Computers in Biology and Medicine 85 (2017) 1–6

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method☆

Teng Zhou[a], Guoqiang Han[a], Bing Nan Li[b,*], Zhizhe Lin[c], Edward J. Ciaccio[e], Peter H. Green[e], Jing Qin[d]

[a] School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China
[b] Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China
[c] Affiliated Shantou Hospital of Sun Yat-sen University, Shantou Central Hospital, Shantou 515000, China
[d] Center for Smart Health, School of Nursing, The Hong Kong Polytechnic University, Hong Kong
[e] Department of Medicine, Celiac Disease Center, Columbia University, New York, USA

ARTICLE INFO

Keywords:
Celiac disease
Videocapsule endoscopy
Deep learning
GoogLeNet
Quantitative analysis

ABSTRACT

Background. Celiac disease is one of the most common diseases in the world. Capsule endoscopy is an alternative way to visualize the entire small intestine without invasiveness to the patient. It is useful to characterize celiac disease, but hours are need to manually analyze the retrospective data of a single patient. Computer-aided quantitative analysis by a deep learning method helps in alleviating the workload during analysis of the retrospective videos.

Method. Capsule endoscopy clips from 6 celiac disease patients and 5 controls were preprocessed for training. The frames with a large field of opaque extraluminal fluid or air bubbles were removed automatically by using a pre-selection algorithm. Then the frames were cropped and the intensity was corrected prior to frame rotation in the proposed new method. The GoogLeNet is trained with these frames. Then, the clips of capsule endoscopy from 5 additional celiac disease patients and 5 additional control patients are used for testing. The trained GoogLeNet was able to distinguish the frames from capsule endoscopy clips of celiac disease patients vs controls. Quantitative measurement with evaluation of the confidence was developed to assess the severity level of pathology in the subjects.

Results. Relying on the evaluation confidence, the GoogLeNet achieved 100% sensitivity and specificity for the testing set. The *t*-test confirmed the evaluation confidence is significant to distinguish celiac disease patients from controls. Furthermore, it is found that the evaluation confidence may also relate to the severity level of small bowel mucosal lesions.

Conclusions. A deep convolutional neural network was established for quantitative measurement of the existence and degree of pathology throughout the small intestine, which may improve computer-aided clinical techniques to assess mucosal atrophy and other etiologies in real-time with videocapsule endoscopy.

1. Background

Celiac disease is one of the most common diseases on the earth, with its incidence reaching to about 1% of the worldwide population [1]. The disease is an autoimmune response to ingested gluten. The immune cascade damages the small intestinal mucosa, i.e. the duodenum and jejunum, as well as the ileum. It manifests as duodenal folds, scalloping of folds, mucosal fissures, crevices or grooves, micronodules in the duodenal bulb, visible submucosal vessels or a mosaic pattern in the small intestinal mucosa [2]. The traditional diagnosis are typically made from the assessment of duodenal biopsies using standard endoscopy, after the serological testing result is confirmed [3]. It is invasive and expensive to use serological testing and endoscopic biopsy to diagnose the celiac disease, particularly in developing countries. More economical methods require consideration for the diagnosis of celiac disease [4].

Video capsule endoscopy is a feasible non-invasive, pain free, and friendly alternative, which can potentially be used to visualized the entire small intestine for detailing the mucosal villous architecture [4]. The visual assessment of villous atrophy is regarded as one of the most

☆ This work was supported partially by the National Natural Science Foundation of China under Grants 61571176 and 61511140099, in part by Anhui Provincial Natural Science Foundation under Grant 1608085J04, and in part by the International Science and Technology Cooperation Plan of Anhui Province under Grant 1503062015.

* Corresponding author.
E-mail address: bingoon@ieee.org (B.N. Li).

http://dx.doi.org/10.1016/j.compbiomed.2017.03.031
Received 11 November 2016; Received in revised form 18 March 2017; Accepted 31 March 2017

critical preliminaries for the diagnosis of celiac disease, which is a great convenience to measure the extent and severity of the mucosal surface at a sufficient resolution in those patients who are suspected of having celiac disease.

For the past decade, several research efforts have suggested the effectiveness and efficiency of video capsule endoscopy in the diagnosis of celiac disease [2,5,6]. Rondonotti et al. [7] evaluated the presence of lesions compatible with celiac disease with a sensitivity of 87.5%, specificity 90.9%, and positive predictive value 96.5%, negative predictive value 71.4%, positive and negative likelihood ratios 9.6 and 0.14, respectively. Rokkas et al. [8] reported a meta-analysis of capsule endoscopy in celiac disease, of which the sensitivity and specificity are up to 89% and 95%, respectively. Ciaccio et al. [4] used capsule endoscopy videos to assess the severity of celiac disease, with sensitivity of 88% and specificity of 80% using an incremental classifier, and a sensitivity of 80% and specificity of 96% when using a threshold classifier, respectively. However, most of these have exploited low-level features, such as brightness or texture, to teach classifier component. These hand-crafted low-level features are incapable of taking full advantage of contextual information to obtain more representative features for more accurate recognition.

Recently, the method of deep learning has drawn a lot of attention in academia and industry [9,10]. This method can be trained to automatically learn rich feature representation of images in a hierarchical scheme [11,12]. Motivated by Szegedy Christian et al. [13], who proposed a deep convolutional neural network architecture named GoogLeNet which was considered state-of-the-art for classification in the ImageNet Large-Scale Visual Recognition Challenge, and the rotation-invariant Fisher discriminative convolutional network (RIFD-CNN) developed by Cheng et al. [14], we introduce a pre-rotation scheme for GoogLeNet to assess celiac disease frames. Quantitative analysis of the videos of the small intestine from celiac disease patients versus controls are performed. Then a measurement method termed evaluation confidence (*EC*) is defined to quantify how confident the classifier indicates a subject to have celiac disease. In the experiments, we find that the classifier not only distinguishes celiac patients with subtle changes in small bowel, but also relates the *EC* calculation to the severity level of the tested patients.

2. Methods

2.1. Capsule endoscopy data description

The analysis of capsule endoscopy videos were approved by the Internal Review Board at Columbia University Medical Center. Informed consent was obtained from all subjects previously. The identifiers of the subjects in the video clips were removed prior to analysis. Only videos reaching the colon were used for further studies, and subjects under 18 years old, pregnant women, and those with a history of intestinal obstruction, presence of a pacemaker, or chronic use of non-steroidal anti-inflammatory drugs (NSAIDs) were excluded.

The capsule endoscopy videos of the small bowel were obtained by the PillCamSB2 video capsule suite(Given Imaging, Yoqneam, Israel), which consists of a recorder, recorder cradle, recorder harness, real-time viewer, real-time viewer cable, antenna lead set, battery pack, and battery charger. The capsule recorder is 26×11 millimeters in size, with the frame rate set to 2 frames per second. All subjects swallowed the capsule early in the morning after a fast 12 h ago with 200 cc of water. After ingestion of the capsule, further water was forbidden for 2 h, and a light meal was forbidden for 4 h. The evaluation period completed after the capsule reached the cecum. The videos were exported using an HIPAA-compliant PC-based workstation equipped with Given Imaging analysis software.

Finally, the capsule endoscopy videos from 11 celiac disease patients and 10 controls, which were obtained at Columbia University Medical Center from May 1, 2008 to July 31, 2009 were retrospectively selected for analysis. All the retrospective videocapsule endoscopy data and duodenal biopsies data were interpreted by three experienced gastroenterologists. The 11 celiac patients included those with suspected celiac disease, suspected Crohns disease, obscure bleeding, iron deficient anemia, and chronic diarrhea who had previously had a diagnostic biopsy on a regular diet, to evaluate for the presence of Marsh IIIA, IIIB, or IIIC lesions, except for one patient with hemophilia. The capsule endoscopy videos were obtained when these patients were either on a regular diet or on a gluten-free diet for no more than 3 months. The 11 celiac disease patients consist of 6 females and 5 males, whose mean ages were 50.5 and 44.0, respectively, while the 10 control patients consisted of 6 females and 4 males, whose mean ages were 50.0 and 51.5, respectively.

2.2. Quantitative analysis using GoogLeNet

The accuracy of image classification has recently been substantially improved thanks to deep learning techniques, such as deep convolutional neural networks (DCNN). In this study, a DCNN based method is investigated for quantitative analysis to distinguish celiac patient vs control data. Without the need for manual rating, the computer-aided analysis is not only user-friendly, rapid, and low cost for learning and operation, but is also immune to subjective results by user bias.

The 22-layer GoogLeNet [13], which won the ImageNet Large-Scale Visual Recognition Challenge, was chosen for this task. The network is 22 layers deep counting layers with parameters without counting pooling layers [13]. The GoogLeNet is a particular incarnation of Inception architectures, which repeat many times leading to the 22-layer model. Readers who are interested in Inception architecture can refer to the latest detailed publication of rethinking Inception architecture [15]. In the Inception architecture, 1×1 convolutions reduce the expensive computation of 3×3 and 5×5 convolutions. A convolution layer of 7×7 with 3×3 pooling stride 2 start the network. Then another convolution layer of 3×3 with 3×3 pooling stride 2 follows. After that there are 9 Inception architectures with occasional max-pooling layers with stride 2 halving the resolution of the grid. The network finally ends with average pooling, dropout layer, linear layer, and softmax. The input of this application is 512×512 rgb image. Another interesting insight of this network is adding auxiliary classifiers connected to the intermediate layers, which boosts the information prorogation within the network and hence improve the training efficacy and enhance the discrimination capability of the network. As a network going deeper, there is a concern for the efficiency to propagate gradients back thought all the layers [13]. It requires the features learned by the intermediate layers to be very discriminative. Therefore, the auxiliary softmax layers are added to the intermediates, which are encouraged discrimination in the lower stages, and propagate back increased gradient [13]. Readers may learn more theoretical and technical details about GoogLeNet in [13], which is visualized in the bottom left panel of Fig. 1.

Six celiac disease and five control subjects were randomly selected from the eleven celiac disease patients and ten control patients to train the GoogLeNet. Four clips from four regions of the small intestine of each subject were extracted from the videoclips for this study. The four regions are: duodenal bulb, duodenum, jejunum, and ileum. Each clip contains 200 frames, at a rate of 2 frames per second. The resolution of each frame of the video clips is 576×576 pixels. The frames containing a large field of opaque extraluminal fluid or air bubbles were removed by using the pre-selection criterion developed by Mamonov et al. [16].

The video frames consist of a square area but were analyzed with a circular shape mask. The area outside the circular mask was filled with a solid color and extra information, as illustrated in the top left panel of Fig. 1. However, this may create an inconsistency in the actual content of the frames. To eliminate the inconsistency, we cropped the frame with an exterior square of the circular mask, as illustrated in the top middle panel of Fig. 1. The exterior square has dimensions of 512×512 pixels.

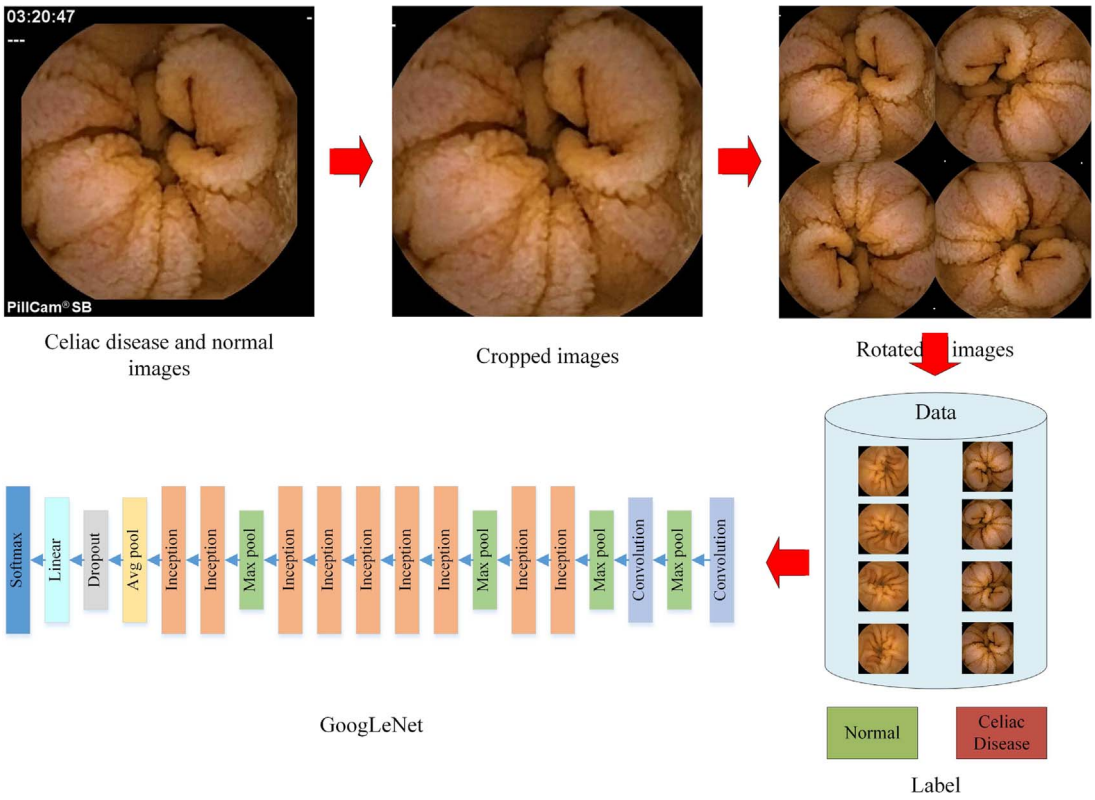

Fig. 1. Overview. The top left panel is the illustration of input images. The top middle panel is the cropped image. The top right panel illustrates the rotated images. The bottom left panel illustrates the GoogLeNet [13]. The bottom right panel is the data and labels prepared for training the GoogLeNet.

The frames captured by the capsule endoscopy were subjected to vignetting which is the fall-off of intensity away from the center of the frame, due to the absence of the ambient light, directional nature of the on-board light, and the optical properties of camera lens [16]. An intensity correction algorithm by Zhang et al. [17] was performed prior to further processing.

The orientation and angle of the camera were undetermined as the capsule passed through the gastrointestinal tract. Since we are more interested in the mucosal texture than the actual orientation of the camera, we have to eliminate the influence of the orientation and angle of the camera. Motivated by Cheng et al. [14], we introduced a pre-rotating scheme. Each frame in the training set was rotated every 15 degrees to form a new candidate proposal for the training set, as illustrated in the top right panel of Fig. 1. Thus, each frame results in 24 proposals. For clarity, only 4 of 24 proposals are plotted in Fig. 1.

Then the GoogLeNet is trained using these frames and the corresponding labels, i.e. celiac disease or normal. The GoogLeNet is first trained with 400 carefully chosen discriminative frames, including 200 from celiac disease patients and 200 from controls. The frames are chosen according to the following rules: 1) the frames do not contain opaque extra-luminal fluid or air bubbles; 2) the frames from celiac disease patients contain evident visual difference include decreased number of mucosal folds, mosaic appearance, and scalloping; 3) the frames cover the four regions of the small intestine. This procedure helps the network to seek a relatively good suboptimal gradient rapidly for parameters initialization. Then all frames are used to fine-tune the snapshot of the trained network. During the training, the output of the softmaxloss layers are weighted by 0.3 and add to the total loss of the whole network in accordance with the suggestion in [13]. Technical implementation of the training procedure of DCNN, i.e. GoogLeNet, has been detailed in [18]. After the GoogLeNet is trained, five celiac disease and five control subjects were used for testing the quantitative analysis.

2.3. Automata diagnosis by expectation of the probability

Four capsule endoscopy clips of the 10 testing subject were preprocessed, the same as during the training sequence, except for the pre-rotating scheme. The pre-rotating scheme was used to force the network to learn rotation invariance features in the training stage. Once learned, this scheme is no long needed in testing stage. Since four video clips were captured sequentially, the frame extracted from these videos should also maintain the sequence order. The ith frame of the video clips of the jth subject is denoted as $\mathcal{I}_i^{(j)}$. In this study, we do not need to distinguish the clips from each other, and this subsequently reduces manual intervention and speeds up the process. At this stage, the auxiliary softmax layers are simply discarded. Each frame $\mathcal{I}_i^{(j)}$ is forwarded into the GoogLeNet, which outputs the probability $\mathbf{P}(\mathcal{I}_i^{(j)})$ that $\mathcal{I}_i^{(j)}$ is classified as suspected celiac disease.

Then we define a measurement term of how confidently we indicate a subject j as having celiac disease. The EC is the expectation of the probability of all frames of subject j, which is denoted as:

$$EC^{(j)} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{P}(\mathcal{I}_i^{(j)}), \tag{1}$$

where N is the number of frames of subject j.

If $EC^{(j)}$ exceeds 50%, then the subject j may be indicated as having celiac disease. If $EC^{(j)}$ is less than 50%, then the subject j is likely to be normal, without celiac disease.

3. Results

Since the frames containing a large field of opaque extraluminal fluid or air bubbles have been removed using the pre-selection criterion, the number of final frames for each subject in the testing set were not equal. To make the bar chart in Fig. 2 more consistent and

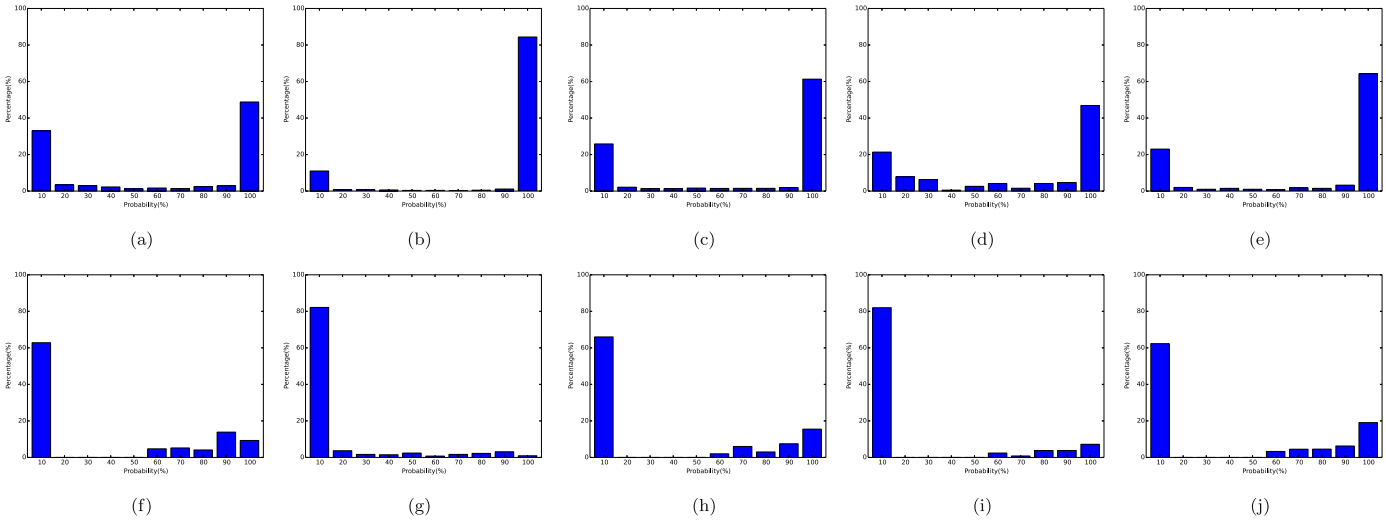

Fig. 2. Bar charts of probability of celiac disease by percentage of images. The bar charts of five celiac disease patients are in a-e, the ones of five control subjects are in f-j.

more comprehensive, the number of frames is shown in a percentage form. For each subject, the probability of the percentage of frames is plotted in the bar chart. The horizontal axis of the bar chart is a certain interval of probability of celiac disease. In the bar chart, a 10 in horizontal axis means the probability interval is $[0, 10)\%$, a 20 means $[10, 20)\%$, and so on. For the last one, 100 indicates that the probability is 100%. The vertical axis designates the percentage of frames that fall into a certain range of probability. The bar charts of the five celiac disease patients and five control subjects are listed in Fig. 2. Fig. 2a-e are five celiac disease patients with celiac disease of Marsh type IIIA - IIIC. Fig. 2f-j are five control subjects.

As is evident in Fig. 2, for patients Fig. 2a-e, more than 40% of the frames have the probability of 100%. In Fig. 2b, the probability of 100% even exceeds 80%. In Fig. 2b and e, this number is over 60%. This may indicate our method has a good sensitive to detect small intestinal pathology, likely from villous atrophy, in celiac disease patients. Moreover, there are only a small number of frames with the probability of interval $[10, 100)\%$ for celiac disease patients, except the one in Fig. 2d.

In contrast, the probability of interval $[0, 10)\%$ accounts for the majority for the control subjects in Fig. 2f-j. Although the minority frames are classified as celiac disease, these bar charts are evidently different from the ones with celiac disease, which can be subsequently verified by the *EC* calculation.

The evaluation confidence for each subject in the testing set is listed in Table 1. As discussed, an *EC* value over 50% is considered suspect for having pathology. In Table 1, the *EC* of the patients from patient 1 to patient 5 correspond to the bar charts in Fig. 2a-e. The *EC* of the control subjects from subject 1 to subject 5 correspond to bar charts in Fig. 2f-j. The *EC* of the celiac disease patients ranges from 57.53% to 86.55%, while the *EC* of the control subjects ranges from 9.58% to 31.79%. For celiac disease patients and control subjects, the *EC* values are significantly different, so that it is easy to distinguish control subjects from celiac disease patients. In order to confirm our hypothesis, a *t*-test performed on the *EC* calculations is listed in Table 2. Notice that the two sample mean values (variance) are 68.91(124.42) and 22.83(99.94). The two-tailed calculated *t*-statistic is 2.78 and the *p*-value for this test is p=0.005. Since the *p*-value is less than 0.05, this means that the *EC* values of the celiac disease patients versus control subjects are significant.

Table 1
Evaluation Confidence for Each Subject.

Patients	patient 1	patient 2	patient 3	patient 4	patient 5
EC	57.53%	86.55%	67.68%	61.77%	71.03%
Controls	subject 1	subject 2	subject 3	subject 4	subject 5
EC	29.52%	9.58%	28.58%	14.70%	31.79%

Table 2
t-test of *EC* (%).

Mean	68.91	22.83
Variance	124.42	99.94
Hypothesized Mean Difference	0	
t Critical two-tail	2.78	
$P(T \le t)$ two-tail	0.005	

In order to further estimate the generalization ability of the deep network, a 7-folds cross-validation is performed on the 11 celiac disease patients and 10 control subjects. The 11 celiac disease patients and 10 control subjects are split into 7 mutually exclusive sets $S_1, S_2, \ldots, S_7$, three for each set. Each set contains at least one celiac disease patient and one control subject. For each fold, 6 sets of samples are used for training, the rest one is for testing. We conducted the experiments on K40 platform with cuDNN enabled. The average forward-backward takes 1327.56 ms and the average forward takes 441.98 ms. The *EC* values are shown in Table 3, where *P* means celiac disease patient and *C* means control subject. From Table 3, the average sensitivity and specificity by cross validation are also 100%.

4. Discussion

In this study, we introduce a deep learning network that could be used for evaluation of celiac disease-associated villous atrophy. The celiac disease is characterized by duodenal folds, mucosal fissures, scalloping of folds, grooves or crevices, visible submucosal vessels, micronodules in the duodenal bulb and a mosaic pattern in the small bowel mucosa. [2] The lesions are visibly different on capsule endo-

Table 3
The *EC* values of 7-fold cross-validation.

Training Set	Testing Set	*EC* (%)
$S_1, S_2, S_3, S_4, S_5, S_6$	S_7 (P_{10}, P_{11}, C_{10})	79.32, 68.46, 27.05
$S_1, S_2, S_3, S_4, S_5, S_7$	S_6 (P_9, C_8, C_9)	81.77, 41.26, 31.12
$S_1, S_2, S_3, S_4, S_6, S_7$	S_5 (P_7, P_8, C_7)	67.43, 50.06, 47.15
$S_1, S_2, S_3, S_5, S_6, S_7$	S_4 (P_6, C_5, C_6)	51.69, 29.54, 17.75
$S_1, S_2, S_4, S_5, S_6, S_7$	S_3 (P_4, P_5, C_4)	52.73, 80.44, 18.20
$S_1, S_3, S_4, S_5, S_6, S_7$	S_2 (P_3, C_2, C_3)	70.68, 6.69, 39.21
$S_2, S_3, S_4, S_5, S_6, S_7$	S_1 (P_1, P_2, C_1)	55.98, 77.86, 25.04

scopy, which therefore makes the deep learning networks suitable for detecting the visual differences of celiac disease patient data versus normal subject data via computerized means. The intermediate parameters of the network is inherited from that of the pre-trained model, whose intermediate parameters are initialize to sensitive values. We continue the network training by the capsule endoscopy frames. This is also coherent with the success of several deep learning algorithms that provide some such guidance for intermediate representations, then the networks are supervised trained [19]. Due to the great power of deep learning network, which is able to learn to extract high-level features automatically, even a subtle visual change in the small bowel mucosa of celiac disease patients can be detected. Thus, our current study achieves a sensitivity of 100% and a specificity of 100% at patient level in the testing set.

In Fig. 2, the majority of frames of all the subjects fall on the probability interval $[0, 10)\%$ and 100%, which means our deep learning network is very confident that a frame has (or has not) a visual difference detected as lesions. A minority of frames fall on the interval $[10, 100)$. This is because even the same celiac disease patient, different regions of small bowel mucosal can manifest a different degree of lesion severity.

Another phenomenon is that the *EC* values of the patients ranged from 57.53% to 86.55%. For example, the *EC* of patient 1 is only 57.53%. We manually checked throughout the assessment of duodenal biopsies. In most biopsies, the severity level of villous atrophy of this patient is subtle and classified as Marsh IIIA (patients with dermatitis herpetiformis), according to the Marsh criterion. Then we also checked the assessment of duodenal biopsies of patient 2, whose *EC* value was 86.55%. We found that, in most biopsies, the severity level of villous atrophy of this patient was type Marsh IIIC (complete villous atrophy). Therefore, the *EC* calculation may also be able to quantify the severity level of villous atrophy. This requires further investigation on additional capsule endoscopy data in the future. To conduct this experiment, the training data can be classified into 6 types according to the biopsies and serological testing results. The 6 types are Marsh 0, I, II, IIIA, IIIB, IIIC. The automatic quantitation of Marsh's types is planning for future works.

A variety of techniques have been well developed for computer vision and image processing [20]. Our result is better then the relative studies. For example, Ciaccio et al. [4] proposed threshold and incremental methods, in which two classifiers, the threshold one and incremental one, were developed using 6 celiac and 5 control patients data as exemplars, and tested on 5 celiacs and 5 controls. In these patients the diagnostic biopsy, taken while on a regular diet, showed Marsh grade II-IIIC lesions, except in one patient with hemophilia who did not undergo endoscopy and biopsy. The sensitivity and specificity of five regions, i.e. duodenal bulb, distal duodenum, jejunum, proximal ileum and distal ileum, by threshold methods were 80% and 96%, respectively, while the sensitivity and specificity of five regions by incremental methods were 88% and 80%, respectively. The automata polling method was developed by Ciaccio et al. [21]. Nine celiac patient data with biopsy proven villous atrophy and seven control patient data lacking villous atrophy were used for analysis. Celiac disease patients had biopsy-proven with scores of Marsh II - IIIC except one hemophiliac patient. Four small intestinal levels (duodenal bulb, distal duodenum, jejunum, and ileum) were analyzed. The overall sensitivity of these levels was 83.9%, while the specificity was 92.9%. Different from the previous studies, the present study need not to manually differentiate the small intestine levels, so the statistic is subject-wise rather than level-wise.

Some meta-analyses were performed in the literature. Three independent studies were investigated in [22]. In the study by Petroniene et al. [23], 10 celiac disease patients and 10 controls were involved. The sensitivity and specificity by two very experienced gastroenterologists' reading of the videos were 100%. Other two less experienced reader achieved less sensitivity. The overall sensitivity and specificity were 70% and 100%. The study by Hopper et al. [24] had sensitivity of 85% and specificity of 100% by one observer. The other study by Rondonotti et al. had four observers, and the sensitivity and specificity were 87.5% and 90.9% by inter-observers agreement in the assessment of mucosal lesions. Another meta-analysis was performed by Rokkas and Niv [25]. A total of 166 individuals included in six studies were pooled. The overall sensitivity was 89% and specificity was 95%. Comparing to manual diagnosis by gastroenterologists or freshers, the deep network achieved competitive results. Thus, computer-aided methods by deep networks may greatly reduce the workload of gastroenterologists to analyze the capsule endoscopy videos.

The cross validation experiment demonstrates the generalization ability of the networks. We think there are several reasons refer to this issue. First, the idea behind Inception architecture is to optimize the local sparse structure in the convolutional network and cover the available dense components [13]. Second, the dropout [26] is a typical way to prevent over-fitting, which remain essential in GoogLeNet though absence of fully connected layers. Third, the batches of examples help estimating the gradient loss over the training set and computation over batches is more effective than individual one [27]. Forth, the auxiliary classifiers help to increase the backward gradient by connecting to the intermediate layer. These classifiers amplify the total loss by adding their weighted loss to the final one.

However, the capsule endoscopy is a relatively new procedure, and particularly, the number of studies of assessment of celiac disease using capsule endoscopy is limited. The number of video clips collected for this studies is relatively small, so a larger data set would be required for validation in a prospective double-blinded study. Yet, in order to confirm celiac disease, positive serologic markers and a biopsy would still be need. The gold standard for the diagnosis of celiac disease remains the assessment of duodenal biopsies using standard endoscopy, after the serological testing result and is used for confirmation. The ability for obtaining biopsies via capsule endoscopy is under development and may be released in the future [2].

5. Conclusions

There are visual differences in the image frames of celiac disease patients versus controls, which can be automatically learned via deep learning techniques. The deep learning network can be used to distinguish the frames of celiac disease patients from those of control subjects. A quantitative measurement for severity level, which we termed evaluation confidence (*EC*), was introduced in this study. The results according to this evaluation confidence were found to be promising, achieving a 100% level of sensitivity and specificity in the testing set. This computer-aided technique may therefore be helpful to eliminate the bias of the clinician during evaluation for the presence of mucosal villous atrophy and also broaden the applicability of capsule endoscopy to other etiologies for real-time assessment. Automatic detection of Marsh types is being developed in further studies by collecting and analyzing larger numbers of video clips from celiac patients and control subjects.

Conflicts of interest statement

The authors have no conflicts of interest.

References

[1] A. Fasano, C. Catassi, Celiac disease, N. Engl. J. Med. 367 (25) (2012) 2419–2426.
[2] C. van de Bruaene, D. de Looze, P. Hindryckx, Small bowel capsule where are we after almost 15 years of use?, World J. Gastrointest. 7 (1) (2015) 13–36.
[3] W.P. Pais, D.R. Duerksen, N.M. Pettigrew, C.N. Bernstein, How many duodenal biopsy specimens are required to make a diagnosis of celiac disease?, Gastrointest. Endosc. 67 (7) (2008) 1082–1087.
[4] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Classification of videocapsule endoscopy image patterns: comparative analysis between patients

with celiac disease and normal individuals, Biomed. Eng. Online 9 (1) (2010) 1.
[5] E.J. Ciaccio, G. Bhagat, S.K. Lewis, P.H. Green, Suggestions for automatic quantitation of endoscopic image analysis to improve detection of small intestinal pathology in celiac disease patients, Comput. Biol. Med. 65 (2015) 364–368.
[6] E.J. Ciaccio, G. Bhagat, S.K. Lewis, P.H. Green, Extraction and processing of videocapsule data to detect and measure the presence of villous atrophy in celiac disease patients, Comput. Biol. Med. 78 (2016) 97–106.
[7] E. Rondonotti, C. Spada, D. Cave, M. Pennazio, M.E. Riccioni, I. de Vitis, D. Schneider, T. Sprujevnik, F. Villa, J. Langelier, et al., Video capsule enteroscopy in the diagnosis of celiac disease: a multicenter study, Am. J. Gastroenterol. 102 (8) (2007) 1624–1631.
[8] T. Rokkas, Y. Niv, The role of video capsule endoscopy in the diagnosis of celiac disease: a meta-analysis, Eur. J. Gastroenterol. Hepatol. 24 (24) (2012) 303–308.
[9] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V.C. Mok, L. Shi, P.-A. Heng, Automatic detection of cerebral microbleeds from mr images via 3d convolutional neural networks, IEEE Trans. Med. Imaging 35 (5) (2016) 1182–1195.
[10] Y. Wang, J.-Z. Cheng, D. Ni, M. Lin, J. Qin, X. Luo, M. Xu, X. Xie, P.A. Heng, Towards personalized statistical deformable model and hybrid point matching for robust mr-trus registration, IEEE Trans. Med. Imaging 35 (2) (2016) 589–604.
[11] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[12] T.-Y. Lin, S. Maji, Visualizing and understanding deep texture representations, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
[14] G. Cheng, P. Zhou, J. Han, Rifd-cnn: Rotation-invariant and fisher discriminative convolutional neural networks for object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2884–2893.
[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.
[16] A.V. Mamonov, I.N. Figueiredo, P.N. Figueiredo, Y.H. Tsai, Automated polyp detection in colon capsule endoscopy, IEEE Trans. Med. Imaging 33 (7) (2014) 1488–1502.
[17] Y. Zheng, J. Yu, S.B. Kang, S. Lin, Single-image vignetting correction using radial gradient symmetry 31 (12) (2008) 1–8.
[18] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: Convolutional architecture for fast feature embedding, arXiv:1408.5093.
[19] Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1798–1828.
[20] B.N. Li, Q. Yu, R. Wang, K. Xiang, M. Wang, X. Li, Block principal component analysis with nongreedy l_1 -norm maximization, IEEE Trans. Cybern. 46 (11) (2016) 2543–2547.
[21] E.J. Ciaccio, C.A. Tennyson, G. Bhagat, S.K. Lewis, P.H. Green, Implementation of a polling protocol for predicting celiac disease in videocapsule analysis, World J. Gastrointest. Endosc. 5 (2013) 313–322.
[22] W. El-Matary, H. Huynh, B. Vandermeer, Diagnostic characteristics of given video capsule endoscopy in diagnosis of celiac disease: a meta-analysis, J. Laparoendosc. Adv. Surg. Tech. 19 (6) (2009) 815–820.
[23] R. Petroniene, E. Dubcenco, J.P. Baker, C.A. Ottaway, S.-J. Tang, S.A. Zanati, C.J. Streutker, G.W. Gardiner, R.E. Warren, K.N. Jeejeebhoy, Given® capsule endoscopy in celiac disease: evaluation of diagnostic accuracy and interobserver agreement, Am. J. Gastroenterol. 100 (3) (2005) 685–694.
[24] A. Hopper, R. Sidhu, D. Hurlstone, M. McAlindon, D. Sanders, Capsule endoscopy: an alternative to duodenal biopsy for the recognition of villous atrophy in coeliac disease?, Dig. Liver Dis. 39 (2) (2007) 140–145.
[25] T. Rokkas, Y. Niv, The role of video capsule endoscopy in the diagnosis of celiac disease: a meta-analysis, Eur. J. Gastroenterol. Hepatol. 24 (3) (2012) 303–308.
[26] N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.
[27] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, arXiv:1502.03167.