DETECTION OF ELECTRICAL EQUIPMENT USING AN ARTIFICIAL NEURAL NETWORK WITH VISUAL SPECTRUM IMAGES AND THERMAL IMAGES

João P.S. Nascimento[1], Bruno V.S. Araujo [1,2], Gustavo A. Rodrigues [1,2], George V. R. Xavier [1,2], Eduardo O. Freire [1,2], George Sales[3], Tarso V. Ferreira [1,2].

[1]Departamento de Engenharia Elétrica, Universidade Federal de Sergipe, São Cristóvão, Brasil
[2] INESC P&D Brasil
[3]Eneva S.A.
tarso@academico.ufs.br

Keywords: Object Recognition, High Voltage Equipment, Thermal Imaging, Database Expansion, YOLO.

Abstract

Automated monitoring of assets in the energy industry, notably high voltage equipment, can greatly benefit from the detection of these equipment through intelligent algorithms. Currently, several topologies and architectures of artificial neural networks are used in the recognition of objects in images. To train these networks, a very large number of images is required, which is not always available. Computer vision and digital image processing techniques are then applied to expand the existing database. This article explores the combination of images at different wavelengths to improve the performance of the YOLOv5 network in recognizing high voltage electrical equipment in a 500 kV substation. With this objective, two image databases, composed by images in the visible spectrum and infrared images were combined in different proportions and used to train the network. The results showed that expanding the training base with images of different wavelengths can improve the network performance.

1 Introduction

The human mind possesses remarkable capabilities to perceive, comprehend, predict, and manipulate the world. Artificial intelligence strives not only to understand these abilities but also to construct intelligent entities [1]. This endeavor has been made possible by technological advancements, which have given rise to powerful tools such as deep learning algorithms. The evolution of these algorithms has facilitated the development of recognition and detection techniques in various fields, including images, videos, audio, characters, and facial recognition.

In the electrical power system, there are several areas in which object detection algorithms can be applied. In monitoring by thermographic inspections, for example, these techniques can be used to identify the equipment, and thus contribute to the detection and diagnosis of defects. An example of automated detection of high voltage equipment, employing convolutional artificial neural networks is shown in Fig. 1. By using artificial intelligence to process images of electrical equipment, it is possible to increase accuracy, reliability, and speed on the identification and segmentation, thus, facilitating the work of further steps in monitoring systems, dedicated to diagnosing, for example. The use of AI also helps reduce human error and the human dependency in monotonous tasks.

These algorithms need a large dataset for training, often with thousands or even millions of data points. Unfortunately, obtaining such a voluminous dataset can be challenging, prompting the need to explore techniques to expand the dataset. For object detection, the database must have images in various sizes, positions, colors, and lighting situations to train the model to perform at their best. Therefore, several image database expansion techniques can be used, highlighting image characteristics manipulation techniques with geometric transformations (rotation, flip-flop, translation, scaling, etc.) and additional processing techniques (photometric transformations, random erasure, noise insertion, etc.); creation of artificial images and use of Artificial Neural Network (ANN) [2-4]. However, the computational routines used to apply the expansion techniques require great computational effort and consequently greater processing time, besides specialized personal to develop the processing algorithms.

https://doi.org/10.1049/icp.2024.0448

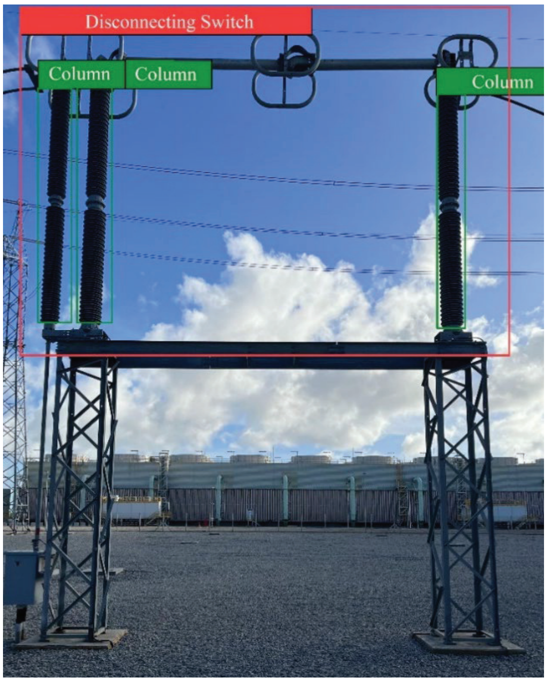

Fig. 1 Example of automated detection of high voltage equipment.

There are several lines of research related to database expansion and among the various studies, some have focused on manipulation of image characteristics [5-6]. Other papers focus on scaling up through deep learning algorithms [7-10]. In addition, there are studies that analyze the efficiency of different forms of database expansion [11-13].

Expanding a database plays a key role in improving the performance of detection and classification algorithms. This artificial expansion improves the network performance in detecting objects without requiring a real expansion of the database. Therefore, efficient database expansion methods can be a versatile alternative, which makes the network more general in object detection. This improved detector becomes more efficient and reliable in its applications, such as diagnosing thermal anomalies in specific parts of high voltage equipment.

In the context of this research, the behavior of an ANN trained with images in the visible light spectrum (IVS) is investigated when infrared images (IR), are inserted in the training database. The inverse situation is also investigated, i.e., IVS images are inserted in IR training database. This form of database expansion has practical advantages, as it is easier to build a large image database composed of IVS than IR, given the high cost of thermal imaging cameras and the cost of the IR inspector work hour.

Given the above and the relevance of the topic, a hypothesis raised in this research is that the expansion of a database of infrared images using images in the visible spectrum, captured via smartphone cameras, may lead to a better identification and detection performance of high voltage equipment in a substation using a convolutional neural network (CNN) from the YOLOv5 family (YOLO stands for “you only look once”).

This paper is structured as follows: section 2 describes the methodology adopted for the construction and development of the study; in section 3 the results obtained are analyzed and discussed and in section 4 the final considerations are made.

2. Methodology

The database used in this study comprises images of high voltage electrical equipment captured at the Porto de Sergipe I Thermoelectric Unit substation, which operates at 500 kV in Barra dos Coqueiros city, Brazil. The substation is owned by Eneva S.A.

The dataset was created using IVS, collected with a smartphone digital camera (3072x4096 pixels), and IR, collected with a FLIR T-840 thermographic camera (464x348 pixels). The resolution and opening angle differences between these two cameras result in images with distinct characteristics.

A total of 1400 images (700 IVS and 700 IR) were taken from various angles and positions of 9 equipment: circuit breaker (CB), disconnecting switch (DS), pedestal insulator (PI), suspension insulator (SI), surge arrester (SA), current transformer (CT), potential transformer (PT), power transformer (TR) and insulating column (IC). It is important to note that the IC is not a standalone piece of equipment, but rather a component of several other equipment. In the database used in this work, all equipment is from the same brand, and all equipment types are from the same model.

2.1. Image labeling

In this research, the object detection algorithm YOLOv5 was used, and the labeling was done using bounding boxes. Nine labels were adopted for this study, corresponding to the eight pieces of equipment mentioned earlier and the IC. Although the images always have a target device highlighted in the center, there are often other devices in the image background or nearby as well. In each image, all fully visible equipment was labeled. The labeling task was performed with aid of the online service present at www.makesense.ai.

2.2. Training arrangements

In this research, Google Collaboratory was used to carry out the training of the YOLOv5 object detection model, which was pre-trained on COCO128 images. All experiments presented in this article were trained for 300 epochs.

This study consists of two experiments, each containing five cases, aimed at evaluating the performance of the YOLOv5 neural network in detecting high voltage electrical equipment. Experiment I (Exp. I) compares the neural network's performance when trained with a database exclusively composed by IR images, and when trained with a database containing both IR images IVS. Meanwhile, Experiment II (Exp. II) evaluates the neural network's performance when trained with a database exclusively composed by IVS images and when trained with a database containing both IVS images

IR. It is important to note that Exp. I used IR images for validation, while Exp. II used IVS images for validation.

In Table 1 it is presented the distribution of images used for Exp. I, and the proportion of IR and IVS images in each case. A total of 500 IR images were randomly picked from the 700 available in the database. For each case, the 500 images were randomly split into training (Train) and validation (Val) sets, with 350 and 150 images, respectively (a 70/30 split). The insertion of IVS was done with in different proportions, as shown in cases 1 to 5, in Table 1. The model was trained five times for each case, and the average value between training sessions was considered as the result.

Table 1 - Training IR images with IVS insertion: number of images per training case and per wavelength.

Case	IT		IVS		Total	
	Train	Val	Train	Val	Train	Val
1	350	150	0	0	350	150
2	350	150	50	0	400	150
3	350	150	100	0	450	150
4	350	150	150	0	500	150
5	350	150	200	0	550	150

In Table 2, it is presented the distribution of images used for Exp. II, and the proportion of IVS and IR images in each case. A total of 500 IR images were randomly picked from the 700 available in the database. For each case, the 500 images were randomly split into training and validation sets, with 350 and 150 images, respectively (a 70/30 split). The insertion with IR was done with in different proportions, as shown in cases 1 to 5, in Table 2. The model was trained five times for each case, and the average value between training sessions was considered as the result.

Table 2 - Training ISV images with IR insertion: number of images per training case and per wavelength.

Case	IT		IVS		Total	
	Train	Val	Train	Val	Train	Val
1	0	0	350	150	350	150
2	50	0	350	150	400	150
3	100	0	350	150	450	150
4	150	0	350	150	500	150
5	200	0	350	150	550	150

2.3. Network performance evaluation

To evaluate the performance of the network, the Mean Average Precision (mAP) was used. It is a parameter that varies from 0 to 1 and is widely used to evaluate the performance of object detection algorithms, since its rigor can be adjusted according to the Intersection over Union (IoU) variation [15]. The IoU is the ratio of the area, in pixels, of the intersection by the union of the predicted and ground-truth bounding boxes, as shown in Fig. 2.

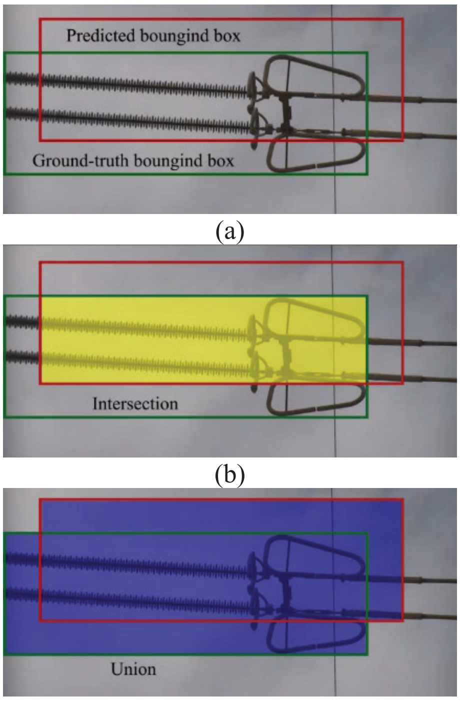

Fig. 2 Intersection over Union. (a) Predicted bounding box and ground-truth bounding box. (b) Intersection. (c) Union.

In general, the performance of an object detection algorithm is evaluated in terms of Precision (p), Recall (r) and mean Average Precision (mAP).

Precision quantifies the percentage of correct predictions and can be calculated by:

$$\boldsymbol{p} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FP}}, \tag{1}$$

where *TP* is the True Positive amount, while FP is the False Positive amount. The recall quantifies the percentage of positives found and can be calculated by:

$$\boldsymbol{r} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}}, \tag{2}$$

where FN is the amount of False Negative. Lastly, the AP is calculated from the area under the curve of the $\boldsymbol{p}\ \boldsymbol{vs}\ \boldsymbol{r}$ graph, which can be calculated by:

$$\boldsymbol{AP} = \int_{0}^{1} \boldsymbol{p}(\boldsymbol{r})\boldsymbol{dr}, \tag{3}$$

The mAP is the average of the various calculated APs. The mAP50 considers an IoU of 0.5, while the mAP50:95 considers an average result of the IoU variation from 0.5 to 0.95, with steps of 0.05.

3. Results

The results for obtained for the proposed methodology are presented in this section.

3.1. Image labeling results

In this research, a total of 1400 images were labeled, comprising of 700 IVS and 700 IR images. In Table 3 are presented the number of instances of each equipment in the IVS dataset. It should be noted that the number of instances is not uniform across the equipment as some equipment are more numerically prevalent in the substation than others and may appear in the background of some images.

In column 1 of Table 3, the first column refers to the abbreviations of each equipment that the ANN was trained to detect.

Table 4 shows the number of instances of each equipment for the IR base. Overall, the ISV base has 5427 instances, while the IR base has 5297. IVS may contain more equipment than IR images, resulting in a larger number of labels, on average. This effect occurs due to the larger resolution and field of view of the IVS camera. This fact may be used as a practical advantage of inserting IVS in the IR database during network training.

Table 3 - Number of IVS Instances per case.

Label	Exp. II (Case 1-5)	Exp. I (case 2)	Exp. I (case 3)	Exp. I (case 4)	Exp. I (case 5)
IC	2535	236	216	280	253
DS	423	33	37	48	46
CB	234	26	20	26	20
IP	94	10	10	12	18
PI	115	17	14	12	20
CT	221	27	19	17	16
PT	140	9	11	11	11
TR	42	8	5	4	3
SA	96	5	7	7	13

Table 4 - Number of IR Instances per case.

Label	Exp. I (case 1-5)	Exp. II (case 2)	Exp. II (case 3)	Exp. II (case 4)	Exp. II (case 5)
IC	2473	267	252	229	252
DS	371	35	46	38	39
CB	224	22	14	16	27
PI	81	5	7	7	9
SI	69	10	13	10	10
CT	274	37	28	25	20
PT	142	12	20	19	12
TR	30	3	3	3	3
SA	111	5	8	5	11

In Fig. 3 and Fig. 4, examples of IR and IVS labeling can be observed, respectively. In Fig. 3, one DS and three IC were labeled, while in Fig. 4, two DS and four IC were labeled.

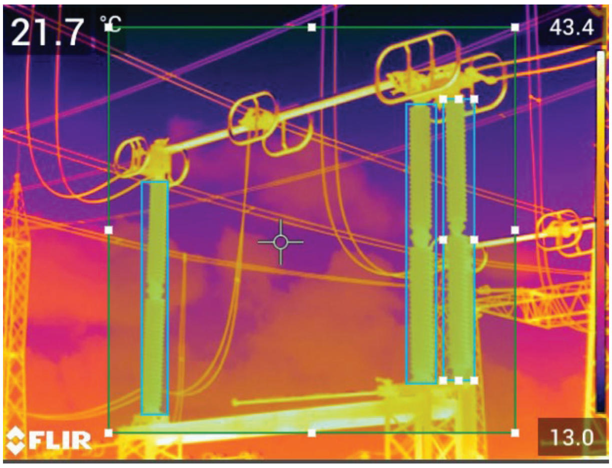

Fig. 3 Example of an IR image labeling.

(b)

Fig. 4 Example of an IVS image labeling.

3.2. Network performance

In Figs. 5 and 6, the average values of mAP.50 and mAP.50:.95 are presented and compared for the detection of all cited high voltage equipment in Exp. I: training with a database exclusively composed by IR images compared to training with a database composed by IR images and IVS.

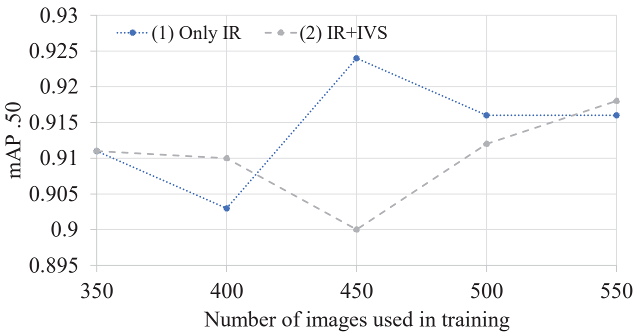

Fig. 5 mAP.50 values for the Exp. I analyzed.

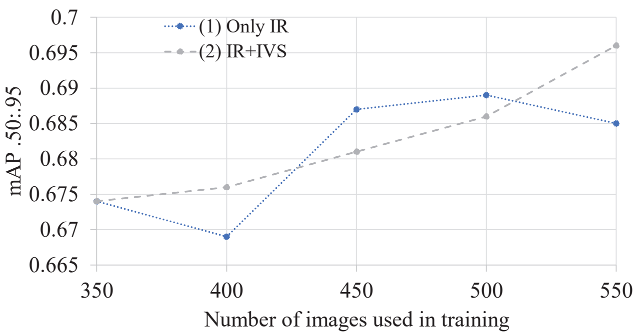

Fig. 6 mAP.50:.95 values for the Exp.1 analyzed.

For mAP.50, which can be seen in Fig 5, it is observed that as the values are close to 1, increasing the number of images at both wavelengths results in only a slight improvement in detection. The point of greatest difference between the curves occurred with 450 images in training, where the use of IR alone was 2.67% better than the mixed case (IVS+IR). However, in inserting 50 images of IVS instead of IR images (resulting in 400 images) resulted in a slightly a better performance. This behavior repeats in the case of 200 images of IVS instead of IR images (resulting in 550 images).

As for mAP.50:.95, which can be seen in Fig 5, there is a clear trend of improvement in detection for both curves. It is noteworthy that the best result was for training with 350 IR with contamination of 200 IVS, in which the value of mAP.50:.95 was equal to 0.696.

In Fig. 7 and 8, the mean values of mAP.50 and mAP50:.95 for Exp. II are presented and compared. This case involves training with a database exclusively composed by IVS images compared to training with a database composed by IVS and IR images.

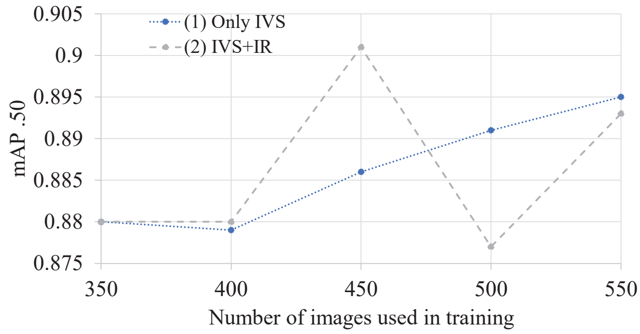

Fig. 7 mAP.50 values for the Exp II analyzed.

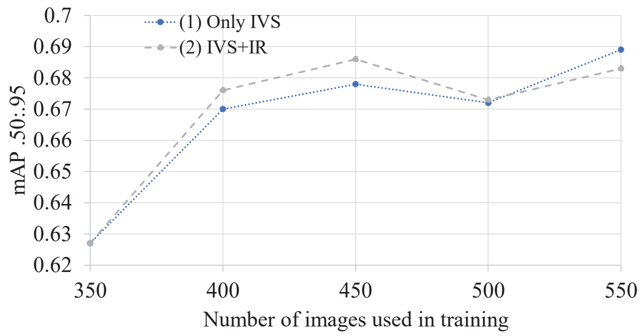

Fig. 8 mAP.50:.95 values for the Exp. II analyzed.

When analyzing Fig. 7, it is noted that the use of IVS alone has a clearer tendency to improve performance, while the insertion of IR images results varying results. This behavior can also be explained by the fact that the values are close to 1, making mAP.50 less sensitive to the increase in the number of images in training.

In Fig. 8 the results of mAP.50:.95 for Exp II are shown. Both curves have a clear upward trend. Expanding the database, whether by IR images or IVS, results in equivalent improvements in network performance. The greatest variation between the curves in Fig. 8 was approximately 1% (case 3, training with 450 images). The results indicate that inserting IR images in a IVS database may not increase or decrease the network accuracy.

4 Conclusion

This article presents an investigation into the performance of YOLOv5 in detecting electrical equipment, using a combination of IVS and IR images for training. Based on the results presented, the following conclusions can be drawn:

- Incorporating IVS into an IR-based training dataset can have advantages in object detection in IR, for example in Exp. 1, case 5 (550 images in training). In this case, numerically, the detection index (mAP.50:.95) were, respectively, 0.685 for infrared training and 0.696 for IR training contaminated with IVS;

- Adding IR to training an IVS-base resulted in an improvement in network performance equivalent to using IVS alone, as shown in Fig. 7 and Fig. 8;

- Replacing IR images with IVS images in an IR image database, keeping the total number of images, can be an economical way to increase the database without necessarily relying on thermal imagers or thermography specialists, without reducing the accuracy of the classifier.

Thus, for practical purposes, as it is easier and less expensive to have a large database of IVS images than IR, it is concluded that it is advantageous to include IVS in training to improve IR detection, especially if the right proportions are considered.

5 Acknowledgements

This work was developed as part of the project GImpSI - Gestão dos Impactos da Salinidade em Isolamentos (Management of Salinity Impacts on Isolations) with INESC P&D Brasil and ENEVA S.A., under the framework of the R&D Program of the Brazilian Electricity Regulatory Agency, code PD-11278-0001-2021.

This study was financed in part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance Code 001.

6 References

[1] Russell, S., & Norvig, P. (2013). Inteligência artificial (R. C. Simille, Trad.). Rio de Janeiro, RJ: Campus Elsevier.

[2] Maharana, K., Mondal, S., & Nemade, B. "A review: Data pre-processing and data augmentation techniques," Global Transitions Proceedings, vol. 3, no. 1, 2022, pp. 91-99. doi: 10.1016/j.gltp.2022.04.020.

[3] Khalifa, N.E., Loey, M. & Mirjalili, S. "A comprehensive survey of recent trends in deep learning for digital images augmentation". Artif Intell Rev 55, 2022, pp. 2351–2377. Doi: https://doi.org/10.1007/s10462-021-10066-4.

[4] Shorten, C., & Khoshgoftaar, T. M.. "A survey on image data augmentation for deep learning". Journal of big data, 2019, 6, (1), pp. 1-48. Doi: https://doi.org/10.1186/s40537-019-0197-0

[5] Gorad, B., & Kotrappa, S. (2021, February). "Novel dataset generation for indian brinjal plant using image data augmentation". In IOP Conference Series: Materials Science and Engineering, 2021, 1065, (1), pp. 1-12. Doi: 10.1088/1757-899X/1065/1/012041

[6] Taylor, L., & Nitschke, G. (2018, November). "Improving deep learning with generic data augmentation". In 2018 IEEE symposium series on computational intelligence (SSCI), 2018, pp. 1542-1547.

[7] Xiao, Y., Decencière, E., Velasco-Forero, S., Burdin, H., Bornschlögl, T., Bernerd, F., ... & Baldeweck, T. (2019, April). "A new color augmentation method for deep learning segmentation of histological images". In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 2019, pp. 886-890, Doi: 10.1109/ISBI.2019.8759591.

[8] Gatys L, Ecker A, Bethge M. "A neural algorithm of artistic style". arXiv, 2015, pp. 1-16 Doi: https://doi.org/10.48550/arXiv.1508.06576

[9] Kollias, D., Cheng, S., Ververas, E., Kotsia, I., & Zafeiriou, S. "Deep neural network augmentation: Generating faces for affect analysis". International Journal of Computer Vision, 2020, 128, pp. 1455-1484, Doi: https://doi.org/10.1007/s11263-020-01304-3

[10] Zhu, X., Liu, Y., Li, J., Wan, T., & Qin, Z. (2018). "Emotion classification with data augmentation using generative adversarial networks". In Advances in Knowledge Discovery and Data Mining: 22nd Pacific-Asia Conference, PAKDD 2018, Melbourne, Australia, 2018, pp. 349-360.

[11] Khosla, C., & Saini, B. S. "Enhancing performance of deep learning models with different data augmentation techniques: A survey". In 2020 International Conference on Intelligent Engineering and Management (ICIEM), 2020, pp. 79-85, doi: 10.1109/ICIEM48762.2020.9160048.

[12] Mikołajczyk, A., & Grochowski, M. (2018, May). "Data augmentation for improving deep learning in image classification problem". In 2018 international interdisciplinary PhD workshop (IIPhDW), 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

[13] WANG, Jason et al. "The effectiveness of data augmentation in image classification using deep learning".

Convolutional Neural Networks Vis. Recognit, 2017, 11, (2017), p. 1-8.

[14] REZATOFIGHT, Hamid et al. “Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression”, Computer Vision and Pattern Recognition, 2019.