A Robust Automatic Meter Reading System based on Mask-RCNN

Abdullah Azeem, Waqar Riaz, Abubakar Siddique, Saifullah, Umar Ajaib Khan
School of Communication and Information Engineering
Chongqing University of Posts and Telecommunications
abdullahazeem06@outlook.com

***Abstract*—Automatic Meter Reading (AMR) is tackled by leveraging the high capability of Region-Convolutional Neural Networks (RCNN). However, license plate recognition is quite popular, but AMR problem is still unexplored. Numerous investigations are still crucial due to some constraints: handcrafted features, low efficiency, blur, rotate digits, low reflection, and poor quality images. With the significance of robust AMR, we have achieved advances of AMR in computer vision and deep learning, but still, these methods lack of some errors. To address these problems, we proposed a method Mask-RCNN (AMR) based on mask region convolutional neural networks (Mask-RCNN) used for counter detection, digit segmentation, and recognition. Experiments demonstrated that our proposed method substantially outperforms the state-of-the-art methods in terms of counter detection and digit recognition on publically available UFPR-AMR dataset.**

Keywords—component, Automatic Meter Reading, Counter Detection, Deep learning, Mask-RCNN

I. Introduction

The term Automatic Meter Reading (AMR) refers to automatically record the consumption of electricity, water, and gas for billing and monitoring [1]. Although, nowadays, we see smart readers [2] in many countries, but still in underdeveloped countries, the meter reading is performed manually [3]. Where a site operator takes a picture as a proof, and another operator needs to verify the image for reading confirmation. This operation is full of human efforts, needs much more time, and efficiency is low: as this operation is prone to error. In this manual system of AMR, an error might be unnoticed as a large number of pictures needed to evaluate, and this evaluation process usually done by sampling [4]. Manual meter reading problems can be resolved by introducing automatical meter reading system. This problem can reduce the error caused by the human factor and save human resources.

Moreover, fully automatic meter reading can be achieved by installing cameras in the meter box [5]. Image-based AMR does not require to install new meters, fast installation, and at a lower cost. AMR overlaps with other Optical Character Recognition (OCR) applications, i.e. robust reading and license plate recognition, as these applications are relying on text information extraction from images under challenging conditions. Though AMR work is not produced in literature [6], such as other OCR applications, some satisfactory work is recently presented in [7].

The unusual challenges in OCR for AMR is rotating digits that are the primary error cause in this system. Though robust methods are used for digit recognition, this problem still exists, and different methods are proposed to address this problem. Rodriguez and Berdug [8] addressed this problem by using the Hausdorff distance-based method. This method achieved good recognition results in real-time, and all images extraction was done by using a single meter. There was no method used for angle correction, and for pre-processing, that is why a controlled environment was required.

Nodari and Gallo [3], they used Multi-layer Perceptron (MLP) without pre-processing and post-processing, for counter detection and digit segmentation. They achieved low F-measures and added some other methods [8] to improve counter detection, and to avoid false positives. They used the watershed method and Fourier analysis for that, and they achieved some excellent results. To check the system efficiency, they use only 100 images that may not be enough. And it was the first time when images were used for publically available experiments.

Although with the significance of robust AMR, we have achieved advances of AMR in computer vision and deep learning. Convolutional Neural Networks (CNNs) [9] are used at all stages of AMR. Prior work depends on handcrafted features that capture the meter color features and some morphological attributes in at least one stage. This is because of noise and might not be robust to other kinds of meters. Deep learning techniques are dependent on big training data to produce high classification accuracy of unseen data. In the AMR work, datasets are not usually available at the public level because these bills belong to related companies. A new database called UFPR-AMR by the author [10] is introduced, which consists of more than 2000 fully annotated images.

In this work, Mask-RCNN (AMR) method based on Mask-RCNN is proposed, which is used for counter detection, digit segmentation, and recognition. Features are extracted by using a network similar to GoogleLeNet, and further fed into two convolutional layers to classify detected or undetected counter. Finally, the RoI-Align layer indicates the probability of each region of interest (RoI) as detected or undetected. Our proposed model produces the best recognition results for the UFPR-AMR dataset.

The remaining part of the paper is ordered as follows. Section. 2 gives methodology, including counter detection, digit segmentation, and digit recognition. Section 3 gives experimental results, and section 4 concludes this work.

II. Methodology

Mask-RCNN (AMR) system based on Mask-RCNN is proposed, which focuses on detecting and finding the meter counter in a challenging situation, as in underdeveloped countries, the meter reading is performed manually. The proposed AMR approach includes counter detection, digit segmentation, and digit recognition. Meters have many textual blocks, in different kinds of meters counter position varies, and usually, it occupies a small space in the image. Counter detection is the fundamental phase because the overall accuracy of the system and the AMR system processing speed depend on the performance of it. Therefore, we propose to locate the counter region first and then perform its recognition in the detected patch.

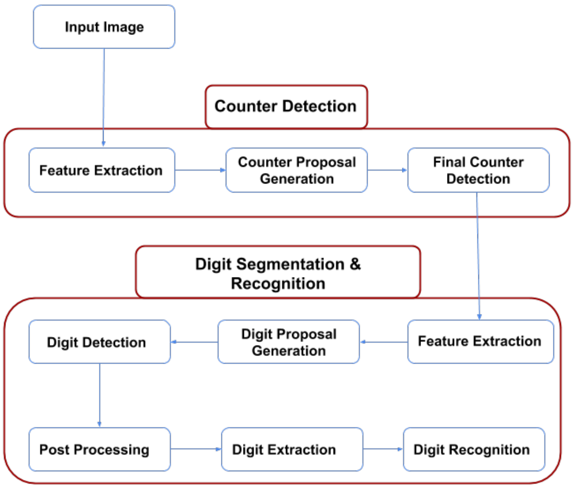


Figure 1. Flow diagram of proposed AMR system

A. Mask-RCNN

Mask-RCNN [11] is employed in our work that is based on faster RCNN. CNN are wildly used for object detection, the first approach, which generates regions to recognize the objects by merging the features and edges, is called RCNN. This approach is based on handcrafted features and does not involve learning. After region extraction, a pre-trained CNN classifies the regions as “nothing” and “object-X.” Regions generated by RCNNs are not accurate enough, and up to a few thousand regions can be generated, which makes this approach slow. This problem is solved by fast RCNN [12] and faster RCNN [13].

The basic idea is to use CNN to learn Regions of interests (ROIs) by using pooling layers to extract region proposals which improve the procedure in two ways:

- Performance of CNN in feature generation.
- Useless steps are removed during feature generation by passing an image through once.

Faster-RCNN uses a CNN to predict bounding boxes on proposed regions of interest. Mask-RCNN improves upon the ROI pooling layers giving much more accurate bounding boxes.

B. Counter Detection

The proposed AMR approach uses Mask-RCNN [11] for counter detection, which classifies the counter region into ‘detected’ and ‘undetected.’ Fig. 2 shows the flow diagram of counter detection.

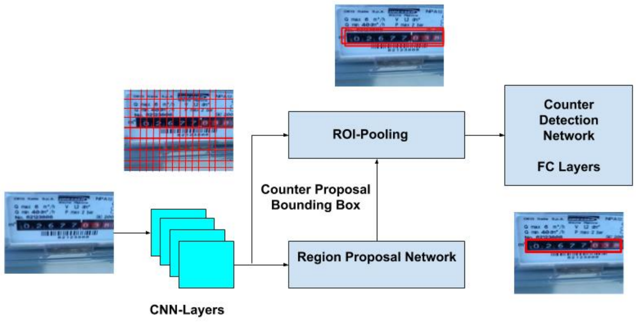


Figure 2. Flow diagram of counter detection

1) Feature extraction

For feature extraction, a network similar to GoogleLeNet is used by removing the inception module [14] and adding several pooling layers by stride two convolutions. Which results in a feature map that is one-eighth size of the original image. This allows us to get fine-grained details that are necessary for automatic meter reading because of low lighting conditions and poor image quality. The configuration of our model for counter detection is described in Table I.

TABLE I. CONFIGURATION OF COUNTER DETECTION NETWORK

Layer	Filters	Size
Conv	64	3x3
Max	64	3x3
Conv	64	1x1/1
Conv	64	3x3/1
Conv	64	3x3/2
Max	64	3x3
Incept	32	3x3
Max	64	3x3
Incept	48	3x3
Conv	128	1x1/1
Conv	128	3x3/1
Conv	128	3x3/2
Incept	64	3x3
Incept	96	3x3
Conv	96	1x1/1
Conv	96	3x3/1
Conv	512	3x3

2) Counter Proposal Generation

For counter-proposal generation, Mask-RCNN is employed without the segmentation stage. RPN, which takes features as an input, produces bounding boxes, is a fully connected classifier. It cannot be used in case meter counters; that is why we slightly modified this network. Considering aspect ratios and scales of counters, we used k = 16 anchor boxes and applied them at each position of the input feature map. These features are further concatenated with a channel axis to form a 512d feature vector, which is fed into two convolutional layers to classify detected/undetected counter and bounding box generation. The anchor boxes are described in Table II.

TABLE II. ANCHOR BOX

Width	Height
64	64
48	128
64	192
96	256
128	352
128	128
192	512
256	704
352	832
256	256
512	512
192	832

3) Final Counter Detection

The final part of counter detection is to classify the proposed regions into a counter and non-counter region. So, this part is similar as described earlier, but here it tries to minimize the false negatives and positives. For this purpose, we utilize the RoI-Align layer and set the pooling layer to 8x7. As the counter tends to be wider, we increase 'fine-grained' in the X-direction. To extract discriminative features, we use fully connected layers containing 2048 neurons with a dropout rate of 0.1. These features are then flattened into vectors and passed through fully connected layers. This layer produces two outputs, indicating the probability of each RoI as detected/undetected.

C. Counter Segmentation and Recognition

Once the counter is detected, we start the digit segmentation. Digit segmentation is an important stage to facilitate the recognition process, which includes the digit extraction from the counter-image. Many factors make the digit recognition process more complex, such as noise, blur, dirt, broken glass, reflections, rotate digits, low resolution, the different colors, and the numbering system. Moreover, digits can be rotating or in some range (e.g., 7-8), in some sort of counters.

In this case, the protocol adopted at Copel is considered. In which lower digits are considered as ground truth. To address these problems, we introduced a system that can effectively segment the digit. For digit recognition, each digit is first segmented from the counter, and it is the most significant phase for recognition. We divide the segmentation mask into two stages. In the first stage, Mask-RCNN is used to sort digits and non-digits, as we want to detect each digit in meter image. And predictions are made for digits and non-digits. In the second stage, digits are removed, which are too small or too wide to be a digit. And regions are extracted, which may vary in y coordinate by amount s where s is the distance between y coordinates. Therefore the height of two successive digits is represented by the formula.

$$s = |y2 - y1| \quad (1)$$

To achieve good results, we need many digits. Therefore we gather 10 classes (0-9) to train our model properly. To increase the number of classes, we add random Gaussian Noise, Change brightness, applying a median filter, and rotating digits in multiple angles (10, -10, +25, -25, +30, and -30). This model is similar to counter detection with different input settings, as shown in Table III.

TABLE III. COUNTER RECOGNITION

Layer	Filters	Size
Conv	64	3x3
Conv	64	1x1/1
Conv	64	3x3/1
Conv	64	3x3/2
Max-pooling	64	3x3
Incept	32	3x3
Conv	128	1x1/1
Conv	128	3x3/1
Conv	128	3x3/2
Max-pooling	64	3x3
Incept	64	3x3
Conv	96	1x1/1
Conv	96	3x3/1
Con-rpm	512	3x3
Fc-RoI	-	-
Fc-RoI	-	-
Fc-RoI	-	-
Fc-RoI	-	-

D. Loss Function and Training

Previously, training parts in faster R-CNN were trained separately, and training weights were merged in the end before full final training. After combining the complete model, we get four different losses: RPN (background), RPN (foreground), R-CNN (final class output), and R-CNN (bounding box regression), which gives us the total loss of:

$$L_{total} = (L_{cls} + L_{BBR})_{RPN} + (L_{cls} + L_{BBR})_{RCNN} \quad (2)$$

Where $(L_{cls} + L_{BBR})_{RPN}$ can be defined as region proposal loss for classifying background and foreground, respectively. And $L_{cls\ RCNN}$ can be described as a classification between backgrounds of different classes, where $L_{BBR\ RCNN}$ is a loss for refining bounding boxes.

III. EXPERIMENTATION

To verify the robustness and effectiveness of the proposed method, we compared our proposed method with three existing methods that include Goncalves et el [15], Vanetti et al [3], and Fast-Yolo [10]. UFPR-AMR dataset is used for the evaluation of the proposed method. Experiments are conducted on NVidia 1080tTi and implemented with Pytorch and OpenCV 3.1.

A. Counter detection on UFPR-AMR dataset

To detect counter, the PASCAL VOC challenge [16] definition of bounding boxes is used, in which the bounding box is considered correct when IOU > 0.5 [3]. According to this definition, our network is able to predict 99.82% of the counter with an average IOU of 80%. With the validation and testing set, a margin of 20% is necessary in order to keep all the digits within ROI. In both cases, all the counters are detected with IOU<0.5 using Mask-RCNN.

The parameters used for training are: 80k iterations, learning rate = [10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4}] with 10k, 25k, and 50k steps.

Figure 3. Example of Counter detection obtained with Mask-RCNN (AMR)

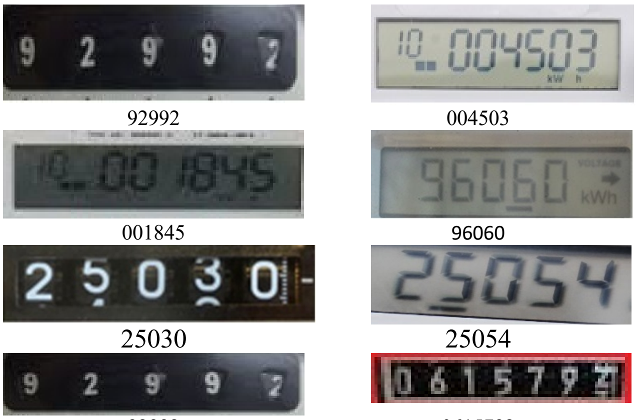
Table IV. Shows that our proposed method performed better among all three methods in terms of F-measure [3], and the highest value is shown in bold. Higher values of (F-measure) represent better results. Moreover, it can be seen that Mask-RCNN (AMR) predicted 100% counters correctly on the UFPR-AMR dataset.

TABLE IV. F-MEASURE OF COMPETING TECHNIQUES COMPARED WITH MASK-RCNN (AMR)

Approach	F-Measure
Gallo et al	95.85%
Vanetii et al	98.1%
Fast Yolo	99.21%
Mask-RCNN (AMR)	**100%**

B. Digit Segmentation

During segmentation, we label each digit into a counter from all training datasets as well as artificially generated datasets. Our method gives much higher recognition accuracy as compared to other methods [3, 10, 15].



92992 004503

001845 96060

25030 25054

92992 0615792

Figure 4. Results obtained with Mask-RCNN (AMR) on the UFPR-AMR

TABLE V. RECOGNITION ACCURACY OF COMPETING TECHNIQUES COMPARED WITH MASK-RCNN (AMR)

Approach	Recognition Accuracy	
	Digits	*Counters*
Gallo et al	92.3%	95.8%
Vanetii et al	96.1%	98.1%
Fast Yolo	98.2%	99.43%
Mask-RCNN (AMR)	**99.86%**	**100%**

IV. CONCLUSION

Though many studies have addressed the automatic meter reading problem, and researchers have implemented several methods to counter this problem. Still, all of these methods carry some advantages and disadvantages. We tackle this problem by leveraging the high capability of Region-Convolutional Neural Networks (RCNN) and propose a method based on mask region convolutional neural networks (Mask-RCNN) for counter detection, digit segmentation, and recognition. Our proposed method Mask-RCNN (AMR), is able to detect digits in all images and produces best recognition results for the UFPR-AMR dataset. Experimental results are compared with three existing techniques in terms of F-measure and recognition accuracy. Experiments demonstrated that our proposed method substantially outperforms the state-of-the-art methods in terms of counter detection and digit recognition on publically available the UFPR-AMR dataset.

ACKNOWLEDGMENT

This work was supported by the Chongqing Research Program of Basic ResearcH AND FRONTIER TECHNOLOGY UNDER Grant cstc2018jcyjAX0227.

REFERENCES

[1] Y. Gao, C. Zhao, J. Wang, et al., "Automatic watermeter digit recognition on mobile devices," in Internet Multimedia Computing and Service, 87–95, Springer Singapore (2018).

[2] Y. Kabalci, "A survey on smart metering and smart grid communication," Renewable and Sustainable Energy Reviews 57, 302 – 318 (2016).

[3] M. Vanetti, I. Gallo, and A. Nodari, "Gas meter reading from real world images using a multi-net system," Pattern Recognition Letters 34(5), 519–526 (2013).

[4] D. Quintanilha et al., "Automatic consumption reading on electromechanical meters using HoG and SVM," in Latin American Conference on Networked and Electronic Media, 11–15 (2017).

[5] V. C. P. Edward, "Support vector machine based automatic electric meter reading system," in IEEE International Conference on Computational Intelligence and Computing Research, 1–5 (2013).

[6] Y. Zhang, S. Yang, X. Su, et al., "Automatic reading of domestic electric meter: an intelligent device based on image processing and ZigBee/Ethernet communication," Journal of RealTime Image Processing 12, 133–143 (2016). Electronic Publication: Digital Object Identifiers (DOIs):

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature 521(7553), 436 (2015).

[8] M. Rodriguez, G. Berdugo, D. Jabba, et al., "HD MR: a new algorithm for number recognition in electrical meters," Turkish Journal of Elec. Engineering & Comp. Sciences 22, 87–96 (2014).

[9] L. Gomez, M. Rusi ´ nol, and D. Karatzas, "Cutting sayre's knot: Reading scene text without ˜ segmentation. application to utility meters," in 13th IAPR International Workshop on Document Analysis Systems (DAS), 97–102 (2018).

[10] Rayson Laroca, Victor Barroso, Matheus A. Diniz, Gabriel R. Gonçalves, William R. Schwartz, David Menotti, "Convolutional neural networks for automatic meter reading," J. Electron. Imag. 28(1) 013023 (5 February 2019)

[11] He, Kaiming, et al. "Mask r -cnn." Computer Vision (ICCV), 2017 IEEE International Conference on. IEEE, 2017.

[12] Girshick, Ross. "Fast r -cnn." Proceedings of the IEEE international conference on computer vision. 2015.

[13] Ren, Shaoqing, et al. "Faster r -cnn: Towards real -time object detection with region proposal networks." Advances in neural information processing systems. 2015.

[14] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[15] J. C. Gonc¸alves, “Reconhecimento de d´ıgitos em imagens de medidores de consumo de gas natural utilizando t ´ ecnicas de vis ´ ao computacional,” Master’s thesis, Universidade Tec- ˜ nologica Federal do Paran ´ a - UTFPR (2016).

[16] M. Everingham, L. Van Gool, C. K. I. Williams, et al., “The pascal visual object classes (VOC) challenge,” International Journal of Computer Vision 88, 303–338 (2010).