Classification of microscopy images of Langerhans islets

Jan Švihlík[a], Jan Kybic[a], David Habart[b], Zuzana Berková[b], Peter Girman[b], Jan Kříž[b], Klára Zacharovová[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

ABSTRACT

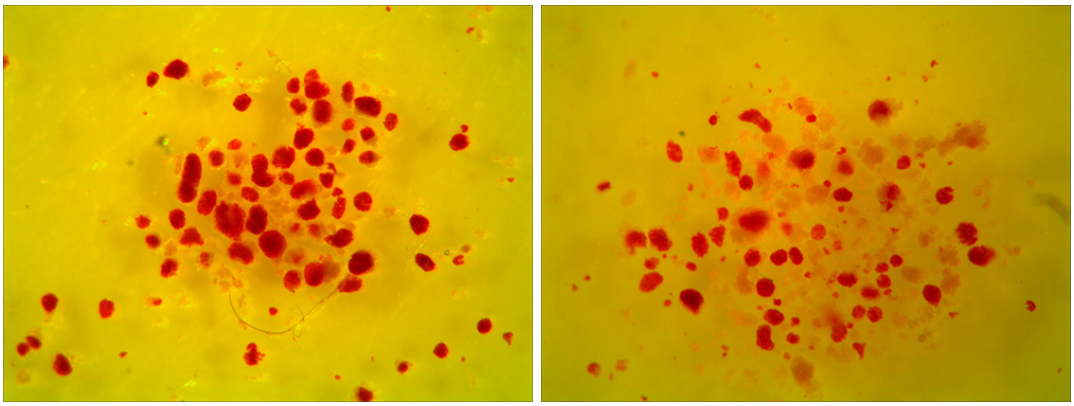
Evaluation of images of Langerhans islets is a crucial procedure for planning an islet transplantation, which is a promising diabetes treatment. This paper deals with segmentation of microscopy images of Langerhans islets and evaluation of islet parameters such as area, diameter, or volume (IE). For all the available images, the ground truth and the islet parameters were independently evaluated by four medical experts. We use a pixelwise linear classifier (perceptron algorithm) and SVM (support vector machine) for image segmentation. The volume is estimated based on circle or ellipse fitting to individual islets. The segmentations were compared with the corresponding ground truth. Quantitative islet parameters were also evaluated and compared with parameters given by medical experts. We can conclude that accuracy of the presented fully automatic algorithm is fully comparable with medical experts.

Keywords: Langerhans islet, linear classifier, support vector machine, IE, segmentation

1. INTRODUCTION

In the Czech Republic, there is an advanced transplantation program for diabetic patients. While the transplantation of pancreas usually results in a longterm cure of diabetes, it has some associated risks and is not available to all patients due to a lack of suitable donors. The alternative and successfully developing method is the microorgan transplantation of isolated islets of Langerhans (Ricordi,[1] 1990). This program was initiated at IKEM in 2006. A laboratory has to isolate the islets from a donor. Before offering it for clinical transplantation, the graft suspension has to meet certain quantitative and qualitative criteria. The basic criterion is the total amount of islets tissue. The graft represents suspension of tissue particles, which is considerably heterogeneous as to the size (25 - 700 μm) and shape (e.g. spheroid, rotational ellipsoid, irregular). For this reason a volumetric unit has been designed and termed the islet equivalent (IE), which is the volume an ideal islet of 150 μm diameter. The additional measures include the isolation index (islet volume in islet equivalents divided by the islet number), the islet size distribution histogram and the graft purity expressed as a percentage of the remaining nonislet tissue. Obtaining a representative sample for islet counting is complicated by the fact that the particles of different sizes, shapes and composition sediment with different velocities. Multiple samples therefore need to be evaluated. However, the time consumption prevents evaluating multiple samples even when assisted by digital image processing. In our laboratory, single sample evaluation takes about 15 minutes whether done manually or assisted by an image analysis software, depending on the number of islets in the sample, their purity and the extent of overlay of individual islets. After extensive training the interoperator variance among four operators using our software dropped to CV 0.08 (Habart,[2] 2013). Other groups reported variation ranging CV 0.02-0.09 among 3-35 persons (Lehman,[3] 1998, Niclauss,[4] 2008, Kissler,[5] 2010, Friberg,[6] 2011). The other problem is that digital image analysis methods are not standardized between laboratories. An example of the acquired image of Langerhans islets could be seen in Fig. 1. The task is to determine the number, area, volume (in islet equivalents), purity and other parameters of the islets, in order to determine the suitability of the graft for transplantation.

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



(a) (b)

Figure 1. Example of images of Langerhans islets, 2048 × 1536 pixels

2. STATE-OF-THE-ART

The gold standard for islet quantification is to use microscope with calibrated grid for islet diameter measurement (Ricordi,[1] 1990) and to estimate their volume from the histogram of the islet diameters. This visual determination is considerably operator dependent (see Fig. 5). The measured diameters of islets are divided into classes with 50 μm increment (Ricordi,[1] 1990). From a distribution of islet diameters, islet equivalents (IE) are evaluated. IE corresponds to islets of an average size of 150 μm. The islets size distribution can be fitted by a probability density function (PDF) model (Buchwald,[7] 2009) such as Weibull and lognormal models. To overcome the problems with manual counting method based on microscope operators, a number of methods utilizing digital image analysis were proposed. The one of the first image analysis approaches (Wile,[8] 1994) exploits analog camera with a grabber and image processing software. The method was to identify islets, eliminate all islets smaller than 20 μm in diameter, and calculate the volumes per islet from pixel area (spherical islet is considered). In (Stegenmann et al, 1998), under bright field illumination, the image was thresholded and the total islet area determined from the binarized image data. Other methods are based on fitting a circle or an ellipse to the islets (Girman,[9] 2003). In most cases, the spherical and ellipsoidal shape of islets are assumed (Stegenmann,[10] 1998), which increases the islet volume estimation error. Then, the ellipse is transformed into a 3D ellipsoid (authors consider prolate ellipsoid, i.e. ellipsoid generated by rotation about the major axis of an ellipse). In (Niclauss et al,[4] 2008), authors utilize prolate volume for 3D islet approximation. In medical literature, the algorithms are mentioned very briefly or not at all (Friberg et al,[6] 2011).

It is possible to directly evaluate the content of extractable insulin or the content of zinc (Stegenmann et al, 1998). However, these techniques turn out to perform poorly in predicting the transplantation success, because the content of the mentioned substances varies from donor to donor. Pisania et al.,[11] 2010, evaluated the number of islet nuclei taking advantage of their homogeneity. Other methods are based on estimation of the amount of tissue from the extracted DNA (Stegenmann et al, 1998) or the total protein contents.

3. METHOD

3.1 Segmentation

We use a pixelwise linear classifier[12] (perceptron algorithm) and a linear SVM (support vector machine)[12] trained using 2420 training samples. Training samples were selected (20 selection for both classes: islet, background) by user in the images using a rectangular selection frames pointed by mouse, see Fig. 2. We used RGB and Lab color space. The vectors of RG, RGB and Lab channels were served as the features. Each rectangular selection contains 11 x 11 pixels, which are transformed to the vector of 121 training samples. For the sake of simplicity, we trained classifiers on one randomly chosen image only. The rest of the images was then segmented.

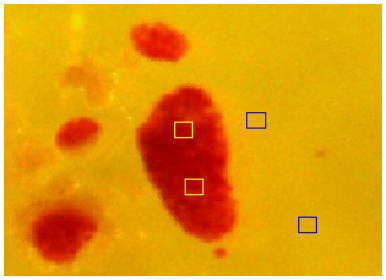


Figure 2. Rectangles used for manual selection of training samples

range [μm]	$(0, 50\rangle$	$(50, 100\rangle$	$(100, 150\rangle$	$(150, 200\rangle$	$(200, 250\rangle$	$(250, 300\rangle$	$(300, 350\rangle$	$(350, \infty)$
$\boldsymbol{\lambda}$	0.000	0.167	0.648	1.685	3.500	6.315	10.352	15.833

Table 1. Conversion factors of the corresponding bins of Ricordi's histogram

3.2 Object detection and counting

To identify individual pixels, we applied method for object labeling and counting to the binary segmentations. The used method is based on the algorithm published in.[13] This algorithm gives the same results as the implemented plugin for FIJI, see section 3.3.1. Every object is labeled using successive integer numbers, background is denoted as "0". We consider 8-connected objects. The areas of islets correspond to the number of constant value labels. The example of image, corresponding segmentation, image overlayed by segmentation and labeled segmentation can be seen in Fig. 3.

3.3 Islet equivalent estimation and ellipse fitting

Islets with diameter smaller than 50 μm were ignored (i.e. substituted islets for background). In accordance with Ricordi method,[7] we calculated diameter from islet areas, whereas the circular shape of islet is assumed. Hence, let A_I be the islet area in pixels and let p be the size of pixel side in image then the islet diameter d_I is given by $d_I = 2\sqrt{p^2 A_I/\pi}$ [μm].

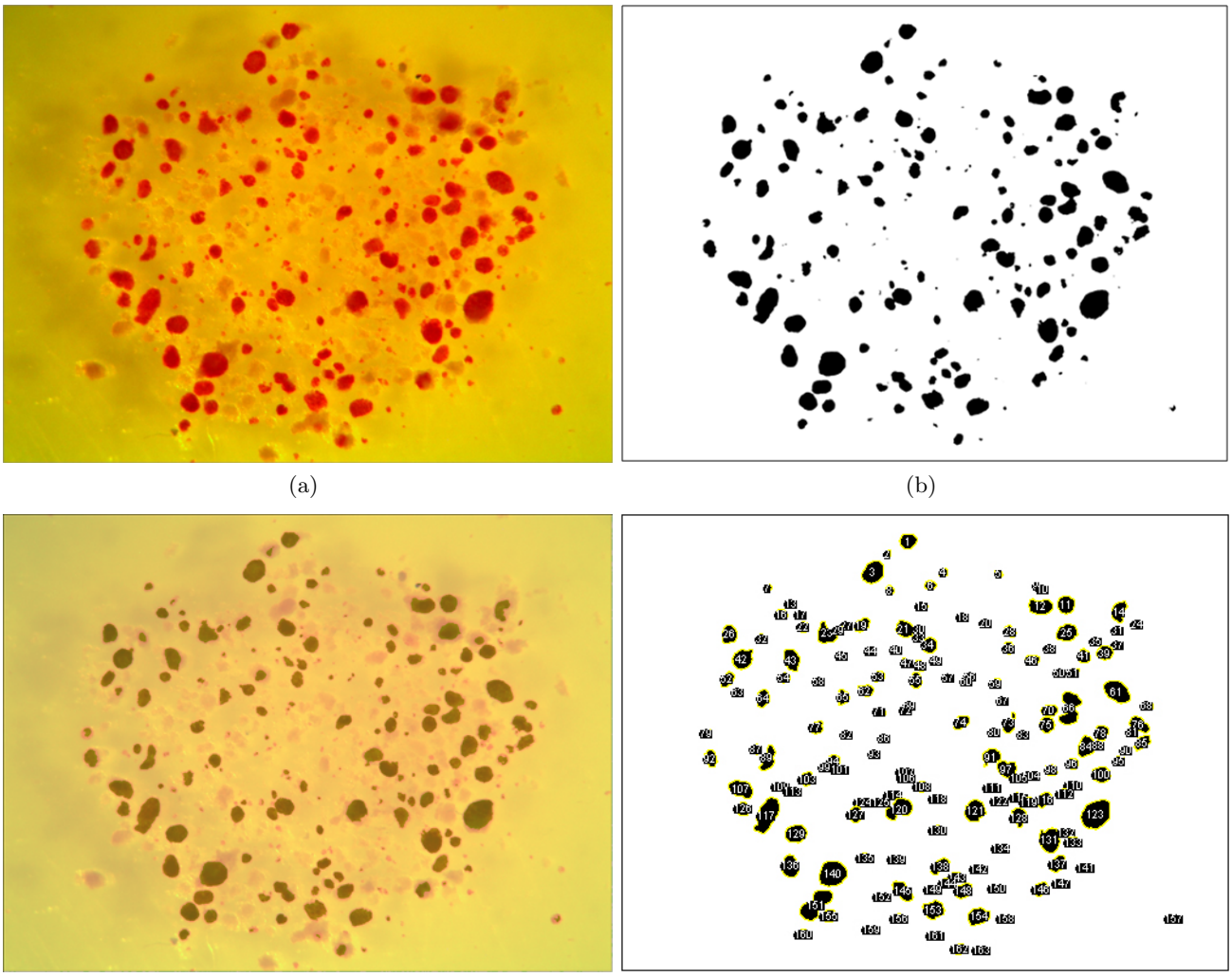
Two approaches for the calculation of IE were applied. The first approach is based on the method proposed by Ricordi.[1] In this method the islets are considered to be circular, therefore the area of every islet is evaluated and then the diameter is derived. These calculated diameters of islets are divided into classes with 50 μm increment, see Tab. 1. From a distribution of islet diameters, islet equivalents (IE) are evaluated. Let N be the frequencies in the Ricordi's bins then the IE_{tot} can be written as

$$IE_{tot} = \sum_k N_k \lambda_k, \quad k = 1, 2, \dots K, \tag{1}$$

where λ_k denotes conversion factor for bin N_k and K is a number of bins. Parameter λ_k is given by

$$\lambda_k = \frac{V_\phi}{V_{150}} = \frac{V_k^{(min)} + V_k^{(max)}}{2 \cdot V_{150}} = \frac{\left(d_k^{(min)}\right)^3 + \left(d_k^{(max)}\right)^3}{2 \cdot 150^3}, \quad k = 1, 2, \dots K, \tag{2}$$

where V_{150} is volume of sphere of diameter 150 μm, V_ϕ denotes mean of volumes, where the first volume $V_k^{(min)}$ is computed using the minimum diameter $d_k^{(min)}$ and the second volume $V_k^{(max)}$ using the maximum diameter $d_k^{(max)}$ of the bin range. For the last bin $(350, \infty)$ of Ricordi's histogram it is assumed that $d_k^{(max)} = 400$. The second approach is based on ellipse fitting,[9] where the ellipse is fitted using method based on moments.[14] Every islet we consider to be ellipse with one-half major axis a and one-half minor axis b then its area is given by $A_e = \pi ab$. Instead of diameter in the case of circle, we use Ricordi's histogram of average length of ellipse's axes $L_{avg} = a + b$. The islet equivalents are then calculated using the same equation (1) as in the previous approach. Correctly, it should be used geometric mean $L_{avg} = \sqrt{ab}$. However, arithmetic mean is implemented in software used by our medical experts.



(a) (b)

(c) (d)

Figure 3. (a) Image, (b) segmentation, (c) image overlayed by segmentation, (d) labeled segmentation

3.3.1 Plugin for FIJI

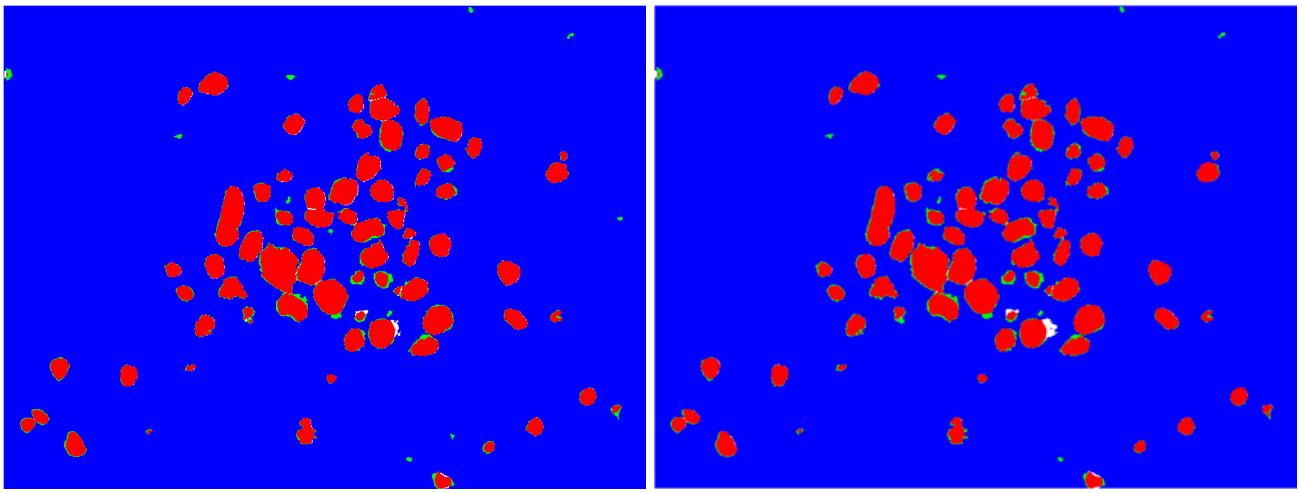
The plugin for FIJI[15] was programmed in Jython language and it allows to analyze of segmented microscopic images of Langerhans islets. The circles and ellipses are fitted[14] to islets and their estimated parameters are used to evaluate volumes of spheres and ellipsoids. Islet equivalents (IE) are estimated using three approaches (two approaches were used in this paper). The plugin is available at*.

4. RESULTS

4.1 Segmentation accuracy

We tested proposed algorithm on the set of 12 microscopy images acquired by camera Bresser: 3.0 MP, CMOS 1/2", 2048 × 1536 pixels. To evaluate the performance of segmentation algorithm, all images of Langerhans islets with ground truth segmentation were segmented and the obtained results were compared with ground truth created by four medical experts. We used the statistical characteristics typical for binary classification, e.g. TP (true positive), FP (false positive), TN (true negative), FN (false negative) and related measures: $specificity$

*https://sites.google.com/site/svihlikja/codes



(a) (b)

Figure 4. Comparison of segmentation with ground truth segmentation, linear classifier (a) RG features, (b) Lab features, True Positive (RED), True Negative (BLUE), False Positive (WHITE), False Negative (GREEN)

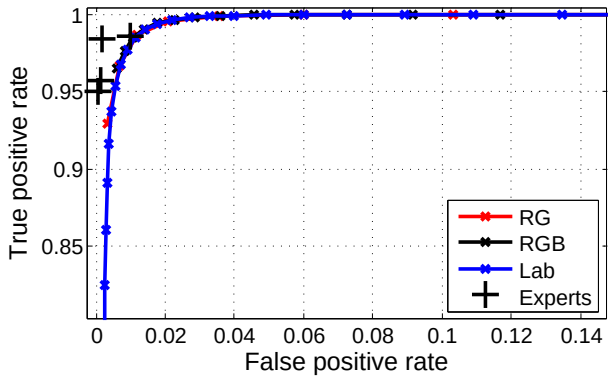
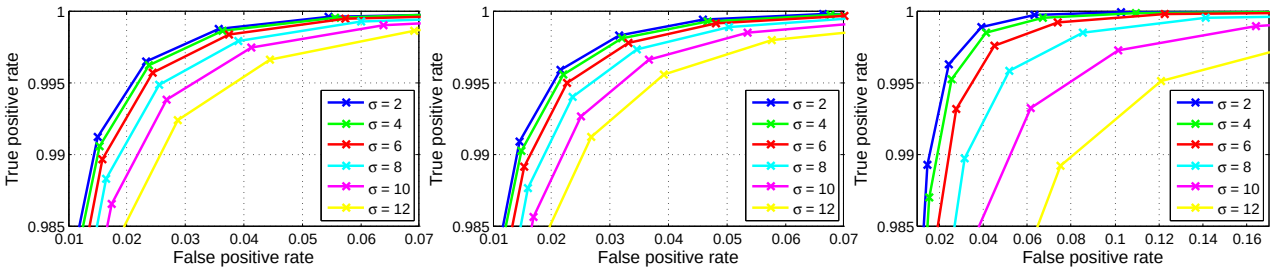


Figure 5. ROC curve computed for all testing images (linear classifier). We show the performance of the automatic system and performance of individual experts

and *sensitivity*. Each of four experts created 3 ground truth segmentations for each image using own software.[9] The ground truth segmentations were obtained via major vote rule.[16] We evaluate expert performance, so that each expert segmentation was compared with major vote rule of the segmentation of others. This shows us if our algorithm has a comparable performance with medical experts. Comparison of segmentation with ground truth segmentation can be seen in Fig. 4. The results presented in this chapter were obtained with linear classifier. ROC curve computed for all testing images, RG, RGB, Lab features and experts can be seen in Fig. 5.

4.1.1 Testing of algorithm robustness: noise contamination

In this section we present the results of segmentation of images of Langerhans islet contaminated by zero mean additive white Gaussian noise with increasing standard deviation ($\sigma = 2 - 12$). ROC curve computed for all testing images, RG, RGB and Lab features can be seen in Fig. 6. From the ROC curves we can concluded that utilization of Lab features causes that the algorithm is less robust in comparison with RG and RGB features



(a) (b) (c)

Figure 6. Contamination of zero mean additive white Gaussian noise, ROC curve computed for all testing images, linear classifier, (a) RG features, (b) RGB features, (c) Lab features

4.2 Comparison of islet parameters

The results obtained with proposed algorithm we compared with results created by four medical experts. All images were evaluated by every expert three times at different times. The IE_{tot} calculated by four experts were averaged and compared with $\widehat{IE_{tot}}$ calculated by our plugin.

We decided to use absolute and relative error to compare the total IE calculated by medical experts IE_{tot} with the total IE calculated by proposed algorithm $\widehat{IE_{tot}}$. The absolute error ε is given by $\varepsilon = |IE_{tot} - \widehat{IE_{tot}}|$ and relative error in percents is given by $\delta = \left|1 - \widehat{IE_{tot}}/IE_{tot}\right| \cdot 100$ [%]. Example of the table of computed characteristics can be seen in Tab. 2. This tables also contains SD (standard deviation) calculated from 4 averaged measurement of Total IE of medical experts.

Test image	1	2	3	4	5	6	7	8	9	10	11	12
IE_{tot} [-]	398	421	798	215	306	539	161	460	646	271	298	549
$\widehat{IE_{tot}}$ [-]	384	400	771	190	293	468	190	391	604	268	290	564
δ [%]	3.5	5.0	3.4	11.5	4.2	13.2	18.3	15.0	6.5	1.3	2.8	2.8
ε [-]	14.0	21.2	27.4	24.7	12.8	70.9	29.4	69.2	41.9	3.4	8.4	15.1
SD [-]	16.8	16.3	75.6	25.7	13.7	21.3	16.4	31.2	77.1	40.4	19.5	48.2
ε/SD	0.8	1.3	0.4	1.0	0.9	3.3	1.8	2.2	0.5	0.1	0.4	0.3

Table 2. Relative and absolute error of total IE, spherical islet, linear classifier, RGB features, mean relative error $\overline{\delta} = 7.2$ [%]

4.2.1 Influence of color space

From the performed measurements we chose the best parameters (method, features, classifier). We always set two parameters to be constant and the third parameters was changing to find its statistical significance.

We tested if the relative error δ and absolute error ε computed for used color spaces differ significantly. From the performed single factor ANOVA test at $\alpha = 0.05$ we can concluded that obtained errors does not differ significantly. We recommend to use RG and RGB features, because Lab features are less robust against the noise. The evaluated errors averaged over all images can be seen in Tab. 3.

4.2.2 Influence of classifier

We tested if the relative error δ and absolute error ε computed from the islet equivalent obtained from segmentation given by linear classifier (perceptron) and linear SVM differ significantly. From the performed t-tests at $\alpha = 0.05$ we can concluded that obtained errors does not differ significantly. Hence, at given α the results obtained by both classifiers are comparable. The evaluated errors averaged over all images are for both classifiers in Tab. 4.

Classifier	Method	Space	δ [%]	ε [-]	SD [-]	ε/SD
SVM	Ellipse	**RG**	8.3	28.3	40.2	0.9
		RGB	8.4	30.0	40.2	1.0
		Lab	8.2	25.0	40.2	0.9

Table 3. Influence of color space: computed mean relative errors δ and mean absolute errors ε for all used color spaces

Space	Classifier	Method	δ [%]	ε [-]	SD [-]	ε/SD
Lab	**SVM**	Ellipse	8.2	25.0	40.2	0.9
	Perceptron		11.4	36.7	40.2	1.2

Table 4. Influence of classifier: mean relative errors δ and mean absolute errors ε computed from errors averaged over all images

4.2.3 Influence of method: circle vs. ellipse

We tested if the relative error δ and absolute error ε computed for islet shapes modeled by circle and islet shapes modeled by ellipse differ significantly. From the performed t-tests at $\alpha = 0.05$ we can concluded that obtained δ errors does not differ significantly. However, at $\alpha = 0.05$, the absolute error ε differs significantly. The evaluated errors averaged over all images can be seen in Tab. 5. We recommend to use modeling of islet volume using the ellipsoid.

Space	Classifier	Method	δ [%]	ε [-]	SD [-]	ε/SD
Lab	SVM	**Circle**	12.0	51.9	33.5	1.8
		Ellipse	8.2	25.0	40.2	0.9

Table 5. Influence of method: mean relative errors δ and mean absolute errors ε computed from errors averaged over all images

5. CONCLUSION

We proposed fast fully automatic method which is able to analyze microscopy images of Langerhans islets and evaluate IE as good as medical experts. A medical experts takes about 15 minutes to evaluate (segmentation and IE computation) a single image. This time depends mainly on experiences of medical expert, light conditions during image acquisition and content of exocrine tissue. In the case of proposed algorithm, the time necessary for image segmentation and islet parameters calculation takes up to 10 seconds. The time necessary for training of classifier depends on the method used for it. In our case it takes about 30 seconds.

ACKNOWLEDGMENTS

This work has been supported by the grant P202/11/0111 "Automatic analysis of light and electron microscopy neuronal data" of the Czech Science Foundation.

REFERENCES

[1] Ricordi, C., Gray, D. W., Hering, B. J., and et al., "Islet isolation assessment in man and large animals," *Acta Diabetologica Latina* **27**(3), 185–195 (1990).

[2] Habart, D., Girman, P., Zacharovova, K., Berkova, Z., Kriz, J., and Saudek, F., "Reproducibility of human islet counting by a digital image analysis," in [*Abstract Book of 3rd joined AIDPIT/EPITA winter symposium*], (2013).

[3] Lehmann, R., Fernandez, L. A., Bottino, R., Szabo, S., Ricordi, C., Alejandro, R., and Kenyon, N. S., "Evaluation of islet isolation by a new automated method (coulter multisizer ile) and manual counting," *Transplantation Proceedings* **30**(7), 373–374 (1998).

[4] Niclauss, N., Sgroi, A., Morel, P., Baertschiger, R., Armanet, M., Wojtusciszyn, A., and et al, "Computerassisted digital image analysis to quantify the mass and purity of isolated human islets before transplantation," *Transplantation* **86**(11), 1603–1609 (2008).

[5] Kissler, H. J., Niland, J. C., Olack, B., Hering, C. R. B. J., Naji, A., and et al., "Validation of methodologies for quantifying isolated human islets: an islet cell resources study," *Clinical Transplantation* **24**(2), 263–242 (2010).

[6] Friberg, A. S., "Quantification of the islet product: Presentation of a standardized current good manufacturing practices compliant system with minimal variability," *Transplantation* **91**(6), 677–683 (2011).

[7] Buchwald, P., Wang, X., Khan, A., and et al, "Quantitative assessment of islet cell products: Estimating the accuracy of the existing protocol and accounting for islet size distribution," *Cell Transplantation* **18**(10-11), 1223–1235 (2009).

[8] Wile, K. J., Fetterhoff, T. J., Coffing, D., Cavanagh, T. J., and Wright, M. J., "Morphologic analysis of pancreatic islets automated image analysis," *Transplantation Proceedings* **26**(6), 3441 (1994).

[9] Girman, P., Kriz, J., and Saudek, F., "Digital imaging as a possible approach in evaluation of islet yield. cell transplantation," *Transplantation* **12**(2), 129–133 (2003).

[10] Stegemann, J. P., O'Neil, J. J., Nicholson, D. T., and Mullon, C. J., "Improved assessment of isolated islet tissue volume using digital image analysis," *Cell Transplantation* **7**(5), 469–478 (1998).

[11] Pisania, A., Papas, K. K., Powers, D. E., Rappel, M. J., Omer, A., Bonner-Weir, S., and et al., "Enumeration of islets by nuclei counting and light microscopic analysis. laboratory investigation," *Journal of Technical Methods and Pathology* **90**(11), 1676–1686 (2010).

[12] Schlesinger, M. I. and Hlavac, V., [*Ten Lectures on Statistical and Structural Pattern Recognition*], Springer Netherlands, first ed. (2002).

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

[14] Mulchrone, K. F. and Choudhury, K. R., "Fitting an ellipse to an arbitrary shape: implications for strain analysis," *Journal of Structural Geology* **26**, 143–153 (2004).

[15] 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).

[16] Lam, L. and Suen, C. Y., "Application of majority voting to pattern recognition: An analysis of its behavior and performance," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* **7**(5), 553–567 (1997).