

Received March 3, 2021, accepted March 22, 2021, date of publication March 30, 2021, date of current version April 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3069747

A Novel Approach for Increased Convolutional Neural Network Performance in Gastric-Cancer Classification Using Endoscopic Images

SIN-AE LEE¹, HYUN CHIN CHO², AND HYUN-CHONG CHO^{1,3}

¹Interdisciplinary Graduate Program in BIT Medical Convergence, Kangwon National University, Chuncheon 24341, Republic of Korea

²Department of Internal Medicine, Institute of Health Sciences, Gyeongsang National University School of Medicine and Gyeongsang National University Hospital, Jinju 52727, Republic of Korea

³Department of Electronics Engineering, Kangwon National University, Chuncheon 24341, Republic of Korea

Corresponding authors: Hyun Chin Cho (hccholuck@gmail.com) and Hyun-chong Cho (hyuncho@kangwon.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2017R1E1A1A03070297). This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2018-0-01433) supervised by the IITP (Institute for Information and communications Technology Promotion).

ABSTRACT Gastric cancer is the third-most-common cause of cancer-related deaths in the world. Fortunately, it can be detected using endoscopy equipment. Computer-aided diagnosis (CADx) systems can help clinicians identify cancer from gastric diseases more accurately. In this paper, we present a CADx system that distinguishes and classifies gastric cancer from pre-cancerous conditions, such as gastric polyps, gastric ulcers, gastritis, and bleeding. The system uses a deep-learning model, Xception, which involves depth-wise separable convolutions, to classify cancer and non-cancers. The proposed method consists of two steps: Google's AutoAugment for augmentation and the simple linear iterative clustering (SLIC) superpixel and fast and robust fuzzy C-means (FRFCM) algorithm for image segmentation during preprocessing. These approaches produce a feasible method of distinguishing and classifying cancers from other gastric diseases. Based on biopsy-supported ground truth, the performance metrics of the area under the receiver operating characteristic curve (i.e. Az) are measured on the test sets. Based on the classification results, the Az of the proposed classification model is 0.96, which is 0.06 up from 0.90 which is the Az of the original data. Our methods are fully automated without the manual specification of region-of-interests for the test and with a random selection of images for model training. This methodology may play a crucial role in selecting effective treatment options without the need for a surgical biopsy.

INDEX TERMS Augmentation, computer-aided diagnosis (CADx), deep learning, gastric cancer, segmentation.

I. INTRODUCTION

According to the worldwide gastric-cancer incidence statistics released in 2018 by the International Agency for Research on Cancer, the world's leading cancer research institute, gastric cancer is the fifth most frequently diagnosed cancer and the third leading cause of cancer death. Fig. 1 shows the world gastric cancer incidence rate for 2018 [1]. As can be seen, eastern Asia and Eastern Europe have the highest incidences.

Ailments, such as gastritis, gastric ulcers, and gastric bleeding have been identified as precancerous lesions that

lead to gastric cancer. Because these gastric lesions are asymptomatic until they progress to cancer, early diagnosis and treatment are the best ways to reduce the incidence, and regular endoscopy is the preferred method for an accurate diagnosis. Then, the lesions can be treated and/or removed [2], [3].

Because the number of images taken by endoscopy equipment is ever-increasing, and the quality of the images is constantly improving, any doctor who relies on looking directly at images to make diagnosis experiences mounting fatigue. Furthermore, different doctors use different mental heuristics based on experience. Thus, owing to subjective biases and occasional lassitude, misdiagnoses sometimes

The associate editor coordinating the review of this manuscript and approving it for publication was Pasquale De Meo.

WORLD GASTRIC CANCER INCIDENCE RATE

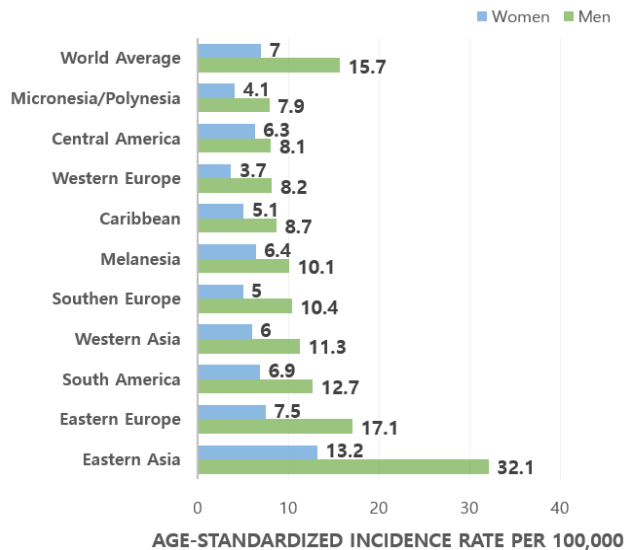


FIGURE 1. World gastric cancer incidence rate in 2018 (Source: GLOBOCAN 2018).

occur. Computer-aided diagnosis (CADx) systems provide doctors with objective information, which tends to increase the accuracy of diagnosis [4], [5].

Currently, various CADx systems are used for the study of gastroscopy. Kanesaka *et al.* [6] used a support vector machine to build a CADx system to facilitate early gastric cancer detection using narrow-band endoscopy images. The study achieves an accuracy of 96.3% for cancer and non-cancer classifications. Khryashchev *et al.* [7] used an endoscopy image-analysis algorithm based on the application of a convolutional neural network single-shot multi-box detector [8]. They classified gastroscopy images into three classes: normal mucosa, non-cancerous pathology, and cancer. The mAP value for detection was 0.875. Li *et al.* [9] developed a model that distinguished early gastric cancers and non-cancerous lesion images based on the inception-v3 network. They compared the performance of two experts with two non-experts and a CNN, showing that the CNN accuracy performed excellently, but the specificity and positive predictive value (PPV) were lower than that of experts. Cho *et al.* [10] collected endoscopic white-light images of pathologically confirmed gastric lesions and classified them into five categories: advanced gastric cancer, early gastric cancer, high-grade dysplasia, low-grade dysplasia, and non-neoplasm. They pre-trained three CNN models using a training dataset. These studies in [6], [7], [9] used the narrow-band image (NBI) instead of the white light endoscopy (WLE) image. The WLE image is a common filming mode for endoscopy and more useful.

CADx systems that use deep learning have developed significantly. However, they typically rely on large labeled datasets and significant expert knowledge and incur a high computational cost. Data collection is time-consuming and costly because patients require privacy and researchers require institutional review board (IRB) approval [11].

A data augmentation method is a representative solution to the data collection problem. Augmentation improves performance while solving the problem of deep-learning overfitting [11], [12]. To augment data, such methods transform existing information into colors and shapes.

Kim *et al.* [13] studied CADx systems and applied two segmentation methods to gastroscopy data. They trained a deep-learning model using the inception-V3 network and classified it as normal and abnormal. They also proposed an augmented CADx system that applied 90°, 180°, and 270° rotations of the same data [14]. Chowdhary *et al.* [15] propose a segmentation model for the detection of cancer in an image. They segment the image using an intuitionistic possibilistic fuzzy c-mean (IPFCM), which combines the intuitive fuzzy c-mean (IFCM) and the possible fuzzy c-mean (PFCM) algorithm.

Ergashev and Cho [16] randomly applied a data augmentation method in the range of 0.9 to 1.1 for brightness and color contrast of the images. Zhu *et al.* [17] performed data augmentation, in which each image was rotated and flipped to expand the amount of data by eight-fold. Asperti and Mastronardo [18] used methods such as rotation, width and height shifting, shear, and zoom in their data augmentation applied randomly within a set parameter range. Frid-Adar *et al.* [19] obtained augmented samples by randomly stretching images in the dataset horizontally or vertically, which became the input of an adapted deep CNN for image recognition. Zhao *et al.* [20] presented an automated data augmentation method for synthesizing labeled medical images. They are trained independent spatial and appearance transform models to generalize a dataset of labeled images based on VoxelMorph, which learns to output a smooth displacement vector field that registers one image to another by jointly optimizing the image similarity loss and the displacement field smoothness term [21]. Their methods essentially generate a homeomorphic mapping domain with CNNs to change the probability distribution of datasets.

The works in [14], [16]–[20] show that the results of applying data augmentation is better. However, the method of augmentation was different for each work. In [14], [17], the augmentation method was chosen by them, and the parameters were also manually adjusted. The studies in [16], [18], [19] set the augmentation method themselves, and the parameter values were randomly applied. These methods can augment numerous images, but results can be biased or lack objectivity.

The purpose of this study is to investigate fully automated methods for classifying abnormal gastric endoscopic images into cancer and non-cancerous lesions using a deep-CNN scheme (i.e., Xception). Two approaches, a SLIC superpixel and FRFCM algorithm for the segmentation and Google's AutoAugment for data augmentation, are developed and applied. The AutoAugment method derives parameters for optimal augmentation policies via reinforcement learning using typical datasets, such as CIFAR-10, which Google has recently established as a data augmentation method [22].

TABLE 1. Types and number of gastroscopy image data.

TYPE		NUMBER
GASTRIC CANCER		63
NON-CANCER	Ulcer	95
	SMT	72
	Gastritis	194
	Bleeding	5
	Polyp	11
	Others	30
TOTAL		470

The major contributions of this work are highlighted as follows:

- A new method for gastric lesion recognition based on a deep learning network that can successfully classify the cancer from pre-cancerous endoscopic images is presented.
- We also propose a novel method that combines image segmentation and augmentation algorithms.
- Our method is fully automated without manual specification of the region of interest for testing and random selection of images for model training.
- Experimental results demonstrate that the efficiency and effectiveness of the proposed task are superior to the basic deep learning model.

The remainder of this article is organized as follows: Section II provides a detailed description of the gastroscopic imaging data set used in the experiment and a proposed methodology. The experimental results are reported in Section III. Section IV describes the discussion of these findings. Finally, there are conclusions in Section V.

II. MATERIALS AND METHODS

A. DATABASE

A dataset was collected with IRB approval from the files of patients who had undergone gastric endoscopic imaging at the Department of Internal Medicine at the Gyeongsang National University Hospital, South Korea. All endoscopic image data were selected by internists and verified via medical examination and biopsy. A total of 470 endoscopic image data were collected from 69 patients. There are 250 images used for training and 220 for testing. The training set included 34 cancer and 216 non-cancer images, and the test set included 29 cancer and 191 non-cancer images. The types of lesions listed in Table 1 include ulcers, gastric cancers, submucosal tumor (SMT), and polyps. The “others” in Table 1 include blood clots, hematin, gastric xanthoma, etc. Example endoscopic images are shown in Fig. 2.

B. DATA AUGMENTATION

It is a typical problem that there are not enough training data to overcome the overfitting of numerous neural network parameters in deep learning-based applications (viz medical image analysis). A typical approach to handling

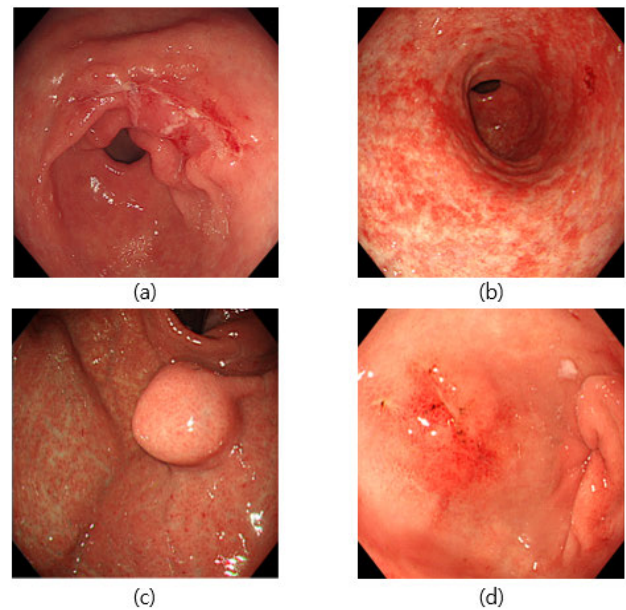


FIGURE 2. Abnormal images of gastric endoscopy: (a) Gastric cancer; (b) Gastritis; (c) Gastric submucosal tumor (SMT); (d) Benign gastric ulcer.

this problem is artificially augmenting datasets using label-preserved transformations, such as random image translation, horizontal/vertical flipping, shearing, rotation, and cropping [18]. A random combination of these transformations can be used to gain a manifold increase in the number of training data, which is useful for minimizing the parameter-overfitting problem. Recently, the Google Brain team released the AutoAugment tool, which is used to search for improved data augmentation policies [22]. The algorithmic process is shown in Fig. 3. A controller recurrent neural network (RNN) samples a policy S , to which the images are to be appended. Using a child network, various augment policies are applied to the dataset. It is an algorithm that obtains performance accuracy R and updates it in the controller RNN to find the best augment policy. By applying the optimized data augmentation policies, high accuracy uses of public data (e.g., CIFAR-10, CIFAR-100, SVHN, and ImageNet) have been achieved. In our experiments, a variant of the policy learned from CIFAR-10 data is used. 25 sub-policies are thus used to expand the training dataset by 25-fold.

A total of 25 augment sub-policies are presented with one sub-policy to progress two ordered operations. The operational techniques applied to the augment policy are defined as Shear X/Y, Translate X/Y, Rotate, AutoContrast, Invert, Equalize, Solarize, Posterize, Contrast, Color, Rightness, Sharpness, Cutout, and Sample Pairing. Two parameters provide the probability values that indicate the likelihood that the operating technique will be applied to the policy so that the size of the operation techniques will be applied to regulate the operation. Table 2 shows the CIFAR-10 augmentation policy we used. The first sub-policy specifies a sequential application of Invert followed by Contrast. The probability of applying Invert is 0.1. The Invert operation does not use magnitude information. Then, we apply a Contrast of 0.2.

TABLE 2. CIFAR-10 policy of augmentation [23].

POLICY	OPERATION 1	OPERATION 2
SUB-POLICY 0	(Invert,0.1,7)	(Contrast,0.2,6)
SUB-POLICY 1	(Rotate,0.7,2)	(TranslateX,0.3,9)
SUB-POLICY 2	(Sharpness,0.8,1)	(Sharpness,0.9,3)
SUB-POLICY 3	(ShearY,0.5,8)	(TranslateX,0.7,9)
SUB-POLICY 4	(AutoContrast,0.5,8)	(Equalize,0.9,2)
SUB-POLICY 5	(ShearY,0.2,7)	(Posterize,0.3,7)
SUB-POLICY 6	(Color,0.4,3)	(Brightness,0.6,7)
SUB-POLICY 7	(Sharpness,0.3,9)	(Brightness,0.7,9)
SUB-POLICY 8	(Equalize,0.6,5)	(Equalize,0.5,1)
SUB-POLICY 9	(Contrast,0.6,7)	(Sharpness,0.6,5)
SUB-POLICY 10	(Color,0.7,7)	(TranslateX,0.5,8)
SUB-POLICY 11	(Equalize,0.3,7)	(AutoContrast,0.4,8)
SUB-POLICY 12	(TranslateY,0.4,3)	(Sharpness,0.2,6)
SUB-POLICY 13	(Brightness,0.9,6)	(Color,0.2,8)
SUB-POLICY 14	(Solarize,0.5,2)	(Invert,0.0,3)
SUB-POLICY 15	(Equalize,0.2,0)	(AutoContrast,0.6,0)
SUB-POLICY 16	(Equalize,0.2,8)	(Equalize,0.6,4)
SUB-POLICY 17	(Color,0.9,9)	(Equalize,0.6,6)
SUB-POLICY 18	(AutoContrast,0.8,4)	(Solarize, 0.2,8)
SUB-POLICY 19	(Brightness,0.1,3)	(Color,0.7,0)
SUB-POLICY 20	(Solarize,0.4,5)	(AutoContrast,0.9,3)
SUB-POLICY 21	(TranslateY,0.9,9)	(TranslateY,0.7,9)
SUB-POLICY 22	(AutoContrast,0.9,2)	(Solarize,0.8,3)
SUB-POLICY 23	(Equalize,0.8,8)	(Invert,0.1,3)
SUB-POLICY 24	(TranslateY,0.7,9)	(AutoContrast,0.9,1)

When applied, it has a magnitude of 6 out of 10. These operations are applied in a specified order. There are 2.9×10^{32} augmentation sub-policies [22], which are randomly selected and applied to learn the training data. We repeat the learning and classification to find the best policy having improved performance.

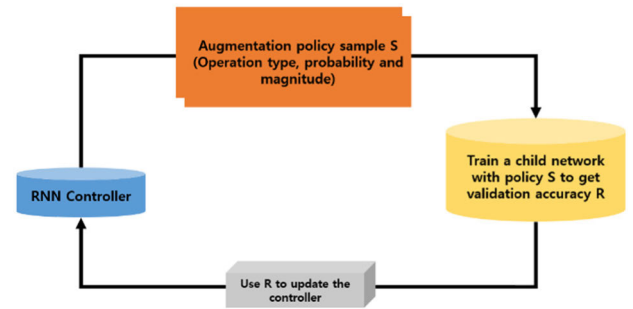
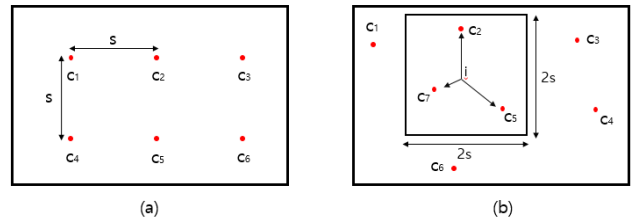
C. SEGMENTATION METHOD: SLIC SUPERPIXEL

The SLIC superpixel is a considerably fast clustering method [24]. The reason for its fast calculation is that clustering is performed using only cluster information within a certain area without requiring the entire image's cluster information. Thus, the pixels in the image are clustered into k superpixels using the same procedure as that of Fig. 4.

When the user inputs k clusters, the SLIC superpixel uniformly arranges the positions of the cluster center coordinates in the image having interval S , as shown in Fig. 4(a). Here, S is calculated using (1), where N is the number of pixels in the image.

$$S = \sqrt{N/k} \quad (1)$$

The advantage of a superpixel algorithm is that it maintains cluster edge information. If the location of the cluster center coordinates is set at regular intervals, the center coordinates overlap with the edge, which can cause the loss of edge information. To solve this problem, we convert the color

**FIGURE 3.** Process of AutoAugment.**FIGURE 4.** SLIC superpixel clustering (a) Interval and setting between clusters, (b) Clustering range.

space of the image to the CIELAB color space, obtain the gradient image using a 1D gradient kernel, $[-1, 0, +1]$ and $[-1, 0, +1]^T$, and move the central coordinates to the lowest of the 3×3 pixel nearest the cluster center coordinate.

When the central coordinate arrangement is complete, we calculate the d_c and d_s of all pixels ((2) and (3)).

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2}. \quad (2)$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}. \quad (3)$$

Here, l_j, a_j, b_j represents the CIELAB color information of the j th cluster center coordinate. l_i, a_i, b_i represents the CIELAB color information of the i th image pixel. x_j, y_j is center coordinate of the j th cluster. x_i, y_i is i th pixel coordinate. d_c is the CIELAB color distance between cluster c_j and pixel i within the $2S$ range, as shown in Fig. 4(b). d_s is the distance between cluster c_j and pixel i within the $2S$ range. If d_c and d_s are obtained, the total distance, D , can be calculated, as with (4).

$$D = \sqrt{(d_c)^2 + \left(\frac{d_s}{S}\right)^2 m^2}. \quad (4)$$

Here, m is the weight parameter. Pixel i is assigned to the cluster having the smallest calculated D value among clusters in the range. After all cluster assignments are completed for all pixels, i , we update the l_j, a_j, b_j, x_j , and y_j information of all clusters to optimize clustering. The l_j, a_j, b_j of the j th cluster is updated to the mean value of l_i, a_i, b_i of the pixel, i , assigned to the j -cluster, and x_j, y_j is changed to the center of gravity of the pixel, i , assigned to the j th cluster. Repeat until the x_j and y_j of all clusters are not changed [24].

D. SEGMENTATION METHOD: FRFCM

The fuzzy c-mean (FCM) algorithm with spatial constraint (FCM_S) is an effective algorithm for image segmentation.

However, this algorithm cannot be robust, especially if it does not have enough knowledge of noise. The fast and robust fuzzy C-means (FRFCM) algorithm is more advanced than other FCM algorithms. Lei *et al.* [25] proposed a new, fast and robust FCM frame work for image segmentation by integrating local spatial and gray information. The objective function of FRFCM is defined as (5).

$$J_m = \sum_{l=1}^q \sum_{k=1}^c \gamma_l u_{kl}^m \|\xi_l - v_k\|^2 \quad (5)$$

The parameter u_{kl} represents the fuzzy membership gray value l with respect to cluster k , and ξ_l is the gray levels. Utilizing the Lagrange multiplier technique, the optimization problem can be converted to an unconstrained optimization problem as (6).

$$\tilde{J}_m = \sum_{l=1}^q \sum_{k=1}^c \gamma_l u_{kl}^m \|\xi_l - v_k\|^2 - \lambda \left(\sum_{k=1}^c u_{kl} - 1 \right) \quad (6)$$

The parameter λ is a Lagrange multiplier. Therefore, the minimization problem is converted to finding the saddle point and taking the derivation of the Lagrange \tilde{J}_m with respect to the parameter, *i.e.*, u_{kl} and v_k . By minimizing (5), we obtained the corresponding solution as follows.

$$u_{kl} = \frac{\|\xi_l - v_k\|^{-2/(m-1)}}{\sum_{j=1}^c \|\xi_l - v_j\|^{-2/(m-1)}} \quad (7)$$

$$v_k = \frac{\sum_{i=1}^q \gamma_l u_{kl}^m \xi_l}{\sum_{i=1}^q \gamma_l u_{kl}^m} \quad (8)$$

We can get a membership partition matrix $U = [u_{kl}]^{c \times q}$ using (7). To obtain a stable U , repeat (7) and (8) until $\max \{U^{(t)} - U^{(t+1)}\} < \varepsilon$ is satisfied, where ε is an exceedingly small error threshold. A new membership partition matrix, $U' = [u_{kl}]^{c \times N}$, that corresponds to the original image can be obtained because $u_{kl}^{(t)}$ is a fuzzy membership of gray value l with respect to cluster k . N is the total number of pixels in the image.

To speed up the algorithm's convergence and obtain a better membership partition matrix, they use the median filter as (9).

$$U'' = \text{med} \{U'\} \quad (9)$$

E. DEEP LEARNING MODEL: XCEPTION NETWORK

Xception is a CNN model that was released in 2016. The network uses the Inception module to reduce connectivity between nodes, further separating finding relationships between each channel from searching local information. Fig. 5 shows an extreme version of the inception module that separately operates 1×1 and 3×3 convolution operations of all channels of the resultant feature map. Thus, this module calculates the feature map per channel. Xception uses the depth-wise separable convolution created by modifying this operation. The depth-wise separable convolution is used to perform a convolution operation on each channel and to apply the 1×1 convolution operation to the result. As shown

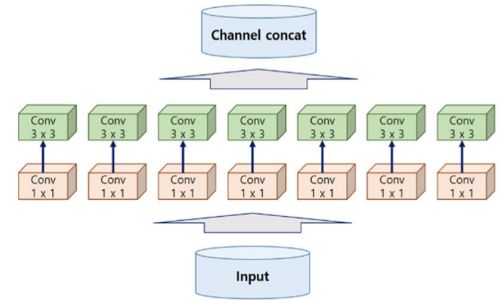


FIGURE 5. Extreme version of inception module.

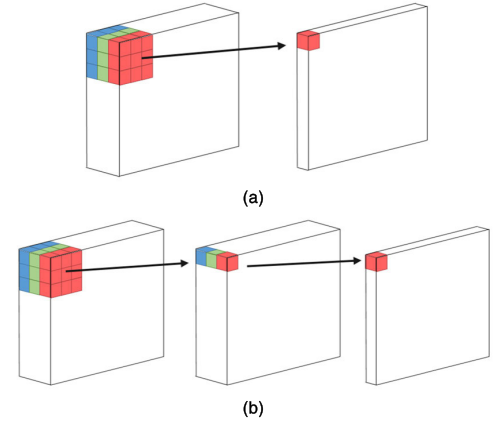


FIGURE 6. Comparison of convolutions (a) Standard convolution, (b) Depth-wise separable convolution.

in Fig. 6(a), if the conventional convolutions create a feature map that takes all channels and local information, the depth-wise separable convolution, shown in Fig. 6(b), creates one feature map for each channel and performs a 1×1 convolution operation to adjust the number of feature maps that are output [26]. The 1×1 convolution operation is a point-wise convolution. The difference from the extreme version of the inception module is the sequence of operations and the presence or absence of an intermediate nonlinear activation function.

In a previous study [27], we found that the results of Xception in the gastric medical image classification were the best among four different deep learning models, Xception, Inception-V3, Resnet-101, and Inception-Resnet-V2. It also found out that the CIFAR-10 augmentation policy is the best for the classification among ImageNet, SVHN and CIFAR-10. Based on the result [27], we selected Xception network for this study.

F. EVALUATION METHODS

As a pre-processing step, images are augmented and segmented using AutoAugment and the SLIC superpixel, respectively. First, each image is augmented into 25 new images using the CIFAR-10 AutoAugment policy. Second, each of these 25 images is segmented into nine areas using the SLIC superpixel and FRFCM algorithm. Ground-truth information of the segmented area is provided by the internist. Fig. 7 shows the process of image augmentation and segmentation.

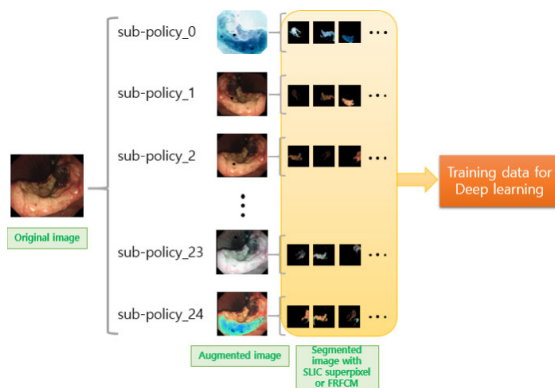


FIGURE 7. Image augmentation and segmentation process.

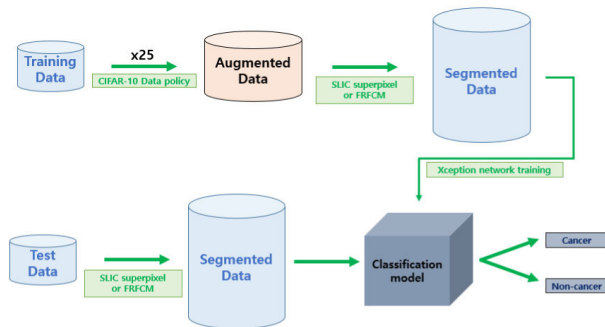


FIGURE 8. Flowchart of the entire procedure.

During the test step, we use the score to represent the results of classification. If the classification model classifies more than 1/3 of the segmented areas as cancer, then it will classify the entire image as cancer. The size of gastric lesions can vary significantly and can include all the segmented regions or only half of them. Others can contain just one region. Therefore, classification is conducted using (10), and the classification threshold value is set through experimentation:

$$\frac{\text{Number of cancerous segmented areas}}{\text{Total number of segmented areas}} \geq \frac{1}{3}. \quad (10)$$

Fig. 8 presents a flowchart of the proposed system. First, the training data are augmented and segmented to generate the classification model. Then, the test data are applied to the segmentation to determine whether the image is cancer or non-cancer.

III. RESULTS

In the field of medical image analysis, a small amount of training data is always a restrictive concern. Lesions of various sizes can be found in endoscopic images. Thus, a model lacking augmentation and segmentation is insufficient for distinguishing and classifying gastric cancer from other gastric diseases. Thus, we proposed using the AutoAugment and two segmentation algorithms to create a CADx system for gastric-lesion diagnosis, as shown in Fig. 8. In models with augmentation, a total of 25 augmented sub-policies are applied, which progresses two tasks in order. In models with segmentation, we used SLIC superpixel and FRFCM to segment images

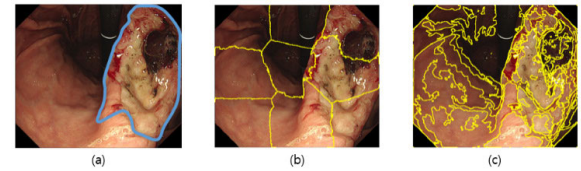


FIGURE 9. Segmented gastric cancer image. (a) Ground truth (b) SLIC superpixel segmentation (c) FRFCM segmentation.

and training was performed based on the internist's ground truth for segmented areas. The number of segmentation areas per image was set to nine. After the segmentation, nine images are generated from one image, and each image has one segmented region which takes the pixel value of the original image and the other region has a pixel value of zero. Fig. 9 shows the segmentation results according to two different segmentation methods. Fig. 9 (a) is the original endoscopy image, and the blue line indicates the area of the lesion as the ground truth as defined by an internist. Fig. 9 (b) and (c) are images with SLIC superpixel and FRFCM segmentation applied, respectively. Clearly, the results of the two segmentation methods are different. The SLIC superpixel has restrictions on clustering space at the segmentation, but FRFCMs have no area limit; therefore, the clustering patches can have very long and wide shapes and may be separated altogether and far away.

The performance comparisons of the application of both algorithms are shown in Fig. 10 and 11. It shows the total results of gastric-lesion classification using original data with augmentation only, segmentation only, and both augmentation and segmentation. The area under the receiver operating characteristic (ROC) curve was 0.90 for the original data, 0.92 with augmentation only, 0.94 with SLIC superpixel segmentation only, 0.92 with FRFCM segmentation only, 0.96 with augmentation and SLIC superpixel segmentation, and 0.94 with augmentation and FRFCM segmentation. As shown by the ROC curves, the proposed algorithm performed the best.

The proposed methods were executed on a 14-core 2.20 GHZ Xeon CPU and 128 GB memory using the CentOS 7.7 Linux operating system. Two NVIDIA Tesla T4 GPUs with 16 GB device memory were used. The network architecture was implemented on MATLAB 2019a using the deep-learning toolbox library. For the original data, the program took ~1 h to train 100 epochs using the training dataset. For augmentation, the program took ~1 d to train 100 epochs using the training dataset. For segmentation, the program took ~3 h to train. For augmentation and segmentation, the program took ~10 d to train 47,000 iterations. The mini-batch size was 60, and the initial learning rate was 0.001.

IV. DISCUSSIONS

The results show that the models trained using augmentation and/or segmentation performed better than the one trained without them. This provides a solution for one of the biggest problems of deep learning in medical image analysis, which relies heavily on large amounts of data. In Fig. 10 and 11,

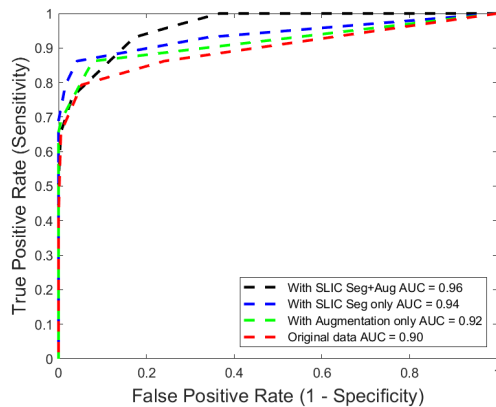


FIGURE 10. The results of ROC curves for the original image with augmentation only, with SLIC superpixel segmentation only, and with augmentation and SLIC superpixel segmentation (the proposed method).

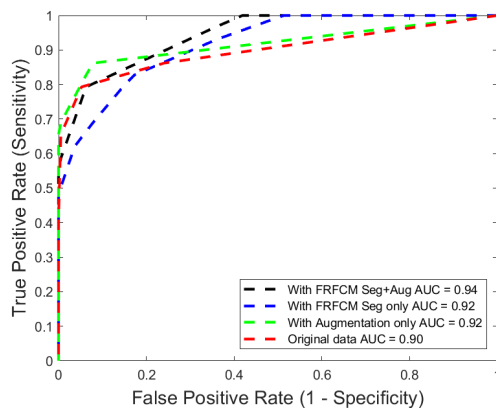


FIGURE 11. The results of ROC curves for the original image with augmentation only, with FRFCM segmentation only, and with augmentation and FRFCM segmentation (the proposed method).

the Az values of the two segmentation algorithms were higher or equal to those of augmentation. One possible reason is that it was more effective to transfer information from the lesion location to the network via segmentation. Fig. 10 and 11 show that the results of the SLIC superpixel are higher than those of the FRFCM. In this study, SLIC superpixel is clearly the best overall performer. It is the most memory-saving and the fastest method, segmenting a 471×407 image in 0.29 s, while FRFCM takes 11.32 s. It shows outstanding boundary adherence, while the segmented shape of the FRFCM method spreads widely in the image. The segmented shape of the SLIC superpixel is more focused on lesions than that of FRFCM. As shown in Fig. 9, SLIC superpixel with restrictions on the clustering area can select about four patches to segregate images with lesions, but at least five patches must be selected in FRFCM. In the gastroscopy image, the SLIC superpixel method was more suitable for detecting cancers because they spread widely from one place rather than from many places. This could be one of the reasons why SLIC superpixel turned out better than FRFCM. SLIC superpixel is easy to use, its only parameter being the number of necessary superpixels.

The classification results obtained after applying segmentation methods are better than the use of only the original image data, demonstrating that learning through area labeling

is effective. It also shows the potential for later detection studies that can provide lesion region information using the probabilities of each patch.

We have proposed a performance improvement in computer-aided diagnosis (CADx) systems for white light gastric endoscopic images instead of narrow-band imaging (NBI) magnification. NBI is an optical technology newly developed and applied in the area of gastrointestinal endoscopy. This technique highlighted to blood vessels using the wavelength, which can be easily absorbed by hemoglobin for maximum contrast. While adding NBI images may improve the performance of the proposed system, they are still used as auxiliary tools in the hospital. We plan to study comparative models using NBI images with different deep CNN networks.

Google's AutoAugment provided the best data augmentation policy for a given dataset using reinforcement learning. A subset of policies achieved by the AutoAugment, which were excellent for the CIFAR-10 dataset, was applied to improve the performance of our gastric cancer classification task. However, several powerful computational resources were required to produce optimal data augmentation on this training dataset. Thus, a fast and flexible data augmentation solution based on Bayesian optimization instead of reinforcement learning should be researched in the future. In addition to methods of finding and applying optimal augmentation policies, a general adversarial network, which is a model architecture for training a generative model, is under investigation. We also plan to study comparative models using other policies of AutoAugment with different deep CNN networks.

V. CONCLUSION

Deep CNNs are powerful and useful algorithms that commonly work well when trained using a large amount of data. The shortage of labeled training data in gastric cancer research is a major problem that limits the application of CNNs in the field of medical image analysis. In this paper, we proposed the CADx system, which can distinguish and classify gastric cancers from other gastric diseases. The system used the Xception network, which involves depth-wise separable convolutions. The proposed method used two algorithms: Google's AutoAugment for augmentation and the SLIC superpixel and FRFCM algorithms for image segmentation. As a result of analyzing the classification performance of the system, we found that augmentation and segmentation enabled the classification model to perform better because the applied data augmentation policy prevented overfitting. Additionally, the segmented images expertly identified as cancer or non-cancer by the internist facilitated our focus on the features of the lesion during training. In this study, we achieved a ROC curve of 0.96, which is 0.06 higher than that achieved when the two methods were not applied, indicating that it is meaningful to train by dividing the area of an augmented image. This methodology may play a vital role in selecting effective treatment options and survival predictions without the need for a surgical biopsy.

Future work includes verifying the results of this study by applying the CADx system to a larger and independent data set, expanding a deep-learning model, comparing other segmentation and augmentation methods. The major question of the impact of the number of segmentations and the classification threshold value in segmented areas as cancer will also need to be addressed in future studies.

REFERENCES

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA A, Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018.
- [2] K.-P. Ko, "Epidemiology of gastric cancer in Korea," *J. Korean Med. Assoc.*, vol. 62, pp. 398–406, Aug. 2019.
- [3] I. J. Choi, "Helicobacter pylori eradication therapy and gastric cancer prevention," *Korean J. Gastroenterol.*, vol. 72, pp. 245–251, Nov. 2018.
- [4] A. Misumi, K. Misumi, A. Murakami, K. Harada, U. Honmyo, and M. Akagi, "Endoscopic diagnosis of minute, small, and flat early gastric cancers," *Endoscopy*, vol. 21, pp. 159–164, Jul. 1989.
- [5] U. Honmyo, A. Misumi, A. Murakami, S. Mizumoto, I. Yoshinaka, M. Maeda, S. Yamamoto, and S. Shimada, "Mechanisms producing color change in flat early gastric cancers," *Endoscopy*, vol. 29, no. 5, pp. 366–371, Jun. 1997.
- [6] T. Kanesaka, T.-C. Lee, N. Uedo, K.-P. Lin, H.-Z. Chen, J.-Y. Lee, H.-P. Wang, and H.-T. Chang, "Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging," *Gastrointestinal Endoscopy*, vol. 87, no. 5, pp. 1339–1344, May 2018.
- [7] V. V. Khryashchev, O. A. Stepanova, A. A. Lebedev, S. V. Kashin, and R. O. Kuvaev, "Deep learning for gastric pathology detection in endoscopic images," in *Proc. 3rd Int. Conf. Graph. Signal Process.*, Jun. 2019, pp. 90–94.
- [8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.
- [9] L. Li, Y. Chen, Z. Shen, X. Zhang, J. Sang, Y. Ding, X. Yang, J. Li, M. Chen, C. Jin, C. Chen, and C. Yu, "Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging," *Gastric Cancer*, vol. 23, no. 1, pp. 126–132, Jan. 2020.
- [10] B.-J. Cho, C. S. Bang, S. W. Park, Y. J. Yang, S. I. Seo, H. Lim, W. G. Shin, J. T. Hong, Y. T. Yoo, S. H. Hong, J. H. Choi, J. J. Lee, and G. H. Baik, "Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network," *Endoscopy*, vol. 51, no. 12, pp. 1121–1129, Dec. 2019.
- [11] B. Sahiner, A. Pezeshk, L. M. Hadjiiski, X. Wang, K. Drukker, K. H. Cha, R. M. Summers, and M. L. Giger, "Deep learning in medical imaging and radiation therapy," *Med. Phys.*, vol. 46, no. 1, pp. e1–e36, Jan. 2019.
- [12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.
- [13] D.-H. Kim, H. Cho, and H.-C. Cho, "Gastric lesion classification using deep learning based on fast and robust fuzzy C-means and simple linear iterative clustering superpixel algorithms," *J. Electr. Eng. Technol.*, vol. 14, no. 6, pp. 2549–2556, Nov. 2019.
- [14] D.-H. Kim and H.-C. Cho, "Deep learning based computer-aided diagnosis system for gastric lesion using endoscope," *Trans. Korean Inst. Electr. Eng.*, vol. 67, no. 7, pp. 928–933, 2018.
- [15] C. L. Chowdhary, M. Mittal, P. A. Pattanaik, and Z. Marszalek, "An efficient segmentation and classification system in medical images using intuitionist possibilistic fuzzy C-mean clustering and fuzzy SVM algorithm," *Sensors*, vol. 20, no. 14, p. 3903, Jul. 2020.
- [16] D. Ergashev and Y. I. Cho, "Skin lesion classification towards melanoma diagnosis using convolutional neural network and image enhancement methods," *J. Korean Inst. Intell. Syst.*, vol. 29, no. 3, pp. 204–209, Jun. 2019.
- [17] Y. Zhu, Q.-C. Wang, M.-D. Xu, Z. Zhang, J. Cheng, Y.-S. Zhong, Y.-Q. Zhang, W.-F. Chen, L.-Q. Yao, P.-H. Zhou, and Q.-L. Li, "Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy," *Gastrointestinal Endoscopy*, vol. 89, no. 4, pp. 806–815, 2019.
- [18] A. Asperti and C. Mastronardo, "The effectiveness of data augmentation for detection of gastrointestinal diseases from endoscopic images," 2017, *arXiv:1712.03689*. [Online]. Available: <http://arxiv.org/abs/1712.03689>
- [19] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, Dec. 2018.
- [20] A. Zhao, G. Balakrishnan, F. Durand, J. V. Guttag, and A. V. Dalca, "Data augmentation using learned transformations for one-shot medical image segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8543–8553.
- [21] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019.
- [22] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, "AutoAugment: Learning augmentation strategies from data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 113–123.
- [23] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019.
- [24] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
- [25] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, and A. K. Nandi, "Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 5, pp. 3027–3041, Oct. 2018.
- [26] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258.
- [27] S.-A. Lee, D.-H. Kim, and H.-C. Cho, "Deep learning based gastric lesion classification system using data augmentation," *Trans. Korean Inst. Electr. Eng.*, vol. 69, no. 7, pp. 1033–1039, 2020.



SIN-AE LEE received the B.S. degree in electrical and electronic engineering from Kangwon National University, South Korea, in 2018, where she is currently pursuing the M.S. degree with the Interdisciplinary Graduate Program in BIT Medical Convergence.



HYUN CHIN CHO received the M.D., M.S., and Ph.D. degrees in internal medicine from Gyeongsang National University School of Medicine, Jinju, South Korea, in 2003, 2008, and 2014, respectively. From 2009 to 2010, she was a fellow with the Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. From 2011 to 2015, she was a Professor with the Samsung Changwon Hospital, Sungkyunkwan University School of Medicine, Changwon, South Korea. She is currently a Professor with the Gyeongsang National University School of Medicine and Gyeongsang National University Hospital, Jinju.



HYUN-CHONG CHO received the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida, USA, in 2009. From 2010 to 2011, he was a Research Fellow with the University of Michigan, Ann Arbor, USA. From 2012 to 2013, he was a Chief Research Engineer with LG Electronics, South Korea. He is currently a Professor with the Department of Electronics Engineering and Interdisciplinary Graduate Program in BIT Medical Convergence, Kangwon National University, South Korea.

...