Automated separation of merged Langerhans islets

Jan Švihlík[a,c], Jan Kybic[a], David Habart[b]

[a]Biomedical Imaging Algorithms (BIA) group, Center for Machine Perception, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, Prague, Czech Republic

[b]Institute for Clinical and Experimental Medicine, Vídeňská 1958/9, Prague, Czech Republic

[c] Department of Computing and Control Engineering, Faculty of Chemical Engineering, University of Chemistry and Technology, Technická 5, Prague, Czech Republic

ABSTRACT

This paper deals with separation of merged Langerhans islets in segmentations in order to evaluate correct histogram of islet diameters. A distribution of islet diameters is useful for determining the feasibility of islet transplantation in diabetes. First, the merged islets at training segmentations are manually separated by medical experts. Based on the single islets, the merged islets are identified and the SVM classifier is trained on both classes (merged/single islets). The testing segmentations were over-segmented using watershed transform and the most probable back merging of islets were found using trained SVM classifier. Finally, the optimized segmentation is compared with ground truth segmentation (correctly separated islets).

Keywords: Merged Langerhans islets, support vector machine, shape descriptors, segmentation

1. INTRODUCTION

Transplantation of isolated pancreatic islets from cadaver donors is a promising therapy for patients with the type 1 diabetes.[1] To determine the quality of isolated islets and their suitability for successful transplantation, microscopy images of islet graft samples are acquired and the quantity and size of the islets is evaluated.

We proposed method for automated separation of touching islets in binary segmentations of microscopy images in order to evaluate distribution of islet diameters correctly. A distribution of islet diameters is useful for determining the feasibility of islet transplantation in diabetes.

In our previous paper we solved the problem of automatic segmentation of microscopy images of Langerhans islets.[2–4]

The islets contained in segmentation often touch and appear as a single object which gives incorrect results. The separation of touching objects can be performed by watershed transform applied in combination with distance transform computed at binary segmentation.[5] In[6] authors solved the problem with touching elliptically shaped grains, where the distance transform failed. They apply skeletonisation which results in open lines. These open lines are then prolonged according to their direction from corresponding end points. If the distance between opposite two lines is under set threshold value then these lines are connected. In[7] authors proposed a region selection-based algorithm for automatic touching cell segmentation on stained muscle images. Firstly, the algorithm generates a set of segmentation candidates and then assigns these candidates proper scores based on a learned cell shape model and local features. Finally, a subset of region candidates is selected as final segmentation based on an Integer Linear Programming scheme under the constraint that no two selected regions can overlap. Arteta et al[8] introduce separate classes for tuples of objects into the detection process. It can be shown that this formulation can be learned within the structured output SVM framework, while the learning only requires weak annotations – a dot on each instance. Funke et al[9] proposed method based on extraction of a merge tree of candidates (i.e., parent nodes are merged segments of their children) that span the whole range from over–

Further author information: (Send correspondence to J. Švihlík)
J. Švihlík: E-mail: jan.svihlik@fel.cvut.cz, Telephone: +420 224 357 590

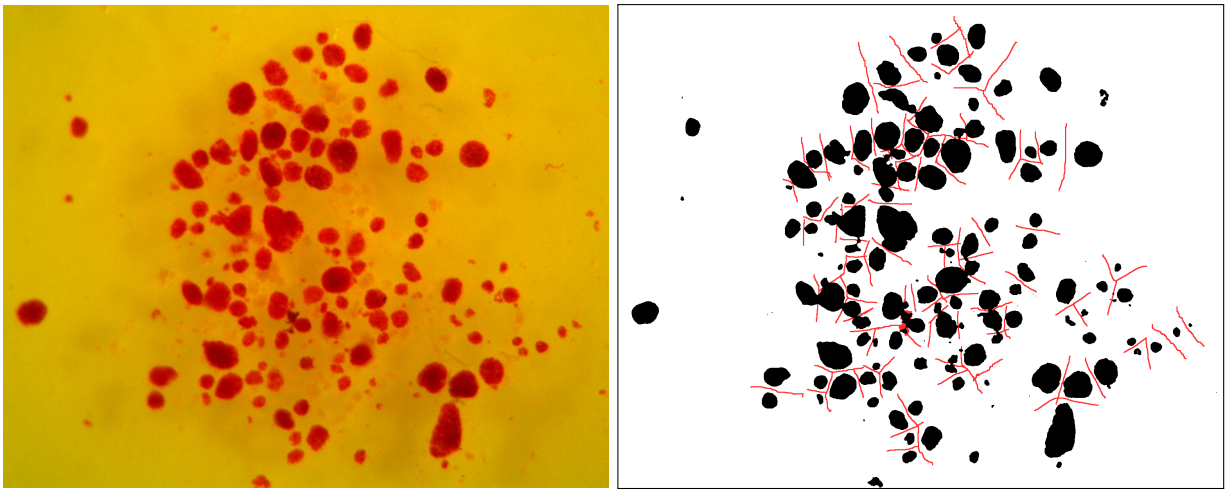
to undersegmentation. A segmentation of the image can now be expressed by a selection of non-overlapping candidates.

Our method consists of several steps. First, we segment islets and background using our algorithm based on logistic regression and RGB features.[2,3] Then, possible divisions of the individual foreground regions are found using the watershed transform. For each possible division, several shape parameters are evaluated. Finally, the most likely division according to the shape model is chosen.

2. METHOD

2.1 Islets separation using watershed transform

First, we segment (logistic regression, RGB features),[2] see Fig. 1 (b), microscopy images (Fig. 1 (a)) of Langerhans islets into 2 classes (background + exocrine tissue/islets) and if necessary we apply color normalization.[3] The exocrine tissue is unwanted part of prepared sample. The ratio between amount of exocrine tissue and amount of islet tissue determines a purity of sample. In this case the purity evaluation is not considered. The binary



(a) Microscopy images (b) Segmentation overlaid by separating mask

Figure 1. Microscopy image and segmentation overlaid by separating mask.

segmentation is then over-segmented using watershed transform. The over-segmentation cannot be seen as a final result, because several single islets are divided into more parts.

2.2 Computation of descriptors and SVM classifier training

Descriptors[10] are selected according to satisfactory scatter-plot. The selected descriptors are described below. Hu invariants L_{23} and L_{24} are based on central moments of inertia. L_{23} could be seen as a measure of circularity, because the minimal value $1/(2\pi)$ is obtained for circular shape.[11] L_{23} and L_{24} are given by

$$L_{23} = \eta_{20} + \eta_{02}, \tag{1}$$

$$L_{24} = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \tag{2}$$

where η_{pq} is given by

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{\frac{p+q}{2}+1}}, \tag{3}$$

and μ_{pq} is given by

$$\mu_{pq} = \sum_{\forall x} \sum_{\forall y} (x - x_c)^p (y - y_c)^q f(x, y), \tag{4}$$

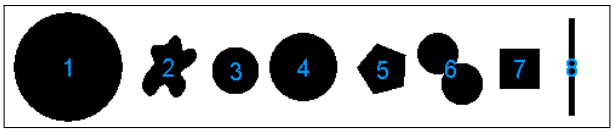


Figure 2. Illustration of L_{23} values obtained for different object shapes, (1) 0.1592, (2) 0.1928, (3) 0.1591, (4) 0.1591, (5) 0.1617, (6) 0.2233, (7) 0.1665, (8) 1.3381.

where x_c and y_c are centroids and $f(x, y)$ denotes an image function (in our case image is binary). The values of L_{23} computed for differently shaped objects can be seen in Fig. 2. From the computed values it can be seen that approximation of circular (objects: 1, 3, 4) areas approaching $L_{23} \approx 1/(2\pi)$. Similarly as L_{23} a compactness[12] is chosen to effectively identify touching islets. The compactness is given by

$$\mathcal{C} = \frac{\mathcal{L}^2}{area}, \tag{5}$$

where $\mathcal{L}$ denotes the length of object border. The shape descriptors are computed on the single islets representing the 1. class and the merged islets representing the 2. class. For every segmentation our medical experts create separating ground truth (see Fig. 1 (b)). It means that merged islets are identified and manually divided into single islets. To overcome unsatisfactory number of merged islets we generate artificial couple and triplet using randomly chosen single islets. Computed shape descriptors are then used in the process of SVM classifier (RBF kernel) training.

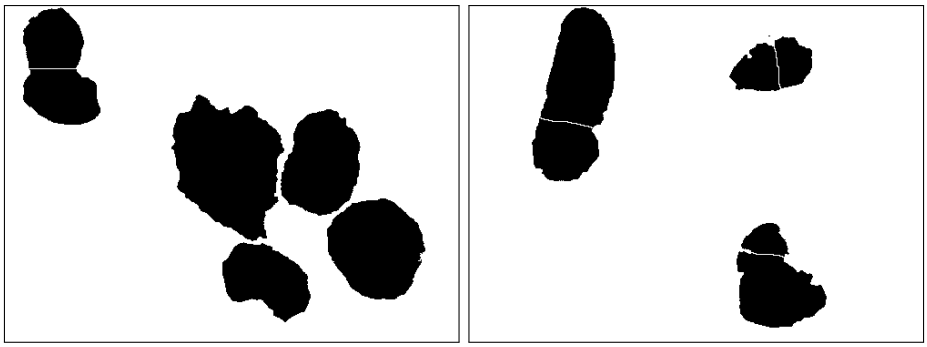
2.3 Object classification

We consider such objects which were before application of watershed transform a member of any group of merged islets. We combine separated objects together and evaluate descriptors for such artificially merged objects. Probability scores (weighted by object area) given by trained SVM classifier are then used to find the most probable merging combination of separated objects (every labeled object in segmentation processed by watershed transform). For instance for the case of 2 objects of areas A_1 and A_2 and probability scores $P_1(object_1|\theta_\mathbf{1})$ and $P_2(object_2|\theta_\mathbf{2})$ we solve the following inequality

$$A_1 P_1(object_1|\theta_\mathbf{1}) + A_1 P_2(object_2|\theta_\mathbf{2}) \lesseqgtr (A_1 + A_2) P_{12}(object_{12}|\theta_\mathbf{12}), \tag{6}$$

where $P_{12}(object_{12}|\theta_\mathbf{12})$ denotes the probability that 2 merged objects form islet and $\theta_\mathbf{1}$, $\theta_\mathbf{2}$, $\theta_\mathbf{12}$ are the vectors of descriptors computed for corresponding object. Based on the solution of equation (6) the proposed algorithm lets the objects separated or join objects together.

An example of correctly separated merged islets and over-segmented single islets can be seen in Fig. 3.



(a) Correctly separated merged islets (b) Over-segmented single islets

Figure 3. Left: correctly separated merged islets using watershed transform, right: application of watershed transform results in over-segmented single islets.

3. RESULTS

3.1 Image data

There are 46 images acquired during the past years in our image database. The first 12 images contain considerable amount of touching islets. These images are segmented using logistic regression classifier and RGB color channel features. Our medical experts create "separating masks" (see Fig. 1 (b)) which enable us to correctly separate islets in the mentioned set of segmentations. The segmentations of separated islets serve as the ground truth and allow us to evaluate the quality of optimized islet separation generated by the proposed algorithm.

3.2 Evaluation of pixel to pixel error

We evaluate pixel to pixel error between segmentations optimized by proposed algorithm and ground truth segmentations (islets separated by medical experts). First, both segmentations are labeled $A: \ \mathbf{x} \rightarrow a(\mathbf{x}) \in \mathbf{L_1}$, $G: \ \mathbf{x} \rightarrow g(\mathbf{x}) \in \mathbf{L_2}$ then a mapping φ between these two label sets $\mathbf{L_1}$, $\mathbf{L_2}$ is found. Islets are labeled using algorithm published in.[13] Every object is labeled using successive integer numbers, background is denoted as "0". We consider 8-connected objects. The pixel to pixel error E can be written as follows

$$E = \min_{\varphi} \sum_{x} [\varphi(a(\mathbf{x})) \neq g(\mathbf{x})]. \tag{7}$$

The values of pixel to pixel error evaluated in accordance with eq. (7) are in Tab. 1. The pixel to pixel error displayed graphically can be seen in Fig. 4 and selected details of segmentations in Fig. 5. There can be seen in Fig. 4 how differently shaped could Langerhans islets be. So that it is not so simple even for experienced medical experts to decide where the separation line is.

Image in Fig. 5 (a) shows 3 pairs of touching islets. The same 3 pairs of touching islets are correctly separated in Fig. 5 (b). However, the specially shaped islet placed at the bottom of image is over-segmented. The over-segmented islet is then merged back using proposed algorithm, see Fig. 5 (c). Due to complicated shapes of certain small islets, there are still some islets over-segmented after the application of the proposed algorithm. However, the smallest islets with diameter under 50 μm are not considered in the process of islet transplantation.

Descriptors	L_{23}, ***area***	L_{23}, L_{24}	L_{23}, C	C	L_{23}	**segmentation**	**watershed**
mean	4.7481	3.2597	3.6795	4.3834	3.4919	5.6016	3.8093
std	3.1585	2.0816	2.8065	4.5185	2.2270	3.5037	2.3767

Table 1. Pixel to pixel error E [%], leave-one-out cross-validation applied to 12 testing images.

3.3 Testing of trained classifier on a new set of images

We also use SVM classifier with L_{23} and L_{24} descriptors trained on 12 testing images to optimize segmentation of another set of 10 images. However these images contain considerably roughly shaped islets in comparison with training set. The values of pixel to pixel error are in Tab. 2. Selected details of the pixel to pixel error displayed graphically can be seen in Fig. 6.

Descriptors	L_{23}, L_{24}	**segmentation**	**watershed**
mean	3.4184	3.8794	3.7463
std	1.3418	2.1806	1.7865

Table 2. Pixel to pixel error E [%], SVM classifier trained on 12 testing images applied to another set of 10 images.

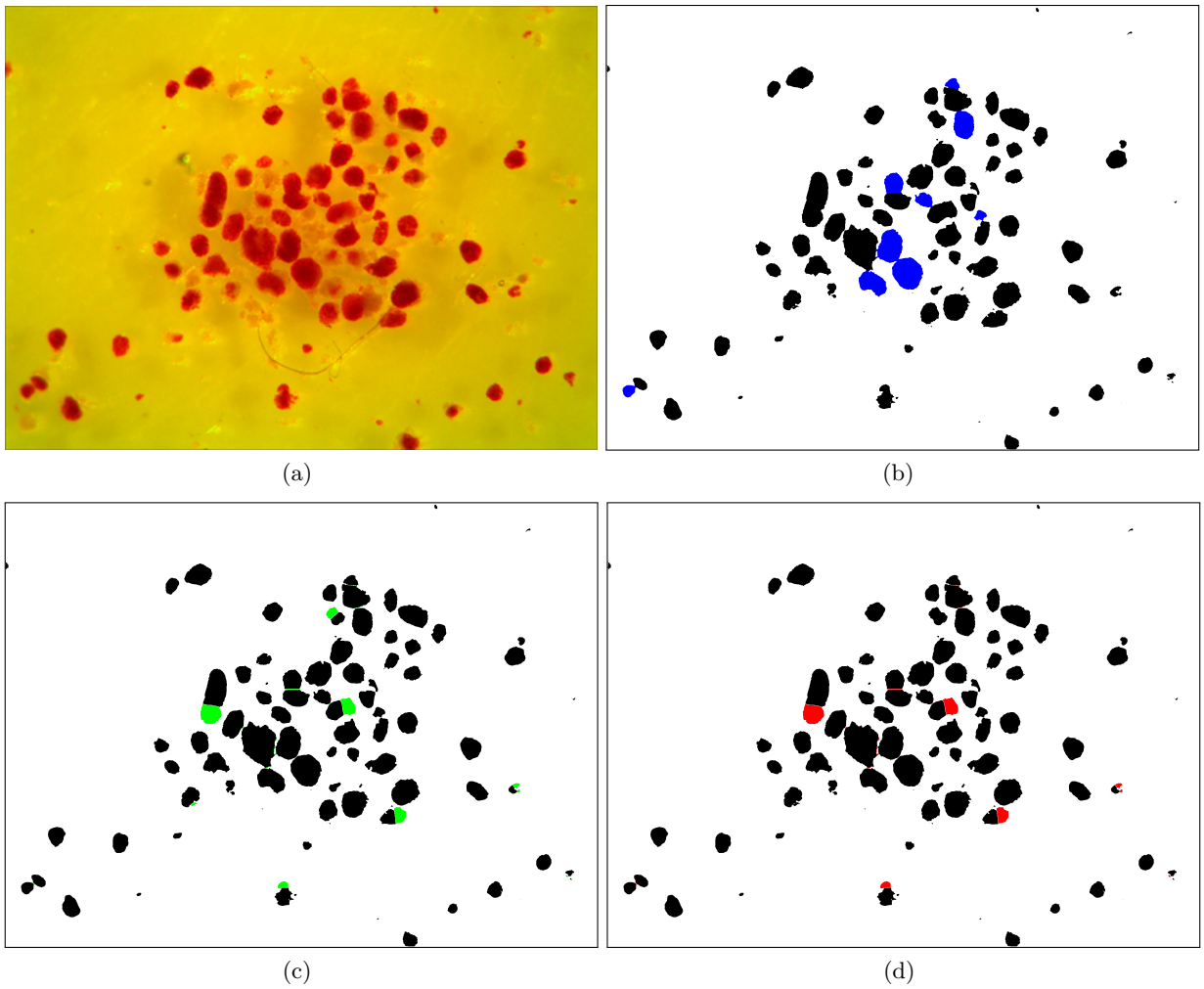


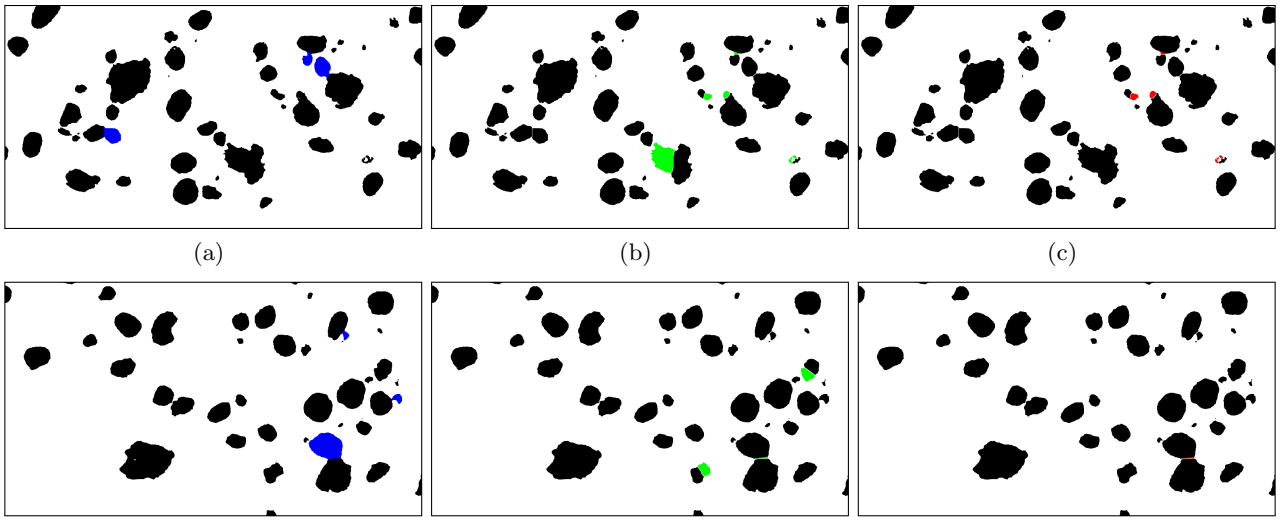
Figure 4. Evaluation of separation quality (a) original image, (b) segmentation given by pixel-wise classifier (error area - blue), (c) watershed transform (error area - green), (d) segmentation optimized by proposed algorithm (error area - red)

4. CONCLUSION

In this paper we presented the fully automatic method which is able to separate touching islets with the same quality as medical experts. The best results were obtained for the descriptors L_{23} and L_{24}. These results are better even in comparison with the results given by watershed transform. The results are also fully comparable with the ground truth created by medical experts. However, we found that islets are considerably differently shaped and sized. Hence, in the future work new set of segmentations where the merged islets will be separated by medical experts must be created. We also plan to test other shape descriptors. For comparison, a medical experts takes about 20 minutes to separate merged islets in a single image, whereas the proposed algorithm takes about 7 minutes (implemented in Matlab, 2 x Intel Xenon 2.53 GHz, 20 GB RAM).

ACKNOWLEDGMENTS

This work has been supported by the grant 14-10440S "Automatic analysis of microscopy images of Langerhans islets" of the Czech Science Foundation.



(a) (b) (c)

(d) Selected detail of segmentation given by classifier

(e) Watershed transform applied

(f) Segmentation optimized by proposed algorithm

Figure 5. Evaluation of separation quality (a, d) detail of segmentation given by pixel-wise classifier (error area - blue), (b, e) watershed transform (error area - green), (c, f) detail of segmentation optimized by proposed algorithm (error area - red).

(a) (b) (c)

Figure 6. Evaluation of separation quality (a) detail of segmentation given by pixel-wise classifier (error area - blue), (b) watershed transform (error area - green), (c) detail segmentation optimized by proposed algorithm (error area - red)

REFERENCES

[1] Alejandro, R., Barton, F. B., Hering, B. J., Wease, S., and Investigators, C. I. T. R., "Update from the collaborative islet transplant registry," *Transplantation* **86**(12), 1783–1788 (2008).

[2] Švihlík, J., Kybic, J., Habart, D., Berková, Z., Girman, P., Kříž, J., and Zacharovová, K., "Classification of microscopy images of Langerhans islets," in [*Medical Imaging 2014: Image Processing*], **9034**, 1–8, SPIE, Bellingham, USA (March 2014).

[3] Švihlík, J., Kybic, J., and Habart, "Color normalization for robust evaluation of microscopy images," in [*Applications of Digital Image Processing XXXVIII*], **9599**, 1–6, SPIE, Bellingham, USA (2015).

[4] Schindelin, ., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., K.Eliceiri, Tomancak, P., and Cardona, A., "Fiji: an open-source platform for biological-image analysis," *Nature Methods* **9**(7), 676–682 (2012).

[5] Gonzalez, C. and Woods, R. E., [*Digital Image Processing*], Prentice-Hall (2008).

[6] Faessel, M. and Courtois, F., "Touching grain kernels separation by gap-filling," *Image Analysis and Stereology* **28**, 195–203 (2009).

[7] Liu, F., Xing, F., and Yang, L., "Robust muscle cell segmentation using region selection with dynamic programming," in [*IEEE 11th International Symposium on Biomedical Imaging, ISBI 2014*], 521–524, SPIE (2014).

[8] Arteta, C., Lempitsky, V., Noble, J., and Zisserman, A., "Learning to detect partially overlapping instances," in [*IEEE Conference on Computer Vision and Pattern Recognition*], 3230–3237, IEEE (2013).

[9] Funke, J., Hamprecht, A., and Zhang, C., "Learning to segment: Training hierarchical segmentation under a topological loss," in [*18th International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI 2015*], **9351**, 268–275 (2015).

[10] Goshtasby, A., [*Image Registration*], Springer-Verlag (2012).

[11] Žunić, J., Hirota, K., and Rosin, P. L., "A HU moment invariant as a shape circularity measure," *Pattern Recognition* **43**, 47–57 (2010).

[12] Šonka, M., Hlaváč, V., and Boyle, R., [*Image Processing, Analysis, and Machine Vision*], Springer (2007).

[13] Haralick, M. R. and Shapiro, L. G., [*Computer and Robot Vision, Volume I*], Addison-Wesley (1992).