Identification of Key Areas in Histology Images for Identification of Collagenous Colitis: a Deep Learning Approach

Vytautas Kiudelis[1,4], Robertas Petrolis[2,3], Rima Ramonaitė[1,4], Dainius Jančiauskas[5], Juozas Kupčinskas[1,4], Povilas Šabanas[1,4], Algimantas Kriščiukaitis[2,3]

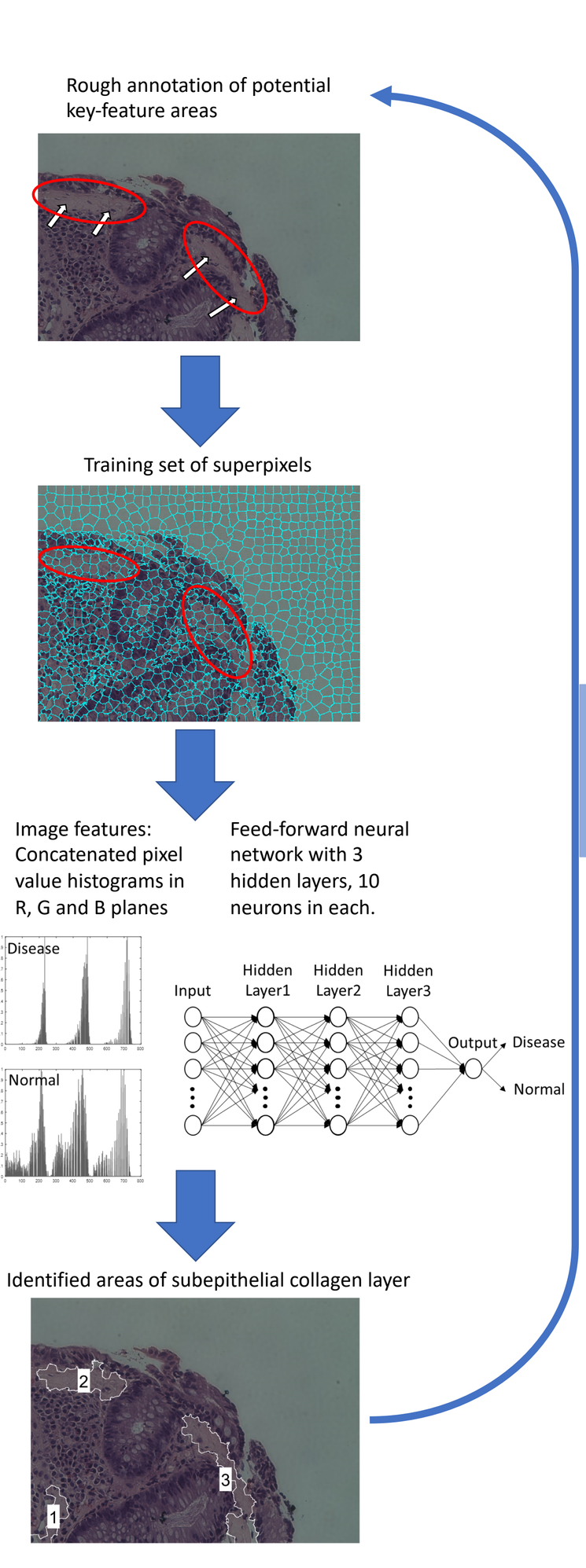
[1] Department of Gastroenterology, Lithuanian University of Health Sciences, Kaunas, Lithuania.
[2] Neuroscience Institute, Lithuanian University of Health Sciences, Kaunas, Lithuania.
[3] Department of Physics, Mathematics and Biophysics, Lithuanian University of Health Sciences, Kaunas, Lithuania.
[4] Institute for Digestive Research, Lithuanian University of Health Sciences, Kaunas, Lithuania.
[5] Department of Pathology, Lithuanian University of Health Sciences, Kaunas, Lithuania.

Introduction

Collagenous colitis (CC) is a form of microscopic colitis (1), an inflammatory disease of the large bowel that causes chronic watery diarrhoea, abdominal pain, faecal incontinence, nightly defecation, and weight loss, resulting in a significantly impaired quality of life (2,3). Incidence of MC has increased significantly during the past decades in some countries and the main reason for this increase is thought to be an enhanced disease recognition (4). The diagnosis of MC is challenging as it can only be diagnosed upon histological examination of colonic biopsies taken from normal or near normal appearing mucosa (1). Histological interpretation of biopsies involves subjective evaluation leading to inter-rater variability discrepancies in diagnosis and treatment plan. Deep learning-based assistance system can objectivise diagnostic key feature selection leading to minimisation of inter-rater variability and improvement of diagnostic accuracy.

The aim of this study was to develop the method and algorithm for robust segmentation of light microscopy images of histological specimen slides (tissue slides) emphasizing and estimating area of key diagnostic features of CC.

Development of machine learning based segmentation algorithms requires large annotated training datasets and preparation of it is time consuming hand-work for the experts. We decided to train our algorithm on just roughly annotated data trying to reach the desired precision of segmentation applying superpixel technique.



The general idea of method was to classify histology image superpixels according the concatenated pixel value histograms in R, G and B planes into disease and normal class. Subsequent connection of the same class superpixels gives potential areas of thickened subepithelial collagen – the key-feature to diagnose collagenous colitis. The result was evaluated by the experts in aim to correct parameters of the algorithm.

Evaluation of detection quality by Expert

Methods

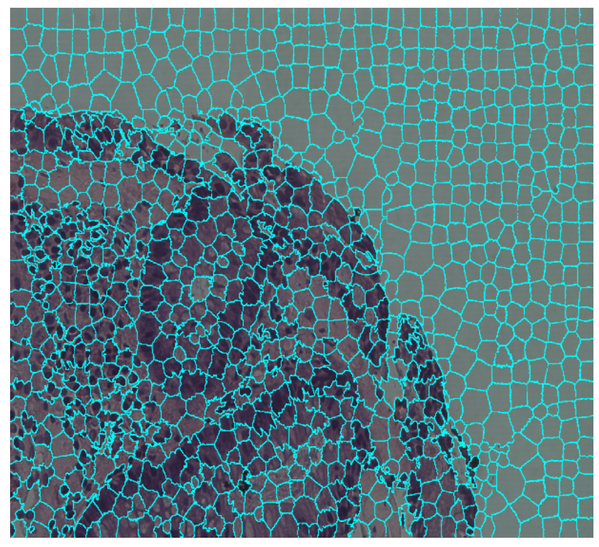
Histological specimen imaging. Histological specimen were fixed with formalin, embedded in paraffin, cut into 3 µm sections, and stained with haematoxylin and eosin (H&E) for histological examination. Histological patient specimen images (10 patients, ~ 60 images per patient) were taken using OLYMPUS IX71 light microscope (x20 magnification) equipped with Q IMAGING EXI aqua camera at (1392 x 1040 px.) resolution.

Image preprocessing. Image pixel values were normalized in RGB colour space by histogram of empty tissue-less area alignment using earlier elaborated algorithm described (5).

Experts roughly annotated image areas, indicating ones containing thickened subepithelial collagen layer as the disease class among the others containing the rest of the tissue indicated as normal class.

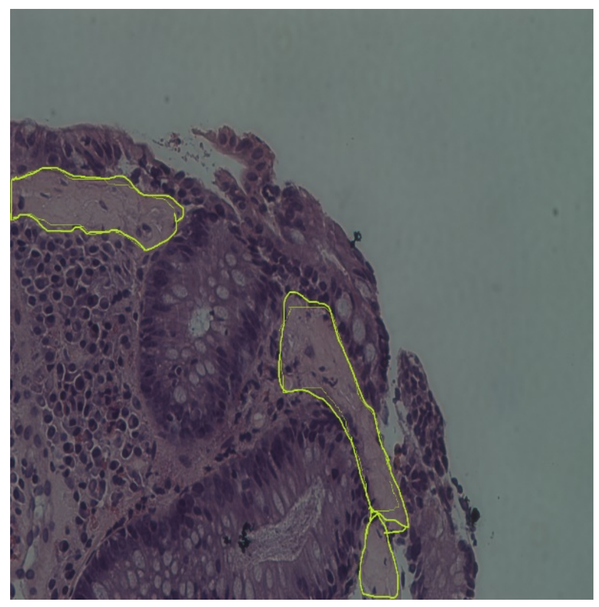
Superpixel technique using Simple Linear Iterative Clustering algorithm (6) was used for initial segmentation of the images. One thousand superpixels was chosen as optimal ammount for image split into areas with maximal similarity. Superpixels falling within the detected backround area of the image (non-tisue area) were ignored.

The number of disease class superpixels in the training set was significantly smaller than the number in normal ones. So, the training set was formed of all disease class and the same ammount of randomly selected normal class superpixels (n_1=n_2=18761). The whole set was splitted into Training, Validation and Testing parts according proportions: 70%, 20% and 10% respectively.



The example of histological image split into superpixels

Algorithm was trained on personal computer with AMD Ryzen Threadripper 3970X 32-Core, 3.70 GHz processor and 128GB of RAM. The training process with data shuffling in every-epoch took ~30 min.



The example of histological image with detected areas of thickened subepithelial collagen layer

Algorithm training results: The algorithm segmentation quality was estimated by count of correctly segmented thickened subepithelial collagen areas, confirmed by the Expert. The algorithm showed 0.807 accuracy, 0.801 sensitivity and 0.813 specificity.

Conclusions:

The elaborated segmentation algorithm could be used for assisted diagnostic process emphasizing areas with candidate key features for identification of collagenous colitis.

The training set of images with only roughly annotated areas of key-features of colagenous colitis is suitable for traiining of the algorithm;

References

1. Miehlke S, et al. European guidelines on microscopic colitis: United European Gastroenterology (UEG) and European Microscopic Colitis Group (EMCG) statements and recommendations. United European Gastroenterol J. 2020;
2. Verhaegh BPM, et al. Course of Disease in Patients with Microscopic Colitis: A European Prospective Incident Cohort Study. J Crohns Colitis. 2021;15(7):1174–83.
3. Nyhlin N, et al. Long-term prognosis of clinical symptoms and health-related quality of life in microscopic colitis: A case-control study. Aliment Pharmacol Ther. 2014;39(9):963–72.
4. Weimers P, et al. Incidence and Prevalence of Microscopic Colitis between 2001 and 2016: A Danish Nationwide Cohort Study. J Crohns Colitis. 2021;14(12):1717–23.
5. Petrolis R., et al. Method of fluorescence imaging for evaluation of membrane potential in cultured neurons using transmembrane voltage sensitive dye. Biomedical engineering 2011: Proceedings of International Conference. 14:16-19, 2011.
6. Radhakrishna A, et al. SLIC Superpixels Compared to State-of-the-art Superpixel Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 34, Issue 11, pp. 2274-2282, May 2012