

# Computer-aided detection of early neoplastic lesions in Barrett's esophagus

10.1055/s-0042-105284
Endoscopy 2016; 48: 617–624



**Publishing House and Copyright:**
© 2016 by
Georg Thieme Verlag KG
Rüdigerstraße 14
70469 Stuttgart
ISSN 0013-726X

# Computer-aided detection of early neoplastic lesions in Barrett's esophagus

**Authors** Fons van der Sommen[1], Svitlana Zinger[1], Wouter L. Curvers[2], Raf Bisschops[3], Oliver Pech[4], Bas L. A. M. Weusten[5], Jacques J. G. H. M. Bergman[6], Peter H. N. de With[1], Erik J. Schoon[2]

**Institutions** Institutions are listed at end of article.

**submitted** 9. July 2015
**accepted after revision** 23. February 2016

**Bibliography**
**DOI** http://dx.doi.org/10.1055/s-0042-105284
Published online: 21.4.2016
Endoscopy 2016; 48: 617–624

ISSN 0013-726X

**Corresponding author**
***Fons van der Sommen, MSc***
Department of Electrical Engineering
Eindhoven University of Technology
P.O. Box 513
Eindhoven 5600 MB
The Netherlands
Fax: +31-40-2473708
f.v.d.sommen@tue.nl

**Background and study aims:** Early neoplasia in Barrett's esophagus is difficult to detect and often overlooked during Barrett's surveillance. An automatic detection system could be beneficial, by assisting endoscopists with detection of early neoplastic lesions. The aim of this study was to assess the feasibility of a computer system to detect early neoplasia in Barrett's esophagus.

**Patients and methods:** Based on 100 images from 44 patients with Barrett's esophagus, a computer algorithm, which employed specific texture, color filters, and machine learning, was developed for the detection of early neoplastic lesions in Barrett's esophagus. The evaluation by one endoscopist, who extensively imaged and endoscopically removed all early neoplastic lesions and was not blinded to the histological outcome, was considered the gold standard. For external validation, four international experts in Barrett's neoplasia, who were blinded to the pathology results, reviewed all images.

**Results:** The system identified early neoplastic lesions on a per-image analysis with a sensitivity and specificity of 0.83. At the patient level, the system achieved a sensitivity and specificity of 0.86 and 0.87, respectively. A trade-off between the two performance metrics could be made by varying the percentage of training samples that showed neoplastic tissue.

**Conclusion:** The automated computer algorithm developed in this study was able to identify early neoplastic lesions with reasonable accuracy, suggesting that automated detection of early neoplasia in Barrett's esophagus is feasible. Further research is required to improve the accuracy of the system and prepare it for real-time operation, before it can be applied in clinical practice.

## Introduction

Over the past four decades, the incidence of esophageal adenocarcinoma has risen rapidly [1]. In particular, prosperous countries in the Western World have seen a drastic increase in incidence [2]. The increase in these countries is partly explained by the prevalence of known risk factors, such as excess body weight and obesity [3], which fuels the expectation that the disease will continue to rise over the coming years.

Most patients with esophageal adenocarcinoma are diagnosed late in the course of the disease, resulting in a bad prognosis for the patient. When detected early, the dysplastic tissue can be treated by endoscopic mucosal resection (EMR) and radio frequency ablation. These local treatments can achieve very high curation rates [2], with only 5 % morbidity, 0 % mortality, and with 93 % of patients achieving complete remission after 10 years [4, 5]. For a long time, the Seattle biopsy protocol has been the standard approach for the detection of early dysplastic lesions in the esophagus. This protocol recommends taking "four-quadrant, 2-cm endoscopic biopsies performed at closely timed intervals" in Barrett's esophagus [6]. With the availability of more advanced endoscopic imaging tools, the paradigm for detecting early neoplasia has shifted from random biopsies toward targeted sampling, based on thorough inspection of the Barrett's segment. As early neoplastic lesions are rare in general Barrett's surveillance, they are often overlooked owing to the subtleness of such lesions, lack of high-definition endoscopes and, most importantly, the inexperience of most endoscopists with such lesions.

Computer-aided detection may support the gastroenterologist in the detection of early neoplasia. Such a system could be a helpful tool during screening of Barrett's esophagus, alerting the endoscopist when a possibly dysplastic lesion is detected; furthermore, it can be used as a learning tool. Although a great body of research on automatic analysis of medical data exists, no studies

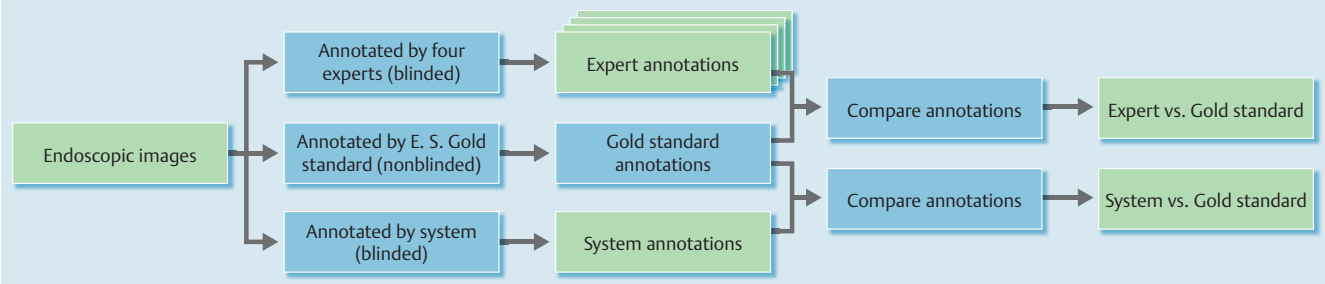


**Fig. 1** Schematic depiction of the study design.

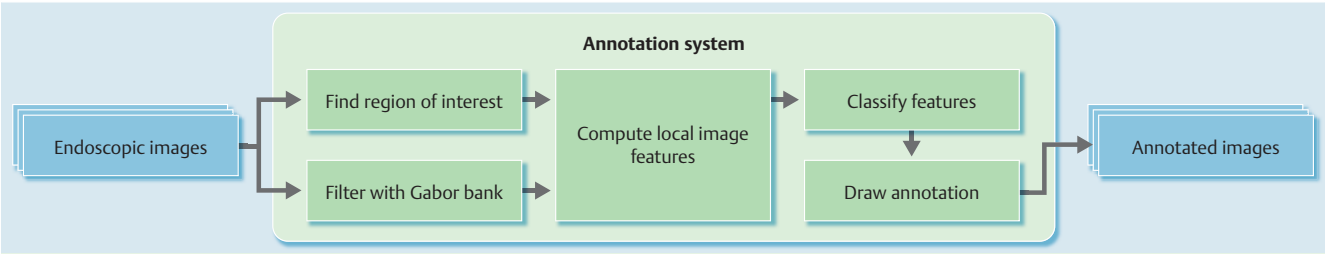


**Fig. 2** Overview of the annotation system.

have yet been performed for a similar analysis of high definition endoscopic imagery. However, comparable studies have investigated the automatic segmentation of images obtained by a wireless capsule endoscope. This pill-like camera takes about 60 000 images during an 8-hour journey through the gastrointestinal tract. Algorithms have been proposed for the detection of tumors [7], bleeding [8], ulcers [9], or abnormal patterns in general [10]. Similar systems have also been proposed for standard video endoscopy [11, 12], with varying success rates. This fluctuating success is, to a clear extent, caused by insufficient picture quality. Automatic detection of early neoplasia is possible only when high definition resolution and associated pre-processing are exploited. The aim of the current study was to investigate the feasibility of a newly developed automated image recognition system for the detection and delineation of early neoplastic lesions in Barrett's esophagus.

## Patients and methods

### Setting

The study was performed with close collaboration between the Department of Gastroenterology and Hepatology, Catharina Hospital, the Netherlands, a tertiary referral center for endoscopic treatment of early neoplasia, and the Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands, which has extensive expertise in image processing and automated image analysis. The requirement for patient consent was waived by the Institutional Review Board of the Catharina Hospital according to Dutch law, as only anonymized images were used.

### Patients, endoscopic procedures, and image selection

All patients included in the study were referred for endoscopic treatment of early Barrett's neoplasia or standard surveillance for Barrett's esophagus at the Catharina Hospital between September 2010 and June 2013.

All procedures were performed by a single endoscopist (E.S.) who had extensive experience in the detection and treatment of early Barrett's neoplasia. During these procedures, high definition images were obtained using a high definition endoscope (EG-590; Fujifilm Inc., Tokyo, Japan) and processor (VP-4450; Fujifilm Inc.), and were stored as JPEG files with a resolution of 1600 × 1200 pixels.

In cases with an early neoplastic lesion amenable to endoscopic resection, high quality images of the lesions were obtained and endoscopic resection was performed according to standard clinical practice using the multiband mucosectomy technique or EMR-cap technique [13]. In patients undergoing standard surveillance who did not have visible lesions, images of normal-appearing Barrett's mucosa were obtained followed by standard random biopsies in four quadrants every 2 cm. All endoscopic resection specimens and biopsies were routinely processed and underwent histological evaluation according to the Vienna classification [14]. In cases with dysplasia or early neoplasia, histopathology was reviewed by a gastrointestinal pathologist with expertise in Barrett's neoplasia.

All images in the study were selected based on the following criteria: a) high perceptual image quality; b) in cases with neoplasia: clear visibility of the lesion and histologically proven high grade dysplasia (HGD) or early carcinoma in the endoscopic resection specimen; c) in cases of nondysplastic Barrett's mucosa: clear absence of visible abnormalities and absence of any dysplasia or carcinoma in the random biopsies obtained during the imaging procedure.

### Annotation of images

All neoplastic lesions in the selected images were subsequently delineated by an expert endoscopist (E.S.) who was not blinded to the endoscopy findings and pathology. Images were delineated using PowerPoint (Microsoft Inc., Redmond, Washington, USA) drawing tool using a predefined color. This annotation was used as the gold standard for the development of the automated detection algorithm and, subsequently, the gold standard for testing

the algorithm for detection and delineation of early neoplastic lesions.

For external validation, four international experts in the detection and treatment of early Barrett's neoplasia (R.B., O.P., B.W., and J.B.), who were blinded to the endoscopy findings and pathology, reviewed all of the images. All reviewers were provided with a separate set of PowerPoint slides containing all images; each slide contained one endoscopic image. Each patient was allocated a unique number, which was displayed next to the image. Each time a reviewer detected an early neoplastic lesion, they were asked to delineate the lesion using the drawing tool in PowerPoint; each expert was assigned a different drawing color for delineation. The annotated endoscopic images were compared according to the study design, as depicted in **Fig. 1**.

## Automated image recognition system

As early carcinoma is commonly associated with deviating color and texture patterns [15], a system was developed that quantifies the image features and discriminates early carcinoma from nondysplastic Barrett's tissue based on this image information. **Fig. 2** shows a schematic overview of the annotation system, which is described briefly below. An in-depth technical explanation of the system has been described previously [16].

## Region-of-interest detection

Not all image regions of an endoscopic image are suitable for analysis; therefore, the upper left module in the annotation system was employed for finding the region of interest in the endoscopic image. An algorithm was designed that automatically detects the lumen, intestinal juices, and specular reflections present in the image. All pixels depicting any of these elements – commonly encountered in endoscopic images – are excluded from further analysis.

## Color and texture features

The module labeled "Filter with Gabor bank" contains an image filter, which is specifically tuned to texture patterns associated with early carcinoma. These texture patterns have been observed in a comparative spectral analysis of image regions that either show or do not show irregular tissue [16, 17]. The middle module in **Fig. 2**, labeled "Compute local image features," quantifies the color and texture information inside the region of interest of the original and the filtered image. The image is divided into square blocks of 50 × 50 pixels, and for each block the sample mean and variance are computed of both the filtered and the original image. This yields texture and color information of each image block.

## Classification and annotation

The next module, labeled "Classify features," employs a machine learning algorithm called Support Vector Machine, which learns the difference between image regions showing neoplasia and image regions showing no neoplasia based on the computed image features. Once trained, it predicts whether or not a certain part (block) of the image exhibits neoplastic tissue. Blocks that show neoplastic tissue are defined as positive training samples, whereas blocks that contain no neoplasia are defined as negative training samples. The last module in **Fig. 2** annotates the endoscopic image, based on the classified image blocks from the previous module.

## System annotations

The automated image recognition system annotated images by delineating each area in an evaluated image that was considered suspicious for early neoplasia according to the developed algorithm. **Fig. 3** shows a schematic overview of the procedure that was employed to generate the system annotations. The computer algorithm was evaluated using leave-one-out cross-validation on a per-patient basis. Prior to feeding the images to the annotation algorithm, the images were randomized and all additional information was stripped from the image data. As the system performance depends on the distribution of positive and negative training samples (image blocks), the experiment was repeated for different positive/negative distributions, ranging from 20 %/80 % to 50 %/50 %. As each combination always sums to 100 %, the results presented here refer only to the fraction of positive training samples (first of the two percentages), which we define as the positive training fraction (PTF). Hence, in this study, the PTF ranges from 20 % to 50 %.

## Outcome measurements

Two outcomes were studied to measure the system performance: (1) How well is the system able to detect a lesion? and (2) How similar is the delineation of the system to that of the expert (gold standard)? These were called detection performance and delineation similarity, respectively.

## Detection performance

A positive detection by the system was recorded when the delineation of the system overlapped with the gold standard delineation. For per-image analysis, sensitivity was defined as the number of correctly annotated images divided by the total number of images that showed a lesion. Specificity was defined as the number of images for which annotation was correctly omitted divided by the total number of images that did not show a lesion.

For per-patient analysis, the number of correctly classified patients was recorded, i. e. patients with lesions vs. patients without visible lesions, where the system classified a patient as having a lesion if it detected a lesion in either one of the images belonging to that patient. Furthermore, the study investigated whether there was a difference in detection performance between patients with HGD and patients with early carcinoma, and compar-

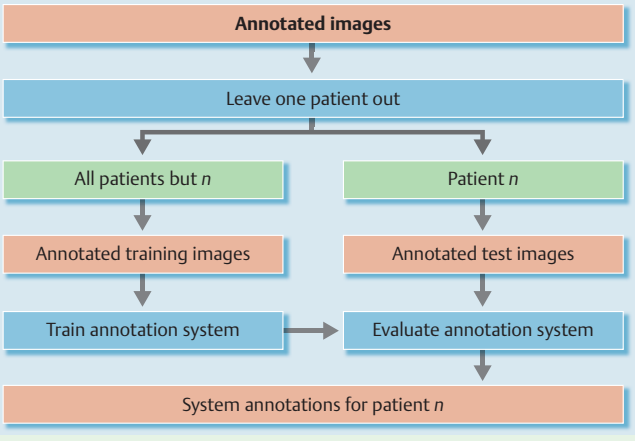


**Fig. 3** Overview of the validation procedure that was employed in the study, where subsequently the system was trained using the images of all but one patient, and evaluated based on the images of the patient who was discarded from the training data.

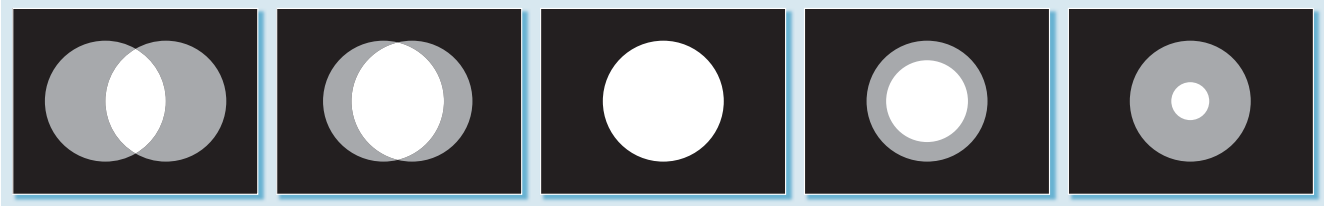

**Fig. 4** Examples of the F score for two annotations: from left to right: F = 0.52, F = 0.77, F = 1.00, F = 0.62, F = 0.20.

ed these individual detection rates between the system and the experts.

### Delineation similarity

The F score, which is defined as the harmonic mean between the pixel-based sensitivity and specificity, was used to measure the similarity between the system annotations and the gold standard annotations. For this, all pixels annotated in the gold standard were considered to be positives and all pixels not annotated were considered to be negatives. The F score was computed for each correctly annotated image, with the average taken as a final score. The same procedure was employed to determine the F score for the expert annotations. **Fig. 4** shows an illustrative example, where the F score is computed for different configurations of two circles. This example is analogous to comparing expert and system annotations, where the white area is annotated by both the expert and the system, the grey area is annotated by either the system or the expert, and the black area is annotated by neither.

### Comparison of the system with the experts

McNemar tests were used to compare the system detections with the detections of the four experts, and paired *t* tests were employed for measuring the difference between the annotations of the system and the annotations of the experts. The difference between the system and the experts was considered to be statistically significant for $P < 0.05$ for both tests. Software package Matlab 2015a (Mathworks, Inc., Natick, Massachusetts, USA) was used to perform the statistical tests.

## Results

### Patients and images

In total, 100 white-light endoscopic images from 44 patients with Barrett's esophagus (30 males, mean age 63 years) were selected. Of these patients, 21 showed a histologically proven early neoplastic lesion (13 early carcinomas, 8 HGD), and 60 images of these lesions were selected for evaluation. In addition, 40 images from 23 patients without dysplasia were selected for evaluation.

### Detection performance

The per-image detection performance of the system showed sensitivities and specificities of 0.20 – 0.95 and 1.00 – 0.40, respectively; a trade-off could be made by adjusting the distribution between positive and negative samples in the training data. To illustrate this trade-off, **Fig. 5** shows a receiver operating characteristic (ROC) plot of the system, where the distribution of training samples ranged between 20% positive/80% negative (PTF = 20%) to 50% positive/50% negative (PTF = 50%). For comparison, **Table 1** highlights five illustrative cases of the system detection performance for different training sample distributions and the detection performance of the experts who annotated the same set of images. The expert scores show a sensitivity of 0.95 – 1.00 and specificity of 0.73 – 0.93 (**Fig. 5**). Using the McNemar test, a statistically significant difference was observed between the system and experts 1 and 3 in all experiments. However, for the experiments with a fraction of positive training samples ranging between 31% and 36%, no statistically significant difference was observed between the system and expert 2 ($P = 0.087$, $P = 0.210$, $P = 0.078$, $P = 0.052$ for a PTF of 31%, 32%, 34%, and 36%, respectively). In addition, the system trained with a PTF of 32% also

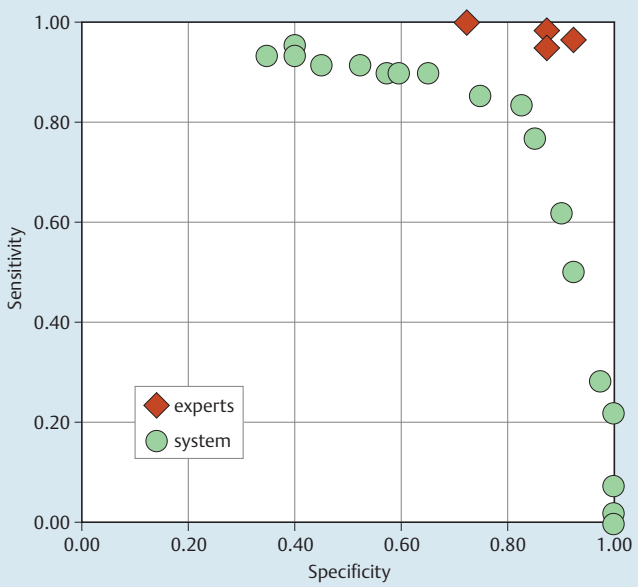


**Fig. 5** Receiver operating characteristic plot of the system for varying parameter settings (green dots) and the expert scores (red diamonds).

**Table 1** Detection performance of the system for different positive training fractions and for the four expert scores.

| | Per-image detection performance | | |
|---|---|---|---|
| | PTF, % | Sensitivity | Specificity |
| System | 30 | 0.62 | 0.90 |
| | 31 | 0.77 | 0.85 |
| | 32 | 0.83 | 0.83 |
| | 34 | 0.85 | 0.75 |
| | 36 | 0.90 | 0.65 |
| Expert 1 | – | 0.98 | 0.88 |
| Expert 2 | – | 0.95 | 0.88 |
| Expert 3 | – | 1.00 | 0.73 |
| Expert 4 | – | 0.97 | 0.93 |

PTF, positive training factor.

**Table 2** Detection of early carcinomas and high grade dysplastic lesions. The detection rates are given at the patient level, i. e. if a lesion was detected in at least one of the images from a single patient, then the lesion was considered to be correctly detected for this patient. False detections were recorded if a patient was falsely classified as having a lesion in at least one of the images annotated by either the system or by the experts.

| | Per-patient detection performance | | | | | |
|---|---|---|---|---|---|---|
| | PTF, % | Sensitivity | Specificity | Early carcinoma | HGD | False detections |
| System | 28 | 0.52 | 0.96 | 8/13 | 3/8 | 1 |
| | 29 | 0.71 | 0.91 | 10/13 | 5/8 | 2 |
| | 30 | 0.86 | 0.87 | 11/13 | 7/8 | 3 |
| | 31 | 1.00 | 0.78 | 13/13 | 8/8 | 5 |
| | 32 | 1.00 | 0.74 | 13/13 | 8/8 | 6 |
| Expert 1 | – | 0.95 | 0.83 | 13/13 | 7/8 | 4 |
| Expert 2 | – | 0.90 | 0.83 | 13/13 | 6/8 | 4 |
| Expert 3 | – | 1.00 | 0.65 | 13/13 | 8/8 | 8 |
| Expert 4 | – | 0.90 | 0.91 | 13/13 | 6/8 | 2 |

PTF, positive training fraction; HGD, high grade dysplasia.

**Table 3** Average annotation similarity between the system and the gold standard annotations.

| | Annotation similarity | | |
|---|---|---|---|
| | PTF, % | Mean F score | SD |
| System | 30 | 0.21 | 0.16 |
| | 31 | 0.30 | 0.20 |
| | 32 | 0.35 | 0.22 |
| | 34 | 0.40 | 0.24 |
| | 36 | 0.42 | 0.24 |
| | 50 | 0.62 | 0.19 |
| Expert 1 | – | 0.85 | 0.15 |
| Expert 2 | – | 0.92 | 0.09 |
| Expert 3 | – | 0.87 | 0.13 |
| Expert 4 | – | 0.81 | 0.18 |

PTF, positive training fraction.

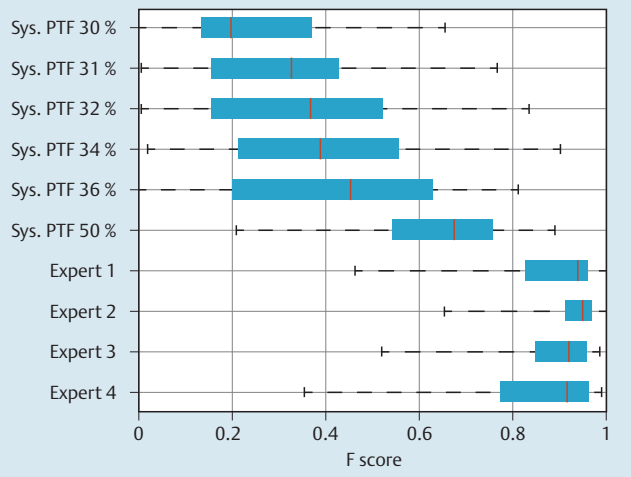


**Fig. 6** Box plot of the annotation scores, where the F scores of the five systems from **Table 3** are incorporated (red bars are medians). To indicate the maximum annotation performance, an additional system 6 has been included, which was trained with an equal amount of positive and negative training windows (positive training fraction 50 %).

did not show a statistically significant difference with expert 4 ($P$ = 0.076).

**Table 2** shows the per-patient detection performance of the system for five interesting cases, where a differentiation between the two pathologies is included. The system achieved a per-patient sensitivity ranging between 0.52 and 1.00, and specificity ranging between 0.74 and 0.96, where again a trade-off could be made by varying the percentage of positive training samples. **Table 2** shows the range in which the system performed optimally on a per-patient level. When trained with 28 % positive training samples, the system showed the least false positives, and detected 8 out of 13 patients with an adenocarcinoma and 3 of 8 patients with HGD. When the system was trained with 31 % positive training samples, it yielded the most false positives but also detected all 13 patients with an adenocarcinoma and all 8 patients with HGD correctly. All experts correctly detected the 13 patients with an adenocarcinoma and between 6 and 8 of the 8 patients with HGD, with a corresponding increase in false detections from 2 to 8. On a patient level, the only statistically significant difference was observed between the system trained with a PTF of 28 % and expert 1 ($P$ = 0.0215). All other comparisons between the system and the experts showed an insignificant statistical difference (minimum $P$ = 0.092).

## Annotation performance

The annotations of the system did not match the gold standard annotations as closely as the annotations of the experts did. **Table 3** shows the F scores of the experts and the system for the same distribution of positive and negative training samples as those shown in **Table 1**. In addition, the best system annotation performance has been included, which was achieved for a PTF of 50 %. To show the variation in annotation performance, **Fig. 6** shows a box plot of the F scores over all of the images corresponding to the results in **Table 3**. From this box plot it is clear that the annotations achieve a higher score when more positive samples are used for system training. Paired $t$ tests showed a statistically significant difference between the annotation scores of the system and the scores of all experts ($P$ < 0.001).

**Fig. 7** depicts four interesting examples of annotated images, where the top row shows the system annotations, the second row provides the gold standard, and the bottom four rows portray the annotations of the four additional experts. The system annotations in **Fig. 7** have been generated using the system when trained with 31 % positive training samples and they have an F score of 0.53, 0.49, 0.69, and 0.34, respectively. This particul-

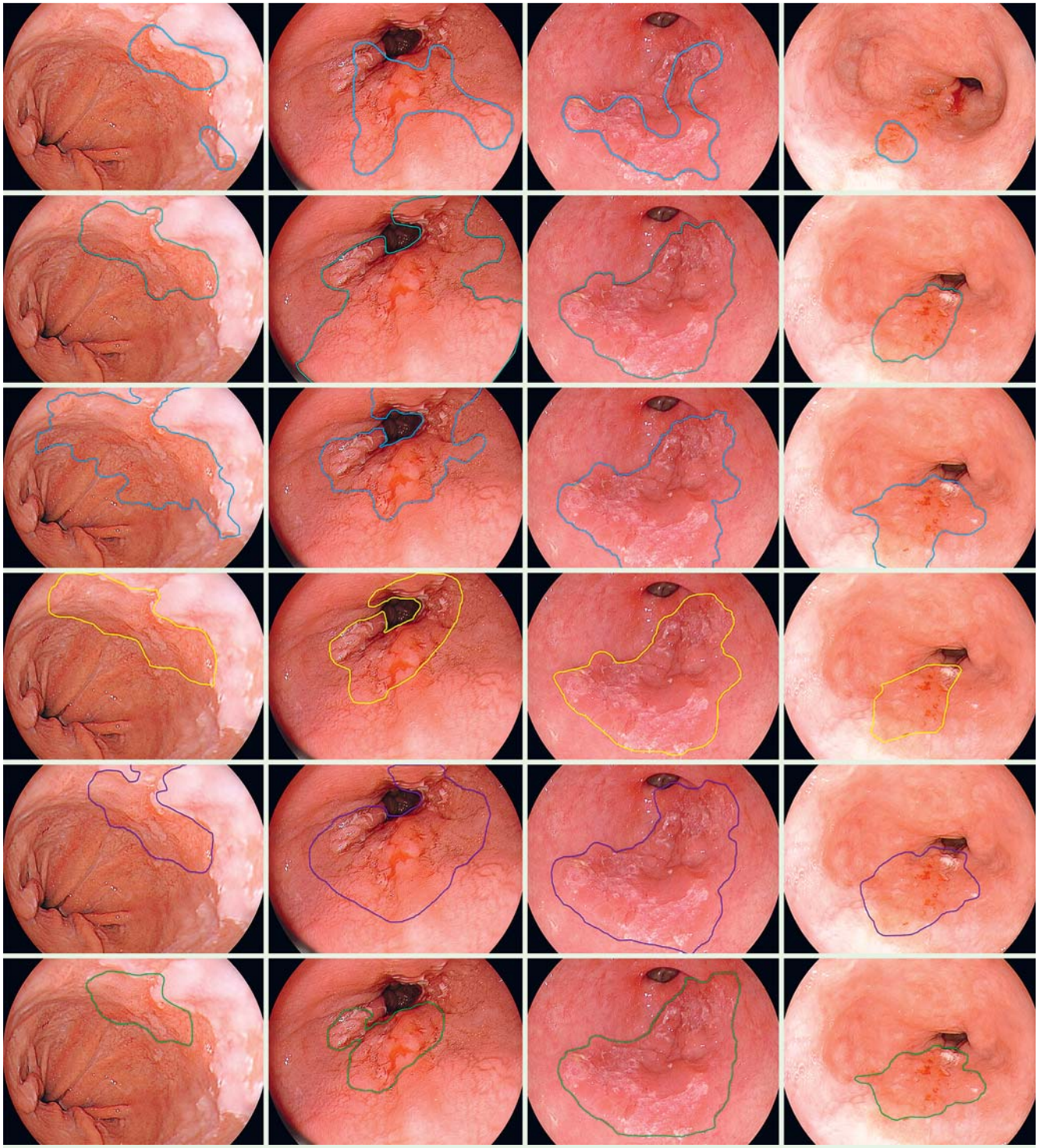

**Fig. 7** System annotations (top row, blue) vs. gold standard annotations (second row, green) and the expert annotations (bottom four rows, blue, yellow, purple, and dark green, respectively). These annotations have been generated by the system that was trained with 31 % positive training windows (see **Tables 1 – 3**). The F scores of the system are 0.53, 0.49, 0.69, and 0.34, respectively.

ar percentage of 31 % positive training samples was chosen for presentation, because the system achieved the most balanced per-image detection performance for this value.

## Discussion

To our knowledge, this is the first study to investigate the feasibility of a supportive detection system for early neoplastic lesions in Barrett's esophagus. A system was trained based on a set of gold standard annotations of 100 high definition endoscopic images. These images were also annotated by four international experts, who were blinded to pathology, in order to validate the system performance.

Although the ROC curve (**Fig. 5**) showed that the experts scored considerably higher than the annotation system on a per-image analysis, the performance of this automated recognition system is very promising for a per-patient analysis, where it performed similarly to the experts (**Table 2**). Preservation and incorporation of valuable endoscopic innovations (PIVI) thresholds suggested in the guidelines of the American Society for Gastrointestinal Endoscopy, are a per-patient sensitivity of 90 % or higher and a negative predictive value of 98 % or higher. Given the results of the current feasibility study, this automated recognition system has the potential to comport to the PIVI standards.

The ROC plot raises the question of where nonexperts would score. Our expectation is that the system will outperform most of the nonexperts. However, this is a subject for subsequent research. Furthermore, the performance gap will be narrowed by improvement of the detection, fuelled by more clinical training data.

It is clear that the delineations of the algorithm did not match the gold standard annotations as closely as the annotations by the experts did; the experts scored significantly better and by a relatively large margin (**Table 3**). However, when more positive training samples were used, the system annotation score approached that of the experts (see **Fig. 6**), although the difference was still large. Furthermore, increasing the fraction of positive training samples also led to a reduction in specificity. This effect can be explained by the observation that when the PTF is increased, the tendency of the system to classify an image block as neoplasia also increases. This results in less conservative annotations and, in turn, to more false positives. This effect is not desirable when keeping a future clinical application in mind, and a high detection performance should be preferred over a high delineation similarity; once a lesion is detected by the system, the endoscopist can further inspect the tissue and determine its boundaries.

Although the initial results are promising, the nature of this first proof-of-principle study includes some important limitations, and several hurdles have to be overcome before the system can be implemented in clinical practice. First, the study is susceptible to selection bias because only the best-quality images, with clear visible lesions in a clear window of view, were selected. Second, the number of images per patient varied; hence, it is possible that a learning effect occurred when the expert annotated several images from the same patient. This has the effect of improving the consistency of the expert results over time, whereas the annotation system could not exploit this capability. It was also noted that the experts may have scored better if a video was provided for each patient, such that a better sense of depth and height of the mucosa was attained. Third, the annotations of the endoscopist who resected the lesions were adopted as the gold standard; the endoscopist was completely informed about the histological outcome. Although pathology results showed the presence of HGD or early carcinoma, the current gold standard cannot guarantee that all the annotated pixels are depicting the lesion. Hence, a consensus-type gold standard is desirable, as it will better reflect the actual annotation accuracy. Fourth, all endoscopic images were acquired using the same type of endoscope. Using a data set acquired by different types of endoscopes would better represent the variation in images that is expected in clinical practice. The system should be accurate for all white-light endoscopy systems, which would validate the robustness of the annotation system. Such a study is currently in preparation. Fifth, at a more technical level, only one machine learning method is explored in this feasibility study, and for only a limited number of parameters. Future studies should include a comparison with alternative methods for a wide range of parameters in order to investigate the robustness and reliability of a computer-aided detection system for detecting early cancer in Barrett's esophagus. Furthermore, the system currently only works for images. As a video is basically a set of images over time, a detection system for endoscopic video is feasible. However, the validation of endoscopic video requires frame-by-frame expert annotations, which would be very laborious to produce.

In summary, this is the first study in which a detection algorithm for early neoplastic lesions in Barrett's esophagus has been proposed and compared with expert annotations. Although the system did not reach the expert level in annotating the images, the detection results were promising, with a per-image sensitivity and specificity of 0.83, and a per-patient sensitivity and specificity of 0.86 and 0.87, respectively, where a trade-off between the system sensitivity and specificity can be made. Based on the results of this proof-of-principle study, we conclude that a detection algorithm for early neoplastic lesions in Barrett's esophagus is feasible. However, additional research is required to improve the accuracy of the system and to prepare it for real-time operation, before it can be applied in clinical practice. Further studies should investigate the detection rates of nonexperts compared with the detection rates of the system. Additional development of the annotation system should include the use of temporal information to further increase the algorithm detection rate and its consistency.

**Competing interests:** None

**Institutions**
[1] Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands
[2] Department of Gastroenterology, Catharina Hospital, Eindhoven, the Netherlands
[3] Department of Gastroenterology, University Hospitals Leuven, KU Leuven, Leuven, Belgium
[4] Gastroenterology and Interventional Endoscopy, St. John of God Hospital, Regensburg, Germany
[5] Department of Gastroenterology, St. Antonius Hospital, Nieuwegein, the Netherlands
[6] Department of Gastroenterology, Amsterdam Medical Center, Amsterdam, the Netherlands

## References


1 *Lagergren J*, *Lagergren P*. Oesophageal cancer. BMJ 2010; 341: c6280
2 *Dent J*. Barrett's esophagus: a historical perspective, an update on core practicalities and predictions on future evolutions of management. J Gastroenterol Hepatol 2011; 26: 11 – 30
3 *Lepage C*, *Rachet B*, *Jooste V* et al. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. Am J Gastroenterol 2008; 103: 2694 – 2699
4 *Behrens A*, *Pech O*, *Graupe F* et al. Barrett's adenocarcinoma of the esophagus: better outcomes through new methods of diagnosis and treatment. Dtsch Arztebl Int 2011; 108: 313 – 319
5 *Phoa KN*, *Pouw RE*, *Bisschops R* et al. Multimodality endoscopic eradication for neoplastic Barrett oesophagus: results of an European multicentre study (EURO-II). Gut 2015; DOI: 10.1136/gutjnl – 2015-309298
6 *Reid B*, *Blount P*, *Feng Z* et al. Optimizing endoscopic biopsy detection of early cancers in Barrett's high-grade dysplasia. Am J Gastroenterol 2000; 95: 3089 – 3096
7 *Barbosa DJC*, *Roupar D*, *Lima CS*. Multiscale texture descriptors for automatic small bowel tumors detection in capsule endoscopy. In: Olkkonen H , ed. Discrete wavelet transforms – Biomedical applications. Rijeka, Croatia: Intech; 2011: 155 – 174
8 *Li B*, *Meng MQ-H*. Computer-aided detection of bleeding regions for capsule endoscopy images. IEEE Trans Biomed Eng 2009; 56: 1032 – 1039
9 *Li B*, *Meng MQ-H*. Texture analysis for ulcer detection in capsule endoscopy images. Image Vis Comput 2009; 27: 1336 – 1342
10 *Kodogiannis VS*, *Boulougoura MG*, *Lygouras J* et al. A neuro-fuzzy-based system for detecting abnormal patterns in wireless-capsule endoscopic images. Neurocomputing 2007; 70: 704 – 717
11 *Karkanis SA*, *Iakovidis DK*, *Maroulis DE* et al. Computer-aided tumor detection in endoscopic video using color wavelet features. IEEE Trans Inf Technol Biomed 2003; 7: 141 – 152
12 *Oh J*, *Hwang S*, *Lee J* et al. Informative frame classification for endoscopy video. Med Image Anal 2007; 11: 110 – 127
13 *van Vilsteren FGI*, *Herrero LA*, *Pouw RE* et al. Radiofrequency ablation for the endoscopic eradication of esophageal squamous high grade intraepithelial neoplasia and mucosal squamous cell carcinoma. Endoscopy 2011; 43: 282 – 290
14 *Schlemper RJ*, *Riddell RH*, *Kato Y* et al. The Vienna classification of gastrointestinal epithelial neoplasia. Gut 2000; 47: 251 – 255
15 *Kara MA*, *Curvers WL*, *Bergman JJ*. Advanced endoscopic imaging in Barrett's esophagus. Tech Gastrointest Endosc 2010; 12: 82 – 89
16 *van der Sommen F*, *Zinger S*, *Schoon EJ* et al. Supportive automatic annotation of early esophageal cancer using local gabor and color features. Neurocomputing 2014; 144: 92 – 106
17 *Setio AAA*, *van der Sommen F*, *Zinger S* et al. Evaluation and comparison of textural feature representation for the detection of early stage cancer in endoscopy. Proceedings of the 8th International Conference on Computer Vision Theory and Applications, 2013 February 21–24 Barcelona, Spain: 2013: 238 – 243