Deep Learning for Image-based Automatic Dial Meter Reading: Dataset and Baselines

Gabriel Salomon*, Rayson Laroca* and David Menotti*
*Department of Informatics, Federal University of Paraná (UFPR), Curitiba, PR, Brazil
Email: {gsaniceto, rblsantos, menotti}@inf.ufpr.br

Abstract—Smart meters enable remote and automatic electricity, water and gas consumption reading and are being widely deployed in developed countries. Nonetheless, there is still a huge number of non-smart meters in operation. Image-based Automatic Meter Reading (AMR) focuses on dealing with this type of meter readings. We estimate that the Energy Company of Paraná (Copel), in Brazil, performs more than 850,000 readings of dial meters per month. Those meters are the focus of this work. Our main contributions are: (i) a public real-world dial meter dataset (shared upon request) called UFPR-ADMR; (ii) a deep learning-based recognition baseline on the proposed dataset; and (iii) a detailed error analysis of the main issues present in AMR for dial meters. To the best of our knowledge, this is the first work to introduce deep learning approaches to multi-dial meter reading, and perform experiments on unconstrained images. We achieved a 100.0% F1-score on the dial detection stage with both Faster R-CNN and YOLO, while the recognition rates reached 93.6% for dials and 75.25% for meters using Faster R-CNN (ResNeXt-101).

Index Terms—automatic meter reading, dial meters, pointer-type meters, deep learning, public dataset

I. INTRODUCTION

Measuring residential energy consumption is known to be a laborious task [1]–[3]. Although smart meters are gradually replacing old meters, there are still many old mechanical meters in operation around the world since their replacement is time-consuming and costly. In many regions, such as remote areas and developing nations, manual on-site readings are still prevalent [4]. Even in developed countries, replacements are still far for complete. For example, in the end of 2018, there were still more than 26 million non-automatic meters in the United States [5].

In the literature, Automatic Meter Reading (AMR) is usually associated with digital and smart meters [6]. In this work, we use this designation exclusively for image-based automatic readings. AMR allows the employees of the service company (electricity/gas/water) or, preferably, the consumers themselves to capture meter images using a mobile device, which is cheaper and more feasible than manual on-site reading, and easier to deploy – in the short/medium term – than the replacement of old meters.

There are two main categories of residential energy meters [7], [8]: (i) analog (with cyclometer and dial displays) and (ii) digital (with electronic display and smart meters), as shown in Fig. 1. This work focuses on dial meters since, although there are numerous dial meters in operation, there are still many open challenges in this context (as detailed further).

(a) cyclometer display

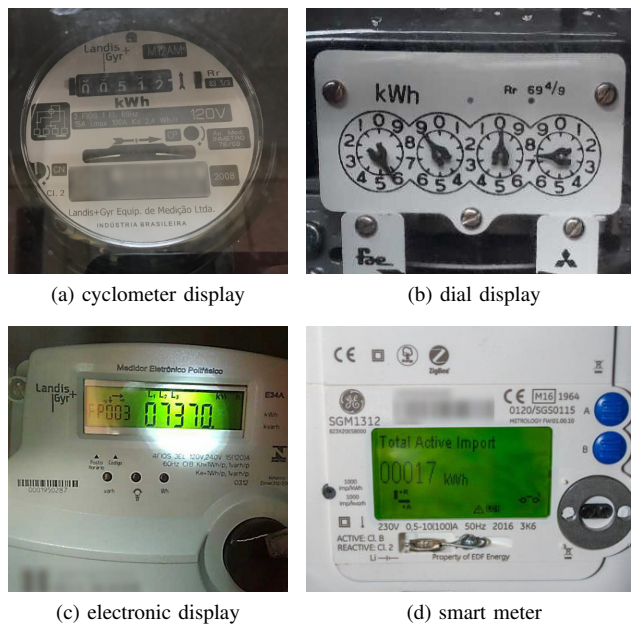(b) dial display

(c) electronic display

(d) smart meter

Fig. 1. The most common types of energy meters.

The Energy Company of Paraná (Copel) [9] measures electricity consumption in more than 4 million consuming units (i.e., meters) per month in the Brazilian state of Paraná. From the images they provided us (see Section III), we estimate that 21% of those devices are dial meters, resulting in more than 840,000 dial meter readings carried out every month.

Most of the dial meter recognition literature is focused on industrial applications, e.g., pressure meters [10]–[12], voltmeter [13] and ammeter [14]. As the meters are generally fixed and indoors, the image quality is strictly controlled.

Although in some cases the conditions are indeed realistic, they are not as unconstrained as in images obtained in outdoor environments, with challenging conditions, e.g., severe lighting conditions (low light, glares, uneven illumination, reflections, etc.), dirt in the region of interest, and taken at a distance. In addition, most approaches are based on handcrafted features [11], [15], and were evaluated exclusively on private datasets [10], [11], [14]–[16]. To the best of our knowledge, there are no public datasets containing dial meter

images in the literature.

Taking into account the above discussion, we introduce a real-world fully-labeled dataset (shared upon request) containing 2,000 meter images, acquired in unconstrained scenarios by Copel employees, with 9,097 individual dials and a well-defined evaluation protocol to assist the development and assessment of new approaches for this task[1] In addition, we conducted experiments using deep learning models in our dataset images to serve as baselines for future work, investigating problems related to dial meter reading and providing guidance for further research through a detailed quantitative and qualitative error analysis.

The remainder of this work is organized as follows. In Section II, we discuss approaches designed for AMR as well as deep learning techniques. The proposed dataset is described in Section III. Section IV presents the evaluated deep learning-based approaches for automatic reading of dial meters, while the results (with a detailed error analysis) are reported in Section V. Lastly, in Section VI, we state the conclusions.

II. RELATED WORK

There are many works in the literature that dealt with AMR. Most of them focus on the recognition of cyclometers and digital meters using Optical Character Recognition (OCR) methods. Recently, deep learning approaches have received great attention in this context [4], [10], [12], [17]–[19]. Dial meter recognition research, on the other hand, is more scarce. Most methods focus on gauges for industrial application [10]–[14], [20], [21]. Although gauges may look similar to energy dial meters, they usually only contain a single dial and one type of dial template, and the image conditions tend to be much more controlled in terms of lighting, dirt, and image quality. In this section, we describe some relevant works on AMR as well as state-of-the-art deep learning approaches for object detection and recognition [22]–[26].

A. Digit-based Meter Reading

Gallo et. al. [2] proposed a method that uses Multilayer Perceptron (MLP) to locate the Region of Interest (ROI) of the meters (also denoted as counter region [1], [2], [4]), Maximally Stable Extremal Regions (MSER) to segment the digits, Histogram of Oriented Gradients (HOG) for feature extraction, and Support Vector Machine (SVM) for digit recognition.

Nodari and Gallo [27] proposed a method named MultiNOD for gas cyclometers reading. It consists of a neural network tree, sharing and resizing features to perform counter detection and digit segmentation. The digit recognition stage was handled using Tesseract. This approach was later improved in [1], with the addition of a Fourier analysis applied to the segmented image, in order to avoid false positives. Finally, SVM was employed for digit classification.

Tsai et. al. [18] employed Single Shot MultiBox Detector (SSD) [28], a deep learning object detector, to locate the counter region in energy meters. The authors reported

an accuracy rate of 100% on their experiments, but did not address the recognition stage.

Yang et. al. [19] proposed a Fully Convolutional Sequence Recognition Network (FCSRN) for water meter analog digit reading, with a novel loss function entitled Augmented Loss (AugLoss). AugLoss addresses the "middle-state" that can occur when the digit accumulator is changing from one display digit to the next one, usually outputting the old displayed digit. Their approach outperformed Recurrent Neural Networks (RNN) and attention-based models on the task of sequence recognition, but the experiments were made in controlled images, with cropped and aligned meters.

Gómez et. al. [17] introduced a segmentation-free approach to perform meter reading. They trained a Convolutional Neural Network (CNN) to yield readings directly from the input images, without the need to detect the counter region. Although their approach has achieved promising results, the authors used a private dataset in the experiments, and only compared their method with traditional algorithms that rely on handcrafted features, which are easily affected by noise and may not be robust to images acquired in adverse conditions [4].

Laroca et. al. [4] designed a two-stage approach for AMR. The Fast-YOLOv2 model [29] was employed for counter detection and three CNN-based models were evaluated in the counter recognition stage. The authors considerably improved their recognition results when balancing the training set in terms of digit classes through data augmentation techniques.

B. Dial Meter Reading

Tang et. al. [15] proposed a complete framework for dial energy meter reading based on binarization, line intersection, and morphological operations. Despite being an interesting approach, the dataset used in the experiments was not published, and the images were obtained in a controlled environment.

In [16], the authors also employed handcrafted features for dial recognition. In addition to binarization and line intersection, the counter region was detected using Scale-Invariant Feature Transform (SIFT) features. Their method was evaluated on a private dataset containing only 141 images taken in a controlled environment.

The following approaches dealt only with single-dial meters (commonly known as gauges) and not with energy meters. Although the problems are similar, there is a fundamental difference: a small error in a multi-dial meter can result in a completely wrong measurement (especially if the error occurs in recognizing the most significant dials). Such a fact needs to be taken into account when evaluating recognition methods.

Several approaches explored handcrafted features, such as Hough Transform (HT), in order to locate the dials [11], [13], [20]. The steps in such works are very similar: image binarization on the preprocessing stage, Hough Circle Transform (HCT) for dial location, and pointer angle detection using Hough Line Transform (HLT) or similar methods. These approaches generally work well in constrained environments, but may not be suitable for real-world outdoor scenarios with uneven lighting and the presence of noise.

---

[1]The UFPR-ADMR dataset is **publicly available** (but upon request) to the research community at *http://web.inf.ufpr.br/vri/databases/ufpr-admr/*.

Mask R-CNN was proposed for pointer recognition in [14], [21]. Fang et. al. [14] used it to find reference key points and the pointer in a gauge scale marks, while He et. al. [21] focused on segmenting the meter dial and pointer. In both works, the angle between the pointer and the dial was explored to retrieve the reading. The datasets used in the experiments were not provided in both works.

Region-based Fully Convolutional Networks (R-FCNs) were used for meter detection [12]. Although the authors used deep learning for detection, the meter reading was performed with handcrafted methods such as binarization, line detection, and skeleton extraction. Liu et. al. [10] evaluated Fast R-CNN, Faster R-CNN, YOLO and SSD for meter detection and concluded that even though Faster R-CNN outperforms the others, YOLO is the fastest. Nevertheless, the recognition was performed by a handcrafted method (i.e., HT) and the images used have not been made publicly available.

C. Deep Learning Methods

ResNet [22] is one of the recent breakthroughs in deep networks. The introduction of residual blocks enabled deeper network architectures while having fewer parameters than shallower networks, such as VGG19 [23]. ResNet also performs better and converges faster. The residual learning process introduces lower level features directly to higher abstraction layers, preserving information. ResNet was later upgraded to ResNeXt [30]. The main difference between them is the concept of "cardinality"; instead of going deeper, ResNeXt uses a multi-branch architecture (cardinality refers to the number of branches used) to increase the transformations and achieve a higher representation power. ResNet and ResNeXt can be employed for recognition (classification) problems.

In order to detect the dials on each image, object detection deep networks will be explored. Faster R-CNN [24] is a state-of-the-art approach that uses attention mechanisms and the sharing of convolutional features between the Region Proposal Network (RPN) and the detection network (originally VGG16) to enhance speed and accuracy. First, the RPN generates region proposals that may contain known objects; then, the detection network evaluates the boundaries and classifies the objects.

Redmon et al. [25] proposed YOLO (You Only Look Once), an object detector that focuses on an extreme speed/accuracy trade-off by dividing the input image into regions and predicting bounding boxes and probabilities for each region. YOLOv2 [29], an improved version of YOLO, adopts a series of concepts (e.g., anchor boxes, batch normalization, etc.) from existing works along with novel concepts to improve YOLO's accuracy while making it faster [31]. Similarly, Redmon and Farhadi [26] introduced YOLOv3 (the latest version of YOLO), which uses various tricks to improve training and increase performance, such as residual blocks, shortcut connections, and upsampling. YOLO-based models have been successfully applied in several research areas [4], [32], [33].

D. Datasets

Most of the referred works do not provide a public dataset to enable a fair comparison of results. There are a few publicly available meter datasets [4], [19], [27], however, none of them have images containing pointer-type meters, only digit-based ones. As far as we know, there is no publicly available dataset containing images of dial meters.

III. THE UFPR-ADMR DATASET

We acquired the meter images from Copel, a company of the Brazilian electricity sector that serves more than 4 million consuming units per month [4], [9]. The images of the meters were obtained at the consuming units by Copel employees using cell phone cameras (note that cell phones of many brands and models were used). All images had already been resized and compressed for storage, resulting in images of 640×480 or 480×640 pixels (depending on the orientation in which the image was taken). To create the UFPR-ADMR dataset, we selected 2,000 images where it was possible for a human to recognize the correct reading of the meter, as the images were acquired in uncontrolled environments and it would not be possible to label the correct reading in many cases.

In each image, we manually labeled the position (x, y) of each corner of an irregular quadrilateral that contains all the dials. These corner annotations can be used to rectify the image patch containing the dials. Fig. 2 shows some images selected for the dataset as well as illustrations of the annotations.

Fig. 2. Examples of the images chosen for the dataset. In the bottom row, there are examples of the annotations provided for each image: in green the irregular surrounding quadrilateral, in blue the bounding boxes around the dials and in red the maximal ellipse contained in the bounding box. Note that the customer meter identification is blurred for keeping subject privacy.

All meters have 4 or 5 dials, being 903 meters (45%) with 4 dials and 1,097 meters (55%) with 5. The values pointed on each dial have an almost uniform distribution of digits, having slightly more 0s than other digits. Information about

the dimensions of the meters and dials in the dataset are shown in Table I. Note the great variability in the size of both meters and dials, for example, the smallest dial (20×29 pixels) is almost 10 times smaller than the largest one (206×201 pixels).

	Min (px) $W \times H$	Max (px) $W \times H$	Mean (px) $W \times H$	Mean Area (px^2)
Meters	96×37	632×336	326×121	42,296
Dials	20×29	206×201	88×86	8,328

Fig. 3 illustrates the distribution of digits per dial. The most prominent bar indicates that the most frequent digit in the first position is 0. Nevertheless, it should be noted that the distribution is not as unbalanced as datasets with digit-based meter images, such as the UFPR-AMR dataset [4], in which the number of 0s in the first position is equal to the sum of 0s in the other positions. This is probably due to the fact that dial meters stopped being manufactured and deployed decades ago, which implies that each dial might have completed many cycles since the installation and may be indicating any value.
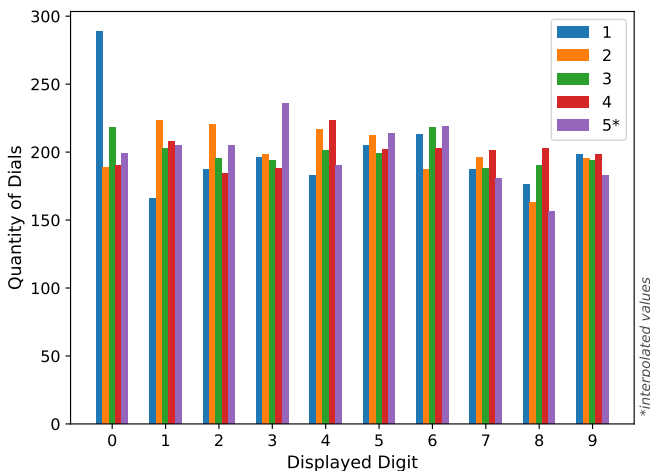

Fig. 3. The distribution of digits according to the dial position on the meter. As 45% of the meters in the proposed dataset do not have a 5th dial, the values of the 5th dial quantities were interpolated proportionally for better visualization.

Table II shows the frequency of digits in the UFPR-ADMR dataset. Unlike datasets containing digit-based meters [4], [19], which were manufactured/deployed more recently, the distribution of the digits is almost uniform across our dataset.

				Frequency / Digit Distribution					
0	1	2	3	4	5	6	7	8	9
996	913	899	906	929	936	942	872	818	886

A. Challenges

The main challenge of the proposed dataset is the quality of the images. Low-end cameras, challenging environmental conditions and high compression are factors that have a high impact on the final image quality. The challenging environmental conditions include: reflections, dirt, and broken glass, and low-quality acquisition may result in: noisy, blurred and low-contrast images. Fig. 4 illustrates the main image-quality issues described above.

In addition to the aforementioned quality issues, there are several types of meter templates and each manufacturer has its own dial model (with variations on the marks) and pointer design. This variations combined with the image capture angle make it difficult to determine the exact pointed value.

Another challenge arises from the presence of clockwise and counter-clockwise dials – for design purposes, each meter has alternating clock directions –, and the direction of the dials may differ depending on the meter model and manufacturer.

B. Evaluation Protocol and Metrics

An evaluation protocol is necessary to enable fair comparison between different approaches. The dataset was randomly divided in three disjoints subsets: 1200 images for training (60%), 400 images for validation (20%) and the remaining 400 images for testing (20%). Following recent works in which datasets were introduced [4], [19], [34], the subsets generated are explicitly available along with the UFPR-ADMR dataset.

To assess the recognition, three metrics are proposed: (i) dial recognition rate, (ii) meter recognition rate, and (iii) mean absolute error. As the main task is to correctly recognize the meter reading, which is a sequence of digits, the meter recognition rate consists of the comparison between the predicted sequence ($pred_m$) and the ground-truth sequence (gt_m), for each of the N meters:

$$MR_{rate} = \frac{1}{N} \sum_{m=1}^{N} match(pred_m, gt_m) \quad (1)$$

$$match(x, y) = \begin{cases} 1, & \text{if } x = y, \\ 0, & \text{if } x \neq y. \end{cases}$$

For the dial recognition rate, we employed the Levenshtein distance (also known as edit distance), a common measurement for computing distance between two sequences of characters. The Levenshtein distance measures the minimum number of edits (addition, removal or replacement of characters) required to transform one sequence in the other. Levenshtein distance is suitable for our evaluation since it can handle small sequence errors in sequences that other metrics would treat as a big error. For instance, if we have a ground-truth sequence $a =$ "1234" and a prediction sequence $b =$ "234", a per-character evaluation metric would consider the error equal to 4, while Levenshtein distance is equal to 1, as the difference between them is a single digit prediction. The Levenshtein distance between the sequences a and b can be determined using:

$$lev_{a,b}(i, j) = \begin{cases} max(i, j), & if \ min(i, j) = 0, \\ lev'_{a,b}(i, j) & \text{otherwise,} \end{cases}$$

(a) uneven lighting   (b) blur   (c) distant capture   (d) reflections

(e) dirt   (f) glare   (g) broken glass

Fig. 4. Samples of the challenging scenarios present on the provided images. We selected for the UFPR-ADMR dataset 2,000 images in which it was possible for a human to recognize the correct reading of the meter. We blurred the region containing the consumer unit number in each image due to privacy constraints.

where:

$$lev'_{a,b}(i,j) = min \begin{cases} lev_{a,b}(i-1,j)+1 \\ lev_{a,b}(i,j-1)+1 \\ lev_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)}. \end{cases}$$

The Levenshtein distance between the prediction and the ground truth is computed and then divided by the longest sequence size between them. This gives us the error. Subtracting the error from 1 gives us the dial recognition rate for each meter. Finally, the mean of all recognition rates yields the total dial recognition error:

$$DR_{rate} = \frac{1}{N} \sum_{m=1}^{N} \left(1 - \frac{lev_{(pred_m, gt_m)}(|pred_m|, |gt_m|)}{max(|pred_m|, |gt_m|)} \right) \quad (2)$$

Considering that the sequence of digits that composes the meter reading is, in fact, a number (integer), correctly predicting the last digit in the sequence is not as important as correctly predicting the first one (i.e., the most significant digit). In order to differentiate and penalize errors in the most significant digits, the mean absolute error is simple yet effective. After converting the sequences to integers ($pred_m$ and gt_m become the integers p_m and g_m, respectively), the mean absolute error can be obtained using:

$$MA_{error} = \frac{1}{N} \sum_{m=1}^{N} |p_m - g_m| \quad (3)$$

IV. EVALUATED APPROACH

We chose two deep networks to evaluate: Faster R-CNN and YOLO. The reason for treating dial meter reading as a detection problem arises from the previous successful approaches to AMR using detection networks [4], [14]. Faster R-CNN presented accurate results in several detection and recognition problems in the literature, while YOLO achieved reasonable results with a high rate of frames per second (FPS), improving the viability of mobile applications. As illustrated in Fig. 5, the proposed pipeline consists of (i) image acquisition, (ii) dial detection and recognition, and (iii) final reading.
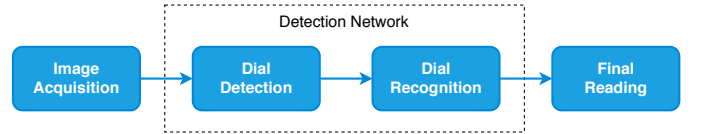

Fig. 5. The main steps to perform dial meter reading.

A. Dial Detection

We perform dial detection directly in the input images, that is, without first detecting the ROI. According to our experiments, presented in section V, this approach achieves the highest F-score value. In other words, our recognition results are not significantly influenced by minor errors on the detection stage, making ROI detection avoidable.

B. Dial Recognition

Faster R-CNN is evaluated with the following residual networks as backbones replacing VGG [23]: ResNet-50 [22] (with 50 convolutional layers), ResNet-101 [22] (with 101 convolutional layers) and ResNeXt-101 [30]. According to [22], ResNets outperform VGG and other several networks in classification tasks; therefore, they are used in our experiments.

For the YOLO-based models, we use the classifiers proposed along with the networks in [26], [29]. YOLOv2 uses the Darknet-19 model as its backbone, which has 19 convolutional layers (hence the name) and 5 max-pooling layers. YOLOv3, on the other hand, uses a network called Darknet-53 (with 53 convolutional layers) for feature extraction; Darknet-53 can be seen as a hybrid approach between Darknet-19 and residual networks [26]. We employed both YOLOv2 and YOLOv3 models in our experiments in order to assess their speed/accuracy trade-off for this task.

C. Final Reading

The final reading is generated according to the position of the detected dial on the image (from leftmost to the rightmost dial). Non-maximum suppression is performed using the Intersection over Union (IoU) metric ($IoU > 0.5$) and considering a maximum of 5 dials per image, keeping only the dials predicted with higher confidence in order to avoid false positives.

V. Experimental Results

We evaluated the performance of the models based on YOLO and Faster R-CNN to detect and recognize the dials simultaneously (note that we used pre-trained weights when fine-tuning both networks). We performed our experiments on a CPU with a Quad-Core AMD Opteron 8387 2.8GHz processor, 64GB of RAM and an NVIDIA Titan Xp GPU. In order to stop the training process and select the best model for testing, we chose the *mean Average Precision* (mAP) evaluation metric, which been commonly employed on object detection tasks [24]–[26]. The mAP can be calculated as follows:

$$mAP = \frac{1}{c} \sum_{i=1}^{c} AP_i \,, \qquad (4)$$

where AP_i stands for the average precision value (for recall values from 0 to 1) of the i-th class.

A. Data Augmentation

We generated new images by creating small variations to the training images to increase the generalization power of the networks. Based on preliminary experiments carried out on the validation set, we generated seven times the number of training images (the combined number of original and augmented images was 9,600). The following transformations were randomly chosen for each image: random scaling [−20%, 20%], random translation [−20%, 20%], random rotation [−15°, 15°] and random shear [−12%, 12%]. The values, which are relative to the original size and position of the images, were chosen randomly within the defined intervals.

B. Evaluation

First, we investigate the performance of the models in the dial detection task. The results are listed in Table III. For comparison, a common method proposed in the literature was evaluated: HCT [11], [20]. F-score was chosen as the evaluation metric, as it is often used to assess detection tasks. As expected, deep learning-based methods (i.e., YOLO and Faster R-CNN) outperformed HCT, reaching very high F-score values. HCT did not cope well with the large variations on lighting, contrast and perspective found in our dataset images.

TABLE III
DIAL DETECTION RESULTS ACHIEVED ON THE UFPR-ADMR DATASET.

Detection Model	Backbone	(%)		
		Prec.	Recall	F-score
Hough Circle Transform	-	53.27	55.28	54.25
Fast-YOLOv3	Darknet	99.94	100.0	99.97
YOLOv3	**Darknet-53**	**100.0**	**100.0**	**100.0**
Faster R-CNN	ResNet-50	100.0	99.94	99.97
Faster R-CNN	**ResNet-101**	**100.0**	**100.0**	**100.0**
Faster R-CNN	**ResNeXt-101**	**100.0**	**100.0**	**100.0**

We performed the recognition (reading) by combining the recognized digits (from the leftmost to the rightmost) and comparing them with the pointed values, using the metrics described in Section III. The recognition results, as well as the FPS rates obtained, are displayed in Table IV.

TABLE IV
RECOGNITION RATE RESULTS OBTAINED ON THE UFPR-ADMR DATASET.

Method	Input Size	FPS	Recognition (%)		Mean Abs. Error
			Dial	Meter	
Fast-YOLOv2	416 × 416	**244**	79.61	42.25	5382.06
Fast-YOLOv2	608 × 608	145	85.24	51.75	3810.34
Fast-YOLOv3	416 × 416	220	83.27	47.75	6098.27
Fast-YOLOv3	608 × 608	120	86.60	54.25	5183.82
YOLOv2	416 × 416	67	91.42	68.00	2615.23
YOLOv2	608 × 608	40	92.51	71.25	1924.98
YOLOv3	416 × 416	35	93.00	73.75	1685.98
YOLOv3	608 × 608	20	93.38	74.75	1591.16
FR-CNN R-50	800 × 800	13	92.56	72.25	1451.81
FR-CNN R-101	800 × 800	11	92.62	71.75	**1343.29**
FR-CNN X-101	800 × 800	6	**93.60**	**75.25**	1591.77

The best performing method was Faster R-CNN (ResNeXt-101) followed by YOLOv3. Faster R-CNN obtained a 75.25% recognition rate per meter and 93.60% per dial, using 800×800-pixel images. After YOLOv3, Faster R-CNN with ResNet-101 performed better than ResNet-50 considering the recognition rate per dial. Interestingly, ResNet-101 presented a lower hit rate considering the recognition at meter level. The lower hit rate is caused by the fact that ResNet-101 errors were better distributed across the images, while ResNet-50 concentrated the errors on fewer images.

The faster method was Fast-YOLOv2, using 416×416 images, achieving 244 FPS. Although YOLOv3 did not surpass Faster R-CNN (ResNeXt-101) in recognition rates, the FPS rates obtained were three times higher (20 FPS and

6 FPS, respectively). Considering that the recognition rates achieved by YOLOv3 were not far behind, this model showed a promising trade-off between accuracy and speed.

The best method regarding *mean absolute error* was Faster R-CNN (ResNet-101) with an error of 1343.29. This means that the method's errors occurred less frequently (or were smaller) on the most significant digits. Table V confirms this statement, as Faster R-CNN (ResNet-101) had fewer errors in the most significant dial (the leftmost).

TABLE V
DISTRIBUTION OF ERRORS BY DIAL POSITION

Dial Position	Frequency (%)				
	1	2	3	4	5
YOLOv3	25.84	15.44	16.94	25.99	15.79
FR-CNN (R-101)	20.70	18.66	23.91	19.83	16.91
FR-CNN (X-101)	23.50	18.55	22.25	18.04	17.66

Fig. 6 presents some correct prediction results. Note that the Levenshtein distance between every correct prediction and its respective ground-truth annotation always equals 0.
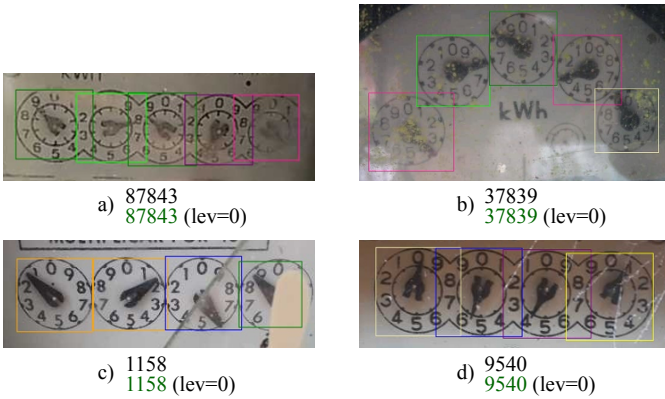

a) 87843
87843 (lev=0)

b) 37839
37839 (lev=0)

c) 1158
1158 (lev=0)

d) 9540
9540 (lev=0)

Fig. 6. Ground-truth and prediction examples of correctly recognized meters, with their respective Levenshtein distance.

C. Error Analysis

The most common errors in the presented approach are caused by:

- **Symmetry:** as there are clockwise and counterclockwise dials, when the digits are blurred, the method can not differentiate the direction and thus may output the mirrored value of the real prediction.
- **Neighbor value:** the most common error. Variables such as angle, lighting, shadows and occlusion (when the pointer is in front of the dial scale mark) can hinder the reading of a dial. Even between the authors, there were some disagreements regarding the correct pointed value in such situations.
- **Severe lighting conditions/Dirt:** shadows, glares, reflections and dirt may confuse the networks, especially in low-contrast images, where those artifacts may emerge

more than the pointer, fooling the network to think that it is a pointer border, resulting in an incorrect prediction.

- **Rotation:** rotated images are harder to predict, as the pointed value is not in the usual position. The predictions may be assigned to neighbor digits that would be in the current angle of the pointer if the image was not rotated.

To illustrate all of the aforementioned causes of errors, some samples are presented in Fig. 7. Table VI summarizes the errors and their frequency on the best two methods: YOLOv3 and Faster R-CNN (ResNeXt-101). Note that most errors are caused by the neighbor values issue, when the pointer is in front of the mark, making it hard to determine if the pointed value is the one after or before the mark.

a) 4062
3061 (lev=2)

b) 01669
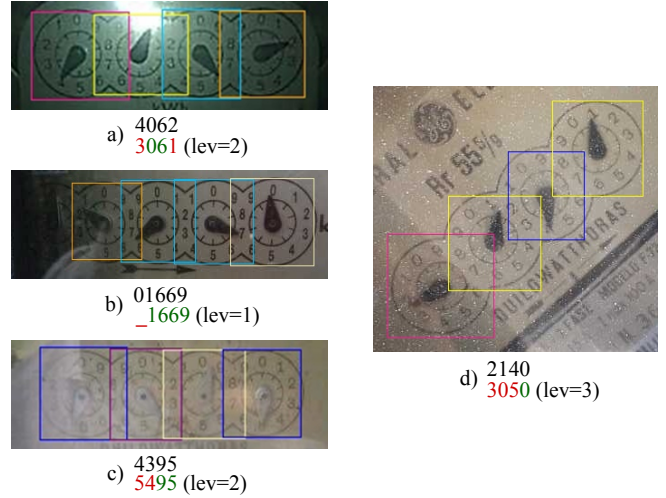_1669 (lev=1)

c) 4395
5495 (lev=2)

d) 2140
3050 (lev=3)

Fig. 7. Ground-truth and prediction examples with their respective Levenshtein distance. The errors are marked in red and include: a) neighbor values, b) severe lighting conditions, c) neighbor value (second dial) and symmetry (first dial); and d) rotation.

TABLE VI
TYPE AND FREQUENCY OF ERRORS OBTAINED ON EVALUATION.

Type of Error	Frequency	
	YOLOv3	FR-CNN (X-101)
Symmetry	2%	3%
Neighbor value	82%	85%
Lighting conditions / Dirt	14%	9%
Rotation	2%	3%

VI. CONCLUSIONS

Imaged-based AMR is a faster and less laborious solution than manual on-site reading, and easier to deploy than the replacement of old meters. In this work, we presented the issues and challenges regarding the automatic reading of dial meters since there are many open challenges in this context.

We introduced a public real-world dataset (shared upon request), called UFPR-ADMR, for automatic dial meter reading, that includes 2,000 fully annotated images acquired on site by employees of one of the largest companies of the Brazilian electricity sector [9]. As far as we know, this is the first

public dataset containing images of dial meters. The proposed dataset contains a well-defined evaluation protocol, which enables a fair comparison of different methods in future works. Considering that the image scenario is challenging in most cases, the deep networks Faster R-CNN and YOLO achieved promising results. This straightforward approach, without ROI detection or image preprocessing, simplified the traditional AMR pipeline [1], [2], [4], reducing the number of steps required to obtain the dial meter readings.

There is a lot of room for improvement, such as new methods to address the boundaries issues between the markers, which should solve most of the errors. In addition, a new loss function that penalizes errors on the leftmost dials should help to reduce the absolute error (minimizing the absolute error is of paramount importance to the service company).

References

[1] M. Vanetti, I. Gallo, and A. Nodari, "Gas meter reading from real world images using a multi-net system," *Pattern Recognition Letters*, vol. 34, no. 5, pp. 519–526, 2013.

[2] I. Gallo, A. Zamberletti, and L. Noce, "Robust angle invariant GAS meter reading," in *International Conference on Digital Image Computing: Techniques and Applications*, Nov 2015, pp. 1–7.

[3] C. Li, Y. Su, R. Yuan, D. Chu, and J. Zhu, "Light-weight spliced convolution network-based automatic water meter reading in smart city," *IEEE Access*, vol. 7, pp. 174 359–174 367, 2019.

[4] R. Laroca, V. Barroso, M. A. Diniz, G. R. Gonçalves, W. R. Schwartz, and D. Menotti, "Convolutional neural networks for automatic meter reading," *Journal of Electronic Imaging*, vol. 28, no. 1, p. 013023, 2019.

[5] U. E. I. Administration. (2019) Electric power annual 2018. [Online]. Available: https://www.eia.gov/electricity/annual/pdf/epa.pdf

[6] Y. Kabalci, "A survey on smart metering and smart grid communication," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 302–318, 2016.

[7] Ausgrid. (2020) Types of meters. [Online]. Available: https://www.ausgrid.com.au/Your-energy-use/Meters/Type-of-meters

[8] Callmepower. (2020) Types of electricity meters. [Online]. Available: https://callmepower.com/useful-information/electricity-meter-types

[9] Copel. (2020) Energy Company Of Paraná. [Online]. Available: http://www.copel.com/hpcopel/english/

[10] Y. Liu, J. Liu, and Y. Ke, "A detection and recognition system of pointer meters in substations based on computer vision," *Measurement*, vol. 152, p. 107333, 2020.

[11] W. Zheng, H. Yin, A. Wang, P. Fu, and B. Liu, "Development of an automatic reading method and software for pointer instruments," in *International Conference on Electronics Instrumentation Information Systems*, June 2017, pp. 1–6.

[12] Y. Huang, X. Dai, and Q. Meng, "An automatic detection and recognition method for pointer-type meters in natural gas stations," in *Chinese Control Conference*, July 2019, pp. 7866–7871.

[13] H. Jiale, L. En, T. Bingjie, and L. Ming, "Reading recognition method of analog measuring instruments based on improved hough transform," in *International Conference on Electronic Measurement Instruments*, vol. 3, Aug 2011, pp. 337–340.

[14] Y. Fang, Y. Dai, G. He, and D. Qi, "A mask RCNN based automatic reading method for pointer meter," in *Chinese Control Conference*, July 2019, pp. 8466–8471.

[15] Y. Tang, C. Ten, C. Wang, and G. Parker, "Extraction of energy information from analog meters using image processing," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 2032–2040, July 2015.

[16] R. Ocampo-Vega *et al.*, "Image processing for automatic reading of electro-mechanical utility meters," in *Mexican International Conference on Artificial Intelligence*, Nov 2013, pp. 164–170.

[17] L. Gómez, M. Rusiñol, and D. Karatzas, "Cutting sayre's knot: Reading scene text without segmentation. Application to utility meters," in *IAPR Intern. Workshop on Document Analysis Systems*, 2018, pp. 97–102.

[18] C. Tsai, T. D. Shou, S. Chen, and J. Hsieh, "Use SSD to detect the digital region in electricity meter," in *International Conference on Machine Learning and Cybernetics (ICMLC)*, July 2019, pp. 1–7.

[19] F. Yang, L. Jin, S. Lai, X. Gao, and Z. Li, "Fully convolutional sequence recognition network for water meter number reading," *IEEE Access*, vol. 7, pp. 11 679–11 687, 2019.

[20] L. Zhang *et al.*, "Pointer-type meter automatic reading from complex environment based on visual saliency," in *International Conference on Wavelet Analysis and Pattern Recognition*, July 2016, pp. 264–269.

[21] P. He, L. Zuo, C. Zhang, and Z. Zhang, "A value recognition algorithm for pointer meter based on improved Mask-RCNN," in *International Conference on Information Science and Technology*, 2019, pp. 108–113.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 770–778.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations (ICLR)*, 2015, pp. 1–12.

[24] S. Ren *et al.*, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[25] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2016, pp. 779–788.

[26] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *arXiv preprint*, vol. arXiv:1804.02767, 2018.

[27] A. Nodari and I. Gallo, "A multi-neural network approach to image detection and segmentation of gas meter counter." in *IAPR Conference on Machine Vision Applications (MVA)*, 2011, pp. 239–242.

[28] W. Liu *et al.*, "SSD: Single shot multibox detector," in *European Conference on Computer Vision (ECCV)*, 2016, pp. 21–37.

[29] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 6517–6525.

[30] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[31] L. Liu *et al.*, "Deep learning for generic object detection: A survey," *International Journal of Computer Vision*, Oct 2019.

[32] R. Laroca, L. A. Zanlorensi, G. R. Gonçalves, E. Todt, W. R. Schwartz, and D. Menotti, "An efficient and layout-independent automatic license plate recognition system based on the YOLO detector," *arXiv preprint*, vol. arXiv:1909.01754, pp. 1–14, 2019.

[33] E. Severo, R. Laroca, C. S. Bezerra, L. A. Zanlorensi, D. Weingaertner, G. Moreira, and D. Menotti, "A benchmark for iris location and a deep learning detector evaluation," in *International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–7.

[34] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz, and D. Menotti, "A robust real-time automatic license plate recognition based on the YOLO detector," in *International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–10.