Image-Enhancement-Based Data Augmentation for Improving Deep Learning in Image Classification Problem

Weihang ZHANG, Yuma KINOSHITA, *Student member*, *IEEE*, and Hitoshi KIYA, *Fellow, IEEE*
Department of Computer Science
Tokyo Metropolitan University, Tokyo, Japan

***Abstract*--In this paper, we propose a novel data augmentation method based on image enhancement. When training CNN models for image classification, it is required to prepare sufficient training data taken in various conditions. However, traditional data augmentation methods are limited to physical transformation, and shooting conditions of images, so exposure conditions lack for data augmentation. Therefore, we utilize an image enhancement method to generate images with different exposures, although enhanced methods are generally used for generating high quality images. Experimental results show that the proposed method improves the classification accuracy of a CNN model. The results also demonstrate that combining the proposed method with other existing data augmentation methods provides further improvement of the classification accuracy.**

I. INTRODUCTION

In recent years, with the growing of convolutional neural networks (CNNs), it has become the most important way to solve various problems such as image classification. CNN models still have many problems that need to be addressed. The most common problem is lack of training samples or uneven class balance within the datasets [1]-[2]. However, collecting a huge amount of images for training is very difficult and costly. Therefore, various data augmentation methods have been proposed. For example, commonly-used methods are random cropping [3], flipping [4], random erasing [5] and GAN-based augmentation [6].

These data augmentation methods focus only on physical transformation, e.g., rotation, shifting and scaling. However, they do not take into account image brightness or contrast despite the fact that images are captured in various shooting conditions including exposure conditions. For this reason, there are some augmentation methods that simply manipulate image brightness or contrast, but these methods do not consider the imaging pipeline of digital cameras.

Because of this situation, we propose a novel data augmentation method on the basis of image enhancement. For the proposed method, we use an image enhancement method: Automatic exposure compensation for single-image-based multi-exposure fusion [7]-[8] method which considers the imaging pipeline. Experimental results show that the proposed data augmentation method can improve the performance of CNN model. In addition, the proposed method is demonstrated to be able to be combined with commonly-used data augmentation methods.

II. PROPOSED METHOD

The framework of the proposed method is shown in Fig. 1. The core idea of the proposed method is to use an image enhancement method using pseudo multi-exposure (ME) images. The use of the proposed method enables us to produce a training set that includes dark, middle, and bright images.

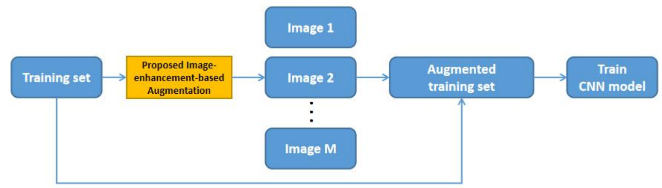

Fig. 1. Framework of proposed data augmentation method.

TABLE I. SIMULATION CONDITIONS

Database	Name	GTSRB	CIFAR-10
	Classes	43	10
	Training set	39209 images	50000 images
	Test set	12630 images	10000 images
	Image size	32×32	32×32
CNN model	Architecture	ResNet-18	
	Batch size	128	
	Learning rate	0.01 (decayed by 0.1 every 40 epoch)	
	Epochs	160	
	Preprocess	z-score normalization	
	Optimizer	SGD	

For the proposed method, we use an conventional image enhancement method: Automatic exposure compensation for single-image-based multi-exposure fusion [7]-[8]. The procedure for generating pseudo ME images by using the method is summarized as follows:

(1) Calculate luminance L from an input image I.

(2) Enhance the local contrast of $L(p)$ at a pixel p by:

$$L'(p) = \frac{L^2(p)}{L_a(p)}, \tag{1}$$

where $L_a(p)$ is the local average of luminance L around pixel p.

(3) Separate an image area based on luminance distribution by

$$\mathrm{P}_m = \{p \mid \theta_m \le L'(p) \le \theta_{m+1}\}, \tag{2}$$

$$\theta_m = \frac{M-m+1}{M}(\max L'(p) - \min L'(p)) + \min L'(p)), \tag{3}$$

where M is the number of separated scene areas and P_m indicates the m-th area.

(4) Calculate scaled luminance L''_m which clearly represents an area P_m by

$$L''_m(p) = \alpha_m L'(p), \tag{4}$$

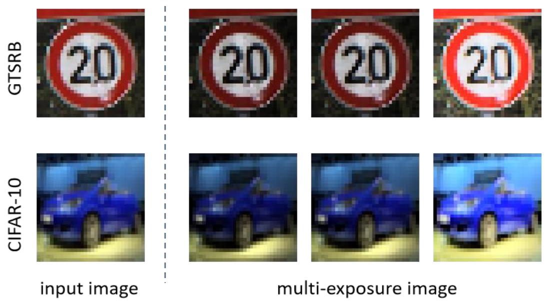

Fig 2. Examples of generated images by proposed method.

$$\alpha_m = \frac{0.18}{g(L' \mid \mathrm{P}_m)}, \tag{5}$$

where $g(L' \mid \mathrm{P}_m)$ is the geometric mean of luminance L' on area P_m.

Tone map scaled luminance L''_m to $\hat{L}_m$ by

$$\hat{L}_m(p) = \frac{L''_m(p)}{1 + L''_m(p)} \left(1 + \frac{L''_m(p)}{{L''_m}^2(p)} \right). \tag{6}$$

Generate a pseudo ME image $\hat{I}_m$ by

$$\hat{I}_m(p) = \frac{\hat{L}_m(p)}{L(p)} I(p). \tag{7}$$

Since parameter M controls the number of separated scene areas, we generated M pseudo ME images by setting M to a desirable number.

SIMULATION

Simulation condition

We trained a CNN model for image classification by using two databases: the GTSRB database [9], and the well-known CIFAR-10 database [10] (see Table I). The CNN model that we used in the simulations was ResNet-18 and it was trained under the conditions shown in Table I. To simulate the case that we have insufficient training samples, we constructed two training sets, called "1/4 set" and "1/8 set", from each of GTSRB and CIFAR-10. For constructing the 1/4 set, a quarter of a set of images in each class was randomly selected. For constructing the 1/8 set, a half of a set of images in each class from the 1/4 set was randomly selected.

The proposed method augmented the original set and the 1/4 set by 4 times, and the 1/8 set by 8 times. Examples of images generated by the proposed method are shown in Fig. 2.

Simulation result

Table II shows the classification accuracy under the use of models trained without any data augmentation (Plain) or with the proposed data augmentation method (Proposed). For calculating the accuracy, we trained the CNN model five times and the mean and standard deviation of accuracy were calculated. From Table II, the proposed method improved the classification accuracy for all six cases, especially for cases with insufficient training samples.

Table III shows the effectiveness of combining the proposed method with commonly-used augmentation methods in terms of classification accuracy for the 1/4 set. Random rotation (in which rotation angle is in [-30°,30°]) and random crop with a size of 32×32 and a padding of 4 were used as the commonly-used methods. From Table III, combining the proposed method with the commonly-used method further improved the performance of the model.

TABLE II. CLASSIFICATION ACCURACY WITH/WITHOUT PROPOSED METHOD (MEAN ± STD.) (%).

Training set		Plain	Proposed
GTSRB	Original set	97.48 ± 0.23	**97.88** ± 0.12
	1/4 set	96.03 ± 0.29	**96.66** ± 0.22
	1/8 set	91.85 ± 0.27	**94.99** ± 0.32
CIFAR-10	Original set	85.92 ± 0.12	**85.94** ± 0.35
	1/4 set	73.48 ± 0.41	**73.80** ± 0.17
	1/8 set	63.36 ± 0.91	**64.92** ± 0.50

TABLE III. CLASSIFICATION ACCURACY UNDER THE USE OF PROPOSED METHOD AND COMMONLY USED AUGMENTATION METHODS (MEAN ± STD.) (%).

Training set		Random rotation & crop	Proposed + Random rotation & crop
GTSRB	1/4 set	96.97 ± 0.46	**97.93** ± 0.34
CIFAR-10	1/4 set	81.33 ± 0.29	**83.21** ± 0.12

IV. CONCLUSION

In this paper, we proposed a novel data augmentation method based on image-enhancement for improving CNN models. The proposed method is based on an image enhancement method using pseudo ME images to generate images with various brightness. Experimental results showed that the proposed augmentation method improved the performance of model. The results also demonstrated that combining the proposed method with commonly used data augmentation methods yielded better performance in image classification.

REFERENCE

[1] A. Mikolajczyk and M. Grochowski, “Data augmentation for improving deep learning in image classification problem,” 2018 International Interdisciplinary PhD Workshop (IIPhDW), pp.117 – 122, 2018.

[2] C. Shorten and T.M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” Journal of Big Data, vol.6, no.1, p.60, 2019.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton. “Imagenet classification with deep convolutional neural networks.” In NIPS, 2012.

[4] K. Simonyan and A. Zisserman. “Very deep convolutional networks for large-scale image recognition.” In ICLR, 2015.

[5] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, “Random erasing data augmentation,” CoRR, 2017.

[6] Maayan F-A, Eyal K, Jacob G, Hayit G. “GAN-based data augmentation for improved liver lesion classifcation.” arXiv preprint. 2018.

[7] Y. Kinoshita and H. Kiya, “Scene Segmentation-Based Luminance Adjustment for Multi-Exposure Image Fusion,”IEEE Transactions on Image Processing, vol.28, no.8, pp.4101 – 4116, 2019.

[8] Y. Kinoshita and H. Kiya, “Automatic exposure compensation using an image segmentation method for single-imagebased multi-exposure fusion,” APSIPA Transactions on Signal and Information Processing, vol.7, 2018.

[9] Sebastian Houben, Johannes Stallkamp, Jan Salmen, Marc Schlipsing and Christian Igel, “Detection of Traffic Signs in Real-World Images: The {G}erman {T}raffic {S}ign {D}etection {B}enchmark”, International Joint Conference on Neural Networks, 2013.

[10] A.Krizhevsky and G.Hinton. “Learning multiple layers of features from tiny images,” 2009.