Intestinal Gland Classification from Colorectal Cancer Tissue Images using Graph-based Methods

AUTHORS:

Institution/Hospital, Department, City, Country.

- Linda Studer, iCoSys, Hochschule für Technik und Architektur Freiburg, Switzerland
- Shushan Toneyan, Institute of Pathology, Faculty of Medicine, University of Bern, Switzerland
- Prof. Dr. phil. nat. Inti Zlobec, Institute of Pathology, Faculty of Medicine, University of Bern, Switzerland
- Prof. Dr. med. Alessandro Lugli, Institute of Pathology, Faculty of Medicine, University of Bern, Switzerland
- Prof. PhD Andreas Fischer, iCoSys, Hochschule für Technik und Architektur Freiburg, Switzerland
- Dr. med. Heather Dawson, Institute of Pathology, Faculty of Medicine, University of Bern, Switzerland

BACKGROUND:

Automated systems can greatly improve the precision and speed with which a diagnosis is established, but implementation in routine can be hampered by a lack of explainability. We propose a graph-based approach for intestinal gland classification based on visual diagnostic features. Graph edit distance (GED) provides an explicit and comprehensible mapping of cells from one gland to cells from another gland using deletion, insertion, and substitution of individual cells (Figure 1).

METHODS:

We evaluated our methods on a dataset of graphs based on normal and dysplastic H&E stained intestinal glands from pT1 stage colorectal cancer.

Epithelial cells were segmented using QuPath and 33 available features (see Table 1) were extracted for each cell. Each cell was considered as a node and was connected to its two spatially closest neighbours with an edge (Figure 2). Out of the 33 node features, the best were selected using forward search. Gland graphs were compared with each other using GED. For an input graph, the GED was calculated to all graphs in the reference set and then classified as either normal or dysplastic based on the label of the three most similar graphs (k-nearest neighbours algorithm).

RESULTS:

Four node features based on the cytoplasm eosin stain and nucleus hematoxylin stain (details see Table 1) were identified for the optimised graph which achieved an accuracy of 83.1%, resulting in an improvement of 16.2% compared to the unlabelled baseline graph.

CONCLUSIONS:

We propose a novel system that combines the representational power of graphs with visual criteria used by pathologists in routine. Future work will involve a deeper evaluation on what kinds of graphs and graph matching methods are most intuitive and comprehensible in order to improve the explainability of our method as well as to include pathologists' feedback into the feature selection process.

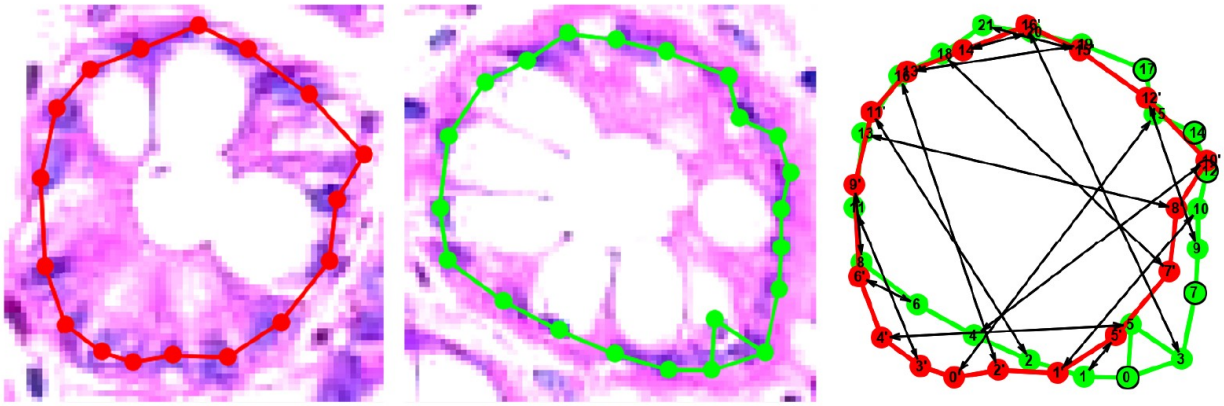
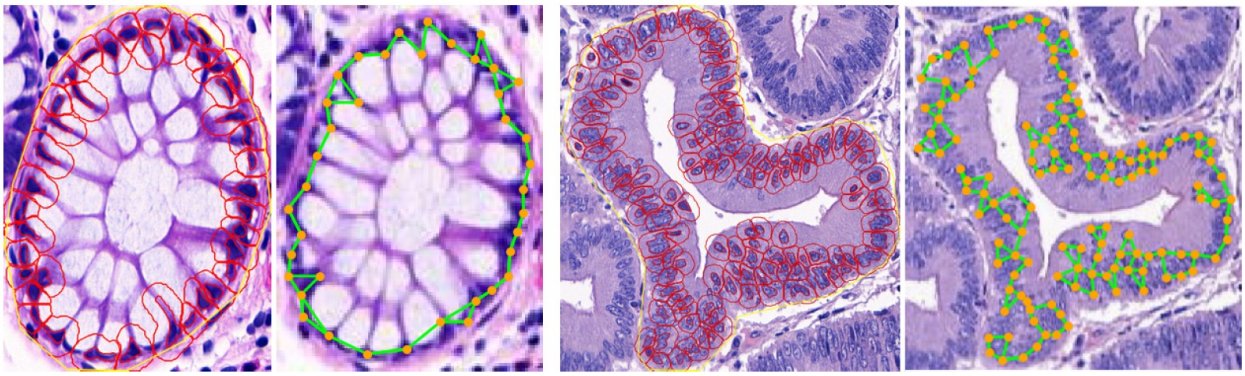


Fig. 1: Example of the GED transformations between two normal glands. The black arrows indicate node label substitution, the nodes circled in black mark deleted/inserted nodes.



(a) Normal glands (b) Dysplastic glands

Fig. 2: Examples of the cell segmentation (left) and graph-representation (right) of a normal and a dysplastic gland. Each detected cell (circled in red) is represented as a node in the graph (in orange). The nodes are connected with edges (in green) based on the physical distance between them.

Table 1: List of available features in QuPath for each cell after cell segmentation. These features can be used as node labels in the gland graphs. Out of the 33 features the forward search selected four for the optimised graph.

	Node Features	
	Based on	Feature
All features from QuPath	Cell	Eosin mean, std, min and max Circularity Eccentricity Perimeter Area Caliper min and max
	Cytoplasm	Eosin mean, std, min and max
	Nucleus	Circularity Eosin sum, mean, std, min, max and rage Hematoxylin sum, mean, std, min, max and range Area Caliper min and max Perimeter Eccentricity
	Misc	Nucleus/cell area ratio
Optimised graph	Cytoplasm	Eosin min
	Nucleus	Hematoxylin mean, min and max