DASHBOARD

DETECTION OF EARLY CANCER IN THE ESOPHAGUS

KEYWORDS: MACHINE LEARNING, PATTERN RECOGNITION, COMPUTER-AIDED DIAGNOSIS, COMPUTER VISION, ESOPHAGEAL CANCER

INTRODUCTION

Esophageal cancers are the fastest rising type of cancer in the Western world. The majority of these cancers are so called adenocarcinomas, which arise in the junction between the stomach and the esophagus. Due to lifestyle and genetic predisposition, acid from the stomach enters the esophagus. As a result, the body starts replacing the esophageal-type cells by small bowel-type cells that are acid resistant, creating a so-called Barrett’s esophagus (Fig. 1). The process of replacing cells by other type cells is called metaplasia. And during this process, there is an increased risk of abnormal cell development, or dysplasia, which is the earliest stage of cancer. At this stage, the gastroenterologist is able to remove the dysplastic tissue locally, using Endoscopic Mucosal Resection (EMR) during a routine screening. Recent studies have shown that this increases the five-year survival rate of the patients from about 10% up to over 95%.

With High Definition (HD) Endoscopy, the specialist physician can identify early cancerous tissue during a routine screening, by closely inspecting the esophageal tissue. However, the identification of these early stages of cancer remains a difficult exercise, even for the trained gastroenterologist. Therefore, a system that automatically detects early cancer during such a screening is highly desirable.

In our research group, we have taken the first steps towards such a detection system. Together with a medical specialist from the Catharina hospital in Eindhoven, we have proven the feasibility of a Computer-Aided Detection (CAD) algorithm that can be trained for the identification of early cancerous tissue in the esophagus, based on HD endoscopic images (Fig. 2). However, there are still a lot of challenges to overcome before the system could help a gastroenterologist in a hospital setting.

CHALLENGES

In order to improve the current system, the following challenges are key topics for an internship or graduation project.

1. **Prevent false classifications**: Due to intestinal juices, specular reflections or other distracting elements, the system yields false detections. Ideally, we want to reduce the number of false positives to zero, while keeping the correct detections as high as positive.
II. **Find and evaluate better features**: In the current system, features are based on Gabor filters and on the color histogram. These features have shown to be quite successful, but there might be even better features for the classification of early esophageal cancer. See our work for a similar study\(^1\).

III. **Optimize the system for real-time implementation**: For the computation of the Gabor-based features, the system needs to filter the input image with a filter bank that contains relatively large filters masks. For a 1200×1600 HD endoscopic image, the filtering operation typically takes 8 to 20 seconds in Matlab. Currently, this is the bottleneck of our system concerning timing issues. For a real-time implementation of the algorithm, the timing properties of the algorithm should be analyzed and optimized.

IV. **Expand the algorithm to endoscopic video**: Currently, the system works on individual endoscopic images, but eventually it should work (real-time) on endoscopic video. The use of temporal information offers great advantages over our current system. Certain elements can be tracked and false predictions can be filtered in temporal domain.

V. **Improve the annotation method**: The system uses a block-based detection method, and filters the result in order to obtain a smooth annotation curve. The smoothening of the detection result is an operation that adds no information, it only makes the annotations more natural. It is a challenging task to build an annotation scheme that actually uses extra information (e.g. statistical region properties) in order to draw smooth delineation curves.

If you are interested in doing your internship or graduation project on one of these topics, please contact:

**Fons van der Sommen**
Email: f.v.d.sommen@tue.nl
Office address: LaPlace building, LG 0.28, VCA office (in the NIRICT lab)

The scope of our research is not limited to the above list. Suggestions for internships/graduation projects that could contribute to our system are very welcome. The current system architecture, shown in Figure 3, could be used as source of inspiration for such suggestions.

---


---

**FIGURE 3: CAD SYSTEM ARCHITECTURE**