koyuncu CF, Sokmensuer C, Gunduz-Demir C. Two-tier tissue decomposition for histopathological image representation and classification. *IEEE Trans Med Imaging.* 2015; 34:275–283. [PubMed: 25203985]
14. Belsare AD, Mushrif MM, Pangarkar MA, Meshram N. Breast histopathology image segmentation using spatio-colour-texture based graph partition method. *J Microsc.* 2015; 262:260–273. [PubMed: 26708167]
15. Bunyak, F., Halfiane, A., Al-Milaji, Z., Ersoy, I., Haridas, A., Palaniappan, K. IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Washington DC: IEEE; 2015. A segmentation-based multi-scale framework for the classification of epithelial and stromal tissues in H&E images; p. 450–453.
16. Basavanthally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, Tomaszewski JE, Bhanot G, Madabhushi A. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. *IEEE Trans Biomed Eng.* 2010; 57:642–653. [PubMed: 19884074]
17. Xu Y, Zhu JY, Chang EI, Lai M, Tu Z. Weakly supervised histopathology cancer image segmentation and classification. *Med Image Anal.* 2014; 18:591–604. [PubMed: 24637156]
18. Chen C, Ozolek JA, Wang W, Rohde GK. A general system for automatic biomedical image segmentation using intensity neighborhoods. *Int J Biomed Imaging.* 2011; 2011:606857. [PubMed: 21760767]
19. Eisses JF, Davis AW, Tosun AB, Dionise ZR, Chen C, Ozolek JA, Rohde GK, Husain SZ. A Computer-based automated algorithm for assessing acinar cell loss after experimental pancreatitis. *PLoS One.* 2014; 9:e110220. [PubMed: 25343460]
20. Eisses JF, Criscimanna A, Dionise ZR, Orabi AI, Javed TA, Sarwar S, Jin S, Zhou L, Singh S, Poddar M, et al. Valproic acid limits pancreatic recovery after pancreatitis by inhibiting histone

- deacetylases and preventing acinar redifferentiation programs. *Am J Pathol.* 2015; 185:3304–3315. [PubMed: 26476347]
21. Ozdemir E, Gunduz-Demir C. A hybrid classification model for digital pathology using structural and statistical pattern recognition. *IEEE Trans Med Imaging.* 2013; 32:474–483. [PubMed: 23204283]
 22. Turkki R, Linder N, Kovanen PE, Pellinen T, Lundin J. Identification of immune cell infiltration in hematoxylin-eosin stained breast cancer samples: Texture-based classification of tissue morphologies. *Proc. SPIE.* 2016; 9791:979110.
 23. He K, Sun J, Tang X. Guided image filtering. *Eruopean Conference on Computer Vision (ECCV).* 2010:1–14.
 24. Gorelick L, Veksler O, Gaed M, Gomez JA, Moussa M, Bauman G, Fenster A, Ward AD. Prostate histopathology: Learning tissue component histograms for cancer detection and classification. *IEEE Trans Med Imaging.* 2013; 32:1804–1818. [PubMed: 23739794]
 25. Doyle S, Feldman M, Tomaszewski J, Madabhushi A. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Trans Biomed Eng.* 2012; 59:1205–1218. [PubMed: 20570758]
 26. McCann MT, Ozolek JA, Castro CA, Parvin B, Kovacevic J. Automated histology analysis, opportunities for signal processing. *IEEE Signal Processing Mag.* 2015; 32:78–87.
 27. Zhang Q, Shen X, Xu L, Jia J. Rolling guidance filter. *European Conference on Computer Vision (ECCV).* 2014:815–830.
 28. Levinshstein A, Stere A, Kutulakos KN, Fleet DJ, Dickinson SJ, Siddiqi K. TurboPixels: Fast superpixels using geometric flows. *IEEE Trans Pattern Anal Machine Intell.* 2009; 31:2290–2297.
 29. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal and Machine Intell.* 2012; 34:2274–2282.
 30. Li, Z., Chen, J. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* Boston, MA: IEEE; 2015. Superpixel segmentation using linear spectral clustering; p. 1356-1363.
 31. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol.* 2001; 23:291–299. [PubMed: 11531144]
 32. Kong H, Gurcan M, Belkacem-Boussaid K. Partitioning histopathological images: An integrated framework for supervised color-texture segmentation and cell splitting. *IEEE Trans Med Imaging.* 2011; 30:1661–1677. [PubMed: 21486712]
 33. Costa, AF., Humpire-Mamani, G., Traina, AJM. 25th SIBGRAPI Conference on Graphics, Pattern and Images. Ouro Preto, Brazil: 2012. An efficient algorithm for fractal analysis of textures; p. 39-46.
 34. Cortes C, Vapnik V. Support-vector networks. *Machine Learning.* 1995; 20:273–297.
 35. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol.* 2011; 2:1–27.
 36. Dice LR. Measures of the amount of ecologic association between species. *Ecology.* 1945; 26:297–302.
 37. Cheng J, Liu J, Xu Y. Superpixel classification based optic disc and opti cup segmentation for glaucoma screening. *IEEE Trans Med Imaging.* 2013; 32:1019–1032. [PubMed: 23434609]

Figure 1.

The challenge with islets segmentation is due to the complex variation in size, color, shape, and texture. For example, islet in (a) is in a regular shape containing dense cells, the islet in (b) is small with only few cells (on the bottom right), and the islet in (c) has an irregular shape and has a similar texture pattern with the objects in the background. The boundary ground-truth of islets are overlapped with the original images. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 2.

Flowchart of proposed segmentation framework with superpixel classification. The framework includes both training procedure (A) and testing procedure (B). Both training procedure and testing procedure contain superpixel segmentation and multi-scale feature extraction, which are two crucial steps in the proposed framework. Blocks in light yellow denote the methods in corresponding steps. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 3.

Segmentation results and comparisons are shown by labeling boundaries of identified islets overlap on the original images. Left column: segmentation results by the proposed framework. Middle column: segmentation results with the methods in Refs. 10 and 11. Right column: corresponding ground-truth labeling provided by a pathologist. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 4.

Examples of superpixel generated with different methods. Original image is overlapped with ground-truth islet boundary in (a); superpixels generated from SLIC (19) and proposed method are shown in (b) and (c) respectively, where the boundaries of all the generated superpixels are overlapped on the original image. In both (b) and (c), an area containing islet boundary is magnified at the left-bottom of the image, in order to show the object adherent performances of different generation methods. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 5. Segmentation results and comparisons of SLIC baseline (left column), proposed framework (middle column), and the ground-truth (right column) are provided by labeling boundaries of identified islets overlap on the original images. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 6.

Segmentation results and comparisons of Gaussian filter baseline (left column), proposed framework (middle column), and the ground-truth (right column) are provided by labeling boundaries of identified islets overlap on the original images. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 7.

Some examples of true positive islet, false positive islet, and false negative islet, which are indicated by green, white, and blue rectangles, respectively. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Comparisons of pixel-wise validation between proposed method and methods in Refs. 10 and 11

Table 1

	ACCURACY			DICE MEASUREMENT		
	MEAN	STD	P VALUES	MEAN	STD	P VALUES
Methods in Refs. (10) and (11)	0.76	0.18	$P < 0.01$	0.31	0.25	$P < 0.01$
Proposed	0.95	0.03		0.66	0.29	

Table 2

Comparisons of computational costs between proposed method and methods in Refs. 10 and 11

	TRAINING TIME	TESTING TIME
Methods in Refs. (10) and (11)	8 min per image	7 min per image
Proposed	1.5 min per image	1 min per image

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3
Comparisons of pixel-wise validation with different superpixel generation methods

	ACCURACY		DICE MEASUREMENT			
	MEAN	STD	P VALUES	MEAN	STD	P VALUES
SLIC baseline	0.90	0.08	$P < 0.05$	0.53	0.26	$P < 0.01$
Proposed	0.95	0.03		0.66	0.29	

Table 4

Comparisons of superpixel-wise classification with different scale definitions with mean and standard deviation values (mean $\pm$ standard deviation)

	SENSITIVITY	SPECIFICITY	ACCURACY	AUC
Gaussian filter baseline	0.64 $\pm$ 0.24	0.84 $\pm$ 0.11	0.82 $\pm$ 0.10	0.89 $\pm$ 0.10
Proposed	0.68 $\pm$ 0.28 _a	0.99 $\pm$ 0.07 _a	0.97 $\pm$ 0.02 _a	0.97 $\pm$ 0.03 _a

Denotes that there is statistically significant difference between framework and proposed framework by *t* test.

Table 5

Comparisons of pixel-wise validation with different scale definitions

	ACCURACY			DICE MEASUREMENT		
	MEAN	STD	P VALUES	MEAN	STD	P VALUES
Gaussian filter baseline	0.82	0.13	$P < 0.01$	0.42	0.27	$P < 0.01$
Proposed	0.95	0.03		0.66	0.29	