Computers in Biology and Medicine 96 (2018) 203–213

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

A survey on Barrett's esophagus analysis using machine learning

Luis A. de Souza Jr. [a,b], Christoph Palm [b,c], Robert Mendel [b], Christian Hook [b], Alanna Ebigbo [e], Andreas Probst [e], Helmut Messmann [e], Silke Weber [d], João P. Papa [a,*]

[a] Department of Computing, São Paulo State University, UNESP, Brazil
[b] Regensburg Medical Image Computing (ReMIC), Ostbayerische Technische Hochschule Regensburg (OTH Regensburg), Germany
[c] Regensburg Center of Biomedical Engineering (RCBE), OTH Regensburg and Regensburg University, Germany
[d] Department of Otorhinolaryngology, São Paulo State University, Brazil
[e] Medizinische Klinik III, Klinikum Augsburg, Germany

ARTICLE INFO

Keywords:
Barrett's esophagus
Machine learning
Adenocarcinoma
Image processing
Pattern recognition
Computer-aided diagnosis

ABSTRACT

This work presents a systematic review concerning recent studies and technologies of machine learning for Barrett's esophagus (BE) diagnosis and treatment. The use of artificial intelligence is a brand new and promising way to evaluate such disease. We compile some works published at some well-established databases, such as Science Direct, IEEEXplore, PubMed, Plos One, Multidisciplinary Digital Publishing Institute (MDPI), Association for Computing Machinery (ACM), Springer, and Hindawi Publishing Corporation. Each selected work has been analyzed to present its objective, methodology, and results. The BE progression to dysplasia or adenocarcinoma shows a complex pattern to be detected during endoscopic surveillance. Therefore, it is valuable to assist its diagnosis and automatic identification using computer analysis. The evaluation of the BE dysplasia can be performed through manual or automated segmentation through machine learning techniques. Finally, in this survey, we reviewed recent studies focused on the automatic detection of the neoplastic region for classification purposes using machine learning methods.

1. Introduction

The adenocarcinoma appearance in Barrett's esophagus (BE) diagnosed patients has increased significantly in western populations. This is mainly explained by obesity, a known risk factor [1–3]. As such, the expectation of this disease to rise in the next years must be considered. The bad prognosis for patients suffering from esophageal adenocarcinoma is related to its late diagnosis. However, when detected at the early stages, the dysplastic tissue can be treated with very successful rates of handling the disease, such as 5% of morbidity and 0% of mortality. Additionally, 93% of patients featured a complete remission of the disease after 10 years treatment [2,4,5]. Developments in interventional therapies, such as endoscopic resection and ablation techniques (radiofrequency ablation, cryoablation) are promising methods for the management of BE, with the potential of reducing the cancer risk in dysplasia diagnosed patients. However, there are limitations of the currently accepted methods for monitoring and evaluating the disease state of BE patients, with the benefit from early diagnosis and additional tools to improve the detection of dysplasia [6–8].

Several endoscopic technologies for image enhancement, such as chromoendoscopy, electronic image enhancement, optical coherence tomography, and confocal laser endomicroscopy have been developed for BE evaluation, enabling endoscopists to conduct a more accurate assessment of the dysplasia with an *in vivo* characterization of esophageal histology [9]. This ability could result in improvements concerning the detection of BE (screening), detection of dysplasia based on BE surveillance, characterizing abnormalities within BE (selecting lesions and delineating margins during endoscopic therapy), and detection of recurrent neoplasia in patients who have received endoscopic treatment (post-treatment surveillance) [9].

BE is often misdiagnosed during endoscopy because of: (1) inability to differentiate columnar mucosa of the proximal stomach (cardia) from metaplastic epithelium in the distal esophagus; or (2) lack of goblet cells in biopsies obtained from columnar lined epithelium in the esophagus. Since dysplasia/BE areas are sometimes not readily perceived with standard white-light endoscopy, the Seattle biopsy protocol is usually recommended, where biopsies are taken for every 1 cm of the BE's mucosa. However, this protocol may be susceptible to sampling errors

* Corresponding author.
E-mail address: papa@fc.unesp.br (J.P. Papa).

https://doi.org/10.1016/j.compbiomed.2018.03.014
Received 19 December 2017; Received in revised form 22 March 2018; Accepted 22 March 2018

because only a small part of the entire BE mucosa is usually considered for sampling purposes, especially in patients with extensive disease area [9]. Besides, the biopsy protocol can be costly and time-consuming, and thus prone to errors. Consequently, the risk of missed dysplasia or cancer diagnosis rises significantly [10]. Other studies have also considered the classification of different esophagus' lesion types based on color and texture information of the injured tissue area as well [11].

Machine learning techniques have benefited from significant improvements in image analysis and artificial intelligence fields. However, related to the automated analysis of BE, we observed one recent work only that attempted to compile relevant articles [12]. This work is a very brief survey to discuss advances in BE Computer-aided diagnosis (CAD) systems in three endoscopy modalities used for esophageal examination: (i) white light endoscopy (WLE), (ii) high-definition white light endoscopy (HD-WLE), and (iii) narrow band imaging (NBI). Focusing on detection methods lately developed for BE detection, the survey is composed of eight papers about automatic detection and evaluation of the BE, compared by its endoscopy modality, number of images and evaluated classifiers applied to the problem, validation method and results. The authors state the challenges for this detection and mention some directions for future research.

Our work aims at reviewing and investigating the feasibility and usage of machine learning techniques in the context of BE evaluation, dysplasia description, and treatment, thus providing more details to the previous brief survey. Next sections present the methodology used to evaluate the compiled articles, as well as some medical background related to the disease.

2. Theoretical background

2.1. Barrett's esophagus

The replacement of squamous cells by columnar cells in the esophagus' mucosa is known as BE. This process is recognized as a complication of gastroesophageal reflux disease, and in some critical stage, it can progress and evolve into esophageal cancer. Fig. 1 illustrates the human esophagus region [2,4,13].

Squamous cells (similar cells to skin or mouth ones) compose the mucosa of the normal esophagus. The normal color of the squamous mucosal surface looks like whitish-pink, while the gastric mucosa goes sharply from salmon-pink to red [2,4]. A demarcation line called squamocolumnar junction or "Z-line" defines the normal esophagogastric junction (Fig. 2), where the squamous mucosa of the esophagus and the columnar mucosa of the stomach meet [13]. BE's mucosa may extend upward in a continuous pattern, making the entire circumference of the distal esophagus covered by columnar mucosa. A difference is stablished among patients with more than 3 cm of BE ("long-segment BE") and those who feature the so-called "short-segment BE", with refers to BE that figures less than 3 cm, as depicted in Fig. 3.

2.2. Machine learning

Machine learning techniques have been paramount in the last decades mainly due to their capability in handling problems non-linearly by nature. Given a dataset composed of samples, the traditional pipeline used for so many years considers partitioning the data into training and testing sets. The former is used to learn the model (i.e., statistics of the data) meanwhile the testing set is used to assess the efficiency of the method.

Depending on the amount of knowledge we have about the training set, machine learning techniques can be categorized into three main groups: (i) supervised, (ii) semi-supervised, and (iii) unsupervised approaches. Supervised techniques usually achieve the best results since they make use of an entirely labeled training set, thus having more information to cope with. Semi-supervised learning approaches make use of both labeled and unlabeled data since only a fraction of the training data is labeled. Different approaches such as active learning-based or reinforcement learning can be referred to as well. In a nutshell, these approaches employ the user knowledge into the learning process, which can thus refine the results and correct possible errors.

Unsupervised learning or clustering stands for the group of techniques that have no information about the training data, which means they must group the data using some heuristic that can get together (i.e., same cluster) samples that share some information. To evaluate such techniques, we usually make use of measures that take into account the compactness and separability of clusters in the feature space, i.e., it is highly desirable to have well-separated and compact clusters at the end

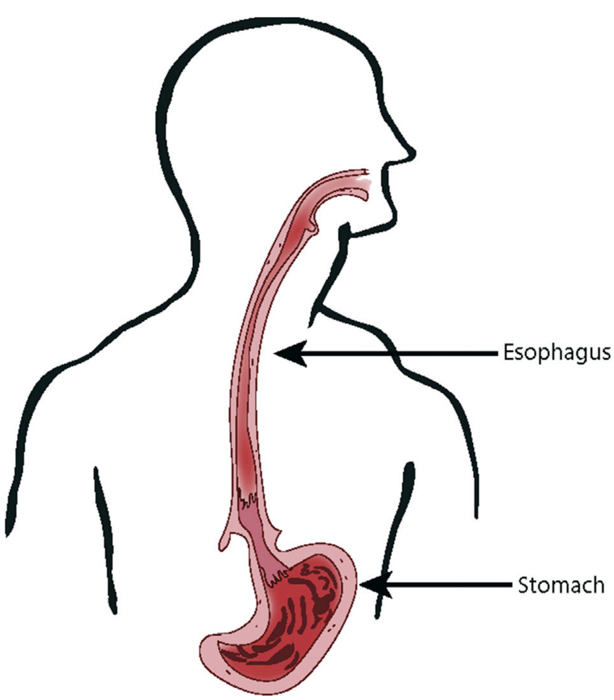

Fig. 1. Esophagus' location in the human body.

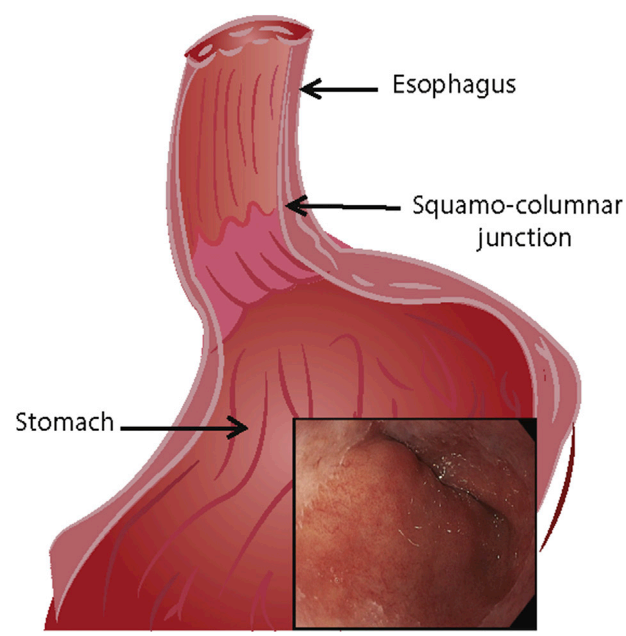

Fig. 2. Squamo-columnar junction and its respective esophagus endoscopic image.

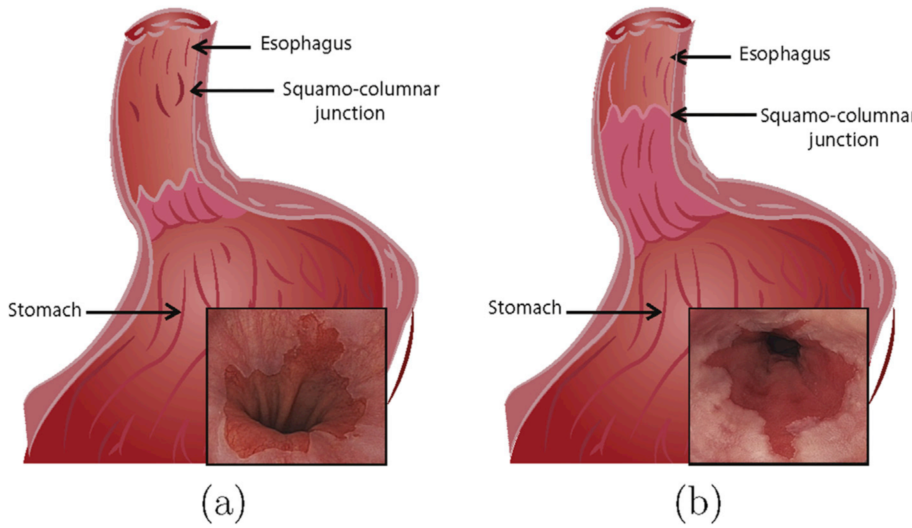

Fig. 3. Endoscopic views from: (a) BE's short-segment and (b) BE's long-segment.

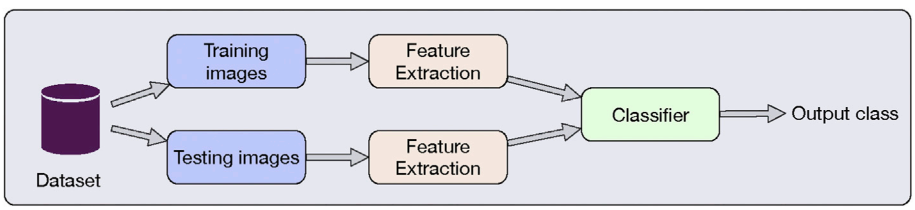

Fig. 4. Standard pipeline used in machine learning-driven applications.

of the clustering process. Fig. 4 depicts a standard pipeline used in applications that use machine learning for solving problems.

3. Surveyed works

In this section, we present the works considered for further study and discussion. The next subsections describe in more in-depth details the works selected by their primary classifier employed.

3.1. Paper selection

To select works within the scope addressed in this systematic review, a search in Science Direct, IEEEXplore, PubMed, Plos One, Multidisciplinary Digital Publishing Institute (MDPI), Association for Computing Machinery (ACM), Springer, and Hindawi Publishing Corporation databases was carried out. To this end, only two keywords were considered for searching purposes: (i) “Barrett's esophagus” and (ii) “Barrett's esophagus machine learning”. The main idea is to provide a fair selection of works and to cover a total of 35 recent works published as follows: before 2011 (5 works), between 2011 and 2014 (6 works), and early 2015–2017 (24 works). Also, the search returned a number of papers not related to machine learning-assisted BE analysis. Therefore, the outcoming of this survey does not consider them all.

3.2. Machine learning analysis of the BE

Machine learning is a branch of computational intelligence dedicated to the development of algorithms that enable a computer program to improve its performance based on learned information. The intense research in this field has motivated a number of works that aimed at using machine learning-oriented techniques to aid BE recognition and distinction between adenocarcinoma and healthy tissue.

In this survey, the works are divided according to the classifiers they have employed to cope with BE identification. Considering the number of works that used Support Vector Machines (SVM) and neural networks, we decided to have a dedicated section for each them. Additionally, a section presenting the comparison of two or more classifiers is also considered, followed by a description of other techniques applied to BE identification, such as *k*-nearest neighbors (*k*-NN), *k*-statistics, and decision trees, among others.

3.2.1. Support Vector Machines-based Barrett's esophagus recognition

Li and Meng [14] presented a new texture-based protocol for ulcer regions using capsule endoscopy (CE) discrimination in endoscopic images. A novel approach based on curvelets and Local Binary Patterns (LBP) was proposed for texture extraction aiming to distinguish ulcer and normal regions. These new features are sensitive to illumination changes, multidirectional features, and feature invariance. Experiments were conducted using two different classifiers on a 4-fold cross-validation procedure: (i) a Multilayer Perceptron Neural Network (MLP) and (ii) Support Vector Machines. The database used for the experiments is private and composed of 100 images from 5 different patients. Regarding the images, $1,800$ patches of normal images and $1,800$ patches of ulcer-diagnosed images were extracted. The authors concluded the proposed textural features were suitable to identify ulcerous regions in CE images, once detection rates of the proposed features with the MLP were 92.37% of accuracy, 91.46% of specificity, and 93.28% of sensitivity.

Rodriguez-Diaz and Singh [15] proposed a computer-based approach that employs the NICE criterion for diagnostic purposes, as well as it provides an on-the-fly interpretation of the histology of the polyps represented by near-focus NBI (NF-NBI) images. The NICE criterion considers three main characteristics when learning the information that may be useful to deal with BE identification: color, vasculature, and surface pattern. The color information was used to encode the tone of neoplastic regions compared to non-neoplastic polyps, which appears to be brown-colored. Regarding the vessels, the authors performed an automatic segmentation of the inter-crypt space and compared its color to the remaining tissue of the polyp to distinguish brown vessels from lighter structures around them. Finally, the authors employed the discrete wavelet transform to describe the local spatial distribution of gray levels (green and blue channels) and then characterize the neoplastic and non-neoplastic regions. Individual features were used as inputs for SVM in a leave-one-out cross validation (LOOCV) protocol. A total of 26 patients and 56 images (16 non-neoplastic and 40 neoplastic polyps) were considered for the database composition. The classification results achieved 86% of sensitivity and specificity.

Nancarrow et al. [16] performed a comparative study to define convincing differences between BE and esophageal adenocarcinoma (EAC) in biopsies from selected patients using SVM classification. A database composed of 54 biopsy specimens from 54 different patients were used for validation purposes, and certified by a pathologist (23 annotated as presenting EAC). The results concluded that BE-affected regions figure a tissue containing an enhanced glycoprotein synthesis mechanism designed to provide mucosal defenses. Such mechanism resists to gastro-esophageal reflux, while EAC exhibits the enhanced extracellular remodeling effects expected in an aggressive form of cancer. Also, evidence of reduced expression of genes associated with mucosal and xenobiotic defenses was also perceived. The authors observed eleven genes that are also represented in at least three other profiling studies used to discriminate among squamous epithelium, BE, and EAC, within the two largest cohorts using an SVM-based LOOCV analysis. The proposed method was considered able to distinguish squamous epithelium and BE reasonably, and it also evidenced that more detailed investigations into profiling changes between BE and EAC are desired. The work mentioned above achieved the following results: sensitivity and specificity higher than 88% concerning the task of discriminating BE from squamous samples, as well as a sensitivity of around 73% when distinguishing EAC (cancer) from BE or squamous (non-cancer).

Veronese et al. [17] proposed a computer-assisted approach to distinguish gastric metaplasia (GM), intestinal metaplasia (IM), and neoplasia (NPL) based on the use of appearance features in confocal images. The database was composed of CLE images obtained from consecutive 29 BE patients undergoing surveillance. In a nutshell, features are extracted based on the division of the image in sub-regions for the further application of LBP for a multiscale evaluation. The evaluation of the method was performed by the comparison of the automatic results with the histological gold standard using SVM-based classifiers. The proposed method identified IM, GM, and NPL in confocal images with accuracy close to the human observer. The validation protocol adopted was the LOOCV, and the overall sensitivity results were: 96% for GM, 95% for IM and 100% for NPL.

Using the technology of High Definition (HD) endoscopy, Muldoon et al. [18] developed a CAD system to help physicians with faster prognosis and decrease the diagnosis miss rate in the context of early-stage cancer detection. The work compared several techniques for texture-oriented feature extraction, including Gabor, co-occurrence matrix features, and LBP. For a better image description, an efficient combination of color and texture features were proposed. A pre-processing step designed for endoscopy images was also considered to improve the classification accuracy. Later on, Principal Component Analysis was used to reduce the number of features for the further usage of SVM. The experimental results were validated by a gastroenterologist and showed a classification accuracy up to 96.48%, in 129 sites calculated in a database composed of 16 HMRE images.

Van der Sommen et al. [19] presented an approach based on HD endoscopic images for the automatic esophagus irregularity identification. They employed the concept of tile-based image processing, so that the system was able to identify early cancer and also locate it in endoscopic images. The identification process was based on the following steps: (i) pre-processing, (ii) feature extraction with dimensionality reduction, and further (iii) classification. The performance detection was evaluated in RGB, HSI, and YCbCr color spaces using the Color Histogram and Gabor features in a database of HD endoscopic images obtained from 66 patients. Other well-known texture features were also considered for comparison purposes. Concerning the classification step, an SVM configured with different parameters and kernel functions were applied. The proposed approach achieved a classification accuracy of 95.9% considering tiles of tumorous and normal tissue of 50×50 pixels, with area under the curve (AUC) of 99%.

Van der Sommen et al. [11] proposed a novel algorithm that calculates local texture and color features based on the original and Gabor-filtered images for the automatic detection of early cancer in high-definition endoscopic images. Appropriate filters based on spectral characteristics of the cancerous regions were designed, and post-processing techniques were further applied to annotate the injured regions and the features were extracted and classified by a trained SVM classifier. For seven evaluated patients, the experiments compared 32 annotations performed by the algorithm with the corresponding annotations made by a gastroenterologist expert using a LOOCV protocol. From 38 lesions highlighted independently by the gastroenterologist, 36 of those lesions with a recall of 95% and precision of 75% were correctly detected by the system.

Hassan and Haque [20] proposed a real-time and computationally efficient bleeding detection technique using Wireless Capsule Endoscopy (WCE) technology. The technique was based on the observation of characteristic patterns present in the frequency spectrum of WCE frames. After these patterns have been defined, the authors developed a texture-based feature descriptor that operates on the Normalized Gray Level Co-occurrence Matrix (NGLCM) in the magnitude spectrum of the images. This descriptor was called Difference Average. The proposed algorithm was validated using a WCE database; the SVM training set was composed of 600 bleeding and 600 non-bleeding frames. Additionally, 860 bleeding and 860 non-bleeding images were chosen from the remaining images to compose the test set. The accuracy, sensitivity, and specificity values achieved were 99.19%, 99.41%, and 98.95%, respectively. The proposed method requires a low computational cost, thus making it suitable for real-time implementations.

Souza Jr. et al. [21] conducted a study to test the feasibility of adenocarcinoma classification in endoscopic images. The 2016 Endovis Challenge database [22] was used for the further extraction of Speed-Up Robust Features (SURF) [23], which were employed together with SVM using the LOPOCV protocol for training and testing purposes. Two classes composed the problem: non-cancerous- and cancerous-annotated images. Two approaches for feature extraction and classification were carried out: using the full images and using the expert-annotated regions of the adenocarcinoma. The results for the “full images approach” were 77% of sensitivity and 82% of specificity. For the “regions approach”, the results were 90% of sensitivity and 95% of specificity.

Zhang et al. [24] conducted a study using endoscopic ultrasonography (EUS) to calculate textural features in a spectral analysis of pixels to provide a quantification of early esophageal carcinoma tissue. A database composed of $1,210$ EUS examination samples was used from 66 patients with early esophageal cancer and 91 without cancer. The textural features of the EUS images were represented as a graph, in which the pixels are the nodes and the similarity between the gray-level or local features of the images are the edges. The similarity were provided by a high-order graph matching of the texture features. Finally, a 10-fold cross-validation approach was considered, and an SVM classifier was applied to calculate the optimal prediction of the esophageal carcinoma samples represented in the graph. As the primary results, the authors obtained 93% of accuracy in the prediction of early esophageal carcinoma, normal and leiomyoma tissues. Considering only the early esophageal carcinoma prediction, the average results of accuracy, sensitivity, specificity and negative prediction were 89.4%, 94%, 95%, and 97%, respectively.

Klomp [25] explored the feasibility in the use of computer vision techniques to correctly predict the presence of dysplastic tissue in VLE images. Three new features based on the classic Haralick features were proposed, and the SVM classifier was applied in a dataset composed of 30 dysplastic BE images and 30 non-dysplastic BE images. Using a 10-fold cross-validation protocol, the authors obtained an area under the Receiver Operating Characteristic (ROC) curve as of 0.95 compared to the 0.81 achieved by the clinical prediction model.

3.2.2. Neural network-based Barrett's esophagus recognition

Seguí et al. [26] introduced a system for small intestine motility identification based on Deep CNN to avoid the time-consuming step of specifying features for each motility event. This study aimed to help physicians with the diagnosis performed by the WCE video screening. Concerning the network training, 100, 000 annotated WCE samples were used, and the remaining 10, 000 samples were employed for testing purposes. The experimental results evidenced the robustness of the new features over others designed using state-of-the-art handcrafted approaches. In particular, the proposed approach obtained a mean accuracy of 96% for six intestinal motility events. Such result allowed the proposed approach to outperform other classifiers trained with classic handcrafted features (a 14% relative performance increase was observed).

Mendel et al. [27] carried out a work in which deep learning was applied in specialist-annotated images containing adenocarcinoma and BE's disease. A dataset provided by the 2016 Endovis Challenge [22] was used for the experimental step, and it comprises 100 annotated endoscopic images (50 presenting BE and 50 presenting adenocarcinoma) from 39 patients (17 not presenting adenocarcinoma and 22 presenting adenocarcinoma). The convolutional neural network was adapted to the set of images by a transfer learning approach in a leave-one-patient-out cross-validation (LOPOCV). With positive results of sensitivity and specificity (94% and 88% respectively), the study demonstrated that it is possible to extend its results to the BE's esophageal segmentation domain itself using deep learning to reach the region affected by adenocarcinoma specifically.

To cope with the problem of time-consuming and cost-inefficient manually ground-truth definition, Georgakopoulos et al. [28] proposed a weakly-supervised learning technique based on CNN that uses only image-level semantic annotations for the training process, instead of using annotations at the pixel's level. The performance of the proposed method was evaluated in the context of CAD system of inflammatory gastrointestinal lesions represented in WCE videos. The results showed the proposed method could be more accurate than the conventional supervised learning with an accuracy of around 90% obtained in a previous proposed data set proposed by Ref. [29].

Chan et al. [30] used an e-nose, a device that utilizes chemical-to-electrical interfaces, to measure the volatile organic compounds (VOC) profiles of disease states. When paired with a machine learning platform, an e-nose can be trained as a canine to serve as a tool for noninvasive diagnostic testing. Such approach was used to perform a cross-sectional study evaluating the breath VOCs of a cohort of 112 patients (66 with BE and 56 without BE) with a history of dysplastic BE to differentiate the differences in BE by dysplasia grade. These VOC profiles were introduced into an artificial neural network in a supervised step to identify data classifiers to discriminate differences in subjects stratified by the presence or absence of BE by biopsies. Optimal models were validated using a leave-some-out cross-validation (LSOCV) approach to generate performance characteristics of BE detection in a CNN classification. The sensitivity and specificity was 82% and 80%, respectively, the accuracy was 81%, and the AUC was 79%. The task of analysis and interpretation of WCE records is complex and require sophisticated CAD systems to assist physicians in the video screening and further diagnosis. Because most of the capsule endoscopy CAD systems share a standard design, each time a new clinical application of WCE appears, a new CAD system has to be structured from scratch.

Yoshida et al. [31] performed a study to evaluate the classification accuracy of gastric biopsy specimens using the *e-Pathologist* image software. A dataset composed of 3062 gastric-biopsy specimen slides were used, being each one evaluated by at least two experts to provide the diagnosis. Finally, the comparison was performed between the experts and the e-Pathologist classification results. A cross-validation protocol was used together with a neural network, which achieved a recognition rate as of 55.6% over a three-class problem: (i) positive for carcinoma, (ii) caution for adenoma, and (iii) negative for a neoplastic lesion. An additional experiment was carried out in a two-class problem: (i) negative for neoplastic regions, and (ii) positive for neoplastic areas. In this analysis, the sensitivity, specificity, and negative predictive value were 89.5%, 50.7%, and 90.6%, respectively, showing a promising direction for the automated classification of injured regions of the intestine.

Hong et al. [32] developed a CAD system to classify endomicroscopy images between gastric metaplasia, intestinal metaplasia, and neoplasia (these last two are sub-classes of BE). A database provided by ISBI 2016 challenge and composed of 155 gastric metaplasia instances, 26 intestinal metaplasia instances and 55 neoplastic samples was considered for experimental purposes together with Convolutional Neural Networks in a cross-validation protocol. The training data were distorted for augmentation purposes as well, providing an accuracy as of 80.77%, thus suggesting that CNN might be useful to this kind of problem.

3.2.3. Comparison among classifiers for Barrett's esophagus recognition

Rajan et al. [33] performed a comparison experiment using several classifiers, such as SVM, *k*-NN, and Boosting on images from different endoscopy modalities (WLE, NBI, Chromoendoscopy). The datasets (125 WLE images, 122 NBI images, and 150 Chromoendoscopy images) have been classified between four categories: Normal Squamous, Gastric Mucosa, BE, and High-grade dysplasia (adenocarcinoma). The classification step was performed using features (i.g., color and texture) obtained from the injured regions of the endoscopic images in a cross-validation protocol. The accuracy ranged from 36.36% up to 89.17% according to the endoscopy modality images and classifier applied.

Considering the use of vibrational spectroscopy for the diagnosis and staging of cancer, Sattlecker et al. [34] conducted a study aiming to corroborate the many promising benefits in the current histopathology methods used in the context of BE identification. To correlate complex multivariate spectral and the disease level, the authors applied machine learning methods, such as SVM, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), and Random Forests to recognize spectral patterns. The validation protocols adopted were LOOCV, bootstrapping, and independent testing. A detailed review of related works was conducted, and the average recognition rates of the surveyed studies were around 90% of sensitivity and specificity, although the majority of the studies used less than 40 samples. The authors concluded that more studies need to be carried out in case we decide to put in practice the combination of spectroscopy and machine learning.

Kandemir et al. [35] performed a study for the diagnosis of BE presented in hematoxylin-eosin stained histopathological biopsies using multiple instance learning (MIL) and Support Vector Machines. Regarding the experiments, the database comprised 214 tissue cores (165 presenting cancer and 69 showing healthy condition) from 97 patients. Rectangular patches of the tissue cores were extracted, and a feature vector was calculated based on a large set of cell- and patch-level features (color features, texture features, and object features such as minimum, maximum, and standard deviation of the cells) for each patch. The tissue core as considered a bag (a group of instances with a unique group-level ground-truth label), while each patch was considered an instance. After many MIL approaches, the authors realized that a graph-based MIL algorithm (mi-Graph [36]) obtained the best performance explained by its inherent suitability to bags with instances that presents spatial correlation. For patch-level diagnosis, the result was around 82% of accuracy and 89% of AUC using Bayesian logistic regression in the distinction of BE and cancer region patches.

Considering the WCE as a promising technology for gastrointestinal disease examination in a non-invasive way, Yu et al. [37] studied the classification problem of the digestive organs for WCE images aiming to save the time of doctors in the image review task. Based on a previous study using Convolutional Neural Networks (CNN), a database composed of 25 real WCE recording samples (approximately 1 million of WCE images) was considered for experimental purposes, and with results nearly to 95% of accuracy. The authors also tried to improve the results by the proposition of a WCE classification system built as a hybrid CNN with an Extreme Learning Machine (ELM). In the new approach, the CNN was designed as a data-driven feature extractor, and the cascaded ELM was designed as a strong classifier instead of the conventionally use of a deep CNN fully-connected classifier. The results showed a performance of around 97% concerning the WCE organ classification accuracy.

Swager et al. [38] conducted a study to identify VLE features from neoplasia areas regarding BE identification, as well as the authors aimed to develop an approach to predict VLE scores. The work used a VLE image database composed of 52 endoscopic resection specimens from 29 BE patients, which were assigned positive and negative to neoplasia. Features potentially significant for the prediction of early BE neoplasia were identified over twenty-five VLE-histology images. In a learning phase, twenty VLE images presenting or not BE neoplasia were scored by two experts blinded to histology. A prediction score was developed by the use multivariable logistic regression analysis, being validated by scoring forty VLE images (50% neoplastic) using the area under ROC curve analysis. The work identified three main VLE features that can be used to assist BE neoplasia identification: (i) lack of layering, (ii) higher surface than subsurface signal, and (iii) presence of dilated glands/ducts. The sensitivity and specificity values obtained were 83% and 71%, respectively, showing promising accuracy.

Another study conducted by Souza Jr. et al. [39] introduced the Optimum-Path Forest (OPF) [40,41] classifier in the adenocarcinoma and BE classification using endoscopic images. The work considered describing endoscopic images (database provided by Ref. [22]) using feature extractors based on key point information, such as the SURF and Scale-Invariant Feature Transform (SIFT) [42], for further designing a bag-of-visual-words that were used as input to both OPF and SVM classifiers in a cross-validation protocol. The OPF classification outperformed the well-known SVM, presenting better results for both feature extractors, with values lying on 73.2% (SURF) - 73.5% (SIFT) for sensitivity, 78.2% (SURF) - 80.6% (SIFT) for specificity, and 73.8% (SURF) - 73.2% (SIFT) for the accuracy.

Pu et al. [43] performed a study to extract cost-efficient biomarkers more efficient than the ones currently available, aiming to provide high sensitivity and specificity in the task of esophageal squamous cell carcinoma (ESCC) diagnosis. The proposed biomarker to be evaluated was the DNA methylation, and 100 samples of ESSC DNA methylation from "The Cancer Genome Atlas" were analyzed along with a particular dataset of 12 samples of the same kind. Candidate CpG sites and their adjacent regions were defined and compared with adjacent normal tissue regions using several machine learning techniques such as Random Forest, SVM, CNN, Logistic Regression, Naive Bayes, LDA and Flexible Discriminant Analysis in a 5-fold cross-validation protocol. The sensitivity, specificity, and area under the curve results of the diagnostic model based on the combination of five genomic regions were 75%, 88%, and 85%, respectively.

Serpa-Andrade et al. [44] proposed a method in which the esophagitis (a condition of chronic BE stage) was described using Fourier Transform on the Z-line signature (esophageal irregularities) for classification purposes. The proposed descriptors were based on statical features and textural information. A database comprising 10 samples of healthy tissue and 16 samples of ill tissue was used, and a cross-validation protocol applied for classification purposes based on k-NN and Random Forests. The best average results obtained were around 81% of precision, 86% of sensitivity, and 72% of specificity.

3.2.4. Additional works

In this section, we present works that make use of classification techniques other than SVM and neural networks. Zopf et al. [45] proposed a study using NBI endoscopy images for the automatic detection of BE using a nearest neighbor classifier. The model extracted features from a proper database of images presenting 326 regions of Interest (ROIs) annotated by experts, and classified between three classes: epithelium, cardiac mucosa, and BE. The features applied to the classification were the co-occurrence matrices, summation and difference of histogram, statistical geometric, and Gabor Filters, leading to a high-dimension vector reduced later. The evaluation has been made measuring the performance of each feature and also by combining them all, using the LOOCV protocol and the Euclidean distance as similarity metric with the following results: accuracy between 85% and 92%. The best BE classification accuracy was around 74%.

In 2016, the BE's International NBI Group (BING) [4] performed a study with the goal of developing a simple and reliable approach to recognize dysplasia as well as esophageal adenocarcinoma (EAC) in individuals affected by Barrett's esophagus. The research group analyzed 60 NBI images containing nondysplastic BE, high-grade dysplasia, and esophageal adenocarcinoma to find out vascular and mucosal patterns visible by NBI to be used as features and then creating the BING criteria. Further, patients that were under supervision or endoscopic treatment for BE were recruited at four institutions in the United States and Europe, performing histologic biopsy analysis and composing a high-quality NBI image database. Experts reviewed 50 NBI images to validate the proposed approach and then evaluated 120 additional NBI images (without previous review) to assess its prediction accuracy. The proposed method identified patients with dysplasia with 85% of overall efficiency.

Pech et al. [46] designed a study to evaluate the potential of endomicroscopy for predicting histology *in vivo* in patients presenting early stage of squamous cells in the esophagus. Twenty-one patients suspected to early squamous cell cancer and recommended for endoscopic therapy were included in this study, being their mucosal areas examined using confocal imaging and resulting in 43 lesion images. Each scanned lesion image was stored and *in vivo* diagnosis was performed during routine endoscopy. Biopsy specimens were extracted from every lesion. The confocal images were reviewed by two personnel blinded to the histology endoscopists. The overall accuracy using the k-statistics was 95%, and the sensitivity and specificity were 100% and 87%, respectively. Intraobserver agreement was close to perfection ($kappa = 0.95$), meanwhile the interobserver agreement was very relevant ($kappa = 0.79$).

Rosenfeld et al. [47] aimed at studying how data mining can be applied to assist the diagnosis of high-risk lesion patients affected by BE in HD videos. As the patient information is open to interpretation, the authors demonstrated that composite rules learned from many experts can be more accurate than that of a single expert. Such fact can be explained because even expert physicians can interpret endoscopy images differently, thus potentially making it relevant to aggregate multiple opinions for the precise interpretation of the endoscopic image. Also, the authors demonstrated that decision trees could learn simple rules to assist the dysplasia diagnosis. The authors employed two decision models in a dataset composed of 47 HD endoscopic videos of the esophagus (23 dysplastic and 24 non-dysplastic): one considering the expert decisions about dysplasia and no-dysplasia, and another without the expert decisions. The overall accuracies concerned the aforementioned models were around 79% (with the experts' decision) and 77% (without the experts' decision).

Li et al. [48] proposed a new learning method based on the multiple instance paradigm to recognize tumor invasion of gastric cancer using computed tomography (CT) imaging. The authors extracted bag-level- and instance-level features for processing and classification purposes using a database composed of 26 patient exams. Since there might be ambiguity when assigning labels to some selected patches in instance-level features, the authors proposed an improved Citation k-nearest neighborhood (Citation-kNN) that achieved recognition rates of around 76.92%.

Curvers and Bergman [49] carried out a study with patients undergoing BE neoplastic or with a suspect of such disease. Microscopic images of the esophagus were obtained from regions with neoplasia suspicious and from other locations randomly sampled for further biopsy analysis.

Table 1
Summarization of the works considered in this survey.

Reference	Classifier	Database	Validation Protocol	Evaluation Method	Results
Li and Meng [14]	SVM	100 endoscopic images from 5 patients	4-folder cross-validation	Presented a study for the application of a new texture extraction scheme in ulcer regions for capsule endoscopy discrimination in endoscopic images.	Detection rate of the proposed features with the MLP was 92.37%, 91.46%, and 93.28% in terms of accuracy, specificity and sensitivity, respectively.
Rodriguez-Diaz and Singh [15]	SVM	16 non-neoplastic and 40 neoplastic NBI images from 26 patients	LOOCV	Conducted a study aiming to explore the feasibility in developing a diagnostic computer algorithm based on the NICE criterion for a real-time interpretation of polyp histology from near-focus NBI images.	The classification based on color resulted in a sensitivity of 86% and specificity of 86% with a high-confidence rate of 77.%
Nancarrow et al. [16]	SVM	54 images from 54 patients (23 presenting EAC)	LOOCV	Performed a comparative study to define convincing differences between BE and EAC using SVM classification.	The results were: sensitivity and specificity higher than 0.88 for discriminating BE from squamous, and the sensitivity for determining EAC (cancer) from BE or squamous (non-cancer) was 0.73.
Veronese et al. [17]	SVM	CLE images from 29 consecutive BE patients undergoing surveillance	LOOCV	Presented a computer-based method for the automatic classification of gastric metaplasia, intestinal metaplasia, and neoplasia on the basis of appearance features of confocal images.	The sensitivity overall results were: 96% of gastric metaplasia, 95% of intestinal metaplasia, and 100% of neoplasia.
Muldoon et al. [18]	SVM	129 sites from a database of 16 HMRE images	cross-validation	Developed a CAD system to help physicians with faster identification of early cancer using color and texture features extracted from HD endoscopy images.	In an SVM-classification approach, the results reached were up to 96.48%.
van der Sommen et al. [19]	SVM	HD endoscopic images from 66 patients	10-fold cross-validation	Proposed an algorithm based on HD endoscopic images for the automatic esophagus irregularity identification and location using color histograms, Gabor features, and SVM-based classification.	The proposed system achieved a classification accuracy of 95.9% with AUC value as of 99%.
van der Sommen et al. [11]	SVM	32 HD endosocopic images from 7 patients	LOOCV	Presented a novel algorithm that computes local color and texture features based on the original and on the Gabor-filtered image for the automatic detection of early cancerous tissue in high defintion endoscopic images.	From 38 lesions, the system detected correctly 36 of those lesions with a recall of 0.95 and a precision of 0.75.
Hassan and Haque [20]	SVM	120,000 WCE frames for the training set and 1720 WCE frames for the test set	cross-validation	Proposed a real-time bleeding detection technique based on the observation of characteristic patterns that appear in the frequency spectrum of the WCE.	The accuracy, sensitivity, and specificity values achieved were 99.19%, 99.41% and 98.95%, respectively.
Seguí et al. [26]	CNN	100,000 WCE images for training the network and 10,000 for testing	cross-validation	Introduced a system for small intestine motility characterization based on Deep Convolutional Neural Networks.	The proposed approach obtained a mean classification accuracy of 96% for six intestinal motility events.
Mendel et al. [27]	CNN	Database provided by [22]	LOPOCV	Carried out a work in which deep learning was applied in specialist annotated images containing adenocarcinoma and BE's disease.	The sensitivity and specificity values of 94% and 88% respectively.
Souza Jr et al. [21]	SVM	Database provided by [22]	LOPOCV	Tested the feasibility of adenocarcima classification in endoscopic images using SURF features and SVM classifier.	The mean results for the "full image approach" were 77% of sensitivity and 82% of specificity. For the "region-based" approach, the results were 89.6% of sensitivity and 95.1% of specificity.
Zhang et al. [24]	SVM	1210 EUS images from 157 patients (66 with early cancer and 91 without)	10-fold cross-validation	A study was conducted using endoscopic ultrasonography (EUS) to calculate textural features in a spectral analysis of pixels to provide a quantification of early esophageal carcinoma tissue.	In the first classification approach, the overall concordance rate was 55.6% with kappa coeficient of 28%. The early esophageal carcinoma average prediction results of accuracy, sensitivity, specificity, and negative prediction were 89.4%, 94%, 95%, and 97%.
Klomp et al. [25]	SVM	30 VLE non-dysplastic images and 30 VLE dysplastic images	10-fold cross-validation	Tested the feasibility in the use of computer vision techniques correctly predict the presence of dysplastic tissue in VLE BE images.	Considering the novel proposed descriptors, the area under ROC curve result was 95%, compared to the 81% of the clinical predicition model.

(continued on next page)

Table 1 (*continued*)

Reference	Classifier	Database	Validation Protocol	Evaluation Method	Results
Georgakopoulos et al. [28]	CNN	Database proposed in Ref. [29]	Patches evaluation (normal or abnormal)	Proposed a weakly-supervised learning method based on CNNs that uses only image-level semantic annotations in the training process for ground truth calculation.	The results achieved by the authors showed the proposed method can be more effective than the conventional supervised learning with an accuracy of around 90%.
Chan et al. [30]	CNN	66 VOCs of patients presenting BE and 56 VOCs of patients without BE	LSOCV	Used an e-nose to perform a crosssectional study evaluating the breath VOCs of a cohort of 112 patients with a history of dysplastic BE to differentiate the differences in BE by dysplasia grade.	The sensitivity result was 82%, the specificity result value was 80%, the accuracy was 81%, and the AUC was 79%.
Yoshida et al. [31]	ANN	3062 gastric-biopsy specimen slides	cross-validation	Using textural features, were performed a study aiming to evaluate the classification accuracy of gastric biopsy spicimens using the *e-Pathologist* image software and expert annotations, in two different comparison approaches.	In the first classification approach, the overall concordance rate was 55.6% with kappa coeficient of 28%. In the second approach, the sensitivity, specificity, and negative predictive vale were 89.5%, 50.7%, and 90.6%, respectively.
Hong et al. [32]	CNN	155 gastric metaplasia instances, 26 instestinal metaplasia instances and 55 neoplastic instances	cross-validation	Developed a CAD system to classify endomicroscopy images between gastric metaplasia, intestinal metaplasia and neoplasia (these last two are sub-classes of BE) using a public database with 262 samples.	The accuracy result obtained was 80.77%, suggesting that CNN could become a good classifier for the task of BE tissue distinction.
Rajan et al. [33]	SVM, *k*-NN, Boosting 1/2	125 WLE images, 122 NBI images, and 150 Chromoendoscopy images	cross-validation	Performed experiments using several classifiers (SVM, *k*-NN, Boosting) in images from different endoscopy modalities (WLE, NBI, and Chromoendoscopy).	The accuracy for detecting BE presented a range of variation from 36.36% up to 89.17% according to the endoscopy modality and classifier.
Sattlecker et al. [34]	SVM, LDA, ANN, and RF	–	LOOCV, Boostraping and independent testing	Conducted a study aiming to cohoborate the many promising benefits over the currently used histopathology methods.	If the combination of spectroscopy and machine learning is mapped into clinical practice, more studies need to be carried out to support the reproducibility.
Kandemir et al. [35]	MIL and SVM	214 tissue cores (165 presenting cancer and 69 showing healthy condition) from 97 patients	cross-validation	Performed a study for the diagnosis of BE's cancer from hematoxylin-eosin stained histopathological biopsy images using multiple instance learning and SVM classifiers.	For patch-level diagnosis, the result was around 82% of accuracy and 0.89 of AUC using Bayesian logistic regression.
Yu et al. [37]	CNN and SVM	25 real WCE recording samples (approximately 1 million of WCE images)	CNN-features compared to SVM-features	Studied the classification problem of the digestive organs for WCE images.	The results showed performance of around 97.25% concerning the classification accuracy.
Swager et al. [51]	SVM, adaBoost, and *k*-NN	52 endoscopic resection specimens from 29 patients	LOOCV	Investigated the feasibility of a computer algorithm to identify early BE neoplasia in VLE images using VLE features and machine learning methods.	AUC of 0.95, and sensitivity and specificity were 90% and 93%, respectively.
Souza Jr et al. [39]	SVM and OPF	Database provided by Ref. [22]	cross-validation	Introduced the OPF classifier in the context of adenocarcinoma and BE classification using SURF and SIFT features combined with a bag-of-visual-words approach for the feature vectors calculation.	The OPF outperformed the SVM, presenting better results for both feature extractors, with values lying on 73.2% (SURF) - 73.5% (SIFT) for sensitivity, 78.2% (SURF) - 80.6% (SIFT) for specificity, and 73.8% (SURF) - 73.2% (SIFT) for the accuracy.
Pu et al. [43]	logistic regression, SVM, CNN, LDA, Naive Bayes and flexible discriminant analysis	100 samples of ESSC DNA methylation and a particular dataset of 12 samples of the same kind	5-fold cross-validation	Performed a study to extract more cost-efficient biomarkers (using DNA methylation) than the ones available until now, aiming to provide high sensitivity and specificity in the ESCC diagnosis.	In the SVM classification approach, the best average accuracy result were reached, with value of 0.82%.
Serpa-Andrade et al. [44]	*k*-NN and Random Forests	10 endoscopic images of healthy tissue and 16 images of ill tissue	cross-validation	proposed a method in which the esophagitis (a condition of cronic BE stage) was described using Fourier Transform on the Z-line signature for classification purposes.	The very best average results obtained were 81% of precision, 86% of sensitivity and 72% of specificity.
Zopf et al. [45]	euclidean distance	326 ROIs annotated by experts	LOOCV	Proposed a study using NBI endoscopy images for automatic detection by classification systems with gastroscopy.	Accuracy in the range of 85% and 92% for the feature combination (BE accuracy as of 74%).
Sharma et al. [4]	NBI classification criteria	50 NBI images plus 120 additional NBI images	Comparison between BING criteria and expert's annotations	Aimed to develop and validate a narrow-band imaging classification system for the identification of dysplasia and cancer in patients with BE.	The criteria identified patients with dysplasia with 85% of overall accuracy, 80% of sensitivity, 88% of specificity, 81% of positive predictive value, and 88% of negative predictive value.

(*continued on next page*)

Table 1 (*continued*)

Reference	Classifier	Database	Validation Protocol	Evaluation Method	Results
Pech et al. [46]	*k*-statistics	43 lesion images	confocal image review (2 experts)	Performed a study to assess the potential of endomicroscopy for predicting histology in patients with early squamous cell cancer in the esophagus.	The results were: accuracy value as of 95%, and the sensitivity and specificity as of 100% and 87%, respectively.
Rosenfeld et al. [47]	Decision Trees	47 HD endoscopic videos (23 dysplastic and 23 non-dysplastic)	–	Studied how data mining can be applied to aid the diagnosis of patients with high-risk lesions within BE.	The overall accuracies concerned the aforementioned models were around 79% (with the experts' decision) and 77% (without the experts' decision).
Li et al. [48]	Citation-*k*-NN	26 patients	LOOCV	Proposed a novel multiple instance learning method for the identification of tumor invasion of gastric cancer with dual-energy computed tomography imaging.	The experimental evaluation was performed using leave-one-out cross validation, obtaining an accuracy of 76.92%.
Curvers and Bergman [49]	Automated image classification	–	independent set of images	Carried out a study with patients undergoing BE neoplastic or surveillance underwent standard gastroscopy.	The results of the quantitative image classification algorithm showed a sensitivity of 84% and specificity of 85% in the learning set, and a sensitivity of 88% and specificity of 85% in the validation data.
Wang et al. [50]	Automatic image analysis	9 VLE volumetric datasets from 7 patients	–	Developed an algorithm that can automatically detect and quantify subsquamous glandular structures in VLE data and RFA in the decrease of the BE extension.	The subsquamous glandular structures depth and eccentricity characteristics were the most significant for the RFA outcome.
Swager et al. [38]	VLE features comparison	30 non-dysplastic VLW images and 30 high-grade dysplasia/early-adenocarcinoma VLE images	cross-validation	Conducted a study aiming to identify VLE features of BE neoplasia and to develop a VLE prediction score.	The sensitivity and specificity values obtained were 83% and 71%, respectively, showing promising results.
Boschetto et al. [52]	Random Forests	116 NBI images	10-fold cross-validation	Presented a CAD system to automate the classification of normal and metaplastic endoscopic NBI images. Eight features were extracted from regions defined as clusters of superpixels.	The overall accuracy, sensitivity -and specificity result values were 83.9%, 79.2% and 87.3%, respectively.
Ghatwary et al. [12]	–	–	–	Presented a brief survey to discuss advances in the development of BE CAD systems WLE, HD-WLE, and NBI endoscopy modalities.	Eight works were listed based on number of images, classifier, validation method, and results.

The database images were further identified as neoplastic or non-neoplastic by two experts (blinded for histology results) with experience in HRME image representation of BE. A tool for the visual interpretation of HRME images was designed for the reviewers to classify each image of the dataset. Also, three endoscopists with HRME experience annotated the entire image set using the developed tool. As a result, an analysis of the HRME images was performed based on the relevant image features selected for the classification step. A sequential and automatic image classification approach was developed and trained in a separate learning set. The results of this learning phase were validated in a selected set of images. The experimental results concerning sensitivity and specificity for neoplasia were around 81% and 76%, respectively, presenting a fair interobserver agreement. The results of the quantitative image classification algorithm achieved were: sensitivity of 84% and specificity of 85% for the learning set, and sensitivity of 88% and specificity of 85% for the validation data. These results corroborate that quantitative analysis of HRME images can provide an accurate classification of neoplastic and non-neoplastic BE tissue, which can be compared to the precision showed in the assessment of experienced endoscopists.

Wang et al. [50] developed an approach for the automatic detection and quantification of subsquamous glandular structures (SGSs) in volumetric laser endomicroscopy (VLE) data sets using automated image processing for the radiofrequency ablation (RFA) in the decrease of the BE extension. There were considered import information of SGSs right before RFA treatment, such as the average number, size, depth, and eccentricity per cross-section, and their correlations with the reduction of maximum BE length at follow-up after RFA were evaluated. After the analysis of nine VLE volumetric datasets from seven patients, there were found strong correlations between the SGS characteristics immediately before RFA, and the change of maximum BE length at follow-up after RFA. The SGS depth and eccentricity characteristics were the most significant for the RFA outcome.

Swager et al. [51] investigated the effectiveness of a computer-assisted tool to identify early BE neoplasia in sixty *ex vivo* VLE image using VLE features and machine learning methods for classification purpose. The database comprises sixty BE patients (30 nondysplastic BE - NDBE- and 30 high-grade dysplasia/early adenocarcinoma images). VLE features from a clinical VLE prediction score for BE neoplasia were used to feed the proposed approach, and novel clinically-inspired algorithm features were developed based on signal intensity statistics and grayscale correlations. The comparison was performed with generic image analysis methods for neoplasia detection. For classification purpose, several machine learning methods were evaluated, such as SVM, adaBoost, and *k*NN, allied to an LOOCV protocol. Three novel clinically-inspired algorithm features were developed as a result of the work, presenting an area under the receiver operating characteristic curve of 95%. Corresponding sensitivity and specificity were 90% and 93%, respectively.

Boschetto et al. [52] presented a CAD system to automate classification of normal and metaplastic endoscopic NBI images. Eight features were extracted from regions defined as clusters of superpixels, which are based on the superpixels of each region: three features are calculated as mean intensities of each color channel, three other features stand for mean intensities of the red-channel with the application of three different morphological filters (top-hat, entropy and range filters), and the last two features are related to the contrast and homogeneity of the superpixels. The classification step was performed using Random Forests in a 10-fold cross-validation approach on a dataset with 116 NBI samples. Following the feature extraction step, the samples were split into training (70% of the instances) and testing (30% of the instances) sets, and the overall accuracy, sensitivity and specificity results were 83.9%, 79.2%, and 87.3%, respectively.

4. Discussions and conclusions

In the last years, the amount of people with BE has increased considerably, mostly in the western countries, turning it in a world's health problem up to date. The use of artificial intelligence and machine learning techniques showed promising results, thus becoming a major aid to cope with BE pattern prognosis.

As it can be noticed in this survey, the application of machine learning has risen in the last years, with high use of SVM, CNN, and other methods for the detection and classification of adenocarcinoma or abnormalities in the esophagus region. In light of that, research in this area becomes very relevant and important for the early, fast, and standardized detection of BE and adenocarcinoma.

In this work, we presented a review concerning BE detection and monitoring using recent studies, being its main contribution to consider very recent works dating from 2011 to 2017 mostly, with the application of machine learning and computer vision for the description and classification between BE and adenocarcinoma. Additionally, the very recent studies for the detection, treatment, and evaluation of the BE are reviewed in this survey, and Table 1 presents a summarization of them all. Currently, based on the works considered in this survey, we can conclude the BE problem assisted by machine learning techniques is not mature yet, and there is a need for even more research to provide solid methods to distinguish the BE and adenocarcinoma regions in endoscopy images and videos.

We have observed the vast majority of works that use machine learning and computer vision for any BE purpose are brand new (between 2015 and 2017), thus highlighting new directions in which the prognosis and treatment of BE will benefit from the technologies to help experts in this task. Also, the definition of a pattern in the BE identification is very important, considering the massive human evaluation of this problem in practice today. We believe the computer learning and classification may help to define markers and identifiers for the best BE description, helping the accurate and fast definition of the injured region in endoscopy images or even in endoscopy videos in a real-time definition way.

The primary challenges related to computer-assisted BE identification are mainly associated with the lack of data since most of the datasets figure a few dozens of patients only. Another problem is related to the absence of public datasets to cope with BE identification, which could foster the research towards more effective approaches to detect early-stage illness from endoscopic images.

Another bottleneck concerns unbalanced data, which may bias the machine learning technique towards the majority class. Data augmentation appears to be an exciting solution together with transfer learning approaches. Also, we believe that combining handcrafted with automatic learned features may be a useful idea, mainly in the context of medical-drive data where the lack of images is of great concern.

Conflicts of interest

None Declared.

Acknowledgements

The authors thank DFG grant PA 1595/3-1, Capes/Alexander von Humboldt Foundation grant number BEX 0581-16-0, CNPq grants 306166/2014-3 and 307066/2017-7, as well as FAPESP grants 2013/07375-0, 2014/12236-1, and 2016/19403-6.

References

[1] J. Lagergren, P. Lagergren, Oesophageal cancer, BMJ 341, ISSN 0959–8138, URL http://www.bmj.com/content/341/bmj.c6280.
[2] J. Dent, Barret's esophagus: a historical perspective, an update on core practicalities and predictions on future evolutions of management, J. Gastroenterol. Hepatol. 26 (2011) 11–30.
[3] C. Lepage, B. Rachet, V. Jooste, Continuing rapid increase in esophageal adenocarcinoma in England and Wales, Am. J. Gastroenterol. 103 (2008) 2694–2699.
[4] P. Sharma, J.J. Bergman, K. Goda, M. Kato, H. Messmann, B.R. Alsop, N. Gupta, P. Vennalaganti, M. Hall, V. Konda, A. Koons, O. Penner, J.R. Goldblum, I. Waxman, Development and validation of a classification system to identify high-grade dysplasia and esophageal adenocarcinoma in barretts esophagus using narrow-band imaging, Gastroenterology 150 (3) (2016) 591–598. ISSN 0016–5085.
[5] K.N. Phoa, R.E. Pouw, R. Bisschops, O. Pech, K. Ragunath, B.L.A.M. Weusten, B. Schumacher, B. Rembacken, A. Meining, H. Messmann, E.J. Schoon, L. Gossner, J. Mannath, C.A. Seldenrijk, M. Visser, T. Lerut, S. Seewald, F.J. ten Kate, C. Ell, H. Neuhaus, J.J.G.H.M. Bergman, Multimodality endoscopic eradication for neoplastic Barrett oesophagus: results of an European multicentre study (EURO-II), Gut 65 (4) (2016) 555–562.
[6] N.J. Shaheen, P. Sharma, B.F. Overholt, H.C. Wolfsen, R.E. Sampliner, K.K. Wang, J.A. Galanko, M.P. Bronner, J.R. Goldblum, A.E. Bennett, B.A. Jobe, G.M. Eisen, M.B. Fennerty, J.G. Hunter, D.E. Fleischer, V.K. Sharma, R.H. Hawes, B.J. Hoffman, R.I. Rothstein, S.R. Gordon, H. Mashimo, K.J. Chang, V.R. Muthusamy, S.A. Edmundowicz, S.J. Spechler, A.A. Siddiqui, R.D. Souza, A. Infantolino, G.W. Falk, M.B. Kimmey, R.D. Madanick, A. Chak, C.J. Lightdale, Radiofrequency ablation in Barrett's esophagus with dysplasia, N. Engl. J. Med. 360 (22) (2009) 2277–2288.
[7] M. H Johnston, J. A Eastone, J. Horwhat, J. Cartledge, J. S Mathews, J. Foggy, Cryoablation of Barrett's esophagus: a pilot study, Gastrointest. Endosc. 62 (2006) 842–848.
[8] B.F. Overholt, M. Panjehpour, D.L. Halberg, Photodynamic therapy for Barret's esophagus with dysplasia and/or early stage carcinoma: long-term results, Gastrointest. Endosc. 58 (2003) 183–188.
[9] P. Sharma, J. Brill, M. Canto, D. DeMarco, B. Fennerty, N. Gupta, L. Laine, D. Lieberman, C. Lightdale, E. Montgomery, R. Odze, J. Tokar, M. Kochman, White paper AGA: advanced imaging in barretts esophagus, Clin. Gastroenterol. Hepatol. 13 (13) (2015) 2209–2218. ISSN 1542–3565.
[10] J.A. Abrams, R.C. Kapel, G.M. Lindberg, M.H. Saboorian, R.M. Genta, A.I. Neugut, C.J. Lightdale, Adherence to biopsy guidelines for Barrett's esophagus surveillance in the community setting in the United States, Clin. Gastroenterol. Hepatol. 7 (7) (2009) 736–742. ISSN 1542–3565.
[11] F. van der Sommen, S. Zinger, E.J. Schoon, P.H.N. With, Supportive automatic annotation of early esophageal cancer using local gabor and color features, Neurocomputing 144 (2014) 92–106.
[12] N. Ghatwary, A. Ahmed, X. Ye, Automated detection of Barrett's esophagus using endoscopic images: a survey, medical image understanding and analysis, in: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings, 2017, pp. 897–908.
[13] J. Hopkins, Barrett's esophagus: introduction, Gastroenterol. Hepatol. (2008) 1–5.
[14] B. Li, M.Q.-H. Meng, Texture analysis for ulcer detection in capsule endoscopy images, Image Vis Comput. 27 (9) (2009) 1336–1342. ISSN 0262–8856.
[15] E. Rodriguez-Diaz, S.K. Singh, Computer-assisted interpretation of the NICE criteria for colorectal polyps using near-focus narrow-band imaging, Gastroenterology 150 (2016). S–434.
[16] D.J. Nancarrow, A.D. Clouston, B.M. Smithers, D.C. Gotley, P.A. Drew, D.I. Watson, S. Tyagi, N.K. Hayward, D.C. Whiteman, Digital pathology: whole Genome expression array profiling highlights differences in mucosal defense genes in barretts esophagus and esophageal adenocarcinoma, PLoS One 6 (7) (2011) 1–15.
[17] E. Veronese, E. Grisan, G. Diamantis, G. Battaglia, C. Crosta, C. Trovato, Hybrid patch-based and image-wide classification of confocal laser endomicroscopy images in Barrett's esophagus surveillance, in: 2013 IEEE 10th International Symposium on Biomedical Imaging: from Nano to Macro, 2013, pp. 7–11.
[18] T. Muldoon, N. Thekkek, D. Roblyer, D. Maru, N. Harpaz, J. Potack, S. Anandasabapathy, R. Kortum, Evaluation of quantitative image analysis criteria for the high-resolution microendoscopic detection of neoplasia in Barrett's esophagus, J. Biomed. Optic. 15 (2010) 026027.
[19] F. van der Sommen, S. Zinger, E. Schoon, P. With, Computer-aided detection of early cancer in the esophagus using HD endoscopy images, Proceedings of SPIE - the International Society for Optical Engineering vol. 8670.
[20] A.R. Hassan, M.A. Haque, Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos, Comput. Biol. Med. 122 (2015) 341–353.
[21] L.A. Souza Jr., C. Hook, J.P. Papa, C. Palm, Barrett's Esophagus Analysis Using SURF Features, Bildverarbeitung für die Medizin 2017: algorithmen - Systeme - Anwendungen, in: Proceedings des Workshops vom 12. bis 14. März 2017 in Heidelberg, 2017, pp. 141–146.
[22] MICCAI, in: 18th International Conference, Medical Image Computing and Computer Assisted Interventions, 2015.

[23] H. Bay, A. Ess, T. Tuytelaars, L.V. Gool, Speeded-up Robust features (SURF), Comput. Vis. Image Understand. 110 (3) (2008) 346–359. ISSN 1077–3142.
[24] Z. Zhang, L. Bai, P. Ren, E.R. Hancock, High-order graph matching kernel for early carcinoma EUS image classification, Multimed. Tool. Appl. 75 (7) (2016) 3993–4012. ISSN 1573–7721.
[25] S. Klomp, F. Van der Sommen, A.-F. Swager, S. Zinger, E. Schoon, W. Curvers, J.J. Bergman, P. With, Evaluation of image features and classification methods for Barrett's cancer detection using VLE imaging, Proceedings of SPIE Medical Imaging 2017 (2017), 101340D.
[26] S. Seguí, M. Drozdzalc, G. Pascuala, P. Radevaa, C. Malageladad, F. Azpirozd, J. Vitriáa, Generic feature learning for wireless capsule endoscopy analysis, Comput. Biol. Med. 79 (2016) 163–172.
[27] R. Mendel, A. Ebigbo, A. Probst, H. Messmann, C. Palm, Barrett's Esophagus Analysis Using Convolutional Neural Networks, Bildverarbeitung für die Medizin 2017: algorithmen - Systeme - Anwendungen, in: Proceedings des Workshops vom 12. bis 14. März 2017 in Heidelberg, 2017, pp. 80–85.
[28] S.V. Georgakopoulos, D.K. Iakovidis, M. Vasilakakis, V.P. Plagianakos, A. Koulaouzidis, Weakly-supervised convolutional learning for detection of inflammatory gastrointestinal lesions, in: 2016 IEEE International Conference on Imaging Systems and Techniques (IST), 2015, pp. 510–514.
[29] A. Koulaouzidis, D.K. Iakovidis, KID: Koulaouzidis-iakovidis Database for Capsule Endoscopy, URL, http://is-innovation.eu/kid/.
[30] D.K. Chan, L. Zakko, K.H. Visrodia, C.L. Leggett, L.S. Lutzke, M.A. Clemens, J.D. Allen, M.A. Anderson, K.K. Wang, Breath testing for barretts esophagus using exhaled volatile organic compound profiling with an electronic-nose device, Gastroenterology (2016) 1–9.
[31] H. Yoshida, T. Shimazu, T. Kiyuna, A. Marugame, Y. Yamashita, E. Cosatto, H. Taniguchi, S. Sekine, A. Ochiai, Automated Histological Classification of Whole-slide Images of Gastric Biopsy Specimens, 2017, pp. 1–9.
[32] J. Hong, B.-Y. Park, H. Park, Convolutional Neural Network Classifier for Distinguishing Barrett's Esophagus and Neoplasia Endomicroscopy Images, vol. 2017, 2017, pp. 2892–2895.
[33] P. Rajan, M. Canto, E. Gorospe, A. Almario, A. Kage, C. Winter, G. Hager, T. Wittenberg, C. Münzenmayer, Automated Diagnosis of Barrett's Esophagus with Endoscopic Images, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-03882-2, 2189–2192.
[34] M. Sattlecker, N. Stone, C. Bessant, Current trends in machine-learning methods applied to spectroscopic cancer diagnosis, Trends Anal. Chem. 59 (2014) 17–25.
[35] M. Kandemir, A. Feuchtinger, A. Walch, F.A. Hamprecht, Digital pathology: multiple instance learning can detect Barrett's cancer, in: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI 2014), 2014, pp. 1348–1351.
[36] Z. Zhou, Y. Sun, Y. Li, Multi-instance learning by treating instances as non-I.I.D. Samples, in: Proceedings of the 26th Annual International Conference on Machine Learning, ICML '09, ACM, New York, NY, USA, 2009. ISBN 978-1-60558-516-1, 1249–1256.
[37] J. Yu, J. Chen, Z.Q. Xiang, Y. Zou, A hybrid convolutional neural networks with Extreme learning machine for WCE image classification, in: Proceedings of the 2015 IEEE Conference on Robotics and Biomimetics Zhuhai, 2015, pp. 6–9.
[38] A.-F. Swager, G.J. Tearney, C.L. Leggett, M.G. van Oijen, S.L. Meijer, B.L. Weusten, W.L. Curvers, J.J. Bergman, Identification of volumetric laser endomicroscopy features predictive for early neoplasia in Barretts esophagus using high-quality histological correlation, Gastrointest. Endosc. 85 (5) (2017a) 918–926.e7. ISSN 0016–5107.
[39] L.A. Souza Jr., L.C.S. Afonso, C. Palm, J.P. Papa, Barrett's esophagus identification using optimum-path forest, 2017 30th SIBGRAPI conference on graphics, Patterns and Images (SIBGRAPI) (2017b) 308–314.
[40] J.P. Papa, A.X. Falcão, C.T.N. Suzuki, Supervised pattern classification based on optimum-path forest, Int. J. Imag. Syst. Technol. 19 (2) (2009) 120–131. ISSN 0899–9457.
[41] J.P. Papa, A.X. Falcão, V.H.C. Albuquerque, J.M.R.S. Tavares, Efficient supervised optimum-path forest classification for large datasets, Pattern Recogn. 45 (1) (2012) 512–520.
[42] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110. ISSN 0920–5691.
[43] W. Pu, C. Wang, S. Chen, D. Zhao, Y. Zhou, Y. Ma, Y. Wang, C. Li, Z. Huang, L. Jin, S. Guo, J. Wang, M. Wang, Targeted bisulfite sequencing identified a panel of DNA methylation-based biomarkers for esophageal squamous cell carcinoma (ESCC), Clin. Epigenet. 9 (1) (2017) 129. ISSN 1868–7083.
[44] L. Serpa-Andrade, V. Robles-Bykbaev, L. Gonzlez-Delgado, J.L. Moreno, An approach based on Fourier descriptors and decision trees to perform presumptive diagnosis of esophagitis for educational purposes, in: 2015 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), 2015, pp. 1–5.
[45] S. Zopf, A.K. Kage, C. Mnzenmayer, T. Wittenberg, E.G. Hahn, M. Raithel, Narrow-band imaging for the computer assisted diagnosis in patients with Barrett's esophagus, Gastrointest. Endosc. 69 (5) (2009). AB376, ISSN 0016–5107, dDW Abstract Issue 2009.
[46] O. Pech, T. Rabenstein, H. Manner, M.C. Petrone, J. Pohl, M. Vieth, M. Stolte, C. Ell, Confocal laser endomicroscopy for in vivo diagnosis of early squamous cell carcinoma in the esophagus, Clin. Gastroenterol. Hepatol. 6 (1) (2008) 89–94. ISSN 1542–3565.
[47] A. Rosenfeld, V. Sehgal, D.G. Graham, M. Banks, R. Haidry, L. Lovat, Using data mining to help detect dysplasia: extended abstract, 2014 IEEE international conference on software science, Technology and Engineering (2014) 65–66.
[48] C. Li, C. Shi, H. Zhang, Y. Chen, S. Zhang, Multiple instance learning for computer aided detection and diagnosis of gastric cancer with dual-energy CT imaging, J. Biomed. Inf. 57 (2015) 358–368.
[49] W.L. Curvers, J.J. Bergman, A new paradigm shift in endoscopy: from interpretation to automated image analysis? Gastrointest. Endosc. 83 (1) (2016) 115–116.
[50] Z. Wang, H. Lee, O.O. Ahsen, K. Liang, M. Figueiredo, Q. Huang, J.G. Fujimoto, H. Mashimo, Novel optical coherence tomography image analysis reveals subsquamous glandular structures as strong predictors of poorer response to radiofrequency ablation in Barrett's esophagus, Gastroenterology 150 (2016). S–434.
[51] A.-F. Swager, F. van der Sommen, S.R. Klomp, S. Zinger, S.L. Meijer, E.J. Schoon, J.J. Bergman, P.H. de With, W.L. Curvers, Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy, Gastrointest. Endosc. 86 (5) (2017b) 839–846. ISSN 0016–5107.
[52] D. Boschetto, G. Gambaretto, E. Grisan, Automatic classification of endoscopic images for premalignant conditions of the esophagus, in: Proceedings of SPIE Medical Imaging 2016 9788, 2016, pp. 1–6.