Grading dysplastic colon polyps using image recognition and machine learning techniques

Pascale De Paepe [1], Jacques Van Huysse[1], Carl Van Paeschsen[1], Ann Driessen[2], Katrien De Wolf[3], Ward Van Laer[3] and Pascal Niville[3]

This study was approved by the ethical committee of AZ Sint Jan Brugge - Oostende AV/Campus Brugge

1. AZ Sint Jan Brugge - Oostende AV / Campus Brugge; 2. UZ Antwerpen; 3. Ixor Mechelen

INTRODUCTION

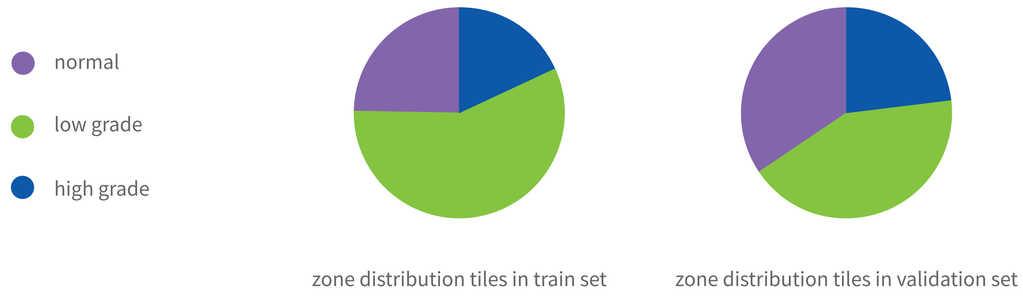
Colorectal cancer is a major health problem. The scale of the systematic screening increases the burden on pathology laboratories, especially in combination with the limited available time frame to process the samples. Whole slide scanners can provide whole slide images (WSI) of classical glass slides. The WSI can be analysed manually by a pathologist or automatically using “Digital image recognition techniques”.

AIM

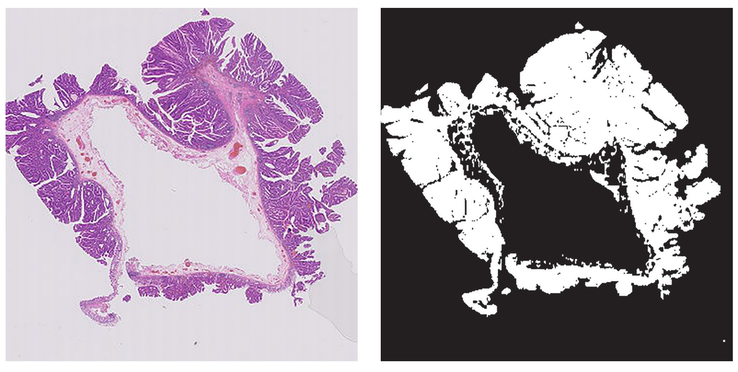
Based on machine learning and image recognition techniques, we design a model that recognises and indicates major zones of interest on WSI of colon polyps as a pre-analytic tool. The zones of interest are: normal tissue, zones with low grade dysplasia and zones with high grade dysplasia as defined by the WHO criteria. (1)(2)(3)

MATERIAL & METHODE

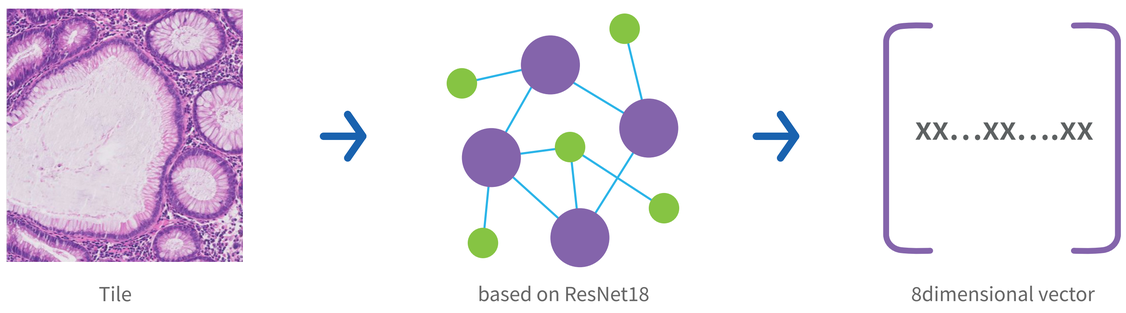
The dataset consists of 70 anonymous WSI’s of colon polyps with dysplastic features, taken with a 40x objective. On each WSI the various zones of interest are annotated by 4 pathologists using the platform designed by IxorThink.



1. Each WSI is divided into tiles, 128 x 128 px. Blank tiles are omitted.



2. Tiles with more than 40% tissue are used as input for the neural network. This network extracts features for each tile.

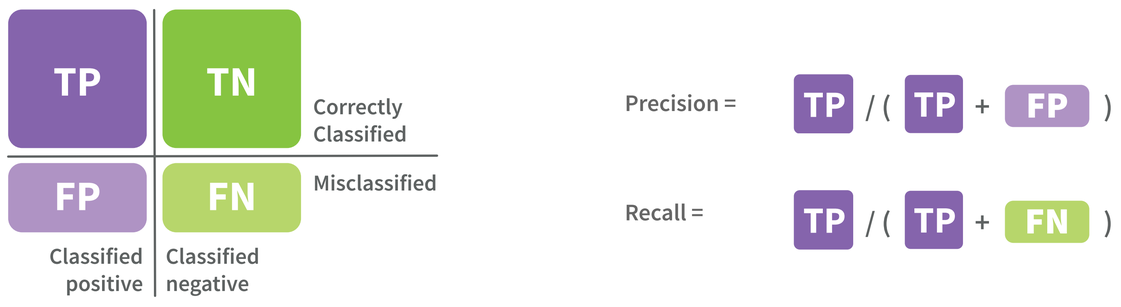


3. A random forest classifier classifies those features taking in to account the features of the 8 adjacent tiles. Thus predicting a tile belonging to a certain class.

RESULTS

Zones of interest	Support	Recall	Precision
High grade	695	65%	77%
Low grade	4397	94%	96%

table 1



TP = True Positive **TN** = True Negative **FP** = False positive **FN** = False Negative

As shown in table 1 and illustrated by figure 1, 2 and 3 we are able to train a model that detects zones of low grade dysplasia with a precision of 96%. The precision for zones of high grade dysplasia reached 77%. The latter less satisfying result might be explained by the much lower incidence of high grade dysplasia in the training set. Expanding our series of WSI will provide an increased number of high grade dysplasia zones and will allow us to improve the model in order to reach similar results for both dysplasia grades.

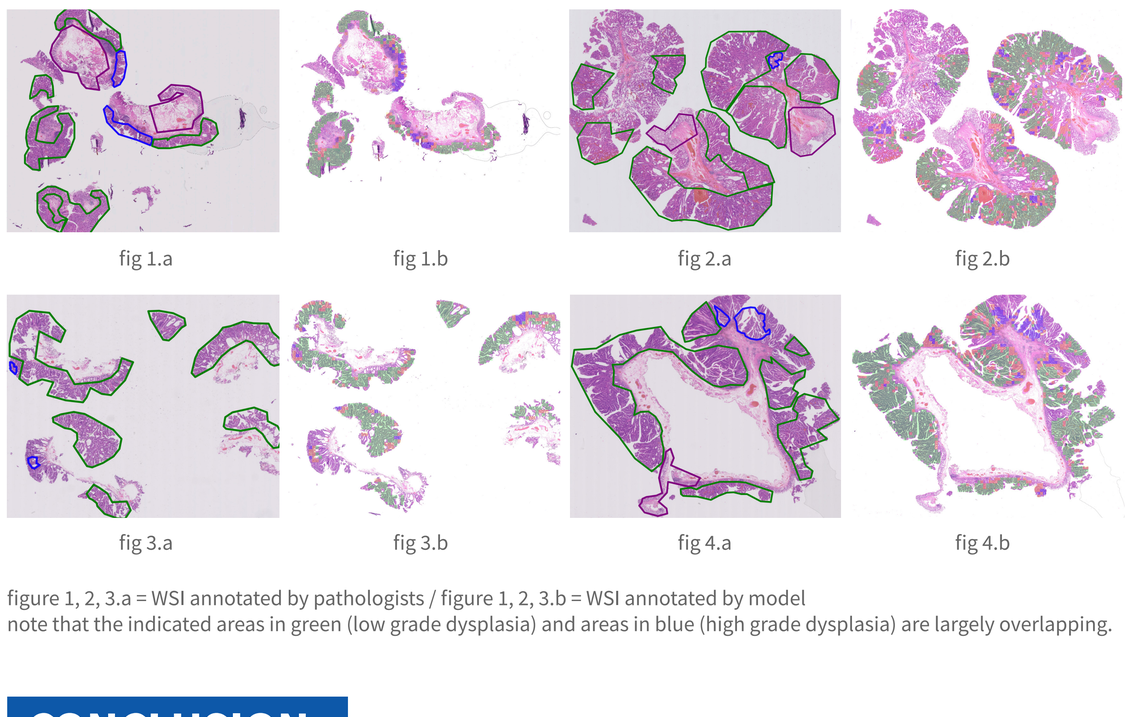


fig 1.a fig 1.b fig 2.a fig 2.b

fig 3.a fig 3.b fig 4.a fig 4.b

figure 1, 2, 3.a = WSI annotated by pathologists / figure 1, 2, 3.b = WSI annotated by model
note that the indicated areas in green (low grade dysplasia) and areas in blue (high grade dysplasia) are largely overlapping.

CONCLUSION

This preliminary study, based on a limited number of WSI of colon polyps, demonstrates that we can develop a model, based on machine learning techniques, that recognises and indicates zones of major interest in these biopsies. The zones of interest are: normal tissue, zones with low grade dysplasia and zones with high grade dysplasia as defined by the WHO criteria. (1)(2)(3)

Based on the results of this study we will further develop our model and construct an end-to-end product in order to streamline histological analysis.

ACKNOWLEDGMENTS

We thank Prof.em. Karel Geboes and Prof. em. Chris De Wolf-Peeters for believing in our project and for reviewing the annotations.

EXTRA INFORMATION

for more information see website www.ixorthink.com

1. Brown LJ, Smeeton NC, Dixon MF: Assessment of dysplasia in colorectal adenomas : an observer variation and morphometric study. J Clin Pathol 1985; 38: 174-179
2. O’Brien MJ, Winawer SJ, Zauber AG e.a.: National Polyp study USA. Gastroenterology 1990; 98(2): 371-379
3. Bosman FT, Carneiro F, Hruban RH, Theise ND: Tumours of the colon and rectum. in: WHO Classification of Tumours of the Digestive System 2010.IARC, 4th Edition, Volume 3, chapter 8