A Chatbot Solution for Self-Reading Energy Consumption via Chatting Applications

Carlos Vinicios Martins Rocha[1] · Arthur Azevedo Lima[1] · Pedro Henrique Carvalho Vieira[1] ·
Carolina Lima Saraiva Cipriano[1] · Anderson Matheus Passos Paiva[1] · Italo Francyles Santos da Silva[1] ·
Simara Vieira da Rocha[1] · Aristófanes Corrêa Silva[1] · Hugo Daniel Castro Silva Nogueira[2] ·
Eliana Márcia Garros Monteiro[2] · Eduardo Camacho Fernandes[2]

Abstract

To mitigate financial loss and follow the recommended sanitary measures due to the COVID-19 pandemic, self-reading, a method in which a consumer reads and reports his own energy consumption, has been presented as an efficient alternative for power companies. In such context, this work presents a solution for self-reading via chatbot in chatting applications. This solution is under development as part of a research and development (R&D) project. It is integrated with a method based on image processing that automatically reads the energy consumption and recognizes the identification code of a meter for validation purposes. Furthermore, all processes utilize cognitive services from the IBM Watson platform to recognize intentions in the dialog with the consumers. The dataset used to validate the proposed method for self-reading contains examples of analogical and digital meters used by Equatorial Energy group. Preliminary results presented accuracies of 77.20% and 84.30%, respectively, for the recognition of complete reading sequences and identification codes in digital meters and accuracies of 89% and 95.20% in the context of analogical meters. Considering both meter types, the method obtains an accuracy per digit of 97%. The proposed method was also evaluated with UFPR-AMR public dataset and achieves a result comparable to the state of the art.

Keywords Energy · COVID-19 · Self-reading · Chatbot · Image processing · Deep learning

1 Introduction

The electric energy distribution process in Brazil has a relevant complexity. Studies done by Empresa de Pesquisa Energética (EPE) show that, in 2018, almost 30% of the national energy consumption is related to residential con-

✉ Italo Francyles Santos da Silva
  francyles@nca.ufma.br

  Carlos Vinicios Martins Rocha
  carlos.martins@nca.ufma.br

  Arthur Azevedo Lima
  arthur@nca.ufma.br

  Pedro Henrique Carvalho Vieira
  pedrocarvalho@nca.ufma.br

  Carolina Lima Saraiva Cipriano
  carol@nca.ufma.br

  Anderson Matheus Passos Paiva
  andersonpaiva@nca.ufma.br

  Simara Vieira da Rocha
  simara@nca.ufma.br

  Aristófanes Corrêa Silva
  ari@nca.ufma.br

  Hugo Daniel Castro Silva Nogueira
  hugo.nogueira@equatorial.com.br

  Eliana Márcia Garros Monteiro
  eliana.monteiro@equatorial.com.br

  Eduardo Camacho Fernandes
  eduardo.fernandes@equatorial.com.br

[1] Federal University of Maranhão, São Luís, Brazil

[2] Equatorial Energy Group, São Luís, Brazil

sumption[1]. This is marked by the frequent occurrence of financial losses, such as energy theft or mistakes in the reading and monthly billing processes.

Electric power companies perform consumption reading and inspection through the meter readers. They are employees responsible for collecting meter readings directly at the consumer's residences. Since it is done manually, this process becomes susceptible to mistakes. Moreover, this practice may bring risks to the health of the meter readers, due to exposure to climate variations and security issues. It is also important to highlight that visits to homes by readers are not recommended according to the social distancing norms elaborated to contain the COVID-19 pandemic.

In this situation, the Equatorial Energy group and other power companies have been searching for a more efficient alternative for the energy consumption reading to mitigate risks, costs and combat irregularities. The use of smart energy meters emerges as a possibility. However, this solution has a high financial cost, and its implementation time becomes impracticable in the short term due to the number of meters to be replaced. Another alternative is self-reading, in witch the reading process is done by the consumers themselves via digital platforms, such as mobile devices, through which the information regarding consumption is sent to the company to be processed and validated. This practice has received support from the Brazilian Electricity Regulatory Agency (ANEEL) through normative resolution[2].

By placing the consumer as an integral part of the reading process, self-reading strengthens the relationship between them and the company. Thus, it is intended to reduce the costs related to this process, while mitigating the occurrence of errors and fraud, especially in areas of difficult access and inspection, such as rural areas. It is worth highlighting that self-reading respects social distance and other public safety measures resulting from the COVID-19 pandemic.

The mass use of mobile devices is another strong motive for the development of a self-reading solution. In Brazil, over 60% of the population uses smartphones. This is one of the highest ratios among the emerging economies (Pew Research Center 2019). Chatting applications are the most popular among the young and adult population, with considerable adherence from the elderly (Nielsen IBOPE 2015). In this context, several companies have developed virtual assistants (also called chatbots) to work via chatting applications aiming to speed up and facilitate service provision through automated customer service (Panorama Mobile Time 2019).

Equatorial Energy group has a chatbot application integrated with WhatsApp[3] through which services like invoice

emission and requesting repairs to the electrical network are offered. This application is already used by a expressive part of the consumers in the Equatorial Energy's coverage areas. So, this motivated the development of a chatbot solution for self-reading to be integrated with that application. As the operation of that is well known by consumers, the self-reading process may be understood and spread more easily.

Thus, this work presents the development of a chatbot solution for the energy consumption self-reading process via chatting applications. It consists of the establishment of a dialogue with the assistant, during which the consumer will be requested to send a picture of his meter. This picture is sent to a server where it will be processed by a method of automatic recognition of the digits that compose the meter reading. This method is based on convolutional neural networks (CNNs) combined with an ensemble of classifiers in order to obtain more accurate results.

And, to guarantee further reliability to the process, the method has another CNN-based stage that performs the recognition of meter identification code for validation of the reading. The chatbot solution integrated with the proposed method can be used in the major chatting applications available, therefore favoring better provision of services related to self-reading. It is also important to mention that the proposed solution is aimed at low tension residential consumers (LT) in the Brazilian states of Maranhão and Pará. The total number of LT consumers is 2,575,871 in Maranhão and 2,445,947 in Pará.

The main contributions presented in this work are: (1) an automatic energy consumption reading method based on image processing, which combines convolutional neural networks and an ensemble of classifiers; (2) a refinement process necessary to reduce false positives and find candidate digits into the display region; (3) an approach to identification code recognition using convolutional neural network, aiming to guarantee safety to the self-reading process; and (4) the integration of the proposed method in a feasible chatbot solution to be used by consumers in the context of self-reading.

Regarding the application scenario, the proposed solution aims to speed up the reading process, mainly in regions more distant from urban centers, such as rural areas. Furthermore, this solution is in the line with the social distance and other sanitary measures due to the COVID-19 pandemic, as it reduces the need for a large number of meter readers on the field, valuing the safety of these employees and also consumers.

This work is structured as follows: Section 2 presents related works; Sect. 3 details the development of the chatbot and the proposed method for reading recognition and validation; Sect. 4 presents the results and their discussion; lastly, Sect. 5 presents the conclusions about the proposed solution.

---

[1] Available in http://www2.aneel.gov.br/arquivos/pdf/atlas_par1_cap2.pdf

[2] Available in http://www2.aneel.gov.br/cedoc/ren2020878.pdf

[3] Available on https://www.whatsapp.com/

2 Related Works

In the literature, some works are found presenting image-based method for automatic energy consumption reading. Generally, these works use fixed cameras pointed to the meter in order to obtain a picture of the *display* region and, afterward, use image processing and machine learning to perform the reading through digit recognition (Parthiban and Palanisamy 2013; Zhang et al. 2016).

Quintanilha et al. (2017) proposed a method for digit recognition based on Histogram of Oriented Gradients (HoG) feature extraction along with the support vector machine (SVM) classifier (Cortes and Vapnik 1995). The authors used an image dataset provided by the Equatorial Energy group, containing only analogical meter samples, obtaining an accuracy of 79,52%. An expanded version of this dataset was used to develop the proposed method, with the addition of digital meter samples.

Shuo et al. (2019) proposed a hybrid approach, combining *deep learning* and traditional algorithms for meter digit recognition. This method is divided into stages. Initially, MobileNet V2 is used (Sandler et al. 2018) to identify the display and crop it. This region is submitted to image smoothing and binarization techniques in order to segment digits, which are then forwarded to a support vector machine (SVM) classifier. These experiments were performed with a private dataset, limiting the method's applicability in another scenario due to the number of parameters and thresholds to be adjusted according to the dataset features. The general accuracy obtained was 88.67% in a test with 300 meter images.

Some works address the automatic meter reading using solely *deep learning* techniques and focus on industry-oriented solutions. These will be addressed next.

Calefati et al. (2019) proposed an end-to-end method based on the use of traditional CNN architectures. It is also divided into two stages: detection and recognition, each one using a specific CNN, aiming, finally, to predict the length of the reading sequence and its value. This method was tested with a private dataset that contains images of different meter models (gas, water and electricity). This method achieves an accuracy of 85.7%, a promising result, given the ample variety of meter types.

Laroca et al. (2019) present an image dataset of analogical and digital meters with varying models, called UFPR-ARM. This dataset contains images captured under unrestrained conditions; therefore, meters may be damaged or they present other noise elements. This work also shows a method for automatic reading via digit recognition with deep learning. The method is composed of two stages: counter detection and recognition. For this task, authors have performed experiments with approaches based on (Redmon et al. 2015) called Fast YOLO and CR-Net, obtaining 94,13% and 98,30% accuracies, respectively, for counter detection and recognition.

Inspired by this last work, Azeem et al. (2020) proposed a method for automatic meter reading using Mask-RCNN. The backbone for feature extraction is a GoogLeNet (Szegedy et al. 2015), and the method was also divided into stages: meter display detection followed by digit segmentation and recognition. For each stage, a network model was trained, in order to reduce the complexity of the process. Experiments were also performed with the UFPR-AMR dataset achieving an a accuracy of 99.86% for digit recognition.

Similarly to the mentioned works, the method proposed by the current work divides the automatic reading process into stages and uses *deep learning* approaches for each one. This method uses different RetinaNet models applied to detect regions of interest in analogical and digital meters, along with a classifier ensemble for digit recognition. However, the related works do not address automatic reading in the context of mobile devices.

Regarding self-reading, there are works focused on the context of mobile devices (Serra et al. 2020a; Mendes et al. 2020; Serra et al. 2020b). However, these methods were elaborated to be incorporated into an application, and their execution requires processing power and storage from the device. On the other hand, in the solution proposed in this work, the image recognition method is executed server side, which is a more robust environment. It is worth highlighting that this solution runs on popular chatting applications, facilitating its use, given that consumers are more used to its operation. Lastly, it is also important to highlight that all the aforementioned works do not have a stage for the recognition of the meter identification code to guarantee the safety of the process.

3 Proposed Solution

This section presents the structural aspects of the proposed solution as well as more information about the developed method for self-reading by chatting apps.

3.1 Structural Aspects of the Chatbot

The proposed solution consists of modules. They are: message management, main module, cognitive services, web services and inference server as seen in Fig. 1.

The message management module is responsible for exchanging messages between the main module and the chatting app. This communication takes place using application-specific RESTFul APIs, for instance, Telegram[4] and Twillio[5] (for Whatsapp). These APIs allow the creation of a communication channel for redirecting messages.

---

[4] Available in https://core.telegram.org/

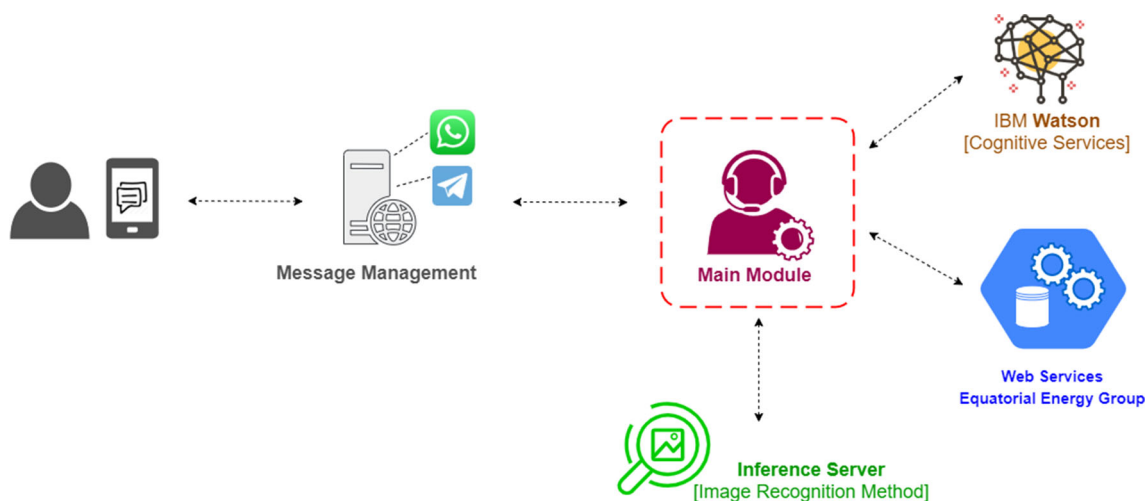[5] Available in https://www.twilio.com/

Fig. 1 Proposed solution: an overview

The virtual assistant's main module receives messages and sends them for the cognitive services module that is responsible for interpreting the dialogue through the IBM Watson platform[6]. This platform contains functions used for creating chatbot solutions that are able to recognize intentions in dialogues with users. In case of the proposed solution, those functions determine what actions the chatbot should perform, for example, to verify or validate the consumer's data by making requests to the web services module that is connected to the Equatorial Group's database. Another example is to request a meter image for a consumer to continue the self-reading process.

The inference server module receives an image sent by the consumer and submits it to the proposed method for reading recognition and validation. This method is based on image processing and deep learning techniques. The environment where this server runs is encapsulated in a Docker container[7] where a Tensorflow Serving[8] is configured. This container is used for running the method for reading recognition and validation whose implementation is based on Tensoflow libraries. Together with Docker, Tensorflow Serving provides services through an API, making the operation of the inference server more simplified and efficient.

In addition, the referred container enables the implementation of the proposed solution, as it maintains the dependencies of the solution and ensures portability to the environment.

Finally, the result generated by the method is sent to the main module to be presented to the consumer, who can correct any recognition errors. The self-reading process ends after the validation step that performs the recognition of the meter identification code (tag). Then, these pieces of infor-

mation will be sent to the Equatorial Group through web services for the billing process and other validations.

The self-reading service for monthly billing will only be available on a specified date. At the beginning of the dialog with the chatbot, the database will be checked if the consumer is on his/her billing date. If so, the self-reading process will be continued with the chatbot asking to the consumer to send the image and the execution of the recognition method on the inference server. Otherwise, the consumer will be notified that the service is not available. It is important to mention that there will be no problem if a consumer does not do the self-reading on the correct date. For these cases, an internal process on Equatorial Energy's server will identify consumers who have not taken the reading so that a reader can visit them and collect their readings in the traditional way.

3.2 Cognitive Analysis of Dialogues

The cognitive analysis of the established dialogues between the actors (consumer and chatbot) is based on the relationship between intentions and entities. The intentions are the set of actions that can be performed. And the entities are complementary information that specify the context to which a given intention refers. In sentences, intentions and entities are represented, respectively, as verbs and verbal complements as seen in Fig. 2.

For the proposed solution, the main intentions to be interpreted in the messages received by the chatbot were mapped. Each intention is associated with a wide set of predefined sentences in order to guarantee more flexibility in the dialogues. Thus, it is possible for consumers to express themselves in different ways to request the same service. At the same time, it expands the chatbot's ability to understand dialogues.

---

[6] Available in https://www.ibm.com/cloud/watson-assistant/
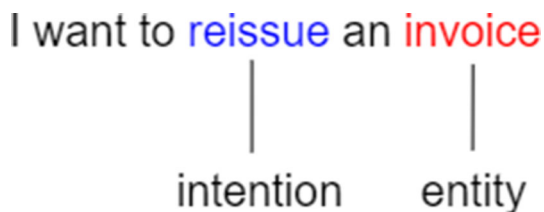
[7] Available in https://www.docker.com/

[8] Available in https://www.tensorflow.org/tfx/guide/serving

Fig. 2 Intentions and entities in a sentence

Table 1 Mapping of intentions and respective examples

Intentions	Examples
Cancellation	"exit", "cancel", ...
Confirmation	"yes", "confirm", "accept", ÿ, ...
Reading	"1248", "3571", ...
Denial	"no", "n", ...
Greeting	"good morning", "hi", "hello"

Fig. 4 Examples of the image dataset

3.3 Image Database

For the development and validation of the reading recognition method, the Equatorial Energy group provided an ample set of energy meter images. These pictures were captured under different orientation and lighting conditions. Meters are separated into two groups: digital and analogical; generally, these meters are protected by a box with a transparent section that allows the readers to see the display digits. Some examples of the images are seen in Fig. 4.

The dataset contains 7513 meter images: 4216 samples for analogical meters and 3297 for digital meters. It is worth mentioning that occasionally, the protective box may have been damaged by the action of external agents. This may hinder the visualization of the meters and other components. However, images with these characteristics were not discarded, as it is necessary to elaborate a product applicable to real scenarios, where such adversities should be considered. Each sample of the dataset is associated with a .xml file that contains the coordinates of bounding boxes which define the objects of interest in the image, which are the meter; the display; the display digits; the identification code (tag); and the tag digits. Figure 5 shows these annotations.

3.4 Recognition Method and Reading Validation

As previously stated, the chatbot requests the consumer to send a picture of their meter. This image is directed to the inference server where the proposed method, whose stages are shown in Fig. 6, is executed.

The first stage is called component detection, in which the meter, and, afterward, the display and tag are detected. These
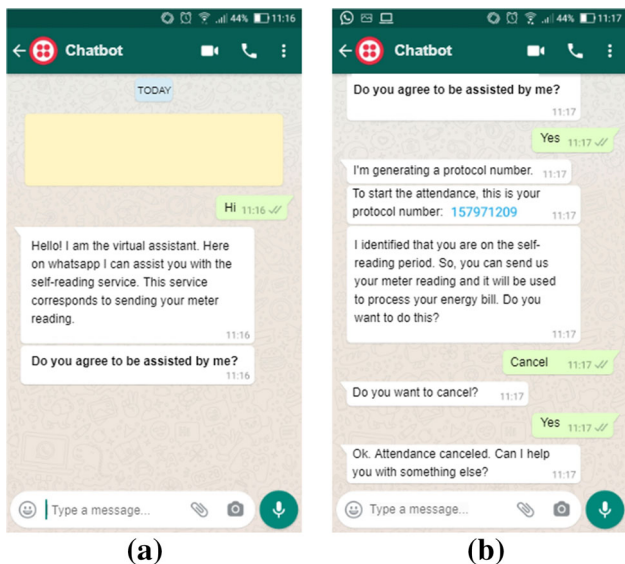

Fig. 3 Example of dialogue between consumer and chatbot when the service is initialized **A** and cancelled **B**

Table 1 shows the mapping of intentions performed, as well as their respective sets of words and sentences. It is important to note that the proposed solution has only one entity (self-reading process); therefore, all intentions refer to it.

Figure 3 exemplifies a conversation between the actors in two scenarios **A** and **B**. In 3(A), a greeting made by the consumer starts the service. 3(B) presents the cancellation of the process. It is worth mentioning that, if the chatbot does not identify the intention of the dialogue, it asks the consumer to send a more enlightening message.
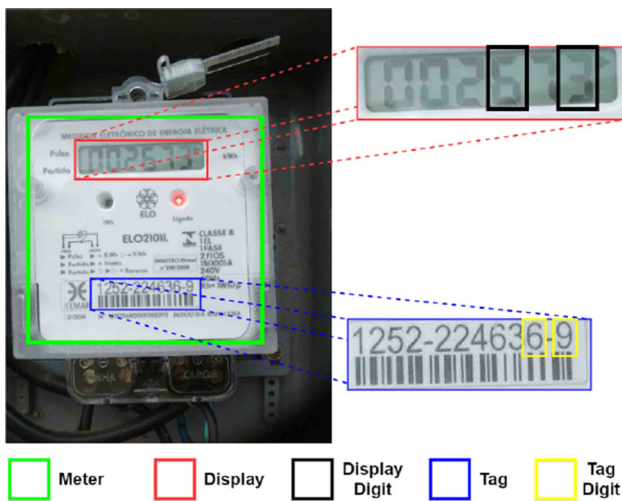
Fig. 5 Examples of the image dataset annotations

last two, once identified, will be submitted, respectively, to the reading and tag digit recognition stages. The proposed method will be detailed shortly.

3.4.1 Component Detection

The component detection stage consists in to verify the presence of the meter, the display, and the tag in the submitted image, using the convolutional network Retina Net (Lin et al. 2017). This network was chosen due to the promising results found in the literature for object detection. This architecture is composed of three parts: backbone; the classification sub-network; and the bounding box generation sub-network.

The backbone is a feature extraction module composed of one or more convolutional networks. Retina Net uses a combination between the ResNet (He et al. 2016) networks and a Feature Pyramid Network (FPN) as its default backbone responsible for generating feature maps of multiple scales. This makes the architecture robust for detecting objects of different sizes, which is desirable in this stage, given the difference in the dimensions of each component (meter, display and tag). Finally, the generated maps are used as input to the classification sub-network, which predicts the probability of an object being present in an image, and to the bounding box generator, which defines the coordinates of the bounding box delimiting the detected object.

To decrease the chances of failure due to different types of meters, two Retina Nets were trained: one for analogical meters and the other for digital meters. Both are activated simultaneously, so that the network that detects the components of interest with the higher confidence ratio leads the remaining processes, which were also separated and specialized for each meter type. In this stage, the backbones chosen are ResNet-152 for analogical meters and ResNet-101 for digital meters.

The result of component detection is decisive to the execution of other stages of the proposed method. If the networks detect the meter, display, and tag, the last two will be cropped from the original image using the generated bounding boxes to produce new images, which will be submitted, respectively, to the reading digit and tag digit recognition stages. Otherwise, the chatbot will send a message to the consumer requesting a new meter image be sent.
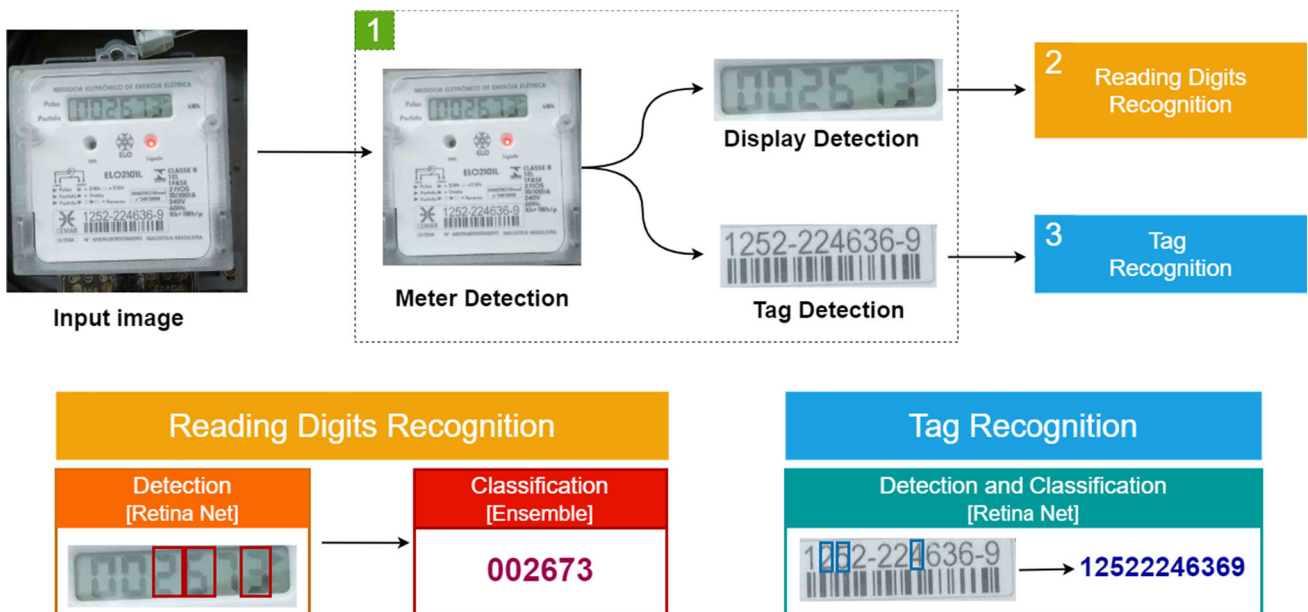

Fig. 6 Proposed method for reading recognition and validation

3.4.2 Reading Digits Recognition

The reading digit recognition stage consists of two steps. First, the display region is submitted to the Retina Net for the detection of digits as a single class (digit). Afterward, each detected digit is used as input to a classification process where, finally, they are classified from 0 to 9.

For digit detection, two networks are trained, one for each meter type. This is necessary due to the difference in characteristics between the digits of both types (Fig. 7). Analogical meters (7.A) present great variability due to the differences between models and manufacturers. However, in the case of digital meters (7.B), digits tend to be more homogeneous.

Once again, ResNet-101 and ResNet-152 backbones are used for analogical and digital meters, respectively. During the execution of the method, the network that obtains the best confidence ratio for the component detection stage determines the type of meter being analyzed and, therefore, which network will be selected for digit detection.

As a result, the network identifies the class digit generating the respective bounding boxes. Nevertheless, the Retina Net output may present certain inconsistencies that need to be treated. Thus, the execution of a refinement process is necessary to obtain more accurate results. This refinement stage is composed of a set of actions that can be visualized in Figure 8.

First, the overlap problem is fixed. This problem is common in detection methods, and it occurs when the network produces two or more bounding boxes for the same object classifying it incorrectly. To solve that, the proposed method executes the following steps:

1. Calculate the area of the smallest bounding box generated by a given digit;
2. Define the intersection between the bounding boxes that delimit the digit;
3. Select the intersection bounding box whose area is N% larger than the area calculated on step 1.

Figure 9 illustrates this process. Based on tests performed, N was defined as 25%. The intersection (green) is considered valid since its area is (at minimum) 25% larger than the bounding box (red) that delimits the digit. At last, those considered invalid are eliminated.

After the overlap removal, the next step is the false-positive (FP) reduction. As mentioned in Sect. 3.1, the chatbot accesses the consumer data, retrieving important meter details such as the number of digits shown on display (Q). Then, the digits detected with over 70% confidence are considered true positives (TP). Otherwise, they are classified as FP. Thus, in the case of a network presenting Q detections classified as TP, these are immediately submitted to the classification process. However, if the number of TP is inferior
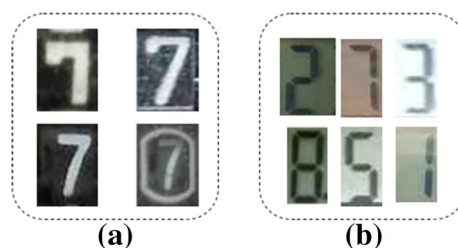

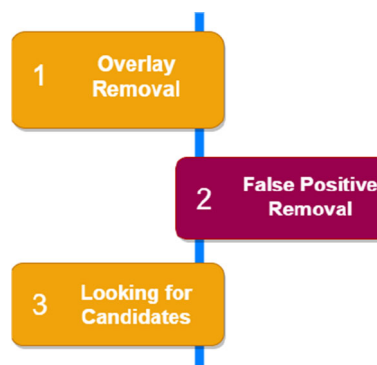**Fig. 7** Examples of differences between digit types in analogical **A** and digital **B** meters

Fig. 8 Stages of the refinement process

Fig. 9 Overlap removal

to Q, the execution flow of the refinement will be directed to the last stage.

The candidate searching is an improvement implemented to localize digits that were not detected, but may be in the display image. Figure 10 shows how that process occurs. Initially, the average width of the bounding boxes is calculated. Afterward, for each consecutively detected pair of digits, the distance between them is compared to the average width. If that distance is greater than the average width, then there is a gap. And so, the number of digits that can fit in that gap is calculated. From this information, bounding boxes are generated in the detected gap, considering the height variation and horizontal distance between consecutive digits when there is no gap.

The candidate searching is also used to locate possible digits missing on the edges of the display. In order to do that, it must be verified if the digits detected on the left and right extremities are not too close to the edges of the display and if the total detection (VP+FP) is less than Q. If those conditions are satisfied, the previously described strategy is
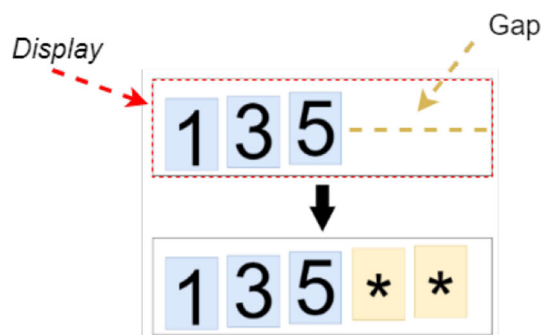
Fig. 10 Example of the execution of candidate searching stage

(a)

(b)

Fig. 11 Examples of meter tags

used to fill the gap, taking into account the height variation and horizontal distance.

Finally, all detected digits are submitted to the classification process that uses an ensemble of classifiers composed of support vector machine (SVM) (Cortes and Vapnik 1995), Xgboost (Chen and Guestrin 2016) and Efficient Net (Tan and Le 2019). The ensemble takes the cropped images of each digit and classifies them from 0 to 9.

As a prior step to classification via SVM and Xgboost, Histogram of Oriented Gradients (HoG) is used to extract features from the digit images (Quintanilha et al. 2017). Efficient Net does not require such treatment, as it is a convolutional network. Thus, the ensemble consists in verifying the majority vote. So, if at least two out of the three classifiers point to a digit belonging to a class, this result will prevail. However, if there is a total divergence, the highest confidence ratio will be considered.

At the end of all of these processes, the proposed method returns a recognized reading sequence to the chatbot, which will send this result to the consumer to confirm the reading or possibly correct mistakes. After this confirmation, the method initiates the reading validation procedure, with the recognition of tag digits.

3.4.3 Tag Digits Recognition

In a scenario where consumers become responsible for reading their own energy consumption, it is necessary to develop mechanisms to prevent ill-intentioned individuals from committing fraud, for example, by sending an image of someone else's meter.

Therefore, to make the self-reading process safer, the proposed method counts on a reading validation stage that consists in the recognition of digits within the detected tag, as shown in Sect. 3.4.1.

As mentioned before, meters may be damaged by external agents. These damages may also affect tags, as shown in Fig. 11. In (A) there is a tag in good conditions, while in (B) damages are present. Existing libraries for codebar recogni-

tion, such as ZBar [9], have high noise sensitivity, which is not desirable for the proposed method.

Thus, an approach for digit recognition was implemented in which the detected tag region is passed as input to a Retina Net with a ResNet-101 as backbone. This network performs simultaneously digit detection and classification. Therefore, in addition to localizing the bounding boxes, the Retina Net classifies each digit from 0 to 9. At this stage, there was no need to incorporate the ensemble, because, despite the different meter types, tag digits have similar characteristics.

Finally, the chatbot searches for the true tag of the consumer on the Equatorial Energy database through web services. This value will be compared to the digits recognized by the proposed method and, if matches, the reading will be validated and sent to the servers of the Equatorial Energy group for the billing process. Otherwise, the consumer will be notified that his reading was not validated and a bill will not be generated.

4 Results and Discussion

The modules of the proposed solution are under development. However, the proposed method for reading recognition and validation integrated with the inference server is at a more advanced stage. Thus, this section presents the results obtained with the experiment carried out with this method.

The 7513 samples from the dataset were divided into two sets as follows: 70% for training and 30% for testing. As mentioned in Sect. 3, the dataset is heterogeneous. In addition to the differences between digital and analogical types, there are also intra-type divergences according to the meter manufacturers. Therefore, it is necessary to compose the training set with a relevant number of examples of each meter type. In

---

[9] Available at http://zbar.sourceforge.net/

Table 2 mAP results obtained for component detection

	Analogical	Digital
Meters	97.43%	96.88%
Displays	99.57%	96.94%
Tags	96.19%	91.01%
Overall	97.73%	94.94%

Table 3 Accuracy results obtained for the reading digits recognition process

	Analogical	Digital
Digits	97.06%	97.83%
Reading sequences	89.00%	77.20%

the case of the test set, a large number of images are ensured in order to simulate the real application scenario which is characterized by the heterogeneity of the meters.

Retina Net models were trained with Adam optimizer (Kingma and Ba 2014) and learning rate of 10^{-5} for 50 epochs using the Early Stopping technique (Prechelt 1998). Data augmentation was not used. The training processes are done offline in an environment apart from the one that keeps the proposed solution running. Therefore, it does not cause any performance problem or interruption of services. Once a network or classifier model is trained and validated, it is necessary incorporate it into the recognition module (inference server).

For the training of the ensemble classifiers, cropped images of digits from the dataset were used. The SVM and Xgboost parameters were estimated using the grid search algorithm. SVM uses the RBF kernel. Regarding the Efficient Net, the proposed method uses the B0 architecture and this network was trained for 50 epochs using the RMSProp optimizer (Igel and Hüsken 2000) and learning rate = $2x10^{-5}$.

For evaluation, mean average precision (mAP) and accuracy were used. The former is widely used to evaluate methods applied for object detection and the latter for classification problems. Table 2 shows the obtained results for the component detection (meters, displays and tags).

The obtained mAP values are grater than 0.9. They show that the proposed method reaches a satisfactory performance for all meter types. For meter detection, for example, there is a small statistical difference related to the mAP values obtained for analogical and digital meters. For the detection of displays and tags, this is not repeated, as the mAP values are better for analogical meters. It was expected because, in the case of analogical meters, these components are not very different from each other despite the different manufacturers. Besides, displays and tags are always found in the same location. On the other hand, digital meters can change drastically according to the manufacturer, which makes the component detection more difficult in this case.

In relation to the reading digits recognition, the classification process was evaluated by means of accuracy by reading sequence and by digit. The reading sequence is the complete digit sequence shown in the meter display. Table 3 shows the obtained results.

As seen, there is an evident difference between the results achieved for analogical and digital meters. The main cause for this is related to some problems in the images acquisition. Digital meters are more prone to interference from external light that may distort the appearance of the digits on the display depending on the angle of capture. These distortions may cause misclassifications.

In Table 3, it is also possible to observe a relevant difference between accuracy per digit and per reading sequence in the two types of meters. This is due to the imbalance of the digit classes in the dataset. So, the classifier learns more about one class from another. And when a complete reading sequence is analyzed, a misclassified digit is enough to indicate that the sequence is entirely wrong. This situation is directly reflected in the accuracy calculation for complete sequences.

Still within the scope of reading recognition, Fig. 12 shows how the chatbot solution presents in the chatting app the self-reading result from an image sent by the consumer.

Regarding the tag recognition process, the obtained results are shown in Table 4. The proposed method achieves a very good result for this process yielding accuracies over 90% for the most cases. Once again, a difference is observed between accuracy per digit and the complete tag sequence. This is also due to the imbalance of the digit classes. However, the classification error is more evident in the case of digital meters.

Table 5 shows a comparative analysis between the proposed method and related works that evaluate their methods using the same dataset provided by Equatorial Energy group. That analysis uses the accuracy per reading digit because those works do not calculate that metric for the reading sequence and tag sequence. The approach proposed by Quintanilha et al. (2017) is directed to analogical meters. Serra et al. (2019) proposed a method applied to all types of meter. Those works use the SVM classifier combined with a feature descriptor. On the other hand, the proposed method uses a combination of a Retina Net model and an ensemble of classifiers that presents better results than the aforementioned works.

To verify the applicability of the proposed method in another context, an experiment was carried out with the UFPR-AMR public data set for reading digits recognition. This dataset was used by Laroca et al. (2019) and Azeem et al. (2020) to evaluate their approaches in contexts similar to the self-reading scenario. It should be noted that, in this

Fig. 12 The chatbot solution presenting in the chatting app the self-reading result from an image sent by the consumer

Table 4 Results for the tag recognition process

	Analogical	Digital
Digits	95.53%	92.49%
Tag sequence	95.20%	84.30%

experiment, only display detection (called counter detection by the aforementioned works) and digit reading recognition were tested, as these works did not include reading validation by tag recognition.

Table 6 shows the obtained results and a comparative analysis between the performance achieved by the proposed method and that presented by the related works. For a fair comparison, accuracy is used to evaluate the results. The UFPR-AMR dataset is divided into three subsets (training, validation and testing) as default. So, this was kept. In addition, the hyperparameters used to train of Retina Net and the ensemble were also maintained.

The proposed method presented expressive results when compared to the related works. For display detection, an accuracy of 98.32% was achieved surpassing the result obtained by Laroca et al. (2019). For the reading digits recognition, the proposed method achieved 97.10%, a result statistically close to the related works, showing a difference of about 2%

Table 5 Comparative analysis between the proposed method and related works that use the dataset provided by Equatorial Energy group. Metric: accuracy per digit

	Analogical	Digital
Quintanilha et al. (2017)	79.52%	-
Serra et al. (2019)	96.72%	97.90%
Proposed Method	**97.06%**	**97.83%**

Table 6 Experiments performed with the UFPR-AMR public dataset: comparative analysis between the proposed method and the related works

	Accuracy	
Methods	Displays (Counters)	Digits
Azeem et al. (2020)	100%	99.86%
Proposed Method	**98.32%**	97.10%
Laroca et al. (2019)	94.13%	98.30%

in relation to the best accuracy which was obtained by Azeem et al. (2020).

The experiment with the UFPR-AMR dataset shows that the proposed method can also be applied in another context, because even though it was not designed based on the features of that dataset, using the same backbones, classifiers and their hyperparameters, the proposed method achieves relevant results. Noting that, it is also possible to affirm that the proposed chatbot solution could be used in the scenario depicted by the UFPR-AMR base, with the exception of the reading validation step, as this base does not have the appropriate annotations.

In general, the proposed method for reading recognition and validation integrated with the chatbot solution achieves good results as seen in the experiments carried out with two datasets (Equatorial Energy group and UFPR-AMR), in which the method outperforms or present results close to the related works. However, as mentioned above, the processes related to the scenario of application of the proposed solution are under development. Then, the achieved results can be improved.

5 Conclusion

In this work, a chatbot solution was presented for self-reading via chatting applications. The proposed method uses Retina Net networks to, initially, detect the meter, the display, and the tag in the scope of analogical and digital meters. Afterward, based on the results of the previous stage, the detected display is passed as input to the reading digits recognition stage. In this stage, another Retina Net is used and its results are submitted to the classification process, performed by an

ensemble composed of the classifiers SVM, Xgboost and Efficient Net. After the reading recognition, the reading validation stage is initiated. It consists in validating the reading via the recognition of tag digits. This process uses the Retina Net for both detection and classification.

The results presented are promising and show the viability of the chatbot solution. The proposed method obtains accuracies of 77.20% and 84.30%, for recognition of reading and tag, respectively, in digital meters, and 89% and 95.20% in the context of analogical meters. From these results, it is observed that the proposed method has more failures in reading recognition in digital meters. This happens because the displays are more susceptible to interference from lighting, which can affect the visualization of the reading sequence. In the case of analogical meters, errors occur when the digits are in transition, being partially occluded. This leads to misclassifications. As previously indicated, all evaluated scenarios consider different meter models. Therefore, despite the limitations indicated, the method showed encouraging results.

In the case of tag recognition, the obtained results show that the idea of classifying digits instead of recognizing the bar code pattern is a viable solution. It is an important result in the self-reading scenario where the bar code can be more damaged than the tag digits. Thus, it is possible to ensure more safety to the self-reading process. This is encouraging, given the urgency for solutions that prioritize the accomplishment of recommended sanitary norms in the COVID-19 pandemic, but that also extend their benefits to the post-pandemic, reflected on the decrease in fraud, reading mistakes, and, consequently, financial losses.

Compared to the related works that experimented with the same dataset provided by the Equatorial Energy group, the proposed method obtains the best accuracy results for reading digits in both analog and digital meters. And, in experiments with the public database UFPR-AMR, the method obtains relevant accuracy results in display (counter) and digit recognition, comparable to the best approaches evaluated in this dataset.

As future works, it is intended to continue experiments to improve results in reading and tag recognition, testing other classifier configurations. It is also intended to evaluate the performance of the proposed method in other datasets to ensure its generalization. Moreover, the proposed chatbot solution will be integrated with the chatbot APIs already used by the Equatorial Energy group.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Azeem, A., Riaz, W., Siddique, A., & Saifullah, U. A. K. (2020). A robust automatic meter reading system based on mask-rcnn. (pp. 209–213)

Calefati, A., Gallo, I., & Nawaz, S. (2019). Reading meter numbers in the wild. (pp. 1–6)

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. (pp. 785–794)

Cortes, C., & Vapnik, V. (1995). *Support-vector networks. Machine learning, 20*(3), 273–297.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. (pp. 770–778)

Igel, C., & Hüsken, M. (2000). Improving the rprop learning algorithm. In: Proceedings of the second international ICSC symposium on neural computation (NC 2000), Citeseer, vol 2000, pp 115–121.

Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization.

Laroca, R., Barroso, V., Diniz, M. A., Gonçalves, G. R., Schwartz, W. R., & Menotti, D. (2019). Convolutional neural networks for automatic meter reading. *Journal of Electronic Imaging, 28*(1), 1–14. https://doi.org/10.1117/1.JEI.28.1.013023

Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. (pp. 2980–2988)

Mendes, P. R. C., Rocha, C. V. M., Paiva, A. M. P., de Souza Oliveira, J. E., da Silva, I. F. S., de Paiva, A. C., da Rocha, S. V., Silva, A. C., Monteiro, E. M. G., da Silva Italo, I. F. S., et al. (2020). Digit recognition of energy meters through voice in the context of an authentication application. *Brazilian Applied Science Review, 4*(3), 1976–1989.

Nielsen IBOPE (2015) 68 milhões usam a internet pelo smartphone no brasil. https://www.nielsen.com/br/pt/press-releases/2015/68-milhoes-usam-a-internet-pelo-smartphone-no-brasil/

Mobile, Panorama. (2019). *Time* Mapa do ecossistema brasileiro de robôs 2019 http://panoramamobiletime.com.br/download/16300/?uid=f312d14021.

Parthiban, K., & Palanisamy, A. (2013). Reading values in electrical meter using image processing techniques. (pp. 1–7)

Pew Research Center (2019) Smartphone ownership is growing rapidly around the world, but not always equally. https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/

Prechelt L (1998) Early stopping-but when? In: Neural Networks: Tricks of the trade, Springer, pp 55–69

Quintanilha, DBP., Costa, RWS., Diniz, JOB., de Almeida, JDS., Braz, G., Silva, AC., de Paiva, AC., Monteiro, EM., Froz, BR., Piheiro, LPA.,& Melho, W. (2017). Automatic consumption reading on electromechanical meters using hog and svm. In: 7th Latin American Conference on Networked and Electronic Media (LACNEM 2017), pp 57–61, https://doi.org/10.1049/ic.2017.0036.

Redmon, J., Divvala, SK., Girshick, RB., & Farhadi, A. (2015). You only look once: Unified, real-time object detection. CoRR arXiv:1506.02640

Sandler, M., Howard, AG., Zhu, M., Zhmoginov, A., & Chen, L. (2018). Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation. CoRR arXiv:1801.04381.

Serra, A., França, J., Marques, R., Figueredo, W., Reis, A., Santos, I., Rocha, S., Silva, A., Monteiro, E., Silva, I., Silva, M., & Santos, J. (2019). *Reconhecimento de dígitos em imagens de medidores de energia no contexto de um aplicativo de autoleitura* https://doi.org/10.17648/sbai-2019-111448.

Serra, A. C., França, J. V. F., da Silva Marques, R. C., Figueredo, W. K. R., Reis, A. B. S., da Silva, I. F. S., da Rocha, S. V., Silva, A. C., Monteiro, E. M. G., da Silva, I. F. S., et al. (2020a). Digit recognition in energy meter images in the context of a self-reading application. *Brazilian Applied Science Review, 4*(3), 1990–2001.

Serra, A. C., França, J. V. F., de Sousa, J. A., de Sousa Costa, R. W., da Silva, I. F. S., da Rocha, S. V., de Paiva, A. C., Silva, A. C., Monteiro, E. M. G., da Silva, I. F. S., et al. (2020b). Semantic segmentation of electricity meters and identification components. *Brazilian Applied Science Review, 4*(3), 2002–2013.

Shuo, H., Ximing, Y., Donghang, L., Shaoli, L., & Yu, P. (2019). Digital recognition of electric meter with deep learning. In: 2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), IEEE, pp 600–607.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. (pp. 1–9)

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In: Chaudhuri K, Salakhutdinov R (eds) Proceedings of the 36th International Conference on Machine Learning, PMLR, Long Beach, California, USA, Proceedings of Machine Learning Research, vol 97, pp 6105–6114, http://proceedings.mlr.press/v97/tan19a.html

Zhang, Y., Yang, S., Su, X., Shi, E., & Zhang, H. (2016). Automatic reading of domestic electric meter: an intelligent device based on image processing and zigbee/ethernet communication. *Journal of Real-Time Image Processing, 12*(1), 133–143.