PTL-AI Furnas Dataset: A Public Dataset for Fault Detection in Power Transmission Lines Using Aerial Images

Frederico Santos de Oliveira
Federal University of Mato Grosso (UFMT)
Cuiabá, MT - Brazil
Email: fred.santos.oliveira@gmail.com

Marcelo de Carvalho
Eletrobras-Furnas
Rio de Janeiro, RJ - Brazil
Email: marcarv@furnas.com.br

Pedro Henrique Tancredo Campos
Eletrobras-Furnas
Rio de Janeiro, RJ - Brazil
Email: tancredo@furnas.com.br

Anderson da Silva Soares
Universidade Federal de Goiás (UFG)
Goiânia, GO - Brazil
Email: andersonsoares@ufg.br

Arnaldo Cândido Júnior
Universidade Estadual Paulista (UNESP)
São José do Rio Preto, SP - Brazil
Email: arnaldocan@gmail.com

Ana Cláudia Rodrigues da Silva Quirino
Eletrobras-Furnas
Rio de Janeiro, RJ - Brazil
Email: anac@furnas.com.br

***Abstract*—We present a new images dataset called PTL-AI Furnas Dataset as a new benchmark for fault detection in power transmission lines. This dataset has 6,295 images, with resolution 1280×720, extracted from the maintenance process of the energy transmission lines at Furnas company. It contains annotations of 17,808 components classified as baliser, bird nest, insulator, spacer and stockbridge. Furnas is a company that generates or transmits electricity to 51% of households in Brazil and more than 40% of the nation's electricity passes through their grid enabling generating the dataset in different backgrounds and climatic conditions. We performed experiments using data augmentation techniques to train Faster R-CNN, Single-Shot Detects (SSD) and YoloV5 models. The benchmark result was obtained using the metrics of Mean Average Precision (mAP) and the Mean Average Recall (mAR) with values mAP=91.9% and mAR=89.7%. The PTL-AI Furnas Dataset is publicly available at https://github.com/freds0/PTL-AI_Furnas_Dataset.**

I. Introduction

Visual inspections are crucial for fault detection in power transmission lines. For power transmission, visual inspection is considered the main preventive maintenance activity in transmission lines. In Brazil, this process complies with the criteria established by the regulatory agent named National Electric Energy Agency (ANEEL) through Normative Resolution No. 906/2020.

Traditionally, the inspection of transmission lines can be carried out in the terrestrial and aerial modalities. In the terrestrial inspection, teams need to observe the defects using binoculars or Unmanned Aerial Vehicles (UAV) with on-board cameras instead of climbing the structures. In aerial inspection, the entire process is done with the use of manned helicopters, and it can be carried out in a much shorter time.

Due to its relevance, the energy sector has been looking for ways to improve the inspection process, for example, using cameras embedded in different platforms, which can capture more accurate images and at a lower cost when compared to traditional inspection. According to Liu et al [1], inspection using UAVs divides traditional inspection into two stages: data collection and analysis. In the first step, the UAV collects the images or videos, and in the second step, they are analyzed by a technical team or a specialized system. Due to its advantages, such as low cost, greater safety and high efficiency, several researches have been carried out on the use of UAVs in the inspection process and, therefore, they have great potential to replace the traditional methods [2], [3].

Automating the inspection of power transmission lines is a challenging task, as it involves locating and classifying highly complex individual components. Several works using object detection techniques based on Deep Learning have been published in recent years for fault detection, looking for components such as insulators [4]–[6], spacers [7], vibration dampers [8], [9], balisers [3], U-bolts [10], transmission cables [11], and also looking for faults from natural causes, such as proximity to vegetation [12], [13], bird nests [14], and the freezing of transmission lines [15], [16]. In Brazil, [17] it is an example of application in the area of power transmission. A review of the status of research can be found at [18].

Although methods based on deep learning present excellent results, public datasets for inspecting transmission lines with high quality images are not common in the literature. Some works, as [6], [19]–[22], present public datasets, but they have either just a few components, or few images, or low resolution. As an alternative to the scarcity of data, data augmentation techniques have been used extensively in this area [4]–[6]. Liu et al [1] carry out an extensive review of the available datasets.

The development of algorithms, conducting research and improving solutions come up against the lack of benchmarks and public datasets. In this sense, we collected, annotated and organized a dataset, named *Power Transmission Lines using Aerial Images Furnas Dataset* (PTL-AI Furnas-Dataset), to be used as a benchmark, which has real data from the power

transmission lines of one of the largest energy generators in the world, the company Eletrobras-Furnas. We also present baseline results considering state-of-the-art deep learning architectures to guide further research. In Section II, we present some works related to fault detection in power transmission lines using deep learning. In Section III, we describe the process of creating the PTL-AI Furnas-Dataset. In Section IV, we detail the experiments performed and, in Section V, the results of these experiments. In Section VI, a conclusion and some final considerations of this work are presented.

II. Related Works

Tao et al [6] present CPLID, a dataset containing 848 images of insulators, with a resolution equal to 1152×864. CPLID was created using data augmentation techniques, mostly segmentation, to increase the volume of images and the diversity of backgrounds. Tomaszewski et al [19] also present a dataset of insulators, however, the images were artificially created, strategically positioning the insulators to obtain lighting and background variations.

Bian et al [20] developed a dataset consisting of 1,300 images of power transmission towers taken from the internet, with different pixel resolutions. Lee et al [23] built two datasets, Dataset 1 for training, with 4000 cable images for training with resolution 128×128, and Dataset 2 for testing, with 200 cable images with resolution 512×512, from Turkish Electricity Transmission Company. The total of annotated cables is not reported.

Abdelfattah et al [21] created TTPLA, a dataset composed of 1,100 images 3,840×2,160, containing annotations of towers (tower-lattice, tower-tucohy, tower-wood) and power transmission cables for instance segmentation. The authors present results using Resnet-50 and Resnet-101 [24] models.

Vieira-e-Silva et al [22] presents a new training pipeline and a new dataset, called *STN Power Line Assets Dataset* (STN-PLAD), containing high resolution images of towers, insulators, spacers, plates and dampers. Experiments are performed using the Single Shot MultiBox Detecton (SSD) [25] and Faster R-CNN [26] models. Table I shows a comparison of the main publicly available datasets.

Dataset	#Imgs	#Obj/img	Img Size	#Objs
STN-PLAD	133	18.1	5472×3078	5
CPLID	848	1.9	1152×864	1
TTPLA	1,100	8.1	3840×2160	4
Bian et al	1300	1.2	-	1
Tomaszewski et al	2,630	1	5616×3744	1
Lee et al	4000	-	128×128	1
PTL-AI Furnas Dataset	6,295	2.8	1280×720	5

TABLE I
Comparison of the main publicly available datasets.

III. PTL-AI Furnas-Dataset

The PTL-AI Furnas-Dataset creation process involves three steps: (1) recording videos, (2) extracting and selecting images, and (3) annotating the components. Next, we present details of this creation process.

Video recording: PTL-AI Furnas-Dataset was obtained from videos of aerial inspections of power transmission lines, using a crewed helicopter that performed inspections during the day in different environments and weather conditions. During the inspections, video recordings with audio are made, in which the components that need maintenance are informed.

Image Extraction and Selection: the images were extracted from the videos, one frame every 3 seconds, with a resolution of 1280×720, using FFmpeg tool. For the selection of images, first the audios of the videos were transcribed using a transcription tool, in order to search for the following keywords: failure detected, maintenance required, verification required, component defective, bird nest found, critical situation. Upon finding a keyword, images within a window of $\pm$ 15s were selected. Subsequently, only the images containing the components were manually chosen. Finally, padding is performed on the images to make them square. In Figure (1), it can be seen some samples from PTL-AI Furnas-Dataset.

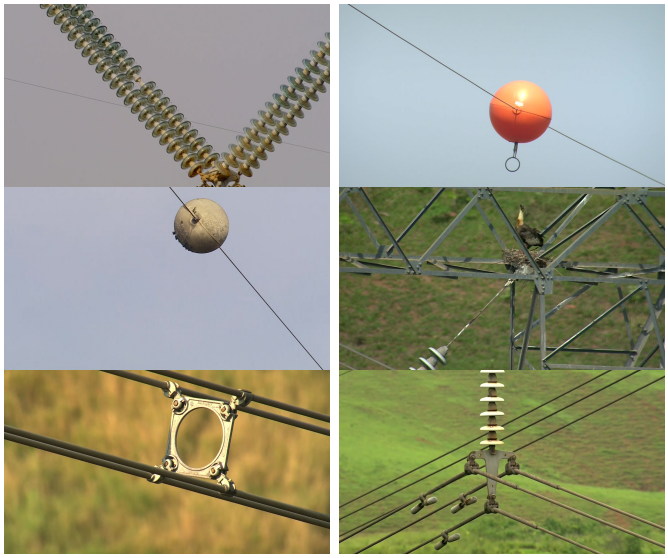

Fig. 1. PTL-AI Furnas-Dataset samples.

Image Annotation: it was defined that the following classes are of interest: *insulator*, *baliser*, *bird nest*, *separator* and *stockbridge*. The objects *insulator*, *baliser* and *separator* also have different states, shown in Table II, which indicate whether or not they need maintenance according to the criteria established by the technical team.

The image annotation process was performed gradually in order to facilitate annotator's task, reducing errors. In this process, a pre-trained model was used as an auxiliary tool, the SSD MobileNetV2 [27] model, which was trained with 2,000 images that were manually annotated using the *Make Sense* tool. The training was performed for 100k steps, fine tuning from a version trained in COCO dataset. SSD MobileNetV2 was chosen because it is a light and fast training model. The annotations were manually verified and corrected when necessary. This process was repeated two more times, including the verified images, until reaching a total of 6,295

Class	Label	#Train	#Test	#Obj/img
Insulator Good	insulator_ok	7,854	821	1.37
Insulator Problem	insulator_nok	1,840	214	0.32
Insulator Unknown	insulator_unk	1,638	181	0.28
Baliser Good	baliser_ok	125	11	0.02
Baliser Problem	baliser_nok	184	20	0.03
Baliser Almost Good	baliser_aok	319	46	0.05
Bird Nest	bird_nest	340	31	0.05
Separator Good	separator_ok	1,016	100	0.17
Separator Problem	separator_nok	21	2	0.01
Stock Bridge Good	stockbridge_ok	2,808	237	0.48
Stock Bridge Problem	stockbridge_nok	0	0	0.00
Total	—	16,145	1,663	2.82

TABLE II
PTL-AI FURNAS-DATASET STATISTICS. IN CLASS AND LABELS COLUMNS, THE SUFFIX *ok* INDICATES A COMPONENT IN GOOD CONDITION, *nok* THERE IS A PROBLEM, *aok* THAT MAINTENANCE MAY BE REQUIRED, AND *unk* IT CANNOT BE DETERMINED. *#Train* AND *#Test* INDICATES THE AMOUNT OF COMPONENTS IN THE SET AND *#Obj/img* THE AVERAGE AMOUNT OF COMPONENTS PER IMAGE.

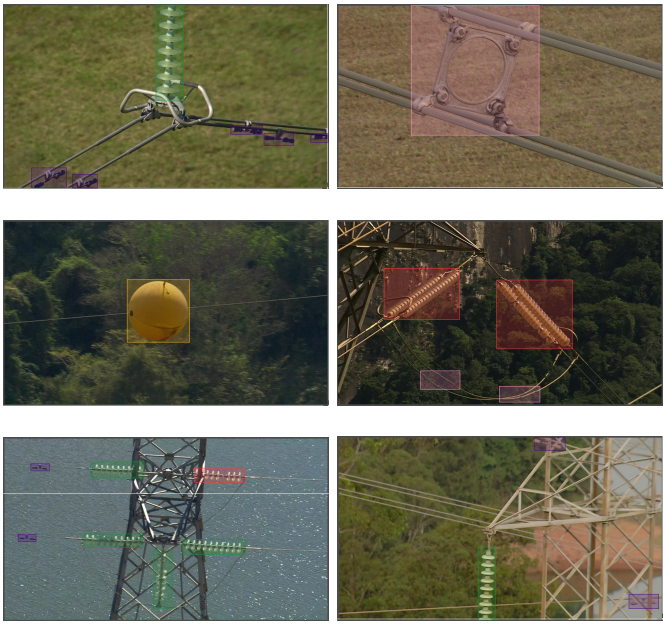

Fig. 2. PTL-AI Furnas-Dataset annotated samples.

images. Examples of the annotated images can be seen in Figure (2).

IV. EXPERIMENTS

In this work, we defined a series of experiments in order to train the state-of-art models for object detection and thus create baselines to evaluate the performance of the models using PTL-AI Furnas-Dataset. The images were randomly divided into the training and testing set, with 90 and 10 percent of the total images, respectively, as shown in Table II. In order to assess whether an object was correctly detected, it was defined that the *Intersection over Union* (IoU) between the bounding box of the ground-truth and the prediction must have a minimum value of 0.5. The experiments were performed on a DGX-1 V100 GPU with 32GB of memory up to the limit of one week of execution. Each experiment was also configured with an early stopping equal to 100 epochs, analyzing Mean Average Precision (mAP) in a validation set, extracted from a selection of five percent of the train set images.

Faster R-CNN [26], SSD [25] and YOLO [28]–[32] models were selected for training because, according to Liu et al [1], they are the main models used in the power transmission lines inspection task. The works [2], [5], [7], [8], [13], [14], [33] confirm this statement. Three different versions of each of these models were used: *Inception* [34], *ResNet-152* and *ResNet-101* [24] backbones for Faster R-CNN model; *ResNet-152*, 101 and 50 for SSD model; v5m6 (medium), v5l6 (large) and v5x6 (extra large) for YOLO model [32].

The following available source-codes were used: Faster R-CNN and SSD source-code from Tensorflow 2.x Object Detection API, and YOLO source-code from Ultralytics, in Pytorch. The parameters used during training can be checked in Table III. All the models were fine tuned from a pre-trained version in the COCO dataset. The training of the Faster R-CNN and SSD models was interrupted when reaching the maximum time of one week, and the training of the YOLOv5 models was interrupted by the early-stopping technique. All the source-code, as well as the checkpoints of each model, are available for download at https://github.com/freds0/PTL-AI_Furnas_Dataset, in order to facilitate the reproduction of the presented experiments.

Faster R-CNN				
Backbone	LR	Momentum	Batch Size	Images Size
Inception	0.008	0.9	4	1280×1280
ResNet 152	0.04	0.9	5	1280×1280
ResNet 101	0.04	0.9	5	1024×1024
Single Shot MultiBox Detector - SSD				
ResNet 152	0.04	0.9	5	1280×1280
ResNet 101	0.04	0.9	5	1024×1024
ResNet 50	0.04	0.9	7	1024×1024
You Only Look Once - YOLO				
YoloV5m6	0.01	0.99	16	1280×1280
YoloV5l6	0.01	0.99	16	1280×1280
YoloV5x6	0.01	0.99	10	1280×1280

TABLE III
PTL-AI FURNAS-DATASET TRAINING PARAMETERS.

In Table III, it can be seen that the models considered more robust, Faster R-CNN Inception, Faster R-CNN ResNet 152, SSD ResNet 152 and all versions of YOLO, were trained using the images in format 1280×1280. This decision was made in order to avoid losses in the quality of the images, and thus obtain better results in these models. The shallower models, Faster R-CNN ResNet, SSD ResNet 101 and 50 were trained with images in the format 1024×1024 in order to perform faster training.

A. Data Augmentation

The training set was expanded using data augmentation techniques, following Song et al [35], through the Imgaug tool. For each image, another five were created using data augmentation. Two functions were applied randomly selected from a set of functions that include basic changes, such as

color, position, scale, noise and dropout, as well as more complex changes, such as deformations, degradations and weather. In Table IV, the data augmentation functions used are shown, grouped according to the technique applied. A sample of the generated images can be seen in Figure (3).

Group	Functions
Color and Contrast	Gamma Contrast, Linear Contrast, Add to Hue and Saturation.
Dropout	Dropout, Course Dropout, Pepper, Coarse Pepper, Salt, Coarse Salt, Salt and Pepper, Coarse Salt and Pepper.
Arithmetic	Add, Add Element Wise, Impulse, Add Gaussian, Add Lapplace,
Geometric	Scale, Translate, Rotate, Shear, Piece Wise, Perspective Transform, Elastic Transformation.
Flip	Horizontal Flip.
Corruption	Jpeg Compression.
Blur	Gaussian, Average, Median, Bilateral, Motion.
Weather	Snowflakes, Rain, Fog, Clouds.

TABLE IV
PTL-AI FURNAS-DATASET DATA AUGMENTATION USED FUNCTIONS.

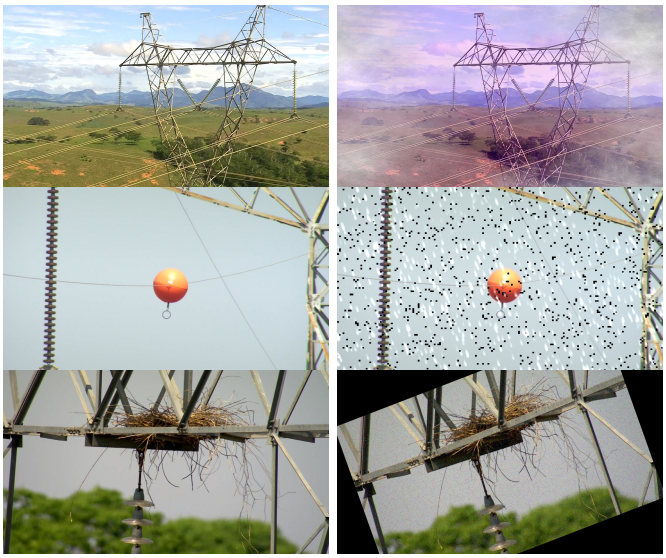

Fig. 3. PTL-AI Furnas-Dataset data augmentation samples.

V. RESULTS

The models are evaluated using the PTL-AI Furnas-Dataset test set. For each of the training performed, it can be verified the results of the metrics *Mean Average Precision* (mAP) and *Mean Average Recall* (mAR) in Table V. The Yolov5l6 model achieved the higher mAP results, with a value equal to 91.9%, and the Yolov5x6 model the higher mAR value, with 89.7%.

Among the family of Faster R-CNN models, the model with *backbone Inception* stood out, presenting 89.6% of mAP and 80.6% of mAR. Among the SSD family models, surprisingly the shallowest model, with *backbone* ResNet 50, presented the best mAP results, with 75.8% and also mAR, with 77.9%.

To assess which classes the models have more difficulty in predicting, in Figure (4), one can visualize the confusion matrices of the two best models: Yolov5x6 and Yolov5l6. In

Faster R-CNN			
BackBone	Image Size	mAP	mAR
Inception ResNet	1280 × 1280	**0.896**	**0.806**
ResNet 152	1280 × 1280	0.808	0.673
ResNet 101	1024 × 1024	0.806	0.632
Single-Shot Detector - SSD			
ResNet 152	1280 × 1280	0.710	0.752
ResNet 101	1024 × 1024	0.688	0.749
ResNet 50	1024 × 1024	**0.758**	**0.779**
You Only Look Once - YOLO			
YOLOv5m6	1280 × 1280	0.913	0.856
YOLOv5l6	1280 × 1280	**0.919**	0.855
YOLOv5x6	1280 × 1280	0.917	**0.897**

TABLE V
MEAN AVERAGE PRECISION (MAP) AND MEAN AVERAGE RECALL (MAR) FOR DIFFERENT MODELS ON PTL-AI FURNAS-DATASET.

this image, it can be seen that the models have problems in classifying mainly *balisers* and *insulators*. This type of error is understandable, since among these classes, in some cases, doubts even occur among the annotators. This occurs mainly in the *baliser* class, where there is no rigid definition between the aok and nok states, since color is the main deciding factor, and between the *insulator* class, given the difficulty in deciding whether the component is in good condition or too far away to define.

The *baliser_nok*, *insulator_nok*, *bird_nest*, and *spacer_nok* classes are critical for the task of fault detection in power transmission lines, as they indicate components failures. In Figure (4), it can be seen that none of these classes had a false-negative percentage (*Background FN* in the image) greater than 5% in any model. About the stockbridge component, because it is very small, it was difficult even for the technical team to detect failures and, therefore, the *stockbridge_nok* label were not annotated.

In Figure (5), it can be seen the Precision × Recall graph referring to the Yolov5x6 (top) and Yolov5l6 (bottom) models of each of the components. For the Yolov5x6 model the classes *baliser_ok*, *spacer_ok*, *insulator_unk* had mAP lower than 0.9, while for the Yolov5l6 model the classes *baliser_ok*, *baliser_aok*, *spacer_ok* and *insulator_unk* had a negative influence on the results.

In Figure (6), there are some examples of the predictions obtained using the Yolov5l6 model.

VI. CONCLUSIONS

In this work, we proposed a new image dataset, named *PTL-AI Furnas Dataset*, publicly available, to be used as a benchmark in the area of fault detection in power transmission line components, which contains images of components of great importance for the inspection activity of power transmission lines. Images are taken in different landscapes, with a wide variety of components, at different angles and positions. A dataset creation process was also presented, in which the SSD-MobileNetV2 model was used as a aid tool during the image annotation process.

PTL-AI Furnas-Dataset aims to fill an existing gap in the area of fault detection in power transmission lines, which is

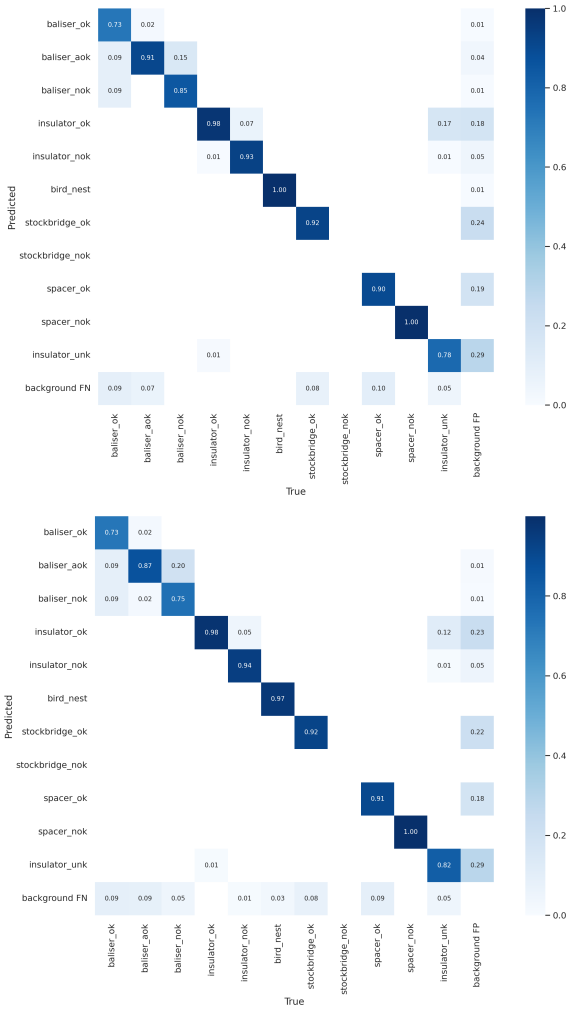

Fig. 4. Confusion matrix of Yolov5l6 model. On the top, results from the Yolov5x6 model, and on the bottom, results from the Yolov5l6 model.

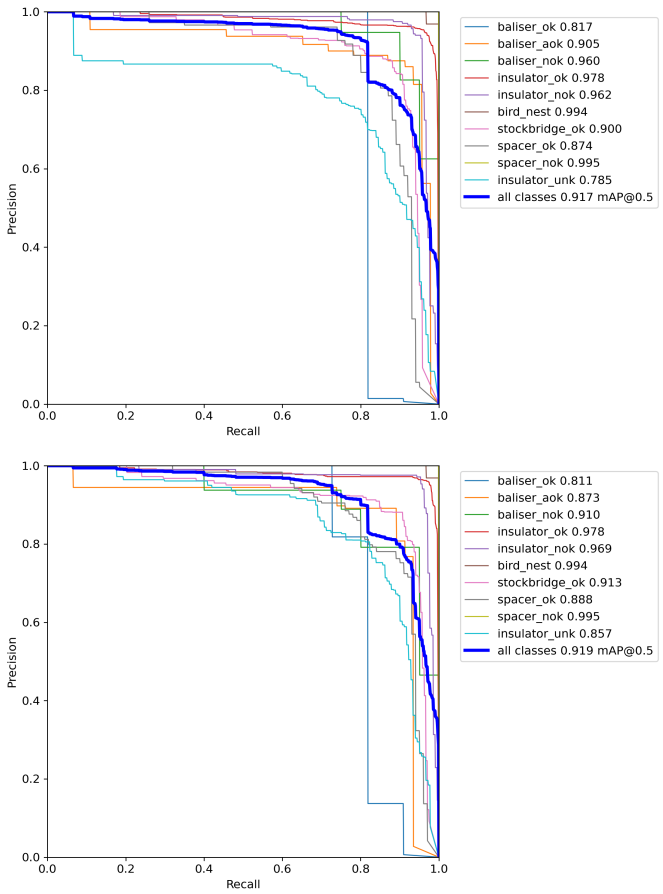

Fig. 5. Precision-Recall Curve of Yolov5x6 (top) and Yolov5l6 (bottom) models.

the lack of public datasets with high quality images and a considerable number of annotated components. In this way, it will be possible to advance the development of research and to improve solutions for the task of fault detection in power transmission lines.

We carried out experiments using state-of-art models in the area of object detection. From these experiments, we found that the Yolov5l6 and Yolov5x6 models showed the highest mAPs. Therefore, with this work we created the baselines to be used as benchmarks for future works. In the future, we should evaluate trained models on our dataset using images from UAVs, in order to verify that models may experience a performance loss due to variations in capture devices, recording angles and image quality.

ACKNOWLEDGMENT

We would like to thank the CyberLabs and CEIA/UFG for financial support for this paper and Eletrobras-Furnas for providing a specialized technical team, support and also the images used to create *PTL-AI Furnas Dataset*.

REFERENCES

[1] X. Liu, X. Miao, H. Jiang, and J. Chen, "Data analysis in visual power line inspection: An in-depth review of deep learning for component detection and fault diagnosis," *Annual Reviews in Control*, vol. 50, pp. 253–277, 2020.

[2] H. Guan, X. Sun, Y. Su, T. Hu, H. Wang, H. Wang, C. Peng, and Q. Guo, "Uav-lidar aids automatic intelligent powerline inspection," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 106987, 2021.

[3] Z. A. Siddiqui and U. Park, "A drone based transmission line components inspection system with deep learning technique," *Energies*, vol. 13, no. 13, p. 3348, 2020.

[4] J. Tan, "Automatic insulator detection for power line using aerial images powered by convolutional neural networks," in *Journal of Physics: Conference Series*, vol. 1748. IOP Publishing, 2021, p. 042012.

[5] D. Sadykova, D. Pernebayeva, M. Bagheri, and A. James, "In-yolo: Real-time detection of outdoor high voltage insulators using uav imaging," *IEEE Transactions on Power Delivery*, vol. 35, no. 3, pp. 1599–1601, 2019.

[6] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 4, pp. 1486–1498, 2018.

[7] X. Liu, Y. Li, F. Shuang, F. Gao, X. Zhou, and X. Chen, "Issd: Improved ssd for insulator and spacer online detection based on uav system," *Sensors*, vol. 20, no. 23, p. 6961, 2020.

[8] W. Chen, Y. Li, and Z. Zhao, "Transmission line vibration damper detection using deep neural networks based on uav remote sensing image," *Sensors*, vol. 22, no. 5, p. 1892, 2022.

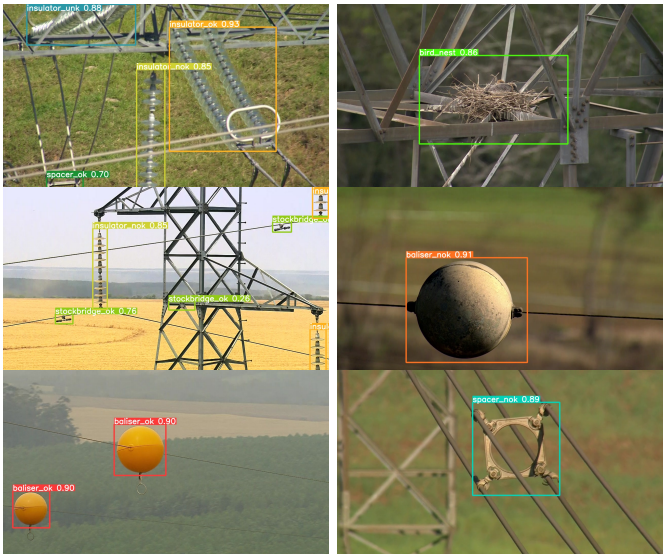

Fig. 6. Yolov5l6 results Samples.

[9] ——, "Transmission line vibration damper detection using multi-granularity conditional generative adversarial nets based on uav inspection images," *Sensors*, vol. 22, no. 5, p. 1886, 2022.

[10] A. Odo, S. McKenna, D. Flynn, and J. B. Vorstius, "Aerial image analysis using deep learning for electrical overhead line network asset management," *IEEE Access*, vol. 9, pp. 146 281–146 295, 2021.

[11] H. Zhang, W. Yang, H. Yu, H. Zhang, and G.-S. Xia, "Detecting power lines in uav images with convolutional features and structured constraints," *Remote Sensing*, vol. 11, no. 11, p. 1342, 2019.

[12] J. I. Larrauri, G. Sorrosal, and M. González, "Automatic system for overhead power line inspection using an unmanned aerial vehicle — relifo project," *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 244–252, 2013.

[13] S. Rong and L. He, "A joint faster rcnn and stereovision algorithm for vegetation encroachment detection in power line corridors," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2020, pp. 1–5.

[14] X. Lei and Z. Sui, "Intelligent fault detection of high voltage line based on the faster r-cnn," *Measurement*, vol. 138, pp. 379–385, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263224119300831

[15] H. Huang, X. Ma, L. Zhao, H. Du, H. Luo, X. Mao, M. Tang, and Y. Liu, "Transmission line icing measurement on photogrammetry method," in *MIPPR 2015: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications*, J. Liu and H. Sun, Eds., vol. 9815, International Society for Optics and Photonics. SPIE, 2015, pp. 221 – 226. [Online]. Available: https://doi.org/10.1117/12.2203579

[16] J. Wang, J. Wang, J. Shao, and J. Li, "Image recognition of icing thickness on power transmission lines based on a least squares hough transform," *Energies*, vol. 10, no. 4, 2017. [Online]. Available: https://www.mdpi.com/1996-1073/10/4/415

[17] R. Y. Olivo, P. L. de Paula Filho, and A. C. Junior, "Uma abordagem neural na identificação de objetos em imagens para auxilio na manutencão de rede elétrica," in *Anais Estendidos do XXXIII Conference on Graphics, Patterns and Images*. SBC, 2020, pp. 179–182.

[18] R. Jenssen, D. Roverso *et al.*, "Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 107–120, 2018.

[19] M. Tomaszewski, B. Ruszczak, and P. Michalski, "The collection of images of an insulator taken outdoors in varying lighting conditions with additional laser spots," *Data in Brief*, vol. 18, pp. 765–768, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352340918302701

[20] J. Bian, X. Hui, X. Zhao, and M. Tan, "A monocular vision–based perception approach for unmanned aerial vehicle close proximity transmission tower inspection," *International Journal of Advanced Robotic Systems*, vol. 16, no. 1, p. 1729881418820227, 2019. [Online]. Available: https://doi.org/10.1177/1729881418820227

[21] R. Abdelfattah, X. Wang, and S. Wang, "Ttpla: An aerial-image dataset for detection and segmentation of transmission towers and power lines," in *Computer Vision – ACCV 2020: 15th Asian Conference on Computer Vision, Kyoto, Japan, November 30 – December 4, 2020, Revised Selected Papers, Part VI*. Berlin, Heidelberg: Springer-Verlag, 2020, p. 601–618.

[22] A. L. B. Vieira-e Silva, H. Felix, T. d. M. Chaves, F. P. M. Simões, V. Teichrieb, M. M. d. Santos, H. d. C. Santiago, V. A. C. Sgotti, and H. B. D. T. L. Neto, "Stn plad: A dataset for multi-size power line assets detection in high-resolution uav images," 2021. [Online]. Available: https://arxiv.org/abs/2108.07944

[23] S. J. Lee, J. P. Yun, H. Choi, W. Kwon, G. Koo, and S. W. Kim, "Weakly supervised learning with convolutional neural networks for power line localization," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017, pp. 1–8.

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

[25] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 21–37.

[26] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., vol. 28. Curran Associates, Inc., 2015. [Online]. Available: https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf

[27] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, jun 2016, pp. 779–788. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.91

[29] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[30] ——, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[31] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[32] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, TaoXie, J. Fang, imyhxy, K. Michael, Lorna, A. V, D. Montes, J. Nadar, Laughing, tkianai, yxNONG, P. Skalski, Z. Wang, A. Hogan, C. Fati, L. Mammana, AlexWang1900, D. Patel, D. Yiwei, F. You, J. Hajek, L. Diaconu, and M. T. Minh, "ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference," Feb. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6222936

[33] C. Liu, Y. Wu, J. Liu, Z. Sun, and H. Xu, "Insulator faults detection in aerial images from high-voltage transmission lines based on deep learning model," *Applied Sciences*, vol. 11, no. 10, p. 4647, 2021.

[34] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[35] C. Song, W. Xu, Z. Wang, S. Yu, P. Zeng, and Z. Ju, "Analysis on the impact of data augmentation on target recognition for UAV-based transmission line inspection," *Complexity*, vol. 2020, 2020.