OBJECT DETECTION ALGORITHMS APPLIED ON LOW VOLTAGE GRID EQUIPMENT

Mohcine EL HARRAS
EDF R&D – FRANCE
mohcine.el-harras@edf.fr

Christophe BIRKLE
Enedis – France
christophe.birkle@enedis.fr

Samuel SALLAUD
EDF R&D - FRANCE
samuel.sallaud@edf.fr

Julien BRUSCHI
EDF R&D – FRANCE
julien.bruschi@edf.fr

ABSTRACT

Distribution system operators (DSO) try continuously to improve the quality of service for customers and increase the safety of third parties and facilities. Yet, failures in technological systems are always bound to happen. An electric distribution box can be broken, a line can touch a tree, etc. This research paper focuses on maintenance. In Enedis, a regional troubleshooting center can be contacted when a grid equipment is considered as damaged or broken. Therefore, Enedis developed FLOE, a mobile application which will allow local authorities to report failures and incidents on the electrical grid. The report is often accompanied with photos to illustrate the situation. Yet, sometimes, it is not easy to identify the network type, users may report non-electric network equipment as gas, water, or telecommunication. As part of this work, an AI-based object detection algorithms have been tested to pre-identify the types of damaged equipment. This paper presents the results obtained.

INTRODUCTION

Enedis has developed a mobile application called FLOE to send automatically report anomalies on its equipment to the Enedis call center. These reports may include several cases reported that do not contain equipment of the network operated by Enedis. Since the appearance of equipment from different networks can be very similar and the people reporting the anomalies may not be specialists, the photos may contain equipment that do not concern Enedis.

The aim of this paper is to present results on **automatic recognition of electric grid equipment** on photos using AI algorithms. The AI component will filter the reports sent to the call center and limit the time wasted on reports related to other types of networks (gas, water, etc.).

An **artificial intelligence** (AI) model is developed and trained to assess if the equipment is electric. It is part of a library developed by Enedis which contains object detection algorithms trained on images of high/low voltage equipment. The work thoroughly compares several deep learning algorithms, ultimately culminating in the selection and implementation of the most effective two which are YOLO [1][6] used for object detection and image segmentation [4] and VGG19 [3] used to perform classification. Mask R-CNN [2] was tested as well for image segmentation, but its results were not as satisfying as YOLO's.

The current work being conducted is a proof of concept, which serves as a demonstration of the potential capabilities and uses of the deep learning model. The goal of this phase is to establish that the model can achieve the desired results, and to gather feedback from our experts to assess the model's performances and make adjustments. Once the results are satisfying, the next step would be to train the model on bigger and more powerful machines to improve its performance.

DATA ANALYSIS

In total, more than 40,000 photos have been annotated. Polygons were drawn around each equipment in each photo. It should be noted that some photos can contain up to 10 equipment. The following figure shows the list of classes and their occurrences.

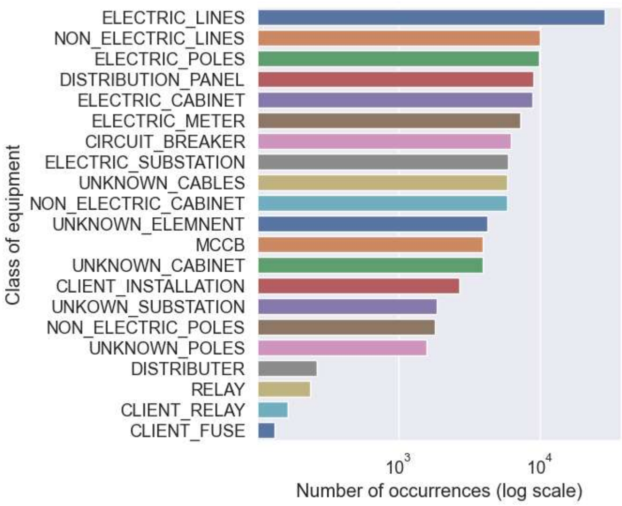

Figure 1- Distribution of classes in annotated images

First, the unknown classes were removed to avoid adding a bias to our model. The classes: Distributer, Relay, Client_Relay and Client_Fuse were also removed due to the lack of data (less than 500 appearances in our dataset). the AI model was trained on the twelve other classes. The results of the training showed that the object detector had difficulties detecting the non-electric equipment. On the one hand, there are a lot of similarities between electric classes and non-electric ones. On the other hand, the model has very few examples of non-electric grid equipment to learn the specific features of these classes. Hence, we decided to merge the electric and non-electric categories into one class. These classes will be disaggregated as soon as the pool of photos will be increased which would provide us with more photos of non-electric equipment. Here is the class distribution after the merge of the dataset that will be used to train the deep learning model:

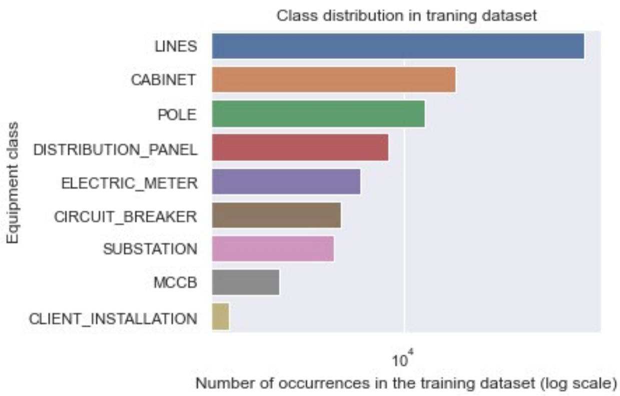

Figure 2- Class distribution on the training dataset (log scale)

APPROACH

The approach used is based on two steps:

- First, object detection and image segmentation: the algorithm generates a rectangle or polygon around the detected objects.
- Then, classification: the algorithm assigns a class to the detected object.

It is to be noted that the **merge** would help **to improve the model's recall** (ability to cover all the objects of a class and not forget any). **The differentiation between electrical and non-electric equipment is made afterwards with a second model**. Hence the approach adopted and illustrated in the figure below.

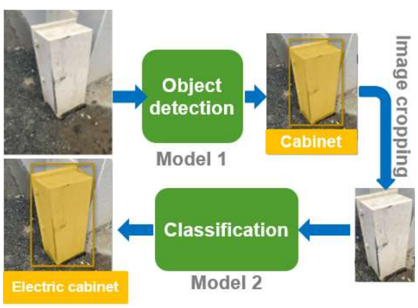

Figure 3- Approach adopted for the differentiation between electrical and non-electric equipment

On the one hand, object detection is a computer vision task that involves identifying the presence and location of objects in an image or video. Several challenges can arise when attempting to perform object detection. Hence, it is important to have images and annotations of a very good quality. On the other hand, image classification refers to the task of assigning a label or class to an input image. To perform image classification, a machine learning model is trained on a large dataset of images that have been labeled with the correct classes. The model learns to recognize patterns and features in the images that are indicative of the class.

TRAINING

Configuration

The models were trained using the following configuration:

- An environment dedicated to object detection with mainly the following libraries: Pytorch 1.7.0 coupled with CUDAToolkit 10.1 and cuDNN 7.6.3.
- Another environment dedicated to performing image classification with Tensorflow 2.3.0, CUDAToolkit 10.1 and cuDNN 7.6.5.
- An Nvidia GRID P40-4Q with a driver version of 418.70 and CUDA 10.1.

Data augmentation

Data augmentation is a technique that can be useful for several reasons. It helps to improve generalization: the model is exposed to a greater number of variations, which can help it to better predict unseen data during testing. In addition, it reduces overfitting which occurs when a model is too closely fit to the training data and does not generalize well to unknown data. It also improves performance: providing the model with additional training examples similar to the original data, but with subtle differences that can help the model learn more robust features. However, if the techniques used are not carefully chosen there is a risk of overfitting the model to the training set.

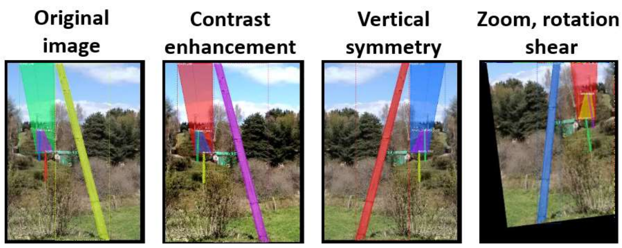

Figure 4- Examples of data augmentation

Transfer learning

Transfer learning is a machine learning technique where a model trained on one task is re-purposed on a second related task. For example, a model that has learned to recognize photos of animals might be repurposed to recognize photos of plants by "transferring" its knowledge of image recognition to the new task. Transfer learning is often used when there is not enough data available to train a model from scratch, or when the task at hand is similar enough to a previous task that it makes sense to start with a model that has already learned some relevant information. Indeed, to do so we followed these steps to train our model:

- We chose a YOLO model with weights pre-trained on the COCO dataset.
- We froze the weights of the pre-trained model, so that they are not updated during training as they have already been optimized on the COCO dataset.
- We added a few layers to the pre-trained model. These layers were trained from scratch and allow the model to adapt to detecting electric grid equipment.
- We transformed the annotations to the YOLO format, then started training. The model learnt to perform the new task by adjusting the weights of the newly added fully connected layers.
- We fine-tuned the model by unfreezing some of the layers of the pre-trained model and training them as well. This allows the model to further adapt to the new task and can sometimes improve performance.

MODEL EVALUATION

The Intersection over Union value (IoU) is used as a threshold to determine if the prediction is correct. It measures the overlap between the model’s prediction and the ground truth bounding box (or polygon).

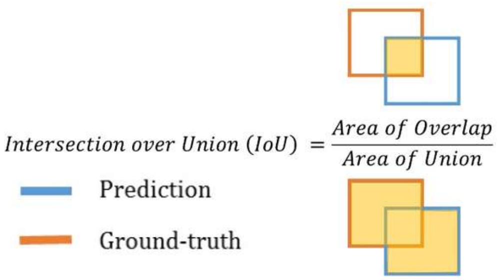

Figure 5- Illustration of IoU

Once, the IoU threshold is fixed (usually fixed at 0.5) we can calculate the metrics below to evaluate the performance of our model:
The precision is the model’s ability to correctly identify positive examples.

$$Precision = \frac{True\ Positive}{True\ positive + False\ Positive} = \frac{TP}{TP + FP}$$

The recall is model's ability to find all the positive examples in the dataset

$$Recall = \frac{True\ Positive}{True\ positive + False\ Negative} = \frac{TP}{TP + FN}$$

The precision-recall (PR) curve plots the precision and recall of a classifier for different thresholds. The area under the PR curve (AUPRC) is commonly used to measure a classifier’s performance. The closer is the AUPRC to one the better. A random classifier would have a 0.5 AUPRC.
Average Precision (AP) measures the model's ability to correctly identify and locate objects of a single class.

$$AP = Area\ under\ PR\ curve\ (AUPRC)$$

Mean Average Precision (mAP) is the mean the AP of all classes in the dataset

$$\text{mAP} = \frac{1}{n\_class} \sum_{i=1}^{n\_class} AP_i$$

It is to be noted that different evaluation metrics may be more or less important depending on the specific task and dataset. In this study, **mAP** was selected to evaluate the object detector and **weighted average recall** and **weighted average precision** were selected for the classifier. These two metrics are weighted by the number of true positives in each class which solves the problem of imbalanced data.

RESULTS ANALYSIS

Object Detection

After training multiple models and tuning the hyperparameters, the model improved from a simple basic one (the orange curve on the next figure) to a more complex one (the red one on the next figure) which is able to catch more details in the images and gives the best performance for the task it was trained on. Images were resized to (640,640) [1][6] which is known to be a high resolution for training computer vision models. A higher resolution of (960,960) was tested but did not improve the results. The batches size was set to 4 during training due to hardware limitations.

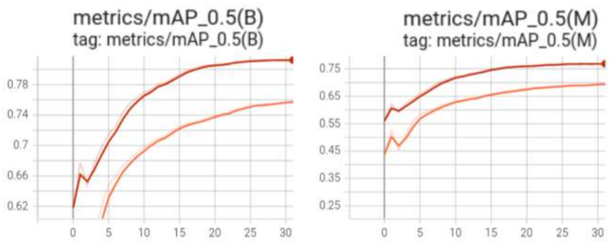

Figure 6- Evolution of mAP@[.5](B:Boudiung box) and mAP@[.5] (M:Mask) for two different models during training

The model was trained on 80% of our data, validated during training on another 10%. Finally, the remaining 10% called testing set was used to evaluate the model by plotting the precision-recall curve and the confusion matrix of the prediction.
It is to be noted that, depending on the quality and the centering, Enedis’ experts sometimes have issues to identify whether the pole is for electric grids.

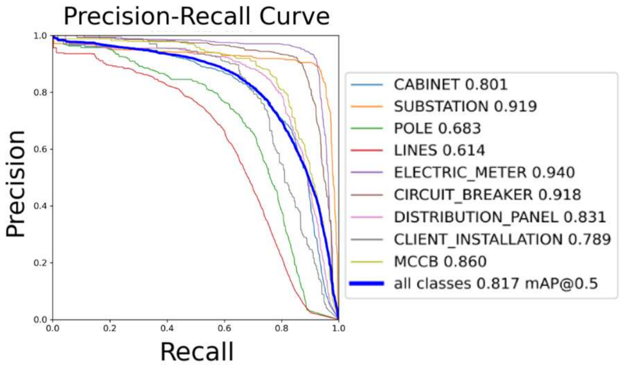

Figure 7- Precision-recall curve of predictions on the testing set

The number next to each class in the legend of the PR curve corresponds the AP of this class at an IoU of 0.5.

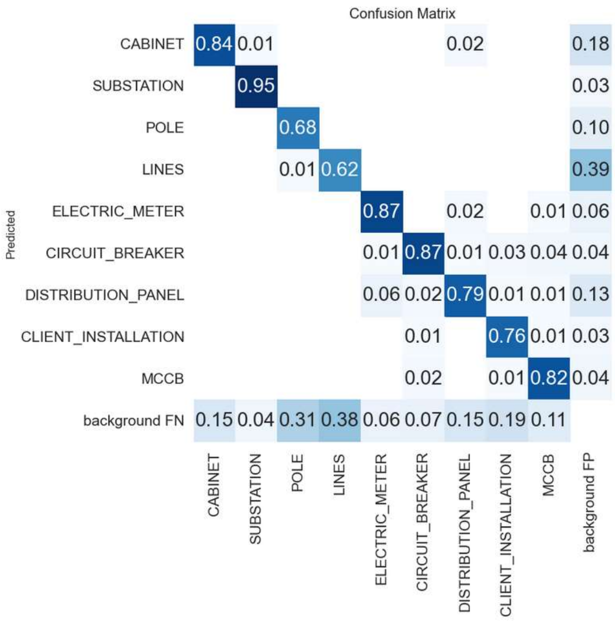

Figure 8- Confusion matrix of predictions on the testing set

The results show that objects with rectangular shape or large surface, such as electrical cabinets, substations or electricity meters are well detected with a mean average precision over **84% (mAP)**. This high mAP score indicates that the object detector is able to **accurately identify and locate objects** within an image, and that it is able to do so across a diverse range of object classes. The mAP score of 0.8 or above is considered a high score and it means that the model has a good balance between recall and precision. Such a high score is a clear indication of the effectiveness and efficiency of the object detector and the training and validation process that was used. This detector can be used in real-world applications with a high degree of accuracy.
Here is an example of predictions performed on the testing dataset:

Figure 9- Predictions on a difficult outside image of electric grid with very noisy background

The number shown beside the box prediction is the probability of presence of the equipment.
On the other hand, the detection of fine objects such as lines and poles is more challenging. Some difficulties were encountered during the object detection phase, here are some examples below:

- Lack of centering: the poles are not centered in the images, the model tends to confuse them with tree branches and thus returns a False Negative, which explains the low recall. In addition, lines are thin and may not have enough pixels to be accurately detected, they also may be occluded by other objects.
- Scale Variations: features in images can appear at different scales.
- Overlapping structures: Distribution network equipment in an image may be partially or completely obscured by other objects.
- Variations of the object in the same class: for example, a line can have several shapes (single phase, three phase, twisted) which makes detection more difficult.
- Limited training data: Object detection algorithms often rely on large amounts of annotated training data to learn how to detect objects.

Classification

After training the object detector, a classification model is trained to differentiate electric grid equipment and non-electric ones. The next figure shows the class distribution of the dataset used for classification. This dataset was built by cropping the objects present in the images using the annotations.

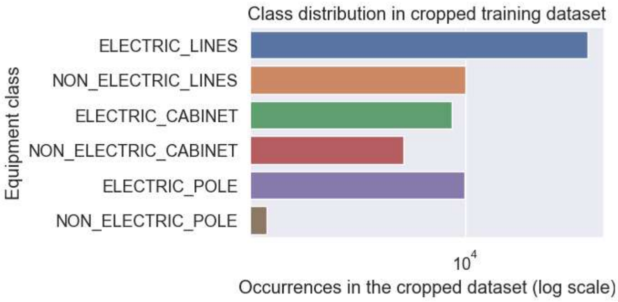

Figure 10- Distribution of cropped images fed to the classification model

3 models were trained to perform classification of electric equipment and non-electric ones.

Cabinet classification

The **cabinets** are **well classified**. Yet, the model has some issues when the door of the cabinets is broken, which is the case on several images. It is to be noted that the specific detection of the symbols: flame, EDF, ERDF or the "telereport" should work, but has not been tested since there are many images of cabinets where the door is broken.

Despite the difficulties encountered, the tests for the classification of the cabinets gave a **weighted average recall of 74%** and a **weighted average precision of 74%**. These results can be explained by:

- A **well-distributed** dataset: 59.6% of electrical cabinets and 30.4% of non-electric cabinets
- An **acceptable volume of data**, 10 772 photos.

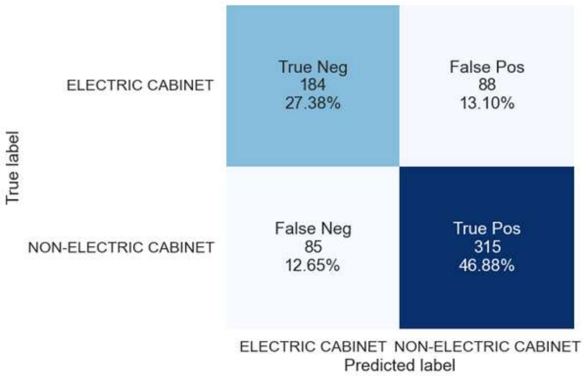

Figure 11-Confusion matrix for the classification of cabinets

Pole classification

The **weighted average precision is at 81%** and the **weighted average recall is at 65%**. The results are **not satisfying** because the **recall is too low.** Some reasons explaining these results are:

- The two distinguishing features are based on poles' heights and the notion of distance or length is difficult to interpret for a model from an image.
- A pole is considered non-electric when no cable is attached. During the classification phase, the image is cropped, so this information is lost.
- There is a lack of data for the class "other pole" or non-electric ones (we only have 1806 instances in our non-electric pole dataset).
- We have an unbalanced dataset: 84.6% of electric poles and 15.4% of non-electric poles.

Line classification

The **weighted average precision is at 62%** and the **weighted average recall is at 65%** for line classification. It is a difficult task even before trying to differentiate between electrical and telecom lines. Moreover, the dataset is unbalanced, 74.2% electrical cables and 25.8% non-electric cables. Thus, the results are not satisfying but might be improved using a specific line detector algorithm.

CONCLUSION

The object detector has proven to be highly effective, with a **mean Average Precision (mAP) of 81.7%** on the testing dataset across all classes. The model was challenged by complex noisy backgrounds when poles and cables are not positioned centrally within the image.

The classifier **performs well for classifying electric and non-electric cabinets**. Nevertheless, when it comes to cables and poles it has encountered issues due to the lack of data, the low resolution, the shape and scale variation of images.

The **upcoming photos from the mobile application called FLOE** will increase the quantity and improve the quality of our dataset. This would allow to push the model performance even further.

As a perspective, in the future, it would be interesting to study the possibility of **anomaly detection [5]** of electrical grid. It can be used to identify malfunctions in equipment. These models would require a pool of annotated photos big and diverse enough to get interesting results.

REFERENCES

[1] Redmon J, Divvala S, Girshick R, Farhadi A, 2016, "You Only Look Once:Unified, Real-Time Object Detection", *IEEE Conference* (pp. 779-788)

[2] K. He, G. Gkioxari, P. Dollar, R. Girshick, 2018, "Mask R-CNN", *arXiv:1703.06870*

[3] K. Simonyan, A. Zisserman, 2015, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556

[4] Souidene Mseddi W, Ghali R, Jmal M, Attia R, 2021, " Fire Detection and Segmentation using YOLOv5 and U-NET", *EUSIPCO 2021.* ISBN: 978-9-0827-9706-0.

[5] S. Akcay, A. Atapour-Abarghouei, T. P. Breckon, 2018, "GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training", *arXiv:1805.06725*.

[6] Redmon J, Farhadi A, 2018, "YOLOv3: An Incremental Improvement", *IEEE Conference (pp. 7263-7271)*