32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Towards a Versatile Automated Cell-Detection System for Science and Diagnostics

Andreas Heindl, Sabine Dekan, Isabella Ellinger, Alexander K. Seewald

Abstract— **Analyzing *in-situ* tissue structures with complex shapes and textures such as multinuclear cells or cells without nuclei is still a challenge for currently available image-processing software. This work aims to provide a versatile system to solve such tasks provided that the structures of interests were detected by immunofluorescence microscopy. Images were automatically acquired using slide-based microscopy. Human domain-experts manually marked up tissue samples to evaluate the performance of the computer generated masks. From precision and recall a balanced F-score was computed to measure the correlation between experts and algorithm output. Exhaustive parameter optimization was conducted to ensure that the optimal input parameters were applied during evaluation of the developed algorithm. This procedure significantly increased the performance compared to manually chosen input parameters. We present an approach that can handle huge tissue areas and does not rely on nuclei detection. Once a mark-up has been created, the algorithm can be parameter-optimized on ground-truth data for the chosen tissue sample. Thereafter, the resulting settings could be applied automatically to the respective stitched image. Concluding, we provide new insights in physiological and pathopysiological cellular mechanisms by automating the *in-situ* analysis of proteins in intact tissues.**

I. INTRODUCTION

Automated microscopy technologies such as slide-based microscopy allow high-throughput screens of large tissue samples in an observer-independent and time saving manner. Subsequently, these acquisitions have to be properly analyzed with adequate software tools that allow an observer-independent and reproducible identification of cells, and quantification of cell-associated markers. Such qualitative and quantitative *in-situ* analysis are of growing importance in both research and diagnostics. Though other biological methods (e.g. Western Blotting) can also verify and quantify protein expression in tissues, they require destruction of tissue structure. Consequently, exact (sub)cellular localization of target proteins is not possible any more, thereby resulting in loss of information. Immunofluorescence microscopy combined with image processing overcomes this problem. However, currently available software tools have a very limited amount of analysis options for biologists. The majority of software tools aim to detect the cell nucleus and then try to grow from this seed until a certain border has been reached (e.g. a cell membrane staining). Unfortunately, not all cell types can be detected using that technique. Among these are multinucleated cells found in tissues such as placenta and bone, but also in many tumors. The challenge of developing software tools for such cells is the variety of shapes and different sizes of these cell types, where most of the common image-processing methods fail. To solve this problem, a system that pursues a more human-like approach is needed.

An important step following implementation of a new algorithm is measurement of its performance. Usually this is done by visual inspection and manual tuning of parameters by the developer(s). A better way is for biological domain experts to mark-up ground-truth data. The resulting mark-ups are compared to automatically generated mark-ups by the algorithm, thus ensuring that the newly developed tools and the tissue experts produce comparable results. The latter approach was used here.

A. Related Research

Automated microscopy-image analysis is a rapidly evolving research field. Huang and Murphy [6] described an impressive image-understanding system that featured image processing, classification, clustering and statistical analysis of fluorescence images to determine the subcellular location of proteins expressed in a certain cell type. Herold et al [5] published a machine learning approach to automatically identify biologically meaningful objects (in their case nuclei) using a support vector machine. A method to detect Clathrin-coated pits using Haar-like features that can deal also with noisy data was published by Jiang et al [8]. Hamilton et al [1] described another approach of an automated subcellular phenotype classification system and also made their image sets publicly available to encourage further research to automated cell-image analysis and classification.

II. MATERIALS

Placental chorionic villi (PCV) represent tissue-substructures that are hard to analyze with commonly available software. PCV exhibit large differences in size and contain a multinucleated cell type, syncytiotrophoblast (STB).

Localization and expression-levels of the vast amount of PCV proteins can give new insights in physiologic protein functions and mechanisms of pregnancy-related diseases [7].

III. AIMS

The aim of this work was to improve the quality of *in-situ* studies concerned with automated localization and quantification of proteins in complex tissues like the PCV. Due to the fact that a typical PCV covers more than 80 fields

A. Heindl and I. Ellinger associated with Medical University of Vienna, Department of Pathophysiology and Allergy Research, Vienna, Austria
S. Dekan is associated with Medical University of Vienna, Department of Clinical Pathology, Vienna, Austria
A. Seewald is associated with Seewald Solutions, Vienna, Austria

978-1-4244-4124-2/10/$25.00 ©2010 IEEE

of view, analysis had to be done on a stitched image instead of separate fields of view.

IV. Methods

Images were acquired with an upright microscope coupled to a fully motorized stage using a 20x lens (nA=0.5) and TissueFAXS as acquisitioning software. Paraffin-embedded placentas were sectioned ($4\mu m$) and antigen retrieval was performed. Unspecific binding sites were blocked and sections incubated with primary anti-Cytokeratin-7 (CK7) antibody (Dako) and the respective fluorescence-conjugated (Alexa Fluor-488) secondary antibody (Invitrogen). Nuclei were labeled with DAPI (Roche) and samples embedded in Fluoromount GTM (SouthernBiotech). Individual images in the respective fluorescent channels were recorded and pictures were sampled with an overlap of 50 respectively 30 pixels for stitching [2]. The images were acquired and stored with lossless compression.

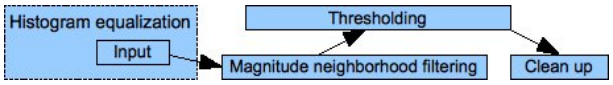

Fig. 1. Algorithm to detect syncytiotrophoblasts in placental tissue. Histogram equalization is done beforehand.

To detect the STB and thereby the delimiting borders of the PCV, specific staining of the STB-associated cytoskeletal protein CK7 was used as starting point. Since CK7 marks the entire STB-associated area, detection of cell nuclei is not required and complex problems like the separation of touching nuclei are avoided. Figure 1 shows the architecture of the STB detector. First, a cumulative histogram equalization over all images of the current acquisition was applied. A histogram of an image with N possible intensity levels is defined as the discrete function $h(r_k) = n_k$ where r_k is the kth intensity level in the interval $[0 \dots N]$ and n_k is the amount of pixels in the image whose intensity level is r_k. These levels will be referred to as bins. Hence a cumulative histogram counts the cumulative number of pixels for each of the bins over all histograms. Let m_j be a histogram and i the total amount of bins then the cumulative histogram M_i is defined as $M_i = \sum_{j=1}^{i} m_j$. M_i is then applied to all images in an acquisition. The goal of this multi-image normalization is remapping the image gray-levels to obtain an uniform histogram which makes it easier to analyze and compare different images.

Afterwards, a magnitude neighborhood filtering was applied that replaces each element by the surrounding neighbor with the highest intensity. This image is then thresholded, resulting in a binary image where 1's mark areas containing STB and 0's represent background pixels. This binarization may yield to small falsely detected spots resulting from dust or unspecific bindings, so a clean-up step is required. This step removes these artifacts of small size and inappropriate shape by morphologically opening the binary image. Each of these image-processing steps has a set of parameters. Tuning them manually is tedious and may result in a suboptimal STB mask. Therefore, all meaningful combinations of parameters were evaluated to find the optimal set of parameters for a given input image.

TABLE I
Possible cases during mark-up (1=STB, 0=BG)

Case	Expert_1	Expert_2	Prop. on ground-truth
1	1	0	8%
2	0	1	5%
3	1	1	87%

The specification of the server is as follows: Dual Quad-Core Intel Xeon E5420 @ 2.50Ghz and 8 GB of main memory running 64-bit Linux. Approximatively 8,100 different masks were generated in five days and automatically compared with the ground truth created by two domain experts. No user interaction was required during the whole process.

V. Results

Ground-truth data had to be generated to evaluate the algorithm. Therefore, two human experts delineated villi borders and STB areas on CK7 stained placental sections on adjacent field of views (each 1392x1024 pixels). Both domain experts have completed a natural scientific training, being pathologist or cell biologist, respectively, with a main focus on placental histology, function and pathophysiology. The mark-up of our experts are shown in Figure 2 and 3. Green indicates the inner STB border whereas red the outer one.

A mask is derived from these annotations of both experts. To compare our computer-generated results with the experts, we combined (using logical OR) both expert masks. An example mask is shown in Figure 4. Table I summarizes all possible scenarios after an expert mark-up. An 1 represents a pixel assigned to the STB by the corresponding expert whereas a 0 indicates that this pixel was not labeled as belonging to the STB. Evaluating the mark-ups of the two domain experts yielded an agreement (Case_3) of 87%. 8% of the pixels from expert 1 were not marked-up by expert 2 respectively 5% of the pixels marked-up by expert 2 were not marked-up by expert 1.

To achieve such an high performance with the developed algorithm, the best parameter combination had to be found. Therefore, an exhaustive parameter-optimization procedure was performed by running the algorithm with a large set of parameters combinations. In this evaluation task, the different impact of each parameter on the final resulting villi mask was measured. Comparison was done on pixel level. A pixel of the output mask was counted as true positive (TP) if at least one of the experts marked this pixel as belonging to the STB. Analogous, a pixel that was not marked by either of the experts nor the algorithm was counted as true negative (TN). If only the algorithm detected the pixel as part of the object then it represents a false positive (FP). If the algorithm did not detect a pixel marked up by at least one of the experts, it represents a false negative (FN). Precision and recall were

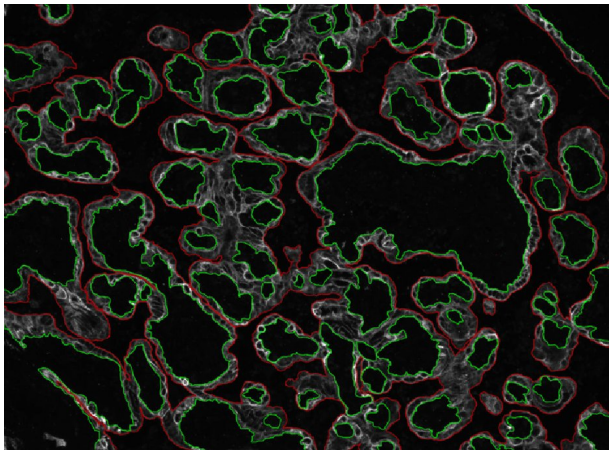

Fig. 2. Example of a domain expert 1 mark-up on a CK7 labeled human placental paraffin-section indicating villi as well as outer STB border in red and inner STB border in green.

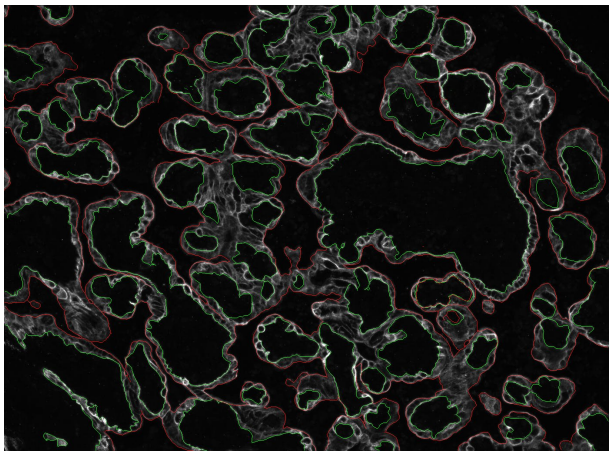

Fig. 3. Corresponding example of domain expert 2 mark-up on a CK7 labeled human placental paraffin-section indicating villi as well as outer STB border in red and inner STB border in green.

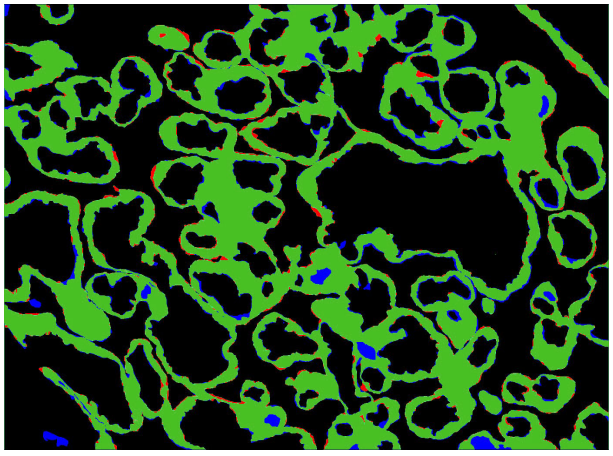

Fig. 4. Overlap of domain expert masks generated by expert 1 (red) and expert 2 (blue). Agreement of both experts is shown in green.

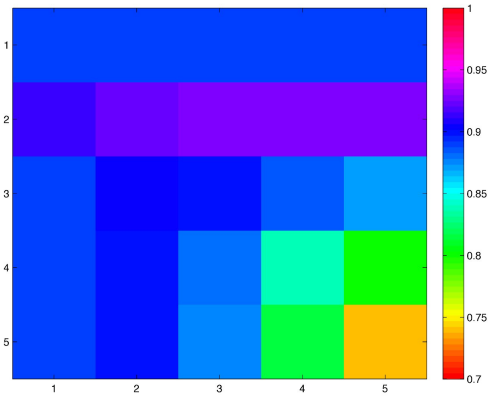

Fig. 5. Plot of Parameter 3 vs Parameter 5.

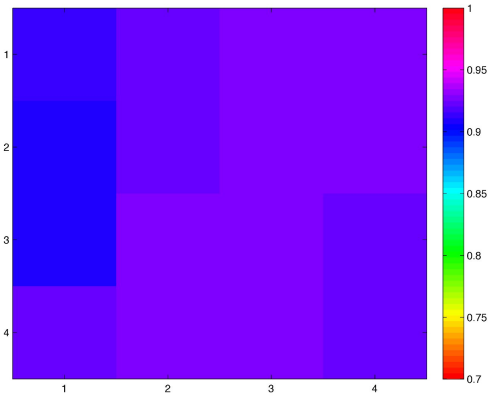

Fig. 6. Plot of Parameter 3 vs Parameter 7.

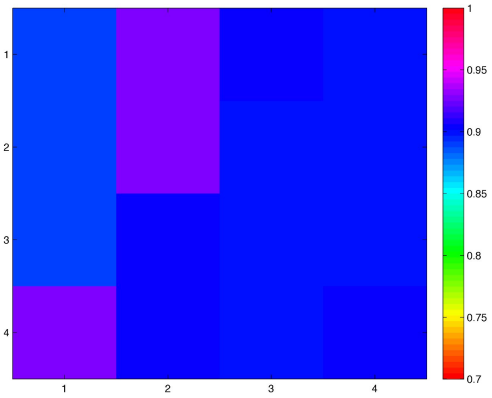

Fig. 7. Plot of Parameter 5 vs Parameter 7.

then defined as $Precision = \frac{TP}{TP+FP}, Recall = \frac{TP}{TP+FN}$. In some fields recall and precision are called sensitivity and specificity. From precision and recall, performance was derived as balanced F-score:

$$F_\beta = (1+\beta^2) * \frac{precision * recall}{\beta^2 * precision + recall}$$

In our case $\beta = 1$ so recall and precision are evenly weighted. The F-score was improved to 0.93 compared to the manually chosen parameter set with 0.78. Reasoning from the exhaustive-parameter optimization, only three (p3, p5 and p7) of the initial seven parameters have a major impact on the final STB mask quality. i/ p3: Size of mxm neighborhood used in the maximum filter, ii/ p5: Threshold used for the channel containing the staining to quantify, iii/ p7: Threshold for morphologically opening the binary image.

We were also interested in the correlation of these three parameters. Bitmap plots visualizing the correlation of these three parameters can be found in Figure 5 to Figure 7 for the respective parameter combinations. A legend matching the colors to the performance is shown next to each figure. From these figures it can be concluded that p3 and p5 have to be chosen well otherwise a mask with a performance lower than 75% will result. In contrast to this parameter tupel the other two shown in Figure 6 and Figure 7 achieve a performance of about 90%, even if we take the combination with the lowest performance. This analysis is important for later integration of this newly developed algorithm into a graphical user-interface (GUI) that will be used by biologists, so a focus on high impact parameters is desirable.

VI. CONCLUSIONS AND FUTURE WORK

A. Conclusion

We presented an algorithm to detect connected tissue structures without previous nuclei detection. As large amounts of images can now be acquired within short time, analysis of huge connected tissue structures is of growing importance and new algorithms are needed to overcome the restrictions of nuclei-based detection. Furthermore, we evaluated this newly developed algorithm by using exhaustive-parameter optimization. This process assured that the highest possible performance of the algorithm was achieved. Additionally, it trimmed the amount of parameters, resulting in a more manageable GUI.

A first application of this novel automated cell-detection analysis was presented at the International Federation of Placenta Associations meeting in 2009 where we analyzed expression of a protein (RAGE, the receptor for advanced glycation end products) in PCV [3][4]. The algorithm was optimized on one subregion per placenta. Afterwards the same settings were automatically used for the whole stitched placenta image. One acquired placental section consisted of a stitched 9x9 set of fields of view, respectively 12,528x9,216 pixels.

B. Future Work

Future work will concentrate on extending cell-detection technologies so that also cell types containing multiple nuclei (e.g. Osteoclasts) can be properly detected, analyzed and quantified. With additional markers and one labeling per structure further classifications (e.g. stem villi) could be accomplished. Another cell type that cannot be detected at present using nuclei based methods are of course cells without nuclei. Erythrocytes are an example for this class and detecting them can improve protein quantification accuracy due to the high autofluorescence associated with erythrocytes in various channels.

During our evaluation we weighted both domain experts equally. In case of n experts (e.g. n$\geq$3) different logical combinations can make sense, e.g. weighting proportional to expert level. Unanimous agreement as well as majority voting are additional possibilities.

VII. ACKNOWLEDGMENTS

This project is funded by FFG (Bridge 818094).

REFERENCES

[1] N. Hamilton, R. Pantelic, K. Hanson, J. L. Fink, S. Karunaratne, and R. D. Teasdale. Automated sub-cellular phenotype classification: an introduction and recent results. In *WISB '06: Proceedings of the 2006 workshop on Intelligent systems for bioinformatics*, pages 67–72, Darlinghurst, Australia, Australia, 2006. Australian Computer Society, Inc.

[2] A. Heindl, R. Ecker, G. Steiner, G. Bises, E. Ellinger, T. Thalhammer, R. Fuchs, H. Uhrova, and A. Seewald. Automated cell-detection technologies for science and diagnostics. Poster, 2009.

[3] A. Heindl, R. Ecker, G. Steiner, G. Bises, T. Thalhammer, R. Rogojanu, H. Uhrova, H. Helmer, E. Ellinger, and A. Seewald. Automated cell-detection technologies for science and diagnostics. *Placenta 30 (9), P06.14*, 2009.

[4] A. Heindl, V. Pohl, R. Rogojanu, R. Ecker, T. Thalhammer, G. Bises, A. Seewald, and I. Ellinger. Towards a versatile automated cell-detection system for science and diagnostics exemplified through receptor for advanced glycated end-products (rage) quantification in placental chorionic villi. Poster, 2010.

[5] Julia Herold, Sylvie Abouna, Luxian Zhou, Stella Pelengaris, David BA Epstein, Michael Khan, and Tim W. Nattkemper. A machine learning based system for multichannel fluorescence analysis in pancreatic tissue bioimages. *8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE)*, 2008.

[6] K. Huang and R.F. Murphy. From quantitative microscopy to automated image understanding. *Journal of Biomedical Optics 9(5)*, pages 893–912, 2004.

[7] T. Jansson, L. Myatt, and TL. Powell. The role of trophoblast nutrient and ion transporters in the development of pregnancy complications and adult disease. *Current Vascular Pharmacology*, pages 521–33., 2009.

[8] S. Jiang, X. Zhou, T. Kirchhausen, and S. T. C. Wong. Detection of molecular particles in live cells via machine learning. *Cytometry Part A*, pages 563–575, 2007.