

Bleeding Classification of Enhanced Wireless Capsule Endoscopy Images using Deep Convolutional Neural Network

ROSDIANA SHAHRIL¹, ATSUSHI SAITO², AKINOBU SHIMIZU²
AND SABARIAH BAHARUN³

¹*Faculty of Computer Systems and Software Engineering
Universiti Malaysia Pahang, Pahang, Malaysia*

²*Tokyo University of Agriculture and Technology, Tokyo, Japan*

³*Malaysia-Japan International Institute of Technology
Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia*

E-mail: {rosdiana}@ump.edu.my;{a-saito}@go.tuat.ac.jp;{simiz}@cc.tuat.ac.jp;{sabariah}@utm.my

This paper investigates the performance of a Deep Convolutional Neural Network (DCNN) algorithm to identify bleeding areas of wireless capsule endoscopy (WCE) images without known prior knowledge of bleeding and normal features of the images. In this study, a pre-processing technique has been proposed to improve the classification accuracy of WCE images into bleeding areas and normal areas by enhancing the WCE images. The proposed technique is applied to WCE images from six cases and divided into one training case and five test cases. To evaluate the effectiveness of the processes, the results were then compared between DCNN, SVM and Fuzzy, and also between DCNN with completely enhanced images and DCNN with normalized images. DCNN has shown to give a better result compared to SVM and Fuzzy logic; and the latter experiment has shown that the WCE images that have undergone the proposed enhancement technique gives better classification result compared to those images that did not go through the technique. The specificity, sensitivity and average are 0.8703, 0.8271 and 0.8907 respectively. In conclusion, DCNN has been proven to be able to successfully detecting bleeding areas from images without having any specific knowledge on imaging diagnosis or pathology.

Keywords: convolutional neural network, wireless capsule endoscopy, deep learning, classification, detection

1. INTRODUCTION

Recently, many research works on the Wireless Capsule Endoscopy (WCE) device have been published. This includes computer aided decision in capsule endoscopy, electronic system etc. WCE is a procedure to examine internal organ using a small camera that can capture images while moving along the GI tract to detect abnormalities and bleeding in colon, esophagus, small intestinal and stomach in which areas flexible endoscope will not get access to. WCE consists of light source, lens, camera, radio transmitter and batteries. First approval of WCE capsule by the FDA (U.S. Food and Drug Administration) was in 2001, and since then, it has become an important tool to detect other abnormalities in the GI tract [1-7]. The patient swallows a capsule like a pill and images are then captured by the camera and sent out wirelessly to a special recorder attached to the patients waist. This process will continue depending on the battery's life. Finally, all recorded images will be uploaded into a personal computer or computer workstation for physicians's analysis.

Received September 5, 2018; revised November 29, 2018; accepted December 3, 2018.
Communicated by Jen-Tzung Chien.

However, the analysis of images by physicians can be very difficult due to the vast number of WCE images produced every time the procedure is conducted. There are about 50,000 images per examination in total for one patient. Many researches have discussed these challenging issues on working with WCE image analysis [8-10]. First and most of the problems identified are due to the long period of time spent for the inspection of huge number of images produced by WCE device. Second problem identified is that different physicians would provide different diagnosis or interpretation of the same image. Third problem faced when dealing with WCE images is that these images are rather dark and vague as such physicians may face difficulties in analyzing and giving diagnosis on. WCE has some drawbacks in terms of its low transmission power and bandwidth constraints, which leads to poor clarity of the images. Due to this constraint of having vague images, detecting features of the abnormalities would also be a challenge in the use of Computer Aided Detection (CAD) system. Recently, the method of deep learning has drawn a lot of attention in computer vision tasks [8, 11-15]. It has been shown to yield very good results in classification tasks. The strengths of DCNN are robustness to small inputs changes, deformation, minimal pre-processing and not necessarily any specific feature extractor choice where it has a great ability to implement in pattern classification tasks include character recognition, object classification. DCNN is also better in classification task that involving large images than other multi-layer network. Architecture of DCNN is more complex than multi-layer network but each layer in DCNN contain hidden neuron which is the output are the weighted sums of inputs where it passes through activation function. The convolution and pooling layers are not fully-connected which makes the relationship between input and outputs at each layer more complex, but it reduces the number of inputs that contribute for every output. Because of these reasons, the method has attracted us to be used in classification task.

A hybrid CNN with Extreme Learning Machine (ELM) was proposed by [11]. The CNN was constructed as a data-driven feature extractor and cascaded ELM acts as a strong classifier. In [12], the authors fine-tuned the layers in CNN on the ImageNet database to diagnose the celiac disease. Binary classification was performed using softmax classifier and Support Vector Machines (SVM). DCNN was used to classify the digestive organs using WCE images [13]. [16] proposed a new automatic bleeding detection strategy based on a DCNN to learn high-level features where rectified linear units (ReLUs) were used as the activation function. They used a softmax function to minimize the cross-entropy loss for prediction. However, the performance in detecting abnormalities of the other classes was poorer than the state-of-the-art technique. Then, the authors improved the complexity of the DCNN [14] by combining handcrafted(HC) features and DCNN with fewer training samples. [15] focused on small-size imbalanced endoscopy images for bleeding detection thus CNN could learn with very limited training data. Data augmentation and image resampling were employed to increase the size of training database. [17] proposed two combined CNNs to avoid the edge feature caching and speed up the hookworm classification in WCE images.

This paper focuses on the development of automated bleeding area detection in a GI tract using WCE images. The impact on the use of DCNN on bleeding classification is investigated through analysis. To deal with the large variety in colours and contrast of WCE images, such as bright red blood in vomit, black stool or dark mixed with stool, this paper proposes an algorithm to detect bleeding areas with such various appearances in WCE images without the need of specific knowledge on bleeding and normal features. To reduce the variation in appearance, normalization process is used to reduce the variation of colours among input images in order to standardize these images. These images are then enhanced using contrast diffusion scheme. Subsequently, classification task to specify

pixels of bleeding areas is carried out using a DCNN. The paper is organized as follows. The related work in the bleeding detection of WCE images is presented in the next section, followed by proposed method, experimental results and discussion. Finally, concluding remarks are presented in the last section of the paper.

2. RELATED WORK

To deal with bleeding detection, colour information is very crucial since colour is the primary property of medical images characterising the bleeding area. Fig. 1 shows the bleeding area in WCE images. The bleeding area is more reddish compared to its neighboring regions. The normal areas generally have pinkish to yellowish colour tissue. Since normal tissue colour is pinkish in colour and is similar to the bleeding colour, it might be a challenge to CAD system to distinguish them.



Fig. 1. Representative bleeding in WCE image.

[5] investigated the colour spaces (RGB, HSB, and YUV) that has the ability to disclose lesion structure while geometry, colour and texture were combined in their analysis. The features were extracted from inside the masked region or also known as Region of Interest (ROI) where the abnormality category then classified to be either ulcer or bleeding by using Support Vector Machines (SVM) and Vector Supported Convex Hull Method (VSCH). [18] proposed colour histogram based on index image to extract the colour texture of bleeding. SVM was used to detect bleeding and normal regions from WCE videos. However, this method has weakness since it is relying on intensity range of MSB (Most Significant Bit) and LSB (Least Significant Bit) in RGB colour spaces. [19] proposed to extract bleeding area using R to G pixel intensity ratio and different statistical parameters. K-nearest neighbour (KNN) was used to classify between bleeding or normal region.

A super-pixel segmentation was proposed by [10] by grouping the pixel to reduce computational complexity. Although these three algorithms are relatively insensitive to colour variation, the computational costs are rather high, which makes it unsuitable for processing a huge number of WCE images. [9] used K-means clustering to make the most of the colour information of the bleeding. Then, colour histogram was utilized to characterize the feature vector from region of interest (ROI). Then, classification based on SVM and K-nearest neighbour (KNN) was employed to detect bleeding. The authors also investigated on which colour spaces (RGB, HSV, YCbCr and LAB) are the most appropriate in describing bleeding characteristic. However, their proposed method focused on easy WCE images, in which bleeding colour is relatively uniform in training and test images.

[20] used Complete Local Binary Pattern (CLBP) to extract the texture information from WCE image and Global Local Oriented Edge Magnitude Pattern (Global LOEMP) descriptor to extract the colour features in parallel. These two feature vectors are combined to get more discriminating one. The result was accuracy 94.07, sensitivity 96.86 and specificity 91.14 when the proposed method was used with SVM as classifier. Compar-

ing the proposed method with MLP classifier, the result was accuracy 93.93, sensitivity 95.50 and specificity 92.29. From the results, SVM is better than MLP classifier. [21] proposed colour histogram of block statistics (CHOBS) to extract local statistical features from each colour plane to against distortions where K-NN was used as a bleeding classification. [6] proposed five colour features in HSV colour space to distinguish between bleeding and non-bleeding frames where SVM was used as a classifier. Prominent efficient Machine Learning (ML) was used to automate video frames to classify bleeding [7]. A deep learning approach via fully convolutional networks (FCNs) was used for bleeding segmentation in WCE images in [22]. They classified the samples into active and inactive subgroups based on the statistical features. [23] proposed block based segmentation where the features were extracted using local binary pattern histogram and average of HSI was taken into account to classify by SVM. [24] proposed to detect inflammatory bowel diseases (IBD) based on textural features, where texture features are from single key point and from neighborhood of keypoints.

From the investigation of these works, it is gathered that WCE images with almost similar bleeding appearance were used in the experiments of most of previous works. Here, the study will focus on detection of the bleeding area of WCE images with various appearances. To overcome the differences in the appearance of the images, a novel approach is proposed where the images have to be enhanced using colour normalization and contrast enhancement, and bleeding detection through classification technique using DCNN. Colour normalization process reduces the variation in colours whereby contrast enhancement of bleeding features in the images is to differentiate bleeding areas from the normal ones.

3. PROPOSED METHOD

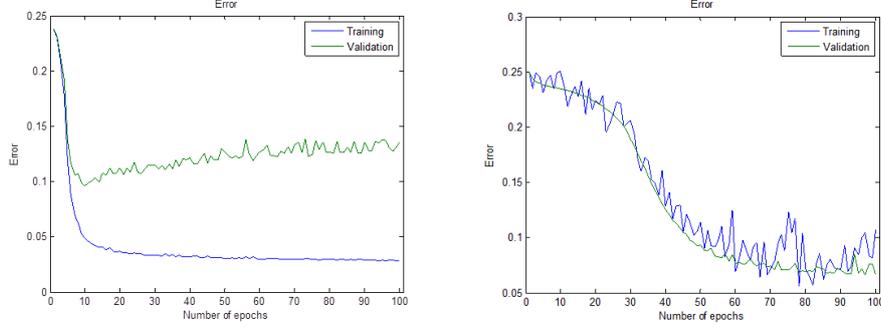
In this section, enhancement method for WCE images is used as a pre-processing technique that employs colour normalization and contrast enhancement method to enhance the quality of WCE images. The aim is to ease the classification task for CAD system. Next, DCNN is proposed as a classification method to classify the bleeding and normal areas in WCE images without known prior knowledge of bleeding and normal features of the images.

3.1 Enhancement of WCE Images

3.1.1 Colour normalization

In classifying the bleeding area of WCE images, we start the experiments on raw images to train the neural network using DCNN. WCE images are divided into three parts: training, testing and validation set. Here, pre-processing technique is not applied on input of WCE images on training data of the neural network. Validation data set is used to estimate how well has the model been trained depending on the size of the training data, the input and etc. The validation and training errors are monitored to avoid over-fitting occurring in the training network using early stopping as its regularization methods. Both these errors should decrease when number of epoch increases. However, these experiments of training neural network without any pre-processing technique shows that training error decreases while the validation error increases after certain point even though it did decrease initially. This problem could be due to the difference in the colour of bleeding and normal areas on both training images and unknown WCE images are very prominent. The colour of bleeding of training images is totally different from the unknown WCE images.

The experiments on training neural network without any pre-processing technique as in Fig. 2 (a). The training of neural network using colour normalization process shows that training error decreases while validation error decreases as in Fig. 2 (b).



(a) Early stopping before colour normalization applied. (b) Early stopping after colour normalization applied.

Fig. 2. Early stopping during train neural network

The objective of the learning algorithm for neural network is to minimize the training and validation errors of each iteration process. After each iteration process, both the training and validation errors were evaluated. By minimizing both the training and validation errors, the performance of neural network using DCNN is maximized and improved to increase the result accuracy in classification. Thus, colour normalization is used which is effective in reducing variation of colours among input images, to minimize errors in training and validation data set. It can also be used to tackle the problem when the learning performance for bleeding classification deteriorates when the colour of unknown images differs from training images. In our proposed method, colour normalization scheme is performed on the images which compute the statistical properties of the distribution to pre-defined values in RGB colour space. The scheme applies principal component analysis (PCA) to an image so as to analyze colour distribution in an image [25, 26]. Following is the procedure of the colour normalization:

Let $f = |f_{ij}|$, $i = 1, \dots, M$; $j = 1, \dots, N$. Cluster center m of all pixels f is calculated by $m = E[f]$ where $E[\]$ is an expectation operation.

Eigenvalues $(\gamma_1, \gamma_2, \gamma_3)$ and eigenvectors of covariance matrix $C = E[(f - m)(f - m)^T]$ are computed.

Rodrigues formula is adapted for the rotation by an angle, θ around a vector n . $R_3(\theta, n) = I - \sin[\theta]U(n) + (1 - \cos\theta)U^2(n)$, where vector n is computed from $v = [(a, b, c)]^T$, or an eigenvector with the largest eigenvalue. $U^2(n) = nn^T - I$, $\|n\| = 1$ and I is an identity matrix. The details can be found in [27]. Gray value of an image is re-scaled by a variable factor, in which overflow above 255 and underflow under 0 are clipped to 255 and 0, respectively.

3.1.2 Contrast enhancement

The normalization process successfully reduces the colour variation of WCE images by normalizing along the orientation of the image. However, the methods also reduce the overall image contrast in WCE images, which might affect the detection task. Contrast enhancement increases the perceptibility of objects information in an image to be observed

by human eyes so the object to be depicted is easy to recognize. We applied a contrast enhancement scheme to the normalized image so as to enhance features of bleeding in each image, in which the original concept was proposed by [4]. The scheme is based on anisotropic contrast diffusion and hessian matrix, in which anisotropic diffusion process reduces image noise whereby hessian matrix describes contrast in an image. There is a drawback to the original technique by [4], where the quality of WCE image is degraded when the number of iteration increases. To solve this problem, variance formula is employed in our proposed method. The variance (σ^2) is a measure of the spread of pixel values around the image mean. By using the variance, it estimates the contrast probability distribution of the image. It will give an idea of how the pixel value spread in an image [3, 26]. The proposed anisotropic diffusion is defined as:

$$u(t) = \text{div}[g(c_{\sigma^2})|\nabla c_{\sigma^2}|] = g(c_{\sigma^2})\nabla c_{\sigma^2} + \nabla g \nabla c_{\sigma^2} \quad (1)$$

$$g(c_{\sigma^2}) = \frac{1}{1 + (\frac{\|f\|}{K})^2} \quad (2)$$

where

$$f = \nabla c_{\sigma^2}.$$

and

$$c(x, y) = v_1^2(x, y) + v_2^2(x, y)$$

Next, after the images are enhanced by both normalization and contrast enhancement techniques, these images are then used for the classification method to detect bleeding using DCNN.

3.2 Bleeding Detection using Deep Convolutional Neural Network (DCNN)

Convolutional neural network (CNN) is known as LeNet-5 and was designed for handwritten digit recognition [28]. Recent high research activity in deep learning had solved several problems of CNN. For example, rapid increase of computational power makes training with large number of images, increase number of layers (deep CNN; DCNN), and/or optimization of the architecture of CNN possible. Such breakthrough shines a light on CNN again [8]. Therefore, a DCNN is employed to classify each pixel in an input of WCE image into either bleeding or not by inputting a local block image whose center is an interest of pixel. The proposed method employs LeNet-5 based architecture which has been proved to have a great performance in classification task. Fig. 3 presents the proposed DCNN architecture, in which the architecture, or number of layer, size of convolution kernel and so on, are optimized in the study.

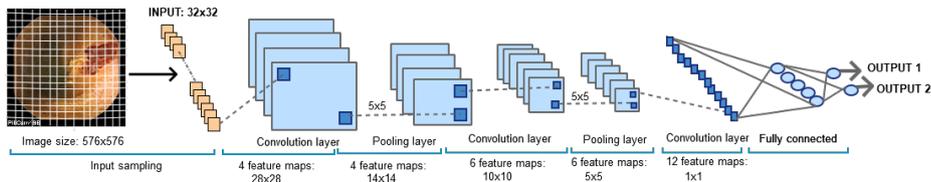


Fig. 3. Architecture of the DCNN used in the study .

The convolutional layers extract patterns that are effective for detecting bleeding from WCE images, meanwhile pooling layers perform a local averaging to suppress undesirable fluctuation of output caused by shift and distortion of an input bleeding pattern. Finally, the fully connected layer merges all features extracted from an input block image to decide whether the centre pixel of the block image is the bleeding area or not. This proposed DCNN architecture is trained using stochastic gradient descent to minimize the following energy function:

$$L(w, b) = \frac{1}{2} \|h_{(w,b)}(x) - y\|^2 \quad (3)$$

where, x is a training sample, ω denotes an weight matrix, b is a bias unit, $h_{(w,b)}(x)$ is an output real number of network and y represents the true class labels. Sigmoid function is used as an activation function.

Let input as $V \in \mathfrak{R}^{s \times r}$ where s and r are the size of input block image, and $\omega \in \mathfrak{R}^{(a \times b) \times n}$ is a weight matrix where symbols $a \times b$ is a size of convolution and n are number of features maps. An input matrix V is convolved with a weight matrix ω . As a result, convolutional procedures yields m feature maps of size $(s - a) \times (r - b)$. Generally, CNN network is composed of number of layer, $l = 1, 2, \dots, n$ where first layer is input layer and layer n is output layer. Each layer l is connected to previous layer $l - 1$. In feed forward propagation, the computation is;

$$x_j^{l,m} = \sigma \left(\sum_{j=1}^n \omega_{ij}^{l-1} x_j^{l-1,m} + b_i^{l-1} \right) \quad (4)$$

where $x_j^{l,m}$ is one of the j^{th} unit of the m^{th} feature map in l^{th} layer, ω_{ij}^{l-1} is a parameter (or weight) between $(i, j)^{th}$ unit of the m^{th} feature map in l^{th} layer with $(i, j)^{th}$ unit in $(l - 1)^{th}$ layer. Also, b is a bias unit and σ is a hyperbolic tangent function used as an activation function. Each parameter ω and b is initialized randomly to very small numbers.

For back-propagation algorithm, the cost function $L(\omega, b)$ is defined as follows:

$$\frac{\partial}{\partial \omega_{i,j}^l} L(\omega, b) = \left[\frac{1}{m} \sum_i \frac{\partial}{\partial \omega_{(i,j)^l}} L(\omega, b; x^i, y^i) \right] \quad (5)$$

$$\frac{\partial}{\partial b_i^l} L(\omega, b) = \frac{1}{m} \sum_i \frac{\partial}{\partial b_i^l} L(\omega, b; x^i, y^i) \quad (6)$$

At each iteration, gradient descent method updates the parameters ω, b as follows to reduce our cost function:

$$\omega_{i,j}^l = \omega_{i,j}^l - \Omega \frac{\partial}{\partial \omega_{i,j}^l} L(\omega, b) \quad (7)$$

$$b_i^l = b_i^l - \Omega \frac{\partial}{\partial b_i^l} L(\omega, b) \quad (8)$$

where Ω is the learning rate.

In the output layer, a pixel corresponding to bleeding area is recognized by applying threshold to the value of (class output 2) – (class output 1). If the value is larger or equal to zero, a centre pixel of an input block image will be recognized as a bleeding area and otherwise it will not. To remedy the over-fitting problem, an early stopping principle is introduced.

4. EXPERIMENTAL RESULTS

In this experiment, MATLAB program on Intel Xeon-2.8GHz platform is used. Endoscopy images are collected from “www.capsuleendoscopy.org”. Bleeding area for the training and test images were manually labeled through each pixel. In this study, all blocks for training, validation and testing were extracted from inside masked areas. In this experiments, nine images from a particular patient were used for the training set while 35 images from five patients were used for the test, or validation of the proposed algorithm and comparison with other algorithms. For training, 90,000 blocks was extracted from nine images, in which 54,000 blocks are from bleeding areas and 36,000 blocks are from other areas. In the validation step, 10,000 blocks from one image were used for early stopping. In the test step, 35 images from five patients with various colour of bleeding is used in which 32x32 size of blocks are extracted from these images and 1,887,533 blocks were extracted from bleeding areas and 6,089,252 blocks were from non-bleeding areas.

As has been mentioned, this DCNN architecture could provide an important factor to achieve high classification performance. We decided that the best parameters for the DCNN architecture depend on the training and validation errors as shown in Table 1. From this table, the lowest training error, $train_{min.err}$ is 0.0102 with the validation error, val is 0.0338 for the DCNN with completely enhanced images (shown in no 9 in the table). While DCNN with normalized images have the lowest training error, $train_{min.err}$ 0.0105 with the validation error, val is 0.0318 (shown in no 9 in the table). To examine the effectiveness of the proposed algorithm which combines contrast enhancement process and DCNN based classification, four different experiments were performed: 1) conventional Fuzzy, 2) SVM, 3) the proposed DCNN based classification, and 4) the proposed DCNN based classification with completely enhanced images. As mentioned above, 35 images were used but for simplicity only selected images were shown to present the results of the experiments. Then, the experiments were run on normalized images based on colour cluster rotation to train the neural network again using DCNN. Training neural network using colour normalization process shows that training error decreases while validation error decreases.

Fig. 4 shows the sequence of images depicting the true bleeding area, bleeding area both detected by classification procedure using DCNN and Fuzzy, each presented by yellow, cyan and green colour respectively. Six patients cases (Patients A-F) are considered in this experiment where the image of patient A is taken to be the training image and images of patients B-F with various bleeding colour are taken to be the testing images. From these figures, it can be seen that DCNN provides a better detection of the bleeding area compared to those performed using Fuzzy based scheme. Then, the comparison result between SVM and DCNN to detect the bleeding are shown in Fig. 5. As we can see from this figure, DCNN provides better detection of the bleeding areas when compared to the one using SVM. SVM has low performance in terms of sensitivity compared to DCNN where it is unable to differentiate between normal and bleeding areas. From Fig. 5 (C1), (C2), (D1), (D2), (F1) and (F2), bleeding areas are mistakenly identified as normal areas.

Since DCNN was shown to have a better performance than Fuzzy and SVM, the study then focuses on the different result on the experiment using DCNN, namely using DCNN with completely enhanced images and with normalized images. Fig. 6 shows the comparison result for the test images as well as the training images between DCNN with completely enhanced images and DCNN with normalized images. Images of patients B-F are the results for the test images of five patients with various colour of bleeding; whereby images of A1, A2 and A3 are the results for the training images of patient A.

Table 1. Comparison between DCNN with completely enhanced images and DCNN with normalized images (Learning rate, $\Omega = 0.01$).

No	Size of Input	No of layer	DCNN with completely enhanced images	DCNN with normalized images
			train _{min_err} , val	train _{min_err} , val
1	28x28	4, 6	0.2400, 0.2400	0.1230, 0.0467
2	28x28	4, 8	0.2400, 0.2400	0.0135, 0.0443
3	28x28	6, 12	0.0135, 0.0437	0.0133, 0.4630
4	28x28	8, 16	0.0153, 0.0570	0.0172, 0.0376
5	32x32	4, 6	0.2400, 0.2399	0.2400, 0.2400
6	32x32	4, 8	0.0134, 0.0519	0.0129, 0.0548
7	32x32	6, 12	0.0149, 0.0497	0.0137, 0.0437
8	32 x 32	8, 16	0.0107, 0.0324	0.0141, 0.0408
9	32 x 32	4, 6, 12	0.0102, 0.0338	0.0105, 0.0318
10	32 x 32	4, 8, 16	0.0166, 0.0455	0.0117, 0.0495
11	32 x 32	6, 12, 24	0.0222, 0.0541	0.0131, 0.0876
12	32 x 32	8, 16, 32	0.0130, 0.0362	0.2400, 0.2400
13	42 x 42	4, 6	0.0124, 0.0410	0.0130, 0.0491
14	42 x 42	4, 8	0.2400, 0.2400	0.0128, 0.0669
15	42 x 42	8, 16	0.0172, 0.0627	0.0149, 0.0520
16	42 x 42	4, 6, 12	0.0199, 0.2460	0.2400, 0.2400
17	42 x 42	4, 8, 16	0.0140, 0.0398	0.2400, 0.2400
18	42 x 42	6, 12, 24	0.0125, 0.0328	0.1250, 0.0855
19	42 x 42	8, 16, 32	0.0156, 0.0389	0.2400, 0.2400
20	64 x 64	4, 6, 12	0.0112, 0.3320	0.2400, 0.2400
21	64 x 64	4, 8, 16	0.0200, 0.0581	0.2400, 0.2400
22	64 x 64	6, 12, 24	0.0113, 0.0322	0.0130, 0.0866
23	64 x 64	8, 16, 32	0.0139, 0.0336	0.2400, 0.2400

The yellow, cyan and green boundaries show the true bleeding area, extracted areas using the DCNN with completely enhanced images and the extracted areas using scheme DCNN with normalized images based on approach by [29]. These results show that the scheme with completely enhanced images detect the bleeding area better when compared to the one with normalized images. As can be seen from Fig. 6, the contrast enhancement scheme is effective in classifying the bleeding areas of the image.

5. DISCUSSION

In this section, Receiver Operating Characteristic (ROC) curves of the different bleeding detection algorithms in which curves were drawn by changing threshold of the classification output are presented. True positive (TP) rate is plotted on the x-axis and false positive (FP) rate is plotted on the y-axis [30]. It is noted that TP rate is the ratio of bleeding pixels correctly classified as bleeding and FP rate is the ratio of non-bleeding pixels incorrectly classified as bleeding. In addition, FP rate is also known as a false alarm rate (FAR), which is equal to (1- specificity).

Fig. 7 shows ROC curves of DCNN with completely enhanced images (blue line) and DCNN with normalized images scheme (red line). ROC curves for WCE images of patients A-F where figures of A1, A2 and A3 are the results for the training images; and B-F are the results for test images. It can be seen that curves with completely enhanced images scheme are superior to curves with normalized scheme for most cases. Area under

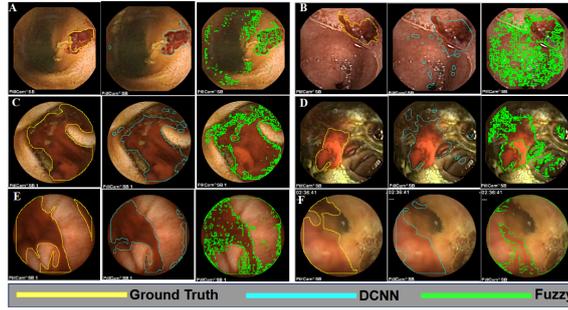


Fig. 4. The sequence of images depicting the true bleeding area, bleeding area both detected by classification procedure using DCNN and Fuzzy.

ROC curve (A_z) showed that A_z 's of the three images out of three training images and 10 images out of 12 test images support the superiority of the contrast enhancement scheme.

The performance indices in the figure are defined as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

$$\text{Average} = \frac{(\text{Sensitivity} + \text{Specificity})}{2} \tag{9}$$

where TP is the number of bleeding pixels correctly classified as bleeding, FN is the number of bleeding pixels incorrectly classified as negative, TN is the number of non-bleeding pixels correctly classified as negative, and FP is the number of non-bleeding pixels incorrectly classified as bleeding.

Fig. 8 summarizes results of DCNN, SVM and Fuzzy Logic after thresholding ROC curves with threshold=0 (default value of the threshold). The boxplots of Fig. 8 prove that DCNN has a great ability in detecting bleeding areas in WCE images compared to SVM and conventional fuzzy logic based approach where it is difficult to optimize the rules, in particular from the view point of high robustness against large variation in colour. As we can see from this figure, the mean of the specificity, sensitivity and average for the DCNN are 0.8907, 0.7363 and 0.8135 respectively whereas the mean of the specificity, sensitivity and average for the SVM are 0.7740, 0.2855 and 0.5297 and the mean of the specificity, sensitivity and average for the fuzzy is 0.5538, 0.4154 and 0.4846. The results showed that DCNN is better than SVM and Fuzzy in terms of specificity, sensitivity and average. Fig. 9 shows boxplots with completely enhanced images and with normalized images process using DCNN based classification. This figure shows the values of mean of the specificity, sensitivity and average for the DCNN with completely enhanced images which are 0.8703, 0.8271 and 0.8907. The mean of the specificity, sensitivity and average for the DCNN with normalized images are 0.8907, 0.7363 and 0.8135. From this figure, it shows that mean and maximum value of specificity for DCNN with completely enhanced images is lower than DCNN with normalized images. However, the minimum value of specificity is higher than DCNN with normalized images. For the minimum, mean, and maximum of the sensitivity and average values, DCNN with completely enhanced images is better than DCNN with normalized images. Although the specificity of DCNN with completely enhanced images is a bit lower than that with normalized, sensitivity and average are

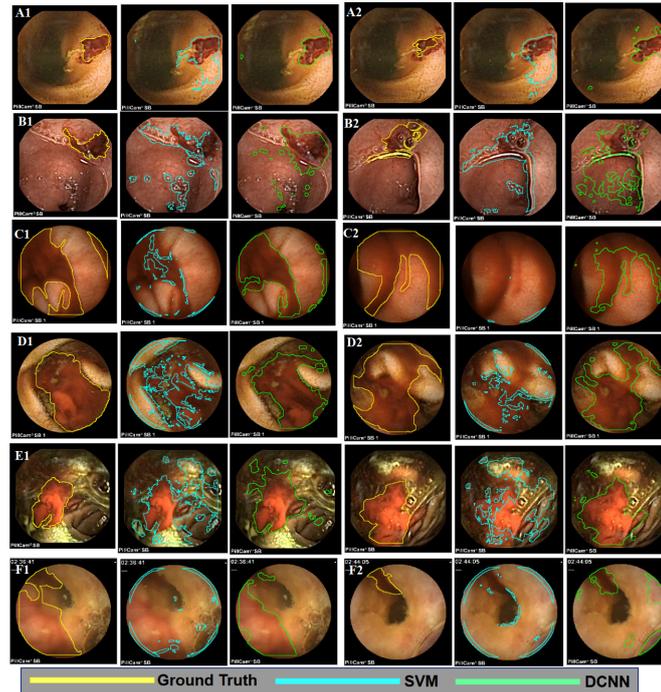


Fig. 5. The sequence of images depicting the true bleeding area, bleeding area both detected by classification procedure using DCNN and SVM.

improved by the proposed enhancement technique. Generally speaking, high sensitivity is more important than specificity in order to prevent false negative which is more severe for not only patients but also to the medical doctors. However the difference should be statistically evaluated and hence, the next section discusses the statistical difference. Fig. 10 is show the sequence of images depicting the true bleeding area and bleeding area detected by classification procedure using Fuzzy, SVM, DCNN, DCNN with completely enhanced images, and DCNN with normalized images in one figure.

Statistical hypothesis test is performed using Wilcoxon signed rank test in terms of performances indices, that are specificity, sensitivity and average of detecting bleeding area after thresholding ROC curves with threshold=0. Although ROC curves are sufficient in showing the performance indices, but in the actual clinical situation, the results of thresholding the output of the classifier could be confirmed using statistical test of these performance indices and could assist the medical officer to affirm his diagnosis. First, statistical tests between DCNN and conventional Fuzzy was carried out under the following null hypothesis H_0 :

1. Specificity of DCNN, SVM and Fuzzy are same.
2. Sensitivity of DCNN, SVM and Fuzzy are same.
3. Average of DCNN, SVM and Fuzzy are same.

The result of statistical test for each performance index are as follows:

1. The null hypothesis in terms of specificity was rejected at the significance level 0.05.
2. The null hypothesis in terms of sensitivity was rejected at the significance level 0.05.
3. The null hypothesis in terms of average was rejected at the significance level 0.05.



Fig. 6. Images of patients of A-F depicting the true bleeding area, using DCNN with completely enhanced images and DCNN with normalized images based approach.

From the results, it can be concluded that the DCNN based classification is effective compared to the SVM and conventional Fuzzy based classification.

Second, we performed statistical tests between DCNN with completely enhanced images and DCNN with normalized images. The null hypothesis H_0 of the Wilcoxon signed rank test are as follows:

1. Specificity of DCNN with completely enhanced images and that of DCNN with normalized images are same.
2. Sensitivity of DCNN with completely enhanced images and that of DCNN with normalized images are same.
3. Average of DCNN with completely enhanced images and that of DCNN with normalized images are same.

Followings are summary of the statistical test:

1. The null hypothesis in terms of specificity was not rejected at the significance level 0.05.
2. The null hypothesis in terms of sensitivity was rejected at the significance level 0.05.
3. The null hypothesis in terms of average was rejected at the significance level 0.05.

As mentioned above, the proposed enhancement technique seems to be ineffective in terms of specificity but the statistical test shows that the difference is not statistically significant. Therefore, we conclude that the proposed enhancement technique is effective for DCNN based classification, in particular from the view point of sensitivity, and average of sensitivity and specificity. Statistical test on area under ROC curve (A_z) between DCNN with completely enhanced images and DCNN with normalized images were also performed. The null hypothesis H_0 of the Wilcoxon signed rank test is A_z of DCNN with completely enhanced images and that of DCNN with normalized images is the same. It can be concluded that the two compared areas of ROC curves (A_z) were rejected which means that they are significantly different at the significance level of 0.05. This indicates that DCNN with completely enhanced images has a better performance than DCNN with normalized images.

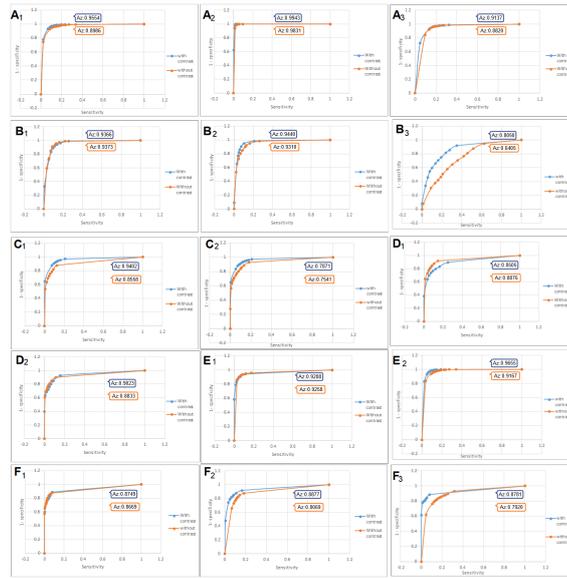


Fig. 7. ROC curves for WCE images of patients A-F.

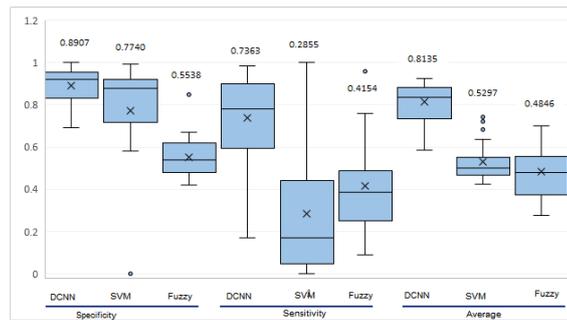


Fig. 8. Performance comparison between DCNN, SVM and the conventional fuzzy logic.

6. CONCLUSION

In this study, a novel method for classification with completely enhanced images scheme by using DCNN to distinguish bleeding and normal areas in WCE images has been presented. In this study, the focus is given to solve the difficult task on detecting bleeding area due to the nature of bleeding that has various colour, pattern and texture. First, colour normalization is used to reduce variation of colours among input images. Second, contrast enhancement scheme to enhance the bleeding colour in order to ease the classification task was performed. Finally, DCNN is used to classify each block of the WCE images without any prior knowledge of bleeding and normal features of the images. The experimental results showed a promising performance of DCNN proposed algorithm with completely enhanced images on bleeding detection of WCE images as compared to DCNN with normalized images, SVM and Fuzzy logic. The specificity, sensitivity and average for DCNN with completely enhanced images are 0.8703, 0.8271 and 0.8907 respectively, whereas the specificity, sensitivity and average for DCNN with normalized images are 0.8907, 0.7363 and 0.8135. These results proved that the pro-

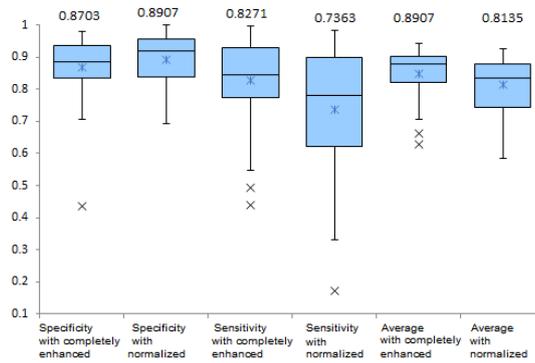


Fig. 9. Performance comparison of DCNN with completely enhanced images and with normalized images.

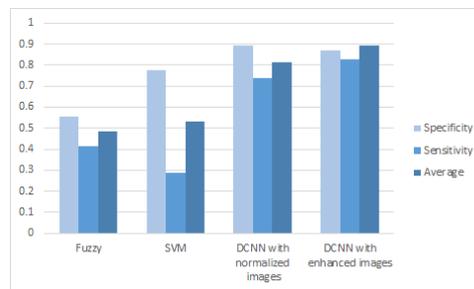


Fig. 10. Performance comparison of Fuzzy, SVM, DCNN with normalized images and with completely enhanced images.

posed enhancement technique gives better performance. The DCNN approach is able to classify bleeding from the normal areas of WCE images. For the overall contribution, the proposed technique, which consists of pre-processing and classification technique to distinguish bleeding and normal areas in WCE images has been developed. The main issue with WCE images is that these images have various colours, patterns and textures in both bleeding and normal areas. Therefore, based on the results, the classification of the enhanced images results clearly contributing to better prediction and evaluation of the images. This can help doctors in giving a better diagnosed. For future works, the higher performance indices in term of specificity, sensitivity and average should be taken into consideration. Some relative works should also be investigated to improve the accuracy of the bleeding detection.

ACKNOWLEDGEMENT

This work is supported by Universiti Malaysia Pahang (UMP) Research Grant Scheme (RDU No. 1803147), Tokyo University of Agriculture and Technology (TUAT) and Japan Student Services Organization (JASSO).

REFERENCES

1. M. Liedlgruber and A. Uhl, "Computer-aided decision support systems for endoscopy in the gastrointestinal tract: A review," *IEEE Reviews in Biomedical Engineering*, Vol. 4, 2011, pp. 73-88.
2. D. K. Iakovidis, "Software engineering applications in gastroenterology," *Global Journal of Gastroenterology & Hepatology*, Vol. 2, 2014, pp. 11-18.
3. R. Shahril, S. Baharun, A. K. M. M. Islam, and S. Komaki, "Anisotropic contrast diffusion enhancement using variance for wireless capsule endoscopy images," in *Proceedings of International Conference on Informatics, Electronics Vision*, 2014, pp. 1-6.
4. B. Li and M. Q.-H. Meng, "Wireless capsule endoscopy images enhancement via adaptive contrast diffusion," *Journal of Visual Communication and Image Representation*, Vol. 23, 2012, pp. 222-228.
5. P. Szczypiski, A. Klepaczko, M. Pazurek, and P. Daniel, "Texture and color based image segmentation and pathology detection in capsule endoscopy videos," *Computer Methods and Programs in Biomedicine*, Vol. 113, 2014, pp. 396-411.
6. S. Suman, F. A. B. Hussin, N. Walter, A. S. Malik, S. H. Ho, and K. L. Goh, "Detection and classification of bleeding using statistical color features for wireless capsule endoscopy images," in *Proceedings of International Conference on Signal and Information Processing*, 2016, pp. 1-5.
7. H. S. Pannu and P. Jarial, "Machine learning techniques for bleeding detection in capsule endoscopy," in *Proceedings of International Conference on Computing Methodologies and Communication*, 2017, pp. 1054-1058.
8. R. Shahril, S. Baharun, A. K. M. M. Islam, and S. Komaki, "Bleeding classification on wireless capsule endoscopy images using deep convolutional neural network," *International Journal of Computer Assisted Radiology and Surgery*, Vol. 10, 2015, pp. S292-293.
9. Y. Yuan, B. Li, and M. Q. H. Meng, "Bleeding frame and region detection in the wireless capsule endoscopy video," *IEEE Journal of Biomedical and Health Informatics*, Vol. 20, 2016, pp. 624-630.
10. Y. Fu, W. Zhang, M. Mandal, and M. Q. H. Meng, "Computer-aided bleeding detection in wce video," *IEEE Journal of Biomedical and Health Informatics*, Vol. 18, 2014, pp. 636-642.
11. J. Yu, J. Chen, Z. Q. Xiang, and Y. Zou, "A hybrid convolutional neural networks with extreme learning machine for wce image classification," in *Proceedings of IEEE International Conference on Robotics and Biomimetics*, 2015, pp. 1822-1827.
12. G. Wimmer, A. Vcsei, and A. Uhl, "Cnn transfer learning for the automated diagnosis of celiac disease," in *Proceedings of the 6th International Conference on Image Processing Theory, Tools and Applications*, 2016, pp. 1-6.
13. Y. Zou, L. Li, Y. Wang, J. Yu, Y. Li, and W. J. Deng, "Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network," in *Proceedings of IEEE International Conference on Digital Signal Processing*, 2015, pp. 1274-1278.
14. X. Jia and M. Q. H. Meng, "Gastrointestinal bleeding detection in wireless capsule endoscopy images using handcrafted and cnn features," in *Proceedings of the 39th*

- Annual International Conference of IEEE Engineering in Medicine and Biology Society*, 2017, pp. 3154-3157.
15. X. Li, H. Zhang, X. Zhang, H. Liu, and G. Xie, "Exploring transfer learning for gastrointestinal bleeding detection on small-size imbalanced endoscopy images," in *Proceedings of the 39th Annual International Conference of IEEE Engineering in Medicine and Biology Society*, 2017, pp. 1994-1997.
 16. A. K. Sekuboyina, S. T. Devarakonda, and C. S. Seelamantula, "A convolutional neural network approach for abnormality detection in wireless capsule endoscopy," in *Proceedings of IEEE 14th International Symposium on Biomedical Imaging*, 2017, pp. 1057-1060.
 17. J. He, X. Wu, Y. Jiang, Q. Peng, and R. Jain, "Hookworm detection in wireless capsule endoscopy images with deep learning," *IEEE Transactions on Image Processing*, Vol. 27, 2018, pp. 2379-2392.
 18. T. Ghosh, S. A. Fattah, C. Shahnaz, and K. A. Wahid, "An automatic bleeding detection scheme in wireless capsule endoscopy based on histogram of an rgb-indexed image," in *Proceedings of the 36th Annual International Conference of IEEE Engineering in Medicine and Biology Society*, 2014, pp. 4683-4686.
 19. T. Ghosh, S. K. Bashar, M. S. Alam, K. Wahid, and S. A. Fattah, "A statistical feature based novel method to detect bleeding in wireless capsule endoscopy images," in *Proceedings of International Conference on Informatics, Electronics Vision*, 2014, pp. 1-4.
 20. S. Charfi and M. E. Ansari, "Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy videos," in *Proceedings of International Conference on Advanced Technologies for Signal and Image Processing*, 2017, pp. 1-5.
 21. T. Ghosh, S. A. Fattah, and K. A. Wahid, "Chobs: Color histogram of block statistics for automatic bleeding detection in wireless capsule endoscopy video," *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 6, 2018, pp. 1-12.
 22. X. Jia and M. Q. H. Meng, "A study on automated segmentation of blood regions in wireless capsule endoscopy images using fully convolutional networks," in *Proceedings of IEEE 14th International Symposium on Biomedical Imaging*, 2017, pp. 179-182.
 23. E. Tuba, M. Tuba, and R. Jovanovic, "An algorithm for automated segmentation for bleeding detection in endoscopic images," in *Proceedings of International Joint Conference on Neural Networks*, 2017, pp. 4579-4586.
 24. C. P. Sindhu and V. Valsan, "A novel method for automatic detection of inflammatory bowel diseases in wireless capsule endoscopy images," in *Proceedings of the 4th International Conference on Signal Processing, Communication and Networking*, 2017, pp. 1-6.
 25. R. Shahril, S. Baharun, and A. K. M. M. Islam, "Pre-processing technique for wireless capsule endoscopy image enhancement," *Journal of Visual Communication and Image Representation*, Vol. 6, 2016, pp. 1617-1626.
 26. R. Shahri, D. Arianti, S. Baharun, A. K. M. M. Islam, and S. Komaki, "Pre-processing technique based on discrete cosine transform (dct) and anisotropic contrast diffusion for wireless capsule endoscopy images," in *Proceedings of IEEE Conference on Biomedical Engineering and Sciences*, 2014, pp. 922-927.

27. D. Paulus, L. Csink, and H. Niemann, "Color cluster rotation," in *Proceedings International Conference on Image Processing*, Vol. 1, 1998, pp. 161-165.
28. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, Vol. 86, 1998, pp. 2278-2324.
29. M. Y. Khachane and R. J. Ramteke, "Article: Fuzzy rule based multimodal medical image edge detection," *IJCA Proceedings on National Conference on Recent Advances in Information Technology*, Vol. NCRAIT, 2014, pp. 28-32.
30. D. J. Hand and R. J. Till, "A simple generalisation of the area under the roc curve for multiple class classification problems," *Machine Learning*, Vol. 45, 2001, pp. 171-186.



Rosdiana Shahril received the Ph.D. degree in Image Processing from Universiti Teknologi Malaysia, in October 2018. She is currently a Senior Lecturer in Universiti Malaysia Pahang, Pahang, Malaysia. Her current research interests include around deep learning, machine learning, computer vision, and image processing.



Atsushi Saito received the Ph.D. degree from Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2016. He is currently a Research Associate with the Tokyo University of Agriculture and Technology. His current research interests include statistical analysis, statistical modelling, multivariate data analysis and machine learning.



Akinobu Shimizu received the Ph.D. degree from Nagoya University in 1995. His research interests include medical image analysis of embryo, adult and cadaver. He is currently a Professor at Tokyo University of Agriculture and Technology. As of 2017 he has published 78 peer-reviewed research papers in journals, 7 book chapters and 96 international conferences, including those published in top medical imaging journals (IEEE Transaction on Medical Imaging, Medical Image Analysis).



Sabariah Baharun received the Ph.D. degree from Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia. Her research interests include graph and fuzzy graph modelling and mathematical thinking. She is currently an Associate Professor at Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia.