Histopathological Image Analysis for Determining Tissue Composition in Gastric Cancer

Harshita Sharma[1], **Norman Zerbe**[2], **Daniel Heim**[2], **Stephan Wienert**[2], **Hans-Michael Behrens**[3], **Olaf Hellwich**[1], **Peter Hufnagl**[2]

[1] Computer Vision and Remote Sensing, Technische Universität Berlin, Germany

[2]Department of Digital Pathology and IT, Institute of Pathology, Charité-Universitätsmedizin Berlin, Germany

[3]Department of Pathology, Christian-Albrechts University, Kiel, Germany

INTRODUCTION

Gastric cancer is the fourth most common cancer, and the second most common cause of cancer-related deaths in the world. According to the WHO, 800,000 cancer-related deaths are caused by gastric cancer each year worldwide [1].

Motivation: Computer-based analysis of histological images of gastric cancer is a prospective challenge in digital pathology. Histological composition of gastric cancer tissues for diagnostic purpose is currently determined by pathologists using visual inspection in routine and research, which is a tedious and time consuming process.

Contribution: We describe an automatic method to determine histological composition of tissues in H&E whole slide images (WSI) of gastric cancer, for heterogeneous datasets with variations in stain intensity and malignancy levels. Such tissue composition analysis can potentially assist pathologists in computer-assisted diagnosis of gastric cancer. The method also provides a basis for automatic differentiation between tumor and non-tumor compartments of the tissue and determination of cancer type, grade or extent.

MATERIALS AND METHODS

Data Acquisition: Her2/neu immunohistochemically stained and H&E stained surgical specimens of 12 cases (one specimen per case) were selected from a previous study of 483 cases of gastric cancer. These were acquired from proximal or distal parts of stomach and scanned with Leica SCN400 microscopic whole-slide scanner at its maximum, nominally 400 times magnification and pixel size 0.0676 μm^2.

Data Annotation: Ten expert pathologists have annotated the WSI areas as:

- **Red polygons:** Her2/neu positive areas marked using the 10% cut-off rule [2].
- **Blue polygons:** Her2/neu negative areas morphologically identified as tumor.

The remaining areas are widely necrotic tissue regions.

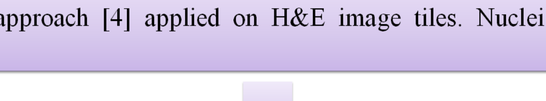


Fig. 1: Example of a Her2/neu stained gastric cancer WSI specimen with pathologists' annotations

Processing Chain: Figure 2 shows the processing chain used in our method.

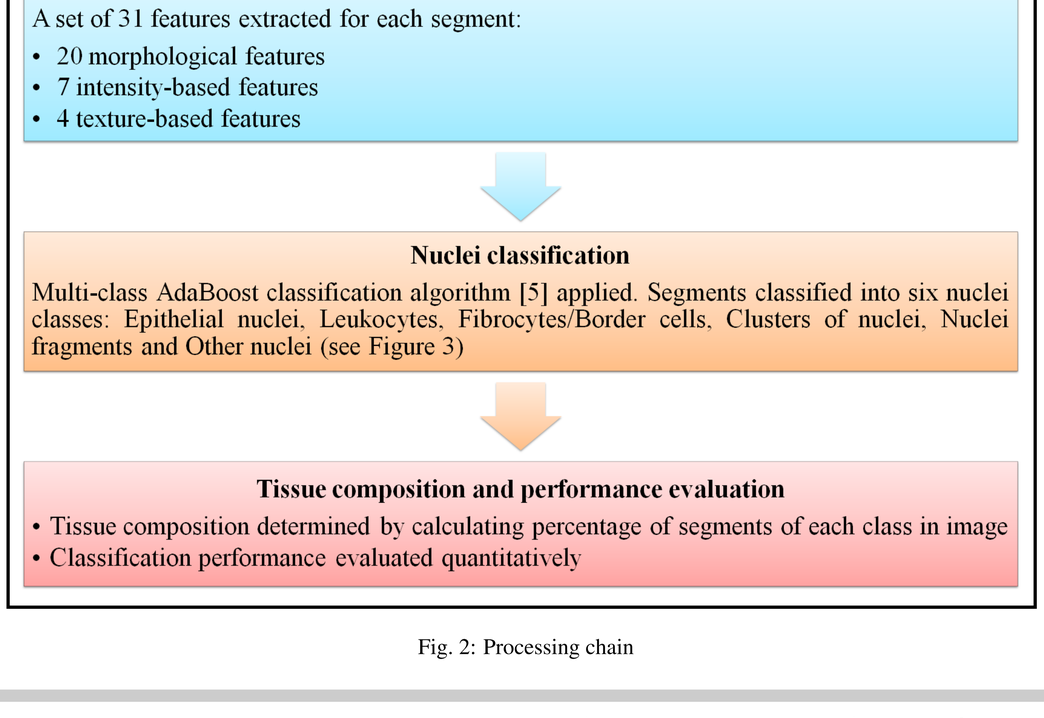


Fig. 2: Processing chain

EXPERIMENTAL RESULTS

Definitions of nuclei classes have previously been approved by expert pathologists and shown in Figure 3.

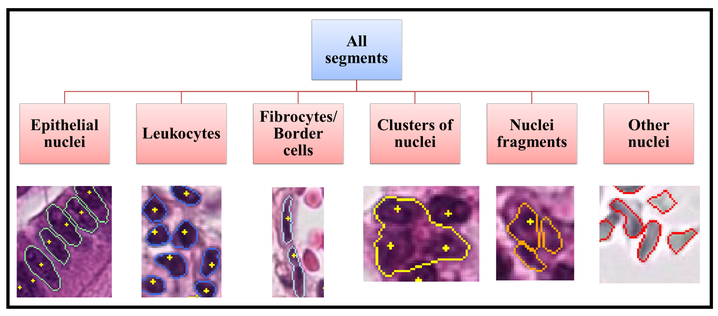
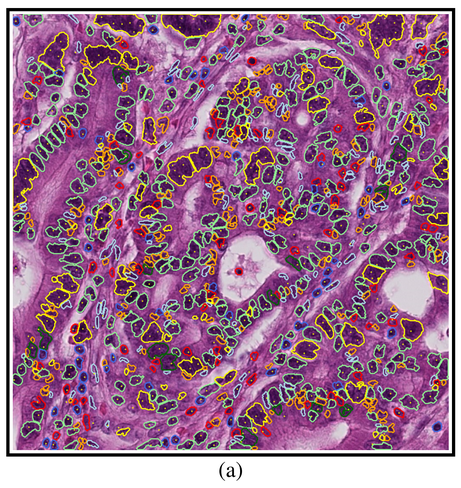
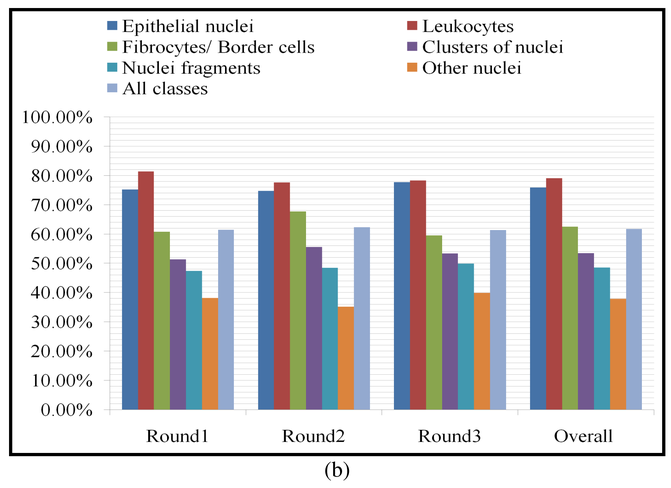


Fig. 3: Class definitions

A **3-fold cross validation** is performed to evaluate the classification. In each round, two-third of the reference data is used for training and one-third for testing without any overlap. Cell nuclei are classified as Epithelial nuclei, Leukocytes and Fibrocytes with good accuracy. Best result is achieved for Leukocyte class with average accuracy of **79.10%**. Fragments of epithelial nuclei (segments created due to overlapping nuclei) and clusters of nuclei are also classified, but with a lower detection rate. Lowest accuracy is obtained for Other nuclei class, because it includes nuclei not clearly visible to be classified as a specific type (ambiguous). Overall multi-class classification accuracy of **61.72%** is achieved.



(a)



(b)

Fig. 4: Experimental results (a) Example of H&E image tile in which nuclei are classified into six classes (b) Classification accuracy using AdaBoost classification method

CONCLUSION AND OUTLOOK

Conclusion:

- A method is proposed to automatically distinguish between various nuclei components and to determine tissue composition in H&E gastric cancer images.
- Overall classification accuracy can be further improved by adding more discriminative features to our current feature set.

Outlook:

- We aim to work towards extraction of high-level topological features based on the graph-theoretic description of tissue.
- We will also explore additional low-level features to describe the information between the nuclei components in the tissue images.

Acknowledgements

This work is supported with funds from the German Academic Exchange Service (DAAD). The authors are grateful to Dr. Christine Böger, Christian-Albrechts University for performing review of training images, without which this study would not have been possible. We thank Mr.Björn Lindequist, Charité for contributing in pre-processing stage and Dr. Fredrick Klauschen, Charité for providing insights into histological architecture and cell types to form classes for subsequent analysis. Moreover, we sincerely thank VMscope GmbH to provide suitable software support for Virtual Microscopy operations required in this research.

References

[1] C. Nordqvist. What is Stomach Cancer? What is Gastric Cancer? *Medical News Today. MediLexicon, Intl.*, March 2013.

[2] V. S. Warneke, H.M. Behrens, C. Böger, T. Becker, F. Lordick, M.P.A Ebert, and C. Röcken. Her2/neu testing in gastric cancer: evaluating the risk of sampling errors. *Annals of Oncology*, 24(3):725–733, 2013.

[3] I. Amidror. Scattered data interpolation methods for electronic imaging systems: a survey. *J. Electronic Imaging*, 11(2):157–176, 2002.

[4] S. Wienert, D. Heim, K. Saeger, A. Stenzinger, M. Beil, P. Hufnagl, M. Dietel, C. Denkert, and F. Klauschen. Detection and segmentation of cell nuclei in virtual microscopy images: A minimum-model approach. *Scientific Reports*, 2, 2012.

[5] C.M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.