Features for Discriminating Normal Cases in Mass Screening for Gastric Cancer with Double Contrast X-ray Images of Stomach

[1]Koji Abe, [2]Hideaki Nakagawa, [3]Masahide Minami and [4]Haiyan Tian
[1]*Department of Informatics, School of Science and Engineering, Kinki University, Japan;*
[2]*Interdisciplinary Graduate School of Science and Engineering, Kinki University, Japan;*
[3]*Kanazawa Gakuin University, Japan;*
[4]*Graduate School of Engineering, Kobe University, Japan;*
[1]koji@info.kindai.ac.jp, [2]nakagawa0623nara@gmail.com, [3]maminami@dream.com,
[4]tian.haiyan2006@gmail.com

ABSTRACT

In a mass screening for gastric cancer, diagnosticians read many stomach X-ray pictures at a time. To decrease the number of reading the pictures in the mass screening, the proposed method discriminates normal cases in stomach X-ray images using the proposed features. In the normal cases, folds on the stomach wall appear in parallel in the images. Considering this characteristic, the proposed method measures parallelism of the folds in the images. Experimental results of the discriminations for 88 images where 13 abnormal cases are included have shown that the proposed features are well effective for recognizing normal cases.

Keywords: Computer-aided diagnosis; Gastric cancer; Double contrast X-ray image of stomach; Medical image processing.

1   Introduction

In mass screenings for gastric cancer, due to financial reasons, double contrast X-ray pictures are generally used on behalf of CT, MRI or photogastroscope. In the screenings, diagnosticians always need a hard labor because they read several hundred X-ray pictures at a time. Besides, since accuracy of the reading is strongly depended on experience of diagnosticians, it is hard for inexperienced doctors to read them well. For the reasons, CAD (computer-aided diagnosis) systems for gastric cancer in X-ray images have been required as a second opinion for diagnosticians [1-5]. The existing systems mark the location of lesions in X-ray images which show an abnormal case [2-5]. In gastric X-ray images, since a cancer part appears as a pattern that lines concentrate at a spot in a stomach area, the existing systems detect the part by measuring degree folds concentrate or applying a filter for emphasizing the concentration part in the images. However, the systems cannot reduce the hard labor due to lack of accuracy in the marking. Besides, to diagnose characteristics of legions, since diagnosticians have to directly read pictures of abnormal cases even though the systems could show the location precisely, the recognition of abnormal cases does not result in decrease of the doctors' labor in the mass screenings.

Therefore, for the sake of reducing the labor by decreasing the number of the readings, a CAD for discriminating normal cases using digital gastric X-ray images was reported [6]. In the images of double contrast X-ray images of stomach, the folds appear in parallel on the stomach wall of normal cases. On the other hand, in abnormal cases, the folds just appear to be irregular in addition to cases where the folds concentrate at a spot. Hence, this system first extracts the folds in the stomach area of the images and then measures degrees of parallelism for the folds. The degrees are measured by extracting characteristics from distribution of the folds in 2D space. After that, using the features, this system discriminates normal cases using a classification machine. However, since the features were strongly depended on the size of stomach area, the system could not always recognize the normal cases correctly. In addition, the extraction of the folds were not automatic. Therefore, this paper proposes a method for extracting the folds automatically and improve the features for measuring the normality.

2   Double Contrast X-Ray Images of Stomach

In radiography to obtain contrast X-ray pictures of stomach, participants drink barium and radiologists take eight X-ray pictures changing direction of their body. Diagnosticians first choose the head-on double contrast X-ray picture among the eight pictures and diagnose the stomach reading it. When it is difficult to diagnose the stomach due to an uncertain stomach or a doubtful case, the other seven pictures are used as the second material. Therefore, as the first step for designing the CAD system of diagnosing normal cases of stomach, the head-on double contrast X-ray images are used throughout this paper. Figure 1 shows double contrast gastric X-ray pictures, where the left is a normal case and the right is an abnormal one. Both of them show folds inside the stomach area. As shown in the left image, folds appear like they are put side by side in parallel in normal cases. On the other hand, in the right image of an abnormal case, folds appear to be irregular. And, Figure 2 shows a typical pattern of gastric cancer, where the folds appear like lines concentration at a spot. The existing CAD systems find the location of gastric cancer such as this by recognizing the concentration pattern of lines in the images. After a lesion appear in the stomach area like gastric cancer or ulcer, the lesion part becomes hard on the gastric wall. Due to this phenomenon, folds are disordered around the part or in the whole of the gastric wall as shown in the right image in Figure 1. These stomachs such as Figure 1 (right) and 2 are diagnosed as abnormal case in mass screenings of gastric cancer and its carriers have to have a re-inspection by another way such as the photogastroscope.

Throughout this paper, all the double contrast X-ray pictures are digitalized as images by CR (Computed Radiography). The size of the images is 1024×1024 pixels with 256 gray levels.
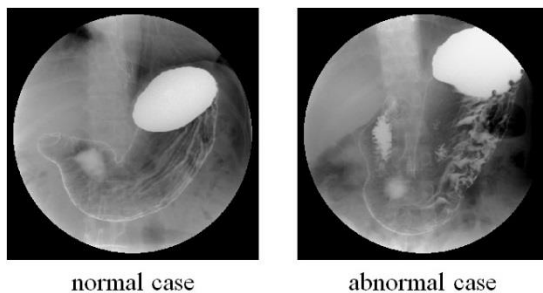

normal case   abnormal case

Figure 1: Double contrast X-ray images of stomach
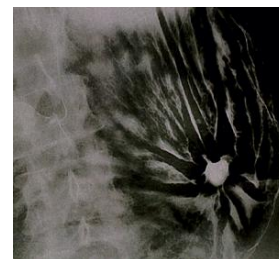

Figure 2: A concentration pattern of stomach folds (a typical case of gastric cancer)

3 Proposed Method

3.1 Preparation

As shown in Figure 1, to extract the characteristic of parallelism for the folds, a lot of noises such as the spinal column, unevenness density, etc. are existed in the images. Generally, in double contrast gastric X-ray images, since the contrast is weak and noises are much around the pyloric part, it would be difficult for even diagnosticians to read the part. To the contrary, since the contrast is strong around the cardia, the folds tend to appear clearly. Hence, as a preprocessing, the target area to extract the parallelism of folds is manually extracted as follows. First, using a tablet PC, a stomach area is extracted by roughly drawing its outline as shown in Figure 3(b). Next, the target area is extracted under the following conditions:

1) The area is the lower part from the horizontal line (b in Figure 3(c)) located at the bottom of the barium-pool, and

2) The area is the righter part from the vertical line (an in Figure 3(c)) located at the connection of the barium-pool to the left side edge in the cardia.

Besides, since the boundary of the stomach area becomes a noise for the extraction of folds, ten pixels from the right side of the boundary of the area to the inner of the area, and thirty pixels from the left side to the inner are cut at every row from the area. Similarly, two pixels from the upper boundary to the inner and ten pixels from the lower boundary to the inner are cut at every column, too. The last area is regarded as the target area. Figure 3 shows the extraction process of the target area, where A is the barium-pool, B is the cardia, C is the pylorus, and Figure 3(d) shows the target area R(x, y) for the image of Figure 3(a). The reason why the target area is restricted as shown above is because the folds appear clear and stable comparatively under the barium-pool, and noises such as shadows of other organs are roughly removed.

A method for extracting the whole stomach area based on the Snake algorism [7] was separately proposed for CAD systems of gastric cancer [8]. However, it is still necessary to enhance its quality for practical use. And, although the target area is extracted by hand as shown above, note that the region for the diagnosis in all the existing CAD systems was extracted manually as well. Therefore, an automated extraction of the target area is designed separately as a future work.
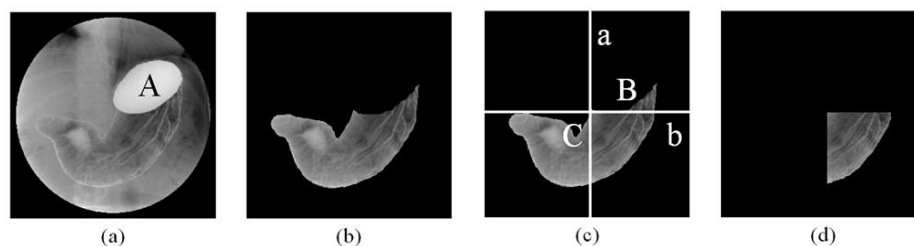

Figure 3: The target area for the normality extraction.

3.2   Extraction of the folds

In the X-ray images, pixel values of the folds are higher than other area and appear remarkably. Considering these characteristics, the brightness gradient strength F(x, y) and the pixel values H(x, y) which is output of the high-pass filter by 7x7 mask [9] are calculated to the target area R(x, y). F(x, y) is obtained by Eq.1. Next, in the 2D feature space created by F(x, y) and H(x, y), every pixel of the folds is classified by k-means clustering [10] into "a candidate of the fold area" and "others". When k-means clustering separates the pixels into two classes, two initial pixels have to be prepared. In this k-means clustering, the couple of two pixels which has the longest distance in the feature space are selected as the initial pixels. For example, after the k-means clustering is conducted to the image of Figure 4(a), the image is converted into Figure 4(b), where the white pixels represents the candidates. Thus, in outputs by the k-means clustering, we can see several regions are appeared. Finally, the folds are extracted by the linear discriminant analysis regarding perimeter of each candidate area as a variable of the discriminant analysis. The discriminant analysis is applied to every candidate area in the image and their discriminant score decides whether the candidate area belongs in the folds or its background. If the score for the region is more than 0, the region belongs in the folds. Figure 4(c) shows the folds extracted from Figure 4(a) via Figure 4 (b).

$$F(x, y) = \sqrt{F_x(x, y)^2 + F_y(x, y)^2}$$
$$\begin{cases} F_x(x, y) = R(x+1, y) - R(x-1, y) \\ F_y(x, y) = R(x, y+1) - R(x, y-1) \end{cases} \tag{1}$$
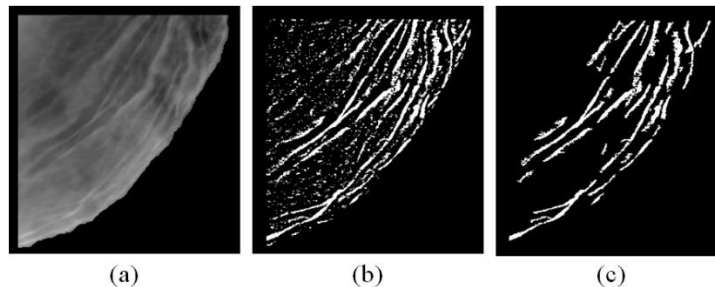

Figure 4: Process for extracting the folds.

3.3   Extraction of features

3.3.1   Extraction of the normalities: f1, and f2

First, the vector convergence index filter [11] is applied to the target area $R(x, y)$. The filter outputs the value $ang(x, y)$ of line angle at every pixel in $R(x, y)$, where $0 \leqslant ang(x, y) < 180$. $AngSum$ (i) is the number of pixels of the angle i in the folds, where $0 \leqslant i < 180$. Feature f_1 is defined as the kurtosis for $AngSum$ (i).

$$f_1 = \frac{1}{180} \sum_{i=0}^{179} \frac{(AngSum(i) - m)^4}{s^4} \tag{2}$$

Where, m is the mean value of $AngSum$ (i) and s^2 is the variance of the $AngSum$ (i). The folds' curves of normal cases appear in parallel along the outline of stomach area, i.e., shape of the curves depends on

the outline. Receiving this phenomenon, f_1 measures parallelism of the folds considering the outline of stomach area by extracting the kurtosis for the one dimensional array *AngSum*.

Next, when *AngSum* (*i_max*) is the maximum number among *AngSum* (0) – *AngSum* (179), regarding the angle *i_max* as *AngMax*, feature f_2 is defined as the proportion of the folds around *AngMax* (*AngMax*±12) to all the folds.

$$f_2 = \frac{1}{area} \sum_{i=0}^{179} F(i)$$

$$F(i) = \begin{cases} AngSum(i), & if \ AngMax - 12 \leq i, \ and \ i \leq AngMax + 12 \\ 0, & otherwise \end{cases}$$

(3)

Where *area* is the number of pixels in the folds, f_2 represents degree of unity for directions of the folds' curve. Since the folds figure with parallel curves in the target area of normal cases, the directions would be concentrated in a small range of angle.

3.3.2   Extraction of the normality: f3

Since folds appear irregularly in abnormal cases and intervals between the folds are smaller than normal cases, the intersections of the folds are much and crowd in zones (which could be lesions) than normal cases. To catch the characteristic, first, the thinning is conducted to the folds. And then, the number of pixels which have are the intersection of at least three neighbors in 8-neighbor are extracted. After that, the folds are removed within the radius *r* at every intersection. Feature f_3 is defined as the proportion of the sum of folds which remained after the folds' removal in the circle of every radius *r* to all the folds.

$$f_3 = \frac{1}{area} \sum_{r=1}^{Rmax} number(r)$$

(4)

where *area* is the number of pixels in all the folds, *Rmax* is the radius in the case when all the folds are removed from the target area, and *number*(*r*) is the number of pixels which represent the remained folds in the target area when the removal is conducted within the circle of the radius *r*. f_3 represents the content of the folds which figure parallel curves to all the folds. In normal cases of the X ray images, since the folds figure parallel curves and crossed curves are hardly perceived, the content in the normal cases could be much than abnormal cases.

Figure 5 shows an extraction result of the intersections, where (a) is the folds, (b) is the thinning for the folds, and (c) is the intersections.
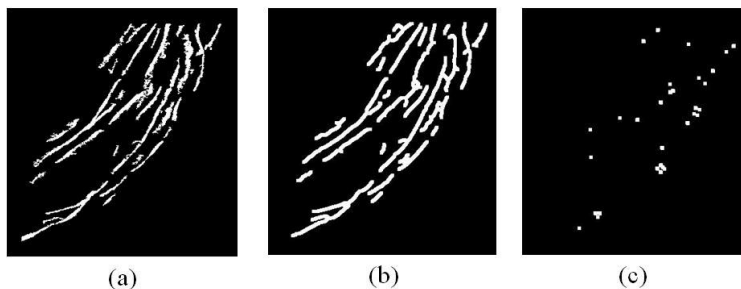

(a)   (b)   (c)

Figure 5: Process for extracting the intersections.

4 Experimental Results

The proposed normalities were extracted from 88 X-ray images (normal case: 75, abnormal case: 13). To examine performance of the proposed normalities, linear discriminant analysis (LDA), linear support vector machine (L-SVM), and nonlinear support vector machine (N-SVM) were applied to the discrimination of the images into normal or abnormal cases by regarding the normalities as variables, respectively. Each of the normalities was normalized to the dataset that the mean value is 0 and the variance is 1. As a kernel function in N-SVM, the Gaussian kernel was used.

The training dataset and test dataset were chosen by the leave-one-out cross validation [12] in all the discriminations. In fact, every discrimination was conducted according to the following procedure:

1.   Choose 1 image from all the images as test data, and use the other images as training dataset.

2.   Discriminate the test data between the class of "normal" and "abnormal".

3.   Changing the test data, repeat from 1 to 2 until every image has been the test data.

4.1   The discrimination of the normal cases including synthetic samples

Performance of machine learning algorithms is typically evaluated using predictive accuracy. However, this is not appropriate when the data is imbalanced and/or the costs of different errors vary markedly. The machine learning community has addressed the issue of class imbalance in two ways: the one is to assign distinct costs to training examples, and another one is to re-sample the original dataset, either by over-sampling the minority class and/or under-sampling the majority class. Here, we apply over-sampling of the minority class (i.e., the class of abnormal cases) using SMOTE [13].

SMOTE is over-sampled by creating "synthetic" examples rather than by over-sampling with replacement. This method creates extra training data by performing certain operations on real data. Synthetic examples are generated in a less application-specific manner by operating in the feature space. The synthetic samples are generated in the following way: Take the difference between the feature vector (sample) under consideration and its nearest neighbor. Multiply this difference by a random number between 0 and 1, and add it to the feature vector under consideration. This causes the selection of a random point along the line segment between two specific features. This approach effectively forces the decision region of the minority class to become more general.

We prepared 75 samples in the normal cases and 75 (i.e. 13 real + 62 synthetic samples) samples in the abnormal cases for the training set used in the leave-one-out cross-validation. The class of abnormal cases was oversampled at 478% of its original size. Table 1 shows discrimination results for normal and abnormal cases respectively obtained by the three discrimination machines of LDA, L-SVM and N-SVM (σ=1.0), where *Precision* and *Recall* represent the discrimination ratio. They are defined as

$$Pr\,ecision = \frac{X_{h \cap c}}{X_c} \times 100 \qquad (5)$$

$$Re\,call = \frac{X_{h \cap c}}{X_h} \times 100 \qquad (6)$$

Where *X* represents the number of images, X_h is the number of the correct answers, and X_c is the number of images discriminated by the proposed method. Figure 6 shows the discrimination ratios

against change of the parameter σ in the Gaussian kernel used in N-SVM. Table 2 shows the discrimination ratios in the case when the discrimination by N-SVM have shown the highest performance (σ = 0.4).

Table 1: Experimental results including synthetic abnormal samples.

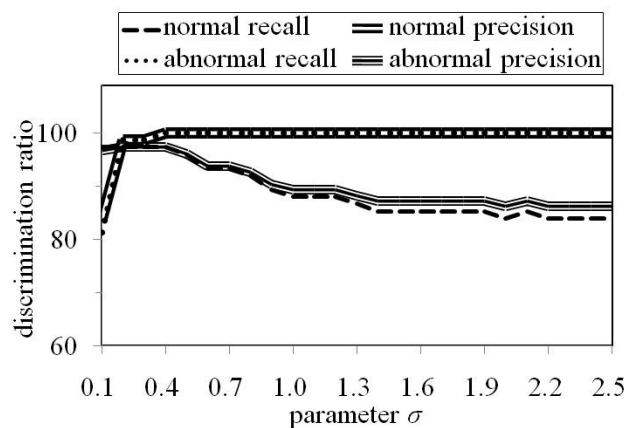| Tool	Normal		Abnormal	
	Recall	*Precision*	*Recall*	*Precision*
LDA	81.33% (61/75)	100% (61/61)	100% (75/75)	84.26% (75/89)
L-SVM	85.33% (64/75)	100% (64/64)	100% (75/75)	87.20% (75/86)
N-SVM (σ =1.0)	88.00% (66/75)	100% (66/66)	100% (75/75)	89.28% (75/84)

Figure 6: The discrimination ratio vs. σ of the Gaussian kernel in N-SVM.

Table 2: Experimental result in the case when σ = 0.4 in N-SVM.

Tool	Normal		Abnormal	
	Recall	*Precision*	*Recall*	*Precision*
N-SVM (σ =0.4)	97.33% (73/75)	100% (61/61)	100% (75/75)	97.40% (75/77)

As shown in Table 1 and Figure 6, the discrimination ratios for each of normal and abnormal cases show more than 80 % and the *Precision* of the normal cases keeps 100 %. Actually, in order to pass the reading test of mammography images in Japan, medical doctors have to correctly read at least 80 % in each of cancer cases and normal cases. Hence, if the experimental results were evaluated from this clinical standpoint, the experimental results would show the proposed normalities are appropriately extracted for the discrimination of normal cases in the gastric X-ray images.

4.2   The discrimination of the normal cases only real samples

Using 75 cases in the normal cases and 13 cases in the abnormal cases (without the over-sampling in 4.1), the discriminations were conducted by the leave-one-out cross-validation as well as 4.1. Table 3 shows discrimination results for normal and abnormal cases respectively obtained by the three discrimination machines of LDA, L-SVM and N-SVM (Gaussian kernel, σ=1.0), where *Precision* and *Recall*

represent the discrimination ratio. Table 4 shows the mean value and the standard deviation of each normality extracted from correct dataset in each case. As shown in Table 4, we can confirm that all the normalities for the normal cases are higher than the abnormal cases.

Table 3: Experimental results using clinical cases.

	Normal		Abnormal	
Tool	*Recall*	*Precision*	*Recall*	*Precision*
LDA	78.66% (59/75)	100% (59/59)	100% (13/13)	44.82% (13/29)
L-SVM	85.33% (64/75)	100% (64/64)	100% (13/13)	51.16% (13/24)
N-SVM (σ=1.0)	86.66% (65/75)	100% (65/65)	100% (13/13)	56.52% (13/23)

Table 4: Results of the extraction of the normalities.

	Normal		Abnormal	
Normality	*Mean*	*SD*	*Mean*	*SD*
f_1	0.109	1.043	-0.632	0.325
f_2	0.178	0.977	-1.027	0.310
f_3	0.206	0.935	-1.193	0.353

Figure 7 shows the result of the extracted folds and intersections for a normal case of discrimination success, where (a) is the target area, (b) is the folds, and (c) is the intersections. Similarly, Figure 8 shows the result for an abnormal case of discrimination success.
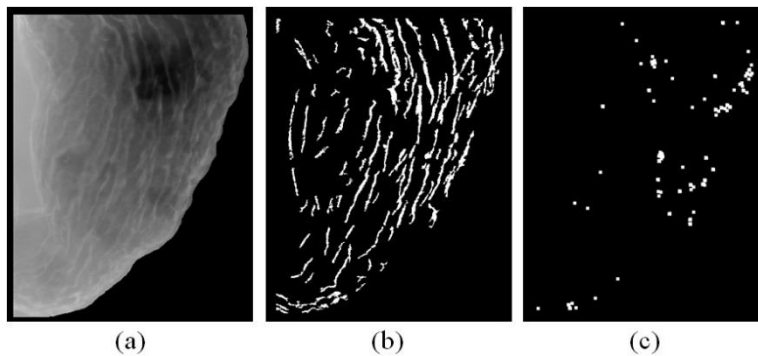

(a) (b) (c)

Figure 7: A result of the extracted folds and intersections for a normal case (discrimination success).
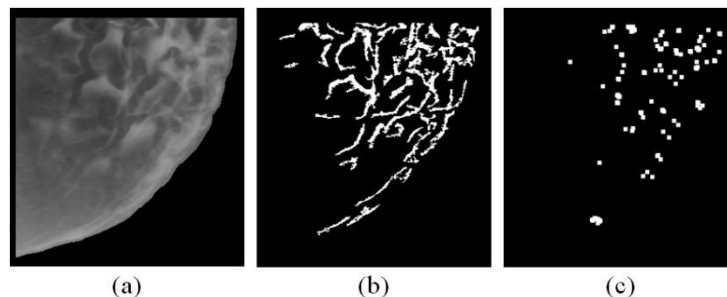

(a) (b) (c)

Figure 8: A result of the extracted folds and intersections for an abnormal case (discrimination success).

5 Discussions

In the experimental results, 8 normal cases were diagnosed as abnormal cases in all the discrimination machines. They can be broadly divided into two kinds. Figure 9 and Figure 10 show a case of each kind, where (a), (b) and (c) are the same as Figure 7, respectively.
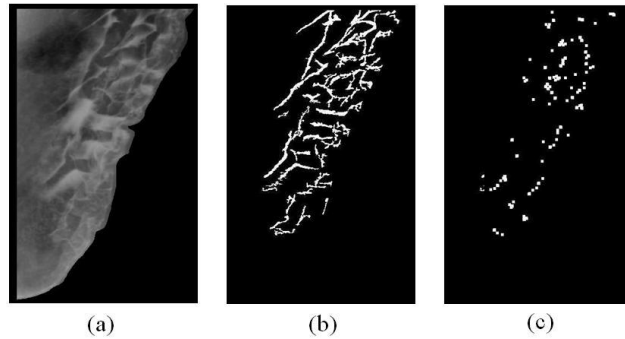

<div style="text-align:center">(a)   (b)   (c)</div>

Figure 9: Failure case 1 (by the proposal: abnormal, correct answer: normal)

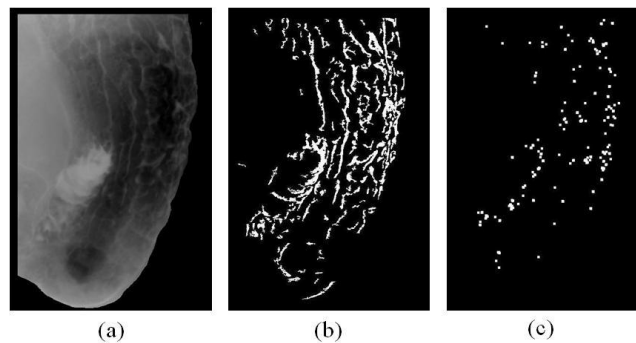<div style="text-align:center">(a)   (b)   (c)</div>

Figure 10: Failure case 2 (by the proposal: abnormal, correct answer: normal)

In the case of Figure 9, the folds could look irregular though this case is a normal case. Besides, the normalities for this case are f_1 = -0.397, f_2 = -0.659 and f_3 = -1.290, i.e., the normalities are closer to the mean value of them in abnormal case shown in Table 3. The reason why the discrimination result of this case is opposite to the diagnosis by a medical doctor is because medical doctors imagines the stomach in the images as a 3D stomach based on their past experience, i.e., they can imagine the depth of the stomach from the X-ray image. On the other hand, in the proposed method, the normalities cannot be extracted by considering the depth of the stomach from the X-ray images. Since the stomach repeats the expansion and contraction contracted movement, there are some cases of stomachs which show distorted wall or have thicker pleats on the wall (i.e., shadow of the folds is thicker and stronger than other cases) in the X-ray images. However, the normalities are not precisely extracted from such cases because all the stomachs are not recognized as a 3D object. In the future, it is necessary to improve the normalities or propose another normality considering the depth between folds in the target area. The other 4 normal cases were discriminated incorrectly for the same reason.

In the case of Figure 10, we can confirm the leakage of barium in the intestine. Since the proposed method cannot distinguish between the folds and the noises, the proposed method cannot extract the features precisely such cases. In the image shown in Figure 10, f_1, f_2 and f_3 are -0.914, -1.086 and -0.609, respectively, and they are closer to the mean values of them in abnormal case than normal case shown

in Table 3. The other 2 normal cases were discriminated incorrectly for the same reason. In the future, it is necessary to consider recognizing the leakage in the target area.

6 Conclusions

Aiming at decreasing the number of gastric X-ray pictures which diagnosticians have to read at a time, this paper has presented features for discriminating normal cases from gastric abnormal cases. The characteristic of the features is to measure the normality of fold patterns in double contrast X-ray images of stomach. The features were used as variables of discrimination machines to identify normal cases. Experimental results by the discrimination trials have shown that discrimination ratios of normal cases by the proposed method have become more than 80% in all the trials.

As future works, it would be necessary to (1) extract the target area automatically, (2) design normalities for the diagnosis considering the depth of stomach in the X-ray images, and (3) remove region of the leakage of barium in the target area.

REFERENCES

[1]. Y. Kita (1996), 'Elastic-model driven analysis of several views of a deformable cylindrical object', IEEE Trans. Pattern Anal. Mach. Intel., 18(12), 1150-1162.

[2]. Y. Mekada, J. Hasegawa, J. Toriwaki, S. Nawano, and K. Miyagawa (1998), 'Automated extraction of cancer lesions from double contrast X-ray images of stomach', Proc. 1st International Workshop on Computer Aided Diagnosis, 407-412.

[3]. J. Hasegawa, T. Tsutsui, and J. Toriwaki (1991), 'Automated Extraction of Cancer Lesions with Convergent Fold Patterns in Double Contrast X-ray Images of the Stomach', Systems and Computers in Japan, 22(7), 51-62.

[4]. J. Hasegawa and J. Toriwaki (1992), 'A new filter for feature extraction of line pattern texture with application to cancer detection', Proc. 11th IAPR Int. Conf. on Pattern Recognition, 352-355.

[5]. Y. Yoshinaga, H. Kobatake, and S. Fukushima (1999), 'The detection and feature extraction method of curvilinear convex regions with weak contrast using a gradient distribution method', Proc. IEEE ICIP 99, 715-719.

[6]. K. Abe, T. Nobuoka, and M. Minami (2011), 'Computer-Aided Diagnosis of Mass Screenings for Gastric Cancer Using Double Contrast X-ray Images', Proc. IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing, 708-713.

[7]. M. Kass, A. Witkin, and D. Terzopoulos (1988), 'Snakes: Active contour models', Int. J. Computer Vision, 1(3), 321-331.

[8].   S. Fukushima, H. Uwai, and K. Yoshimoto (2000), 'Optimization-Based Recognition the Gastric Region from a Double-Contrast Radiogram', IEICE Trans. (Japanese Edition), J83-D-II (1), 154-164.

[9].   R. A. Schowengerdt (1983), 'Techniques for Image Processing and Classfication in Remote Sensing', ACADEMIC PRESS, 72-83.

[10].   M. J. Canty (2014), 'Image Analysis, Classification and Change Detection in Remote Sensing', CRC Press, 329-332.

[11].   Y. Yoshinaga and H. Kobatake (2000), 'The line detection method with robustness against contrast and width variation applied in gradient vector field', Systems and Computers in Japan, 31(3), 49-58.

[12].   M. Stone (1974), 'Cross-Validatory Choice and Assessment of Statistical Predictions', J. of the Royal Statistical Society, Series B (Methodological), 36(2), 111-147.

[13].   NV. Chawla, KW. Bowyer, LO. Hall, and WP. Kegelmeyear (2002), 'SMOTE: Synthetic Minority Over-sampling Technique', J. of Artificial Intelligence Research 16, 321-357.