Bulletin of Networking, Computing, Systems, and Software – www.bncss.org, ISSN 2186-5140
Volume 12, Number 1, pages 38–39, January 2023

Preprocessing Scheme of input images for Efficient Classification with Deep Learning

JunWoo Park, Youngduk Kim*
ICT Research Institute
DGIST
Daegu, Korea
{ qkrwnsdn233, ydkim*}@dgist.ac.kr

***Abstract*—Recently, deep learning-based artificial intelligence technology has significantly improved object recognition ability and provides efficient classification and decision results in various fields. For example, by photographing a food using a smartphone, it is possible to easily check the information of the food. However, the recognition performance varies highly depending on the quality and size of the picture taken by the user. In this paper, we propose a method to convert food packaging paper recognition into input data suitable for network calculation by performing a pre-processing process rather than the conventional detection method. This provides a result of improving the recognition rate by emphasizing the region to be detected through the image. For performance verification, we compared recognition rates with existing methods after training with different amount of datasets.**

Keywords— Object recognition, VGGNet, Image preprocessing

I. INTRODUCTION

Recently, deep learning research has been actively conducted worldwide. In particular, due to its characteristics, deep learning is used widely and efficiently in image processing and object recognition in such images. For example, for visual identification of vehicles, people, animals, foods, and other objects, a human generally judged with the naked eye, but all processes of recognition can be automated through deep learning algorithms with data training. If it is applied to medical services, there is a situation where malignant tumors such as cancer cells may be missed due to a doctor's mistake. Deep learning can be an alternative solution to provide accurate judgment and treatment [1]. However, deep learning may produce unpredictable output results when different types of data are entered for input data from the previously trained data set, which finally degrades performance of object recognition.

In this paper, to tackle above problem, we propose an efficient recognition method to identify a food product based on its visual characteristics of the food's package. The proposed method solves the problem through a noble preprocess of removing unnecessary elements using image processing [2] and guide-line based image editing. This preprocessed image is optimized for deep learning network input data, which results in recognition performance enhancement.

A. System Architecture for Experiments

As shown in Fig 1, there a smartphone application which can transmit input images[3] and receive detection results for object recognition improvement tests. An AI server detects images using preprocessing techniques, and a gateway server transmits information to smartphone application to facilitate communication between application and AI server.

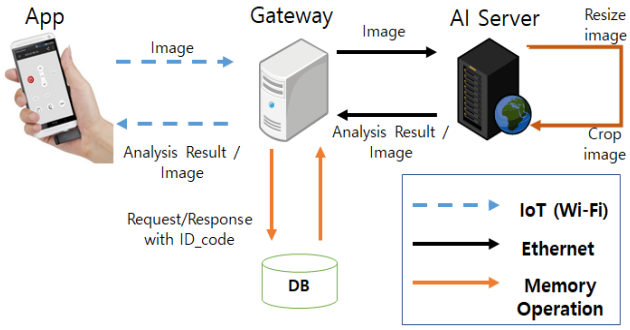

Fig. 1. System configuration for experiments

The AI server with object recognition algorithm was built as a server using the VGG16Net [4] CNN (Convolutional Neural Network) and it receives input images from user application. Then, it analyzes and classify in real time. The detailed specifications on the systems used in the experiment are shown in Table 1.

Table 1 Application and server information

	Application	**Gateway**	**AI**
Langue	Kotlin	Kotlin	Python
Tool	Android Studio	IntelliJ	Python IDLE
OS	Android	Window 10	Ubuntu 18.04
Equipment	Smart Phone	Desktop	Desktop

Through the above processing process, the user takes a picture (image) to be identified, matches the DB information input based on the result detected by the gateway server, and returns it to the client to show the food information.

B. Image Preprocessing Techniques

In general, the performance of detection result varies depending on whether the characteristics of the target object in the input image are properly displayed on the image and where

the object is located in the image. Also, the performance varies depending on the conditions (e.g. illuminance, angle, height) under which the target image was captured. Thus, it is effective to remove unnecessary elements except objects in one image to detect the correct results using these characteristics. As shown in Fig 2, the operation of removing a specific object area was implemented by applying it to an Android application that transmits an input image from the test system. It is important to note that the design of each food package is not constant and all different. Thus, if the package is introduced into the guidelines (the solid white line in Fig 2), only the corresponding part is reconstructed into an image. The unexpected background image displayed in the reconstructed image is unavoidable to impose a common condition on non-uniform food products. Nevertheless, the proposed method is very effective because the target image (object) to be detected is highlighted.

Fig. 2. Input image processing technique Course

Then, the input image from which only a specific area is extracted is transmitted to the AI server through the gateway server. The input image received is resized at a resolution of 300 X 400 in real time manner. Since the constructed AI server is composed of VGG16Net, it receives 224 x 224 x 3 input images and convolutionizes 3 x 3 filters of the same size to represent a feature map. If an image with unexpected resolution is entered, a detection error may occur during the resizing process. Therefore, by considering such facts, it was fixedly converted to a size of 300 x 400, and the packaged product data was equally readjusted at a ratio of 300 x 400 during deep learning training. If training is performed at the same rate, there will be no data corruption when resizing feature maps in the VGG network.

C. *Experiment*

The performance was verified by applying the technique [5] to above mentioned system and conducting an experiment. As for the food package, '*Haribo Goldberen 100g*' product was selected, and the test was conducted with 10 image data of different environmental conditions, respectively, and the detection results were compared when the proposed method was applied and not applied. The learning file to be used for the test used a total 20 class of 16,000 images as learning data, and the hyperparameters were set to 3 for Epoch and 32 for batch size. We prevented the overfitting problem through the experimental environment and confirmed that the proposed algorithm accurately recognizes each input data. However, when the user's input data including the unnecessary background was entered without applying our pre-processing method to the network, only two images were successfully detected. The overall test results are shown in Figure 3 below.

Fig. 3. System-based detection tests

II. Conclusion

The main contribution of this study is to implement deep learning algorithm to recognize food's package and provide information about target food to clients based on the results. In this case, it is important to ensure an accurate recognition rate by the fact that if the recognized data is inaccurate, it may cause side effects (e.g. allergic reactions) during food intake. However, in general, since the specific size and shape of the input image for the deep learning network were not explicitly determined, the recognition rate was degraded. Our proposed method significantly increases the recognition rate of deep learning by providing new guidelines and preprocessing operations.

Future research is to conduct research on more kinds of food images by increasing the number of classes of data and increasing learning data, and then finally provides performance enhancement for the object detection even if the number of classes increases. In addition, we plan to deliver various information on food can be inquired in various ways in addition to image recognition.

Acknowledgment

This work was supported by a grant (no. 21163MFDS518) from MFDS (Ministry of Food and Drug Safety) of Korea.

Reference

[1] ISMAIL, Nur Syahmi; SOVUTHY, Cheab. Breast cancer detection based on deep learning technique. In: *2019 International UNIMAS STEM 12th engineering conference (EnCon)*. IEEE, 2019. p. 89-92.

[2] VAN HEEL, Marin, et al. A new generation of the IMAGIC image processing system. *Journal of structural biology*, 1996, 116.1: 17-24.

[3] TZOU, Kou-Hu. Progressive image transmission: a review and comparison of techniques. *Optical engineering*, 1987, 26.7: 581-589.

[4] SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[5] SUN, Youcheng, et al. Testing deep neural networks. *arXiv preprint arXiv:1803.04792*, 2018.