Object Detection Model Training Technique for Numeric Text Detection in Analog Power Meter Images

Young Hwa Cha[1] and Byung Joon Park[2]

[1]*Student, Department of Computer Science, Kwangwoon University, Seoul, Korea*
[2]*Professor, Department of Computer Science, Kwangwoon University, Seoul, Korea*
[1]*yhcha@kw.ac.kr,* [2]*bjpark@kw.ac.kr*

Abstract

Recently, with the expansion of the demand response market, the need to automatically read power meters is increasing. In addition, computer vision using deep neural networks has rapidly developed, and object detection, the task of finding a specific object in an image, has now reached high accuracy. There are several studies on the subject of automatic power meter reading. However, studies that apply object detection to reading meters are difficult to find. Automatic power meter reading is generally divided into the task of finding the number corresponding to the usage and the task of recognizing the number. In this paper, we only deal with finding numbers using object detection. Training is needed to create an object detection model that finds only the numeric area in the meter. The important factors in training a model are the amount of training data and the number of training epochs. Training a lot of epochs with a lot of training data will show high detection performance, but it takes a lot of time to prepare for training and training data. In this paper, we present the amount of training data and the number of epochs that can take the least time to train an object detection model with a detection performance of over 99%. Also, the results of various experiments performed to find this parameter are recorded in the paper.

Keywords: Object detection, Computer vision, Machine learning, Deep neural networks, Automatic power meter reading

1. Introduction

In recent years, demand response has emerged as an alternative to increasing the supply by increasing power generation facilities as a way to solve the imbalance in power supply and demand. Demand response [1] is an act to maintain the balance between power demand and supply by inducing a decrease in power use by raising electricity rates or instructing load reduction when a power supply and demand crisis in which the maximum value of power demand exceeds the supply amount is expected. In order to realize the demand response, it is necessary to analyze the power load pattern by season, month, day, or time period to predict power demand. In order to analyze the power load pattern more accurately, it is necessary to measure the power consumption for each time zone according to the place or target where the power is used. In recent years, an environment in which power consumption can be measured in real time has been created with the spread of digital power meters, but analog power meters are being used in more places than digital meters, the power consumption of the analog power

meter is recorded by manpower. For this reason, we conducted a study to automatically read analog power meters. The automatic reading of the power meter requires two processes. The first is to find the part that displays the amount of power in the meter image, and the second is to recognize the number. This paper only deals with finding the numeric area. Several studies have been conducted to automatically read power meters. Many other studies of automatic power meter reading use image processing to detect numeric area. However, we use an object detection model composed of deep neural networks. The object detection model can expect higher accuracy than image processing in detecting numeric area. Generating an accurate detection model requires more than a certain amount of training data and epochs. This is a time consuming task. In this paper, we present the number of epochs and training data that can achieve the accuracy of over 99% for a model that detects a numeric area in an analog power meter. In addition, we tested several experiments to reduce the time to create a model, and propose optimal training parameters obtained from the experimental results.

2. Related work

Elrefaei et al [2] is a study to recognize the Arabic number of power meters used in Saudi Arabia. They cropped the numeric area by image processing such as gray scaling, binarization, and noise reduction on the meter image. Their experimental results showed that the accuracy of cropping a numeric area from an image was 90.47%, and 21 images were tested.

Anis et al [3] is a study that recognizes the number of displays in digital power meters. They used RGB to YCbCr color convert for the meter image to find the display area using the threshold of display color. Their experimental results showed that the accuracy of detecting a reading region from an image was 96.30%, and 135 images were tested.

Sun et al [4] is a study that recognizes the number of displays in digital power meters used in China. In Sun et al, displays were found using image processing techniques. The difference from this study is that in Sun et al, a digital power meter equipped with a liquid crystal display that displays numbers in the form of a 7-segment is targeted, and no deep neural network is used to detect the display.

Object detection [5] is a problem that has been dealt with for a long time in computer vision, a field of artificial intelligence, and active research has been conducted, and the main goal is to check whether an object to be searched for exists in the image and, if so, the location and range of each object is to return. Recently, there have been studies using deep neural network models as a method for object detection, and representative models are YOLO [6] and Faster R-CNN [7].

3. Method

3.1. Problem definition

In this paper, we deal with the problem of detecting only the area that displays the power consumption in the meter image with various texts. In particular, we are targeting a picture of an analog meter where the numbers for each digit of usage are displayed by moving the dial. The dial indicating the amount of usage in the photo corresponds to the part to be detected, and the text is limited to numbers. In addition, the meter image includes texts such as a manufacturing company name, a product name, and a product number.
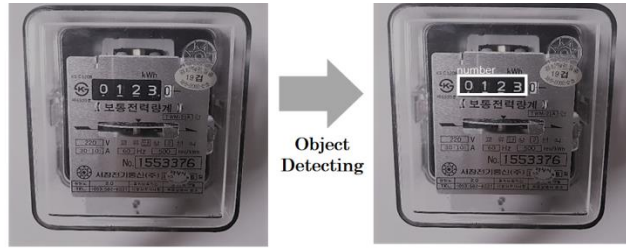
Figure 1. Example of detecting usage text in an analog power meter image

3.2. Model

In this paper, two object detection models are used to detect the power consumption numerical value in the analog power meter image. One is YOLO and the other is Faster-RCNN. In this paper, both models are modified and used to detect only one class. Both models used a model implemented with TensorFlow [8], It is not fine-tuned by adding a text part corresponding to the usage of the power meter to the previously learned weight, the weights were learned to solve only the problems defined in the paper.

3.3. Dataset

To test the two models used in the experiment with the power meter, data sets were created manually. The datasets are composed of a pair of XML files in Pascal VOC [9] format with location information on a single image and the part to be detected in the image. The images used in the experiment are images of two types of analog power meters currently used in Korea. The resolution of the image is 2160×2160, and it is an RGB color image. Two types of analog power meters were taken from different directions, 6000 images per meter, 12000 were combined, of which 1200 were used for training, and the remaining 10800 were tested. We will refer to the meter datasets of one type as meter1 and the datasets of other meter types as meter2.

Table 1. Datasets of electric power meter images

Dataset	Image Resolution	Number of Images	Example of Images
meter1	2160×2160	6000	
meter2	2160×2160	6000	

3.4. Measurement

We use mAP (mean Average Precision) as a measure to evaluate model performance. And if the IoU (Intersection over Union) is 0.6 or more, it was evaluated as accurately detected. This means that all four numbers in the numeric area are included in the bounding box.

4. Experiment

The experiment was conducted in an environment using four NVIDIA gtx-1080 ti with 11 GB of memory. Section 4.1 compares the performance of YOLO and Faster-RCNN. Section 4.2 compares the performance of training YOLO, which shows better performance, by reducing the resolution of the training image. Section 4.3 compares the performance of training with different numbers of YOLO training images.

4.1. Performance comparison between YOLO and Faster-RCNN

The performance of YOLO and Faster-RCNN was compared, and an experiment was conducted on the datasets that combine both the datasets meter1 and meter2 and both meter1 and meter2. For each dataset, one tenth was used as training data and the rest was used for testing. Learning rate was 0.0001, batch size was 12, and images were trained and 400 epochs were trained. Both YOLO and Faster-RCNN generate the output image by adding the bounding box for the detected part to the input image. Detection performance was measured by accurately detecting if all four numbers were included in the added bounding box, and those that contained no or fewer than three numbers inside the bounding box were incorrectly detected. Detection speed was the average of the set used in the test as the time taken to process a single image. As a result of training and testing both the datasets meter1 and meter2, YOLO was 99.84%, which is 0.43% more accurate than 99.41% of Faster-RCNN, and the detection speed is 109ms for YOLO and Faster- RCNN was 213ms, so YOLO was almost twice as fast as Faster-RCNN. Models trained only on meter1 or meter2 showed slightly higher performance in both YOLO and Faster-RCNN than models trained together. [Table 2] provides data on the detection performance of models trained for 200, 300, and 400 epochs. When both models were trained for 100 epochs, the bounding box could not be drawn. The boundary box appeared in the image after 150 epochs, and the detection performance increased as shown in [Table 2], and did not increase any more after the 400 epochs.

Table 2. Detection performance and detection speed of YOLO and Faster-RCNN

Model	Dataset	Number of Training Images	Number of Test Images	Detection Performance (%)	Detection Speed (ms)
YOLO	meter1	600	5400	99.90	110
	meter2	600	5400	99.85	107
	meter1+meter2	1200	10800	99.84	109
Faster-RCNN	meter1	600	5400	99.64	212
	meter2	600	5400	99.42	212
	meter1+meter2	1200	10800	99.41	213

Table 3. Detection performance according to the training epoch of YOLO and Faster-RCNN

Model	Dataset	Epochs	Detection Performance (%)
YOLO	meter1+meter2	200	79.35
		300	91.52
		400	99.84
Faster-RCNN	meter1+meter2	200	82.12
		300	93.16
		400	99.41

4.2. YOLO detection performance experiment according to the training image resolution

The model was trained by reducing the resolution of the images from 2160×2160 to 416×416. Only the image of the training set was reduced, not the image of the test set. As a result of the experiment, the model trained with images of 416×416 resolution had a low accuracy of 0.03% compared to the model trained with 2160×2160 resolution images. We trained 400 epochs, images at 2160×2160 resolution took 49.68 hours, and images at 416×416 resolution took 3.2 hours. As it shows a slight level of difference in accuracy, training with 416×416 resolution images does not affect the detection of numeric text parts. The reason we specifically trained the image to be reduced to 416×416 is because it is the size of the input layer. If the images to be used as training data are larger than the size of the input layer, we suggest reducing the resolution of the training images to the size of the input layer as a pre-processing process.

Table 4. Training time and detection performance according to the training image resolution

Model	Dataset	Number of Training Images	Training Image Resolution	Detection Performance (%)	Training Time (hour)
YOLO	meter1+meter2	1200	2160×2160	99.84	49.68
			416×416	99.81	3.2

4.3. YOLO detection performance experiment according to the amount of training data

[Table 5] shows the results of training the model with 1200, 600, 300, and 200 training images. When the training image was reduced from 1200 to 600 and the same generation was trained, there was no deterioration in performance and the training time was cut in half. The model trained with 300 images showed significantly lower detection performance in the 400th generation and 99.85% performance in the 1000 epochs. The model trained with 200 images showed 96.52% accuracy in 1500 generations, and the detection performance did not increase afterwards. As a result of the experiment, it is necessary to train more than 1000 epochs with 300 samples in order to obtain a model with a detection performance of 99% or more.

Table 5. Training time and detection performance according to the number of training images

Model	Dataset	Resolution	Number of Training Images	Epochs	Detection Performance (%)	Training Time (hour)
YOLO	meter1+meter2	416×416	1200	400	99.81	3.2
			600	400	99.83	1.6
			300	1000	99.85	2
			200	1500	96.32	2

5. Conclusion

In this paper, we proposed a method to train an object detection model to detect the numerical domain in an analog power meter. Our experiments show that using the object detection model to solve the problem of detecting a numeric region in a power meter shows higher accuracy than using image processing. However, training a high-accuracy object detection model is a time-consuming task. Therefore, we proposed the training parameters that are the least time consuming to train the model. Using the proposed method, it would be possible to solve not only the power meter, but also the problem of detecting a numeric area in a gas meter or water meter image of a similar shape.

Acknowledgements

References

[1] P. Palensky and D. Dietrich, "Demand side management: demand response, intelligent energy systems, and smart loads," IEEE Transactions on Industrial Informatics, vol.7, no.3, pp.381-388, (2011)

[2] L. A. Elrefaei, A. Bajaber, S. Natheir, N. AbuSanab, and M. Bazi, "Automatic electricity meter reading based on image processing," IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, pp.1-5, (2015)

[3] A. Anis, M. Khaliluzzaman, M. Yakub, N. Chakraborty, and K. Deb, "Digital electric meter reading recognition based on horizontal and vertical binary pattern," 2017 3rd International Conference on Electrical Information and Communication Technology (EICT), Khulna, pp.1-6, (2017)

[4] J. Sun, Yuzhong Ma, Han Yang, Ning Wang, Yu Ding, Aiping Song, Yongwei Zhu, and Juntong Xi, "Character recognition method for low-contrast images of numerical instruments," 2018 IEEE 3rd International Conference on Image, Vision and Computing, Chongqing, China, pp.157-162, (2018)

[5] Li Liu, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, and Matti Pietikäinen, "Deep learning for generic object detection: a survey," International Journal of Computer Vision, vol.128, pp.261-318, (2020)

[6] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You only look once: unified, real-time object detection," The IEEE Conference on Computer Vision and Pattern Recognition, pp.779-788, (2016)

[7]   R. Girshick, "Faster R-CNN," The IEEE International Conference on Computer Vision, pp.1440-1448, **(2015)**

[8]   https://www.tensorflow.org/

[9]   http://host.robots.ox.ac.uk/pascal/VOC/

Authors

Young Hwa Cha
PhD student
Department of Computer Science
Kwangwoon University
Seoul, Korea

Byung Joon Park
Professor
Department of Computer Science
Kwangwoon University
Seoul, Korea

This page is empty by intention.

Young Hwa Cha and Byung Joon Park