2019 8th Brazilian Conference on Intelligent Systems (BRACIS)

Image-based Electric Consumption Recognition via Multi-task Learning

Ricardo Costa da Silva Marques
Computer Applied Group (NCA)
Federal University of Maranhão (UFMA)
São Luís, Brazil
ricardo.marques@nca.ufma.br

Arthur Costa Serra
Computer Applied Group (NCA)
Federal University of Maranhão (UFMA)
São Luís, Brazil
arthursrr@nca.ufma.br

João Vitor Ferreira França
Computer Applied Group (NCA)
Federal University of Maranhão (UFMA)
São Luís, Brazil
jvitorfranca@nca.ufma.br

João Otavio Bandeira Diniz
Computer Applied Group (NCA)
Federal University of Maranhão (UFMA)
São Luís, Brazil
joao.bandeira@nca.ufma.br

Geraldo Braz Junior
Computer Applied Group (NCA)
Federal University of Maranhão (UFMA)
São Luís, Brazil
geraldo@nca.ufma.br

João Dallyson Sousa de Almeida
Computer Applied Group (NCA)
Federal University of Maranhão (UFMA)
São Luís, Brazil
Jdallyson@nca.ufma.br

Marcia Izabel Alves da Silva
Maranhão Energy Company (CEMAR)
Equatorial Energia S.A.
São Luís, Brazil
marcia.silva@cemar-ma.com.br

Eliana Marcia Garros Monteiro
Maranhão Energy Company (CEMAR)
Equatorial Energia S.A.
São Luís, Brazil
eliana.monteiro@equatorialenergia.com.br

***Abstract*—This work presents an approach to detect and recognize digits in meter displays by applying the multi-task learning technique. Two convolutional networks are used and divided as tasks to reach different goals. One to detect digits and another to recognize them from different sources, and at the end to return display prediction. In order to validate this methodology, a displays images dataset was labelled. In order to use multi-task learning to reduce the number of failures associated with the reading process. For this, we propose a two-task approach, the first one to detect digits using the networks Faster R-CNN and RetinaNet, and the second stage for recognizing them using Resnet152 network. The initial tests yielded promising results, with 1360 displays, divided into 70% for training and 30% for the test, obtained the following values of Mean Average Precision (mAP), 0.91 (Faster R-CNN) and 0.90 (RetinaNet) on detection and 98.2% accuracy in the classification.**

***Index Terms*—Automatic Meter Reading, Deep Learning, Convolutional Neural Networks, Multi-Task Learning.**

I. Introduction

The study of Automatic Meter Reading (AMR) technologies [20], is becoming increasingly widespread. This approach is a form of automatic data collection of consumption, diagnosis and status of water, gas, electricity, and other meters. Through its use, irregularities, flaws and expenses with periodic visits can be avoided for each location where a meter is located.

In Brazil, much electricity measurement consumption, in general, is done manually. Monthly, electricity distribution companies send their readers (professional responsible for measuring consumption) to their customers' homes to calculate consumer spending. The value registered by the meter is entered into an application that calculates the consumption and generates the invoice, however, in violation cases, the application requests an image for auditing [16].

In this context, the reading process currently uses an image as a way of analyzing read errors. With the intention of avoiding errors in the consumption charged, losses, thefts or frauds. In Brazil, the occurrence of non-technical losses represents a total of 44% of the total annual loss that is of the order of 52 Terawatts/hour (TWh) [22].

Thus, technological solutions with an emphasis on market-oriented AMR systems aiming at overcoming inconsistencies in the consumption measurement in energy distribution services, as well as to reduce the losses caused by frauds [22]. Through this reading process automation human interaction is decreased, as an outcome of computer intelligent systems, with a high level of reliability in the reading by image.

Therefore, the purpose of the present work is to construct a multi-task learning methodology for digits detection and recognition in analogical and digital power consumption meters. For this, we investigate a two-task approach, the first one to detect digits using the Faster R-CNN and RetinaNet networks, and the recognition stage using Resnet152 network.

Due to the great image database diversity of ownership of Maranhão Energy Company (CEMAR), the main contributions of this work are the following: a) the use of multi-task approach in deep networks, b) application of this technique in the problem of digits detection and recognition from electric meters, c) create a model able to infer in and d) investigate the use of Faster R-CNN and RetinaNet networks in this problem

2643-6264/19/$31.00 ©2019 IEEE
DOI 10.1109/BRACIS.2019.00080

Authorized licensed use limited to: Universita degli Studi di Napoli Federico II. Downloaded on September 15,2025 at 01:32:17 UTC from IEEE Xplore. Restrictions apply.

with the ResNet152 network as a backbone.

II. Related Work

A. Image Processing

To solve problems related to AMR, the approaches encountered in the literature generally make use of image processing and machine learning techniques, respectively, for the tasks of detecting and classifying digits in analogic meters.

Threshold techniques for the purpose of extracting digits are commonly used in feature set form [14], [1] and [15]. In [6] and [26] thresholding methods are also used but based on the seven-segment digits characteristics mixed with Histogram of Oriented Gradients and Template Matching techniques for digits detection and subsequent classification. In the work of [9] a solution with adaptive thresholding is proposed, and the classification is done by using mathematical morphology.

B. Deep Learning

Convolutional neural networks (CNN) [8] also emerge as an alternative for detection and recognition. [4] and [7] respectively use CNN techniques with Bidirectional Long Short-Term Memory (BLSTM) [24] and Tesseract OCR [13], both for analog-type meters. Following this line [23] are based on two segmentation methodologies using CNN and Recurrent Neural Network (RNN) [19], to develop a technique called Fully Convolutional Sequence Recognition Network (FCSRN) that recognizes digits in consumption.

The subject of digit recognition in meters is relatively new and driven mainly by the associated industry. Because it is a problem that embodies concepts of detection and recognition, solutions proposed in literature are usually validated on a restricted dataset. Among the works cited, three stand out that use CNN models as a solution.

In this work, we mean to use the multi-task learning technique through the combination of two networks, one for digit detection in meters of electrical consumption and another for classification. In this way, it is aiming to create a generic recognition model that can handle meters of both types (analogic and digital). For this purpose, the objective would be to apply networks briefly explored in AMR problems. Aiming to generate a model with good discriminatory power.

III. Multi-Task Learning

This section presents multi-task learning definition, detection and recognition networks used by this work, as well as the key points for understanding the proposed methodology.

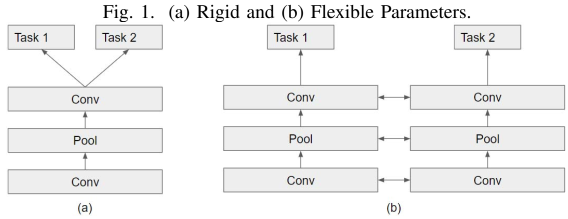
Multi-Task Learning (MTL) is a paradigm in machine learning, and its purpose is to intensify useful information contained in various related tasks to help increase the generalization performance of all tasks [25].

The MTL configuration is analogous to Transfer Learn [12], however, the main difference is which MTL gives the same priority level for all tasks. Since in Transfer Learn, only a target task is prioritized and the others serve as resource providers.

A. Rigid and Flexible Parameters

In deep learning, there are two most widely used ways to accomplish multi-task learning in deep neural networks. In this context, multitasking learning is usually done by sharing rigid or flexible layer parameters (Figure 1).

Fig. 1. (a) Rigid and (b) Flexible Parameters.



Rigid parameter sharing is the most commonly used approach for MTL in neural networks [3] and this approach was used to create our model. It is usually applied by sharing hidden layers across all tasks while maintaining multiple job-specific output layers.

In order to understand the MTL characterization, following the definition, given *m* learning tasks $\{\tau_i\}_{i=1}^{m}$ where all tasks or a subset of them are related, MTL seeks to improve the learning model for τ_i using knowledge contained in all or some tasks.

By the MTL definition, with a focus on supervised learning, usually a task τ_i is accompanied by a dataset D_i which consists of n_i training examples, that is:

$$D_i = \{x_j^i, y_j^i\}_{j=1}^{n_i} \tag{1}$$

where $x_j^i \; \epsilon R^{d_i}$ is the *j*-th training instance at τ and y_j^i its label. The X^i denotes the training data matrix for τ_i, ie $X^i = (x_1^i, ..., x_{n_i}^i)$. When different tasks share the same training data, $X^i = X^j$ for $i \neq j$, MTL can be reduced to multi-label learning or multi-output regression.

IV. Object Detection with CNN

In general, there are two different approaches to performing detection: single-stage and two-stage detection. Using single-stage, a fixed amount of bounding boxes predictions take advantage of a proposition model to find objects. In the second type, another network is used to adjust propositions and produce a final prediction, this pattern is known as two-stage detection [10].

The main goal of this type of task is to detect instances in a group of objects with a predefined location and to discover all objective locations of all objects in image, for later validation with markup using a comparison between bounded box and ground truth.

A. Faster Region-based Convolutional Network (Faster R-CNN)

Region Proposal Network (RPN) directly generate region proposals, predict bounding boxes and detect objects. Faster R-CNN combines RPN and Fast R-CNN [18], and detects objects on a wide variety of scales and proportions.

Authorized licensed use limited to: Universita degli Studi di Napoli Federico II. Downloaded on September 15,2025 at 01:32:17 UTC from IEEE Xplore. Restrictions apply.

To train RPNs, we assign binary labels to differentiate objects from non-objects in each anchor. Thus, labels are assigned to two types of anchor: (1) anchors with the highest Intersect over Union (IoU) value overlapping a bounding box, (2) anchors with IoU greater than 0.7 compared to any ground truth.

Note that a single ground-truth box can assign positive labels to multiple anchors. Thus, a negative label is assigned to an anchor if its IoU is less than 0.3 for all ground-truth boxes. Anchors that are neither positive nor negative do not contribute to training purpose.

From these definitions, the multi-task function loss [18] is minimized. Thus, loss is defined by the equation 4.

$$A = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) \tag{2}$$

$$B = \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \tag{3}$$

$$L(\{p_i\}, \{t_i\}) = A + B \tag{4}$$

since, i is the index of an anchor in a mini-batch and p_i is the probability of an i anchor being an object. The ground-truth p_i^* class will be 1 if the anchor is positive and 0 if negative. The t_i vector represents the four coordinates that fit the predicted bounding box and (t_i^*) is a ground-truth associated with a positive anchor. Therefore, the classification error L_{cls} is used to differentiate the classes (object and non-object).

For the regression loss, it is used $L_{reg}(t_i, t_i^*) = R(t_i - t_i^*)$. Such that, R is the loss defined in [18]. The term $p_i^* L_{reg}$ refers to the regression loss that will be activated only for positive anchors ($p_i^* = 1$) and deactivates for the opposite ($p_i^* = 0$). Therefore, the outputs layers $class$ and reg are $\{p_i\}$ and $\{t_i\}$ respectively. The two terms are normalized with N_{class} and N_{reg}, and with weight balancing λ.

B. RetinaNet

RetinaNet is a network composed of backbone and two subnets with specific tasks. The backbone is responsible for computing a feature map of an entire input image. The first subnet performs object classification at the backbone output, the second subnet performs bounding box regression [11]. The components of RetinaNet will be described below.

As a backbone to RetinaNet, the Feature Pyramid Network (FPN) [10] network is used by default to generate a multi-scale convolutional pyramid. The FPN is used to apply two subnets, one for anchors classification and another for bounding boxes regression.

In short, the backbone of this network increases the standard convolutional capacity because the top-down structure and side connections efficiently construct a multi-scale resource pyramid from a single resolution input image [10]. Each pyramid level can be used to detect objects at different scales.

Bounding boxes of anchors similar to those of the RPN variant in [10] are used. Anchors have areas of 32^2 for 512^2, which increase as pyramid levels rise. At each pyramid anchors level are used in three aspect ratios {1: 2, 1: 1, 2: 1 }. In total, there are A = 9 pyramid anchors and at all levels they cover the range of 32-813 pixels in relation to network input image.

The classification subnet predicts the probability of object presence in each spatial position of each anchor A and K object classes. This subnet is a fully connected layer, inserted at each pyramid level. The parameters of this subnet are shared across all pyramid levels.

The bounding boxes regression subnet is used in parallel with the object classification subnet. Thus, another fully-connected layer is added at each pyramid level in order to regress the compensation of each anchor bounding box to an object near the ground-truth, if one exists. The regression sub-network is identical to the classification subnet, except that it has a linear output layer. Classification and regression networks, although they share a common structure, use separate parameters.

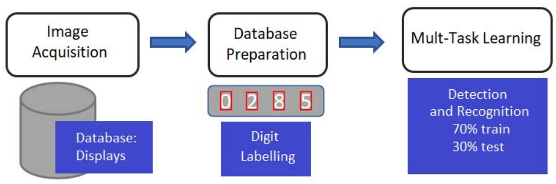
The main advantage of using this network comes from the Focal Loss concept. For this thought is designed to deal with classes imbalance, occurrence of descriptive examples taking stock. This method concentrates the training in a more heterogeneous set, the model has two main parameters to be adjusted:

- α: balancing of focal loss on unbalanced data;
- γ: smooth setting at the rate at which easy examples are weighted down. If $\gamma = 0$, the focal loss is equivalent to cross validation, when γ increases weighting factor effect also increases, since this factor is very sensitive to the data set.

V. Methodology

This section presents the methodology that must be followed to fulfil the proposed objectives. This methodology is summarized in 3 steps, starting with the acquisition of electric meter displays, database preparation through a process of marking bounding boxes of digits in displays, followed by the multi-task learning stage that is divided into two fronts, digit detection and recognition (2).

Fig. 2. Stages of proposed methodology.



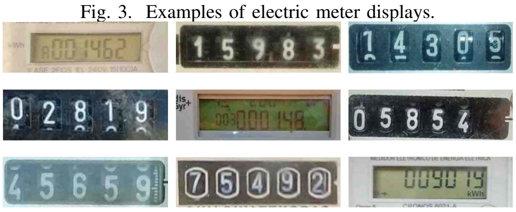
A. Image Acquisition

The database used is composed of analogic and digital meter display images. A set of 10 thousand images of displays were acquired through the SIVAL system , which was developed by Applied Computer Core (NCA) and Cemar Electric Power Plant (CEMAR) [17]. This dataset used is private, with the appropriate permission, for the purpose of constructing and improving recognition algorithms. In dataset there are 63

Authorized licensed use limited to: Universita degli Studi di Napoli Federico II. Downloaded on September 15,2025 at 01:32:17 UTC from IEEE Xplore. Restrictions apply.

different meter models with divided 12 digit sources. Some examples of images are shown in Figure 3.

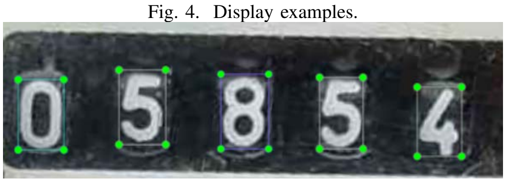
Fig. 3. Examples of electric meter displays.



B. Database Preparation

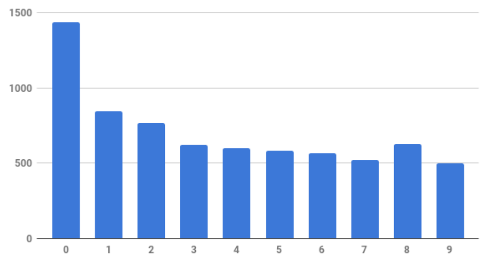
For experiments accomplishment, it is necessary to annotate the displays, locating the digits and informing which digit is in that location. Initially, 1360 displays were annotated according to the standard used by the proposed detection networks, using the graphics annotation tool LabelImg [21]. Annotations are saved in an XML file in Pascal VOC format, the format used by ImageNet [5]. An example of marking with the tool can be seen in Figure 4.

Fig. 4. Display examples.



The data set was divided into 10 labels, 0 to 9, numbered from 1 to 10. Note that there is a zero digit imbalance in relation to another classes, this is due to a large number of meters where the first two or three digits are number zero sequences. The digits are separated into 10 different classes: according to the digits arrangement, as shown in the figure graph 5.

Fig. 5. Class layout by digit.



It is assigned to expand the annotated database as a way of validating the proposed methodology with a wider and varied range of cases.

C. Multi-Task Learning

Once the base is labelled and prepared for the experiments, this methodology analyzes the use of Faster R-CNN and RetinaNet routines for digit detection and recognition.

In the experiment using the Faster R-CNN network, the Resnet152 network was used as the backbone (classification layer). First, the images pass through the convolution layer and the feature maps are extracted. Then, a sliding window is used in the RPN for each location in the resource map. For each site, anchors k (k = 9) are used in 3 scales (128, 256 and 512) and 3 aspect ratios (1: 1, 1: 2, 2: 1) to generate region proposals. A class layer generates $n = 100$ detections, whether or not it has an object for n bounding boxes. The reg regression layer generates four values (coordinates of box center, width and height) of n bounding boxes. Thus, the feature map size $W \times H$, has WHk anchors in total. Finally, the loss is calculated as described in Section IV-A, in Equation 4.

In the experiment using RetinaNet, it is composed of a backbone network and two subnets specific to the task of detection and recognition. The backbone is responsible for calculating a feature map of an input image of a complete input image. The first subnet makes the classification backbone output and the second one regresses the bounding boxes resulting from the convolutions.

The backbone was constructed with a Resnet152, the classification subnet predicts the probability of object presence in each spatial for each anchors position A and object classes K. And it uses an input feature map with C channels of a pyramid level, the subnet applies four 3x3 conv layers, each with C filters and each one followed by ReLU activations. Finally, sigmoid activations are attached to outputs. Focal loss is applied as the loss function.

In training, network initialization is very important. The authors use an initial error probability of 0.01 for all anchors and assign it to the last convolution layer of the classification subnet.

With this, the weight initialization is set at 0.0001, learning rate of 0.01, it is used for the first 60 thousand iterations. In this way, the learning rate reduces by 10 after 60 thousand iterations and again at 80k iterations [11]. Inference, the main predictions from all network levels are merged and non-maximum suppression is applied with a limit of 0.85 to generate the least bounding boxes without overlapping.

VI. Results

This section presents the results obtained with the methodology applied in the digits detection and recognition in electric meter displays. Experiments were performed using the 70% training ratio because in general deep neural networks require a large number of training examples and 30% in the test. Five tests replicates were performed.

To validate the detection results we measure the mean average precision and the intersections over union metrics. IoU measure is the overlap between two regions. This metric is used to predict boundary overlaps with a ground truth (Equation 5).

$$IoU = \frac{AreaOfOverlap}{AreaOfUnion} \tag{5}$$

Authorized licensed use limited to: Universita degli Studi di Napoli Federico II. Downloaded on September 15,2025 at 01:32:17 UTC from IEEE Xplore. Restrictions apply.

Average precision is calculated from precision (frequency at which an object is correctly classified) and recall (frequency at which examples of a class are found), i.e. the precision in all recall values (Equation 6).

$$AP = \frac{1}{N} \sum_{recall_i} precision(recall_i) \tag{6}$$

Methodology execution on the annotated database obtained as results in the detection values presented in Table I. This table shows the methodology results in terms of mAP, on the total digits and in relation to the total of completely detected displays, using different IoU values. The accuracy in the classification is presented in Table II, this presents the comparison with other methodologies. The inference time of the Faster R-CNN and RetinaNet networks are, respectively, 86ms and 193ms per image in average.

TABLE I
INITIAL METHODOLOGY RESULTS WITH TWO IOU VALUES (0.50 AND 0.75) TO EVALUATE NETWORKS PERFORMANCE PRESENTED PER DIGIT AND PER DISPLAY.

Digit	Faster R-CNN (IoU=0.5)	Faster R-CNN (IoU=0.75)	RetinaNet (IoU=0.5)	RetinaNet (IoU=0.75)
0	0.929	0.9583	0.809	0.9195
1	0.926	0.9574	0.832	0.8188
2	0.934	0.9365	0.800	0.9114
3	0.915	0.8981	0.827	0.8820
4	0.934	0.9283	0.792	0.9115
5	0.892	0.9246	0.807	0.9114
6	0.926	0.9249	0.931	0.9159
7	0.919	0.9479	0.819	0.9092
8	0.920	0.9405	0.800	0.9137
9	0.927	0.9550	0.814	0.9483
By Display	**0.913**	0.9020	0.802	**0.9042**

TABLE II
COMPARISON WITH OTHER WORKS.

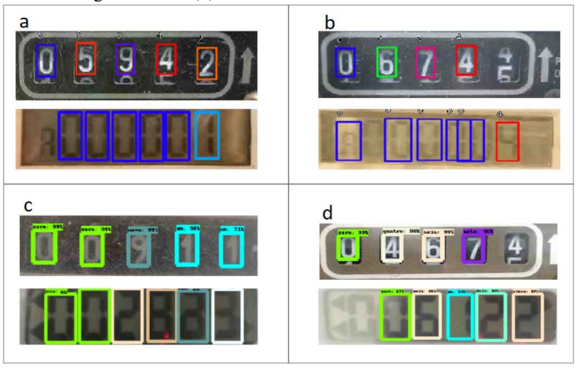
Work	Technique	Study	Database	Accuracy
[4]	CNN + Tesseract OCR	Analogic Water Meters	Private	80.33%
[7]	CNN + BLSTM	Analogic Water Meters	Private	96%
[6]	Template Matching + HOG	Digital Electrical Meters	Private	73%
[23]	FCSRN	Analogic Water Meters	Water-Meter-Number-DataSet	96%
Our Method	**MTL (ResNet152)**	**Digital and analogic electric meters**	**Private**	**98.2%**

From the results presented, it was noticed that the network Faster R-CNN obtained the best result when using the value of IoU in 0.5. With the IoU value increased to 0.75, there was a drop in accuracy in the first experiment. On the other hand, in the second experiment, using the RetinaNet network, the results obtained remained more stable, in terms of display, even with IoU variation. The best results were 0.91 mAP with IoU equal to 0.5, with the first experiment, and 0.90 mAP with IoU equal to 0.75 in the second.

The variation in the value of IoU directly influenced the result by controlling the number of bounding boxes detected, by increasing the value of IoU more regions can be captured, thus the more chances of false detections occurring. With this, in the first there was a decrease in the result because it is more accurate, with IoU 0.5, 1k bounding boxes are generated, by increasing the IoU to 0.7 the amount of regions increases to 2.5k, the second network generates more regions of interest in relation to the first, respectively 100k and 120k for the two defined IoU values.

However, the first network was more stable with the IoU variation, while the second one had a significant increase in the result. Note that the inference by the Faster R-CNN network is more accurate. In this case, it is possible to verify the results obtained in this work. The results show that some images were generated with different situations present in the base (Figure 6) in analogical and digital displays.

Fig. 6. Methodology results for cases of error and correctness with the best results used:(a) RetinaNet hit cases, (b) RetinaNet error cases, (c) Faster R-CNN matching cases and (d) Faster R-CNN error cases.



VII. CONCLUSION AND FUTURE WORK

This work presented a methodology with the use of the multi-task learning strategy, for the tasks of digits detection and recognition from digital and analogic meters. The initial tests were obtained as the best result under the mAP metric, popularly used to validate the accuracy of object detection networks. The values obtained were 0.91 with the Faster R-CNN network and 0.90 for RetinaNet.

In digit recognition, separate approaches are commonly used to classify analogical and digital digits. However, with the use of multi-task learning to share information between detection and reconnaissance networks, a single model was able to infer with good precision against digital and analogical meters, even with Class 0 imbalance in relation to others.

Authorized licensed use limited to: Universita degli Studi di Napoli Federico II. Downloaded on September 15,2025 at 01:32:17 UTC from IEEE Xplore. Restrictions apply.

As future tasks, the proposed methodology intends to use variations of the techniques used, through the search for better network hyper-parameters with HyperOpt optimizer [2]. Examine the use of other detection networks (YOLO, SSD (Single Shot Detector)) and recognition (Inception, MobileNet) to observe the results in relation to the already used approaches. Test approaches to control the bounding boxes of networks with a feature detector such as Maximally Stable Extremal Regions (MSER), and finally, use base-balancing techniques such as Augmentation to increase the amount of relevant data in the data set.

VIII. Acknowledgment

The authors are grateful to Coordination of Superior Level Staff Improvement (CAPES) and also CEMAR/CELPA for the financial support provided through the ANEEL project PD-00037-0036/2019.

References

[1] Al-Ali, A., El-Hag, A., Bahadiri, M., Harbaji, M., and El Haj, Y. A. (2011). Smarthome renewable energy management system. Energy Procedia, 12:120–126.

[2] Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., & Cox, D. D. (2015). Hyperopt: a python library for model selection and hyperparameter optimization. Computational Science & Discovery, 8(1), 014008.

[3] Caruana, R. (1997). Multitask learning. Machine learning, 28(1):41–75.

[4] Cerman, M., Shalunts, G., and Albertini, D. (2016). A mobile recognition system for analog energy meter scanning. In International Symposium on Visual Computing, pages 247–256. Springer.

[5] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., and Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09.

[6] Diniz, P., Marques, R., Figueredo, W., Junior, G. B., de Almeida, J., Silva, A.,de Paiva, A., and Monteiro, E. (2018). Automatic power consumption reading on digital meters based on adaptive thresholds and multiresolution templates. Anais do Computer on the Beach, pages 317–324.

[7] Gao, Y., Zhao, C., Wang, J., and Lu, H. (2017). Automatic watermeter digit recognition on mobile devices. In International Conference on Internet Multimedia Computing and Service, pages 87–95. Springer.

[8] LeCun, Y., Kavukcuoglu, K., and Farabet, C. (2010). Convolutional networks and applications in vision. In Proceedings of 2010 IEEE International Symposium on Circuits and Systems, pages 253–256. IEEE.

[9] Lei, H., Li, L., Zhang, P., Wang, M., and Li, X. (2010). A binarization algorithm for digital meter image based on grayscale morphology. In 2010 3rd International Congress on Image and Signal Processing, volume 5, pages 2427–2429. IEEE.

[10] Lin, T.-Y., Dollar, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017a). Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2117–2125.

[11] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollar, P. (2017b). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision, pages 2980–2988.

[12] Pan, S. J. and Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10):1345–1359.

[13] Patel, C., Patel, A., and Patel, D. (2012). Optical character recognition by open source ocr tool tesseract: A case study. International Journal of Computer Applications, 55(10):50–56.

[14] Popa, M. (2011). Gateway design and implementation in an automatic meter reading system based on power line communications. In Networked Computing and Advanced Information Management (NCM), 2011 7th International Conference on, pages 295–298, Gyeongju, Korea (South). IEEE.

[15] Prapasawad, C., Pornprasitpol, K., and Pora, W. (2012). Development of an automatic meter reading system based on zigbee pro smart energy profile ieee 802.15.4 standard. In Electron Devices and Solid State Circuit (EDSSC), 2012 IEEE International Conference on, pages 1–3, Hong Kong, Thailand. IEEE.

[16] Quintanilha, D. B. P., da Silva, G. L. F., Eduardo, C., Falcão, M., Silva, A. C., de Paiva, A. C., & Márcia, E. DETECCAO AUTOMÁTICA DE MEDIDORES ELÉTRICOS EM IMAGENS UTILIZANDO UMA COMBINAÇÃO DE SVM E CNN.

[17] Reis, A. B. S., de Sousa, J. A., da Silva Neto, O. P., Froz, B. R., da Silva, I. F. S., do Maranhão - CEMAR, C. E., & do Para SA-CELPA, C. E. SEGMENTAÇÃO AUTOMÁTICA DE MEDIDORES ELETROMECÂNICO E ELETRÔNICO BASEADO EM IMAGENS.

[18] Ren, S., He, K., Girshick, R., and Sun, J. (2015a). Faster R-CNN: Towards real-time object detection with region proposal networks. In Cortes, C., Lawrence,N. D., Lee, D. D., Sugiyama, M., and Garnett, R., editors, Advances in Neural Information Processing Systems 28, pages 91–99. Curran Associates, Inc.

[19] Schuster, M. and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45(11):2673–2681.

[20] Shu, D., Ma, S., and Jing, C. (2007). Study of the automatic reading of watt meter based on image processing technology. In 2007 2nd IEEE Conference on Industrial Electronics and Applications, pages 2214–2217. IEEE.

[21] Tzutalin, L. (2019). Git code (2015).

[22] VIDINICH, R. and NERY, G. A. L. (2009). Pesquisa e desenvolvimento contra o furto de energia. Revista Pesquisa e Desenvolvimento da ANEEL–P&D, (3).

[23] Yang, F., Jin, L., Lai, S., Gao, X., and Li, Z. (2019). Fully convolutional sequence recognition network for water meter number reading.IEEE Access, 7:11679–11687.

[24] Zhang, S., Zheng, D., Hu, X., and Yang, M. (2015). Bidirectional long short-termmemory networks for relation classification. In Proceedings of the 29th Pacific Asia conference on language, information and computation, pages 73–78.

[25] Zhang, Y. and Yang, Q. (2017). A survey on multi-task learning. arXiv pre print arXiv:1707.08114.

[26] Zhang, Z.-J., Li, Y., and Yuan, W.-Q. (2010). Meter character recognition method based on gray template matching. In Proceedings of the 29th Chinese Control Conference, pages 2987–2990. IEEE.

Authorized licensed use limited to: Universita degli Studi di Napoli Federico II. Downloaded on September 15,2025 at 01:32:17 UTC from IEEE Xplore. Restrictions apply.