Ain Shams Engineering Journal xxx (xxxx) xxx

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: https://www.sciencedirect.com

Full Length Article

Hypertuned-YOLO for interpretable distribution power grid fault location based on EigenCAM

Stefano Frizzo Stefenon [a,b], Laio Oriel Seman [c], Anne Carolina Rodrigues Klaar [d], Raúl García Ovejero [e,*], Valderi Reis Quietinho Leithardt [f,g]

[a] Digital Industry Center, Fondazione Bruno Kessler, Trento, Italy
[b] Department of Mathematics, Computer Science and Physics, University of Udine, Udine, Italy
[c] Department of Automation and Systems Engineering, Federal University of Santa Catarina (UFSC), Florianópolis, Brazil
[d] Graduate Program in Education, University of Planalto Catarinense, Lages, Brazil
[e] Expert Systems and Applications Laboratory, E.T.S.I.I. of Béjar, University of Salamanca, 37700 Salamanca, Spain
[f] Lisbon School of Engineering (ISEL), Polytechnic University of Lisbon (IPL), 1549-020, Lisboa, Portugal
[g] Center of Technology and Systems (UNINOVA-CTS) and Associated Lab of Intelligent Systems (LASI), 2829-516, Caparica, Portugal

ARTICLE INFO

Keywords:
EigenCAM
Convolutional neural networks
You only look once
Power grids

ABSTRACT

Ensuring the reliability of electrical distribution networks is a pressing concern, especially given the power outages due to surface contamination on insulating components. Surface contamination can elevate surface conductivity, thereby resulting in failures that can lead to power shutdowns. Addressing this challenge, this paper proposes an approach for real-time monitoring of electrical distribution grids to prevent such incidents. A hypertuned version of the you only look once (YOLO) model is tailored for this application. We refine the model's hyperparameters by integrating a genetic algorithm to maximize its detection performance. The EigenCAM technique enhances the visual interpretability of the model's outcomes, providing operators with actionable insights for maintenance and monitoring tasks. Benchmark tests reveal that the proposed Hypertuned-YOLO outperforms Detectron (Masked R-CNN), YOLOv5, and YOLOv7 models. The Hypertuned-YOLO achieves an F1-score of 0.867 and a mAP@0.5 of 0.922, validating its robustness and efficacy.

1. Introduction

The expansion of the power supply has been increasing over the years, due to technological advances that help provide power with quality and continuity [1]. Telecontrolled systems help maintain the power grid working in hard-to-access places, reducing the time needed to reestablish electrical power when a shutdown occurs [2]. For the power grid to be stable and reliable, insulators need to be inspected to ensure they are in working condition. The main faults found in insulators are contamination over their surface [3], foreign objects close to them (such as bird's nests), fissures, and cracks [4].

The use of deep learning for the classification of adverse conditions is proving to be a promising approach [5] due to its capability to identify non-linearly separable patterns [6]. Considering that pixel-by-pixel classification is impractical using big data, convolutional neural networks (CNN) excel at using kernels that reduce the computational effort of the evaluation [7]. A major challenge in using CNN models is that in some cases there may be classification due to the background of the image [8], differences in clarity caused by sunlight variations, and classification due to objects that are not being evaluated [9].

Interpretive models are a solution because they generate a result that can be validated by the operator [10], and specialized object detection methods can be used for accurate fault identification. By highlighting the source of failure, interpretive models can be used by less specialized operators, making it easier for maintenance teams to use these approaches. Besides interpretive models, there is a tendency to use hybrid methods that use different backbones than the standard models or combine preprocessing techniques when the data set is limited [11].

The you only look once (YOLO) model performs object identification in a single step, making it a more efficient framework for object detection and classification, being typically superior to region-based CNN (R-CNN), fast or faster R-CNN models, which are widely applied for fault diagnosis [12]. One of the challenges in using YOLO is to define which version to use, as there is a trade-off between computational ef-

* Corresponding author at: Expert Systems and Applications Laboratory, University of Salamanca, Salamanca, Spain.
E-mail address: raulovej@usal.es (R.G. Ovejero).

https://doi.org/10.1016/j.asej.2024.102722
Received 14 June 2023; Received in revised form 8 September 2023; Accepted 16 February 2024

fort and desired speed [13], and for specific images, CNN models can perform well [14].

Considering the advantages of the YOLO structure, this paper proposes an optimization of the model based on hypertuning using a genetic algorithm and the application of the eigenvector-based class activation mapping (Eigen-CAM) to have interpretability of the results. The main contributions of this paper are:

- The first contribution is related to the dataset, in which the conditions of the components of the electrical network after inspections in the distribution system were annotated. This dataset is based on photos of distribution networks taken from the ground, a condition in which operators typically perform power system inspections.
- The second contribution is regarding the hypertuning of the model based on a genetic algorithm that optimizes its structure. This approach results in better performance than detectron (masked R-CNN), YOLOv5, YOLOv7, and YOLOv7x.
- The third contribution is given by the interpretability of the results which helps in understanding why the failure was identified, thus helping the maintenance team in the specific correction of the issue, here the Eigen-CAM is used for this goal.

The remainder of this paper is organized as follows: Section 2 presents a discussion on power system failures based on related works and applications of Eigen-CAM and YOLO for other fields. In Section 3 the proposed model is explained, in Section 4 the results and discussions about the application of the method are presented, and finally, in Section 5 final considerations are presented.

2. Study background

Electric power distribution systems are exposed to environmental variations because they are generally in an external environment [15], these variations have an impact on the occurrence of breakdowns and consequently on the quality of power distributed to consumers [16]. The contamination of insulators increases their surface conductivity resulting in increased leakage currents and as a consequence increases the chance of a disruptive failure [17].

The proximity of the mains to trees or of insulators to foreign objects (such as a bird's nest) can reduce the distance from the electrical insulation to the ground, resulting in a greater chance of an electric shutdown [18]. The main indicators that measure the quality of the power supply are based on calculating how many faults occur and how long it takes to restore power. Based on these indicators it is possible to assess the time and frequency of outages, which are used to determine whether the power supply is being carried out in accordance with the stated specifications [19].

Regarding contaminated insulators, the study of the time series of the leakage current has been evaluated in [20] (using a long short-term memory network) and [21] (based on the group method of data handling) to determine if there is a chance of a disruptive failure occurring. In [22] the Fourier transform from the ultrasonic time series of insulators in different conditions is evaluated. Other studies have been evaluating the conditions of the electrical system using the finite element method [23], besides the use of particle swarm optimization [24].

To identify anomalies in the power grid before shutdowns occur, power system inspections are conducted [25]. Inspections can be performed either by land, with a team passing by the power line and assessing its condition, or by air [26]. Currently, the use of unmanned aerial vehicles (UAVs) have been widely used [27], due to the difficulty in assessing hard-to-reach places and the cost of inspection when using manned aircraft [28], which can be used in combination with YOLO [29]. The main focus of the airborne inspection is the registration of images that are subsequently analyzed using image processing techniques [30].

In electrical system inspections, specific equipment is used to improve the understanding of the condition of the electrical power network [31]. Popular are devices based on sound capture (typically that cannot be heard by the human being) and techniques based on image classification [32]. One of the challenges in using specific equipment such as ultrasound is the need for advanced training and operator experience to interpret the results. For this reason, the use of deep learning models for image processing of electrical system inspections shows promise for identifying network failures [33].

CNNs have a high potential to be used for power system fault identification, given their capability to process large datasets it is promising to be applied in big data (case of electrical power system inspections) [34]. The main advantage of using models based on CNNs is the lower computational effort in using kernels for feature extraction (compared to classification pixels to pixels, which would be impractical), which makes it possible to evaluate the model using deep learning [35]. In which it is possible to apply models with background knowledge to improve the performance of the approach [36].

Despite well-established CNN models might have promising results for classification, in the case of power system images, there may be brightness or luminosity variation in the dataset which leads to false positives, i.e., to classify an object as a failure when it is not because the image background has properties that are present in images of defective grid components [30]. For this, object detection-based models are frequently used as presented by Ben Atitallah et al. [37], Li et al. [38], and Cai et al. [39] since they have the ability to identify exactly the location of the failure [40].

The YOLO stands out for performing detection in a single step [41], being more efficient than classic sliding window methods that require more computational effort [42]. Especially for fault detection, variations of the YOLO model are successfully applied to identify adverse conditions [43]. In the work by Sadykova et al. [28], augmentation techniques are applied to avoid overfitting when training the model. In [44] the YOLO model is improved by combining with other CNN classifiers.

Wu and Li [45] showed that, with an accuracy rate of 93.5%, the use of an improved version of YOLOv3 can be more accurate than Faster R-CNN. Besides YOLOv3, there are many versions of YOLO that can be used for object detection such as YOLOv4 [46], YOLOv5 [47], YOLOv6 [48], and more recently YOLOv7 [49]. Among these versions, one of the most popular is YOLOv5 [50], because it is a version that has major differences from its previous version and is developed based on the PyTorch framework, which facilitates the evaluation of this model, the latest versions are similar variations of this version.

Based on YOLO's high capability in identifying faults, adopting an interpretable approach to the model can further aid the understanding of the reason for the fault, thus helping the operator to identify the exact location where there may be a potential failure to be repaired [51]. Interpretability in CNNs refers to the capacity to comprehend the inner workings and decision-making processes of the network. This is a key feature of CNNs since it allows us to understand how the network is processing incoming data and making predictions, and it can also help us discover any potential network flaws or limitations [52].

Visualizing the feature maps generated by the convolutional layers is one technique for interpreting CNNs. As pictures, these feature maps represent the output of each convolutional layer. By evaluating these feature maps, we can learn how the network processes input data and what features it extracts [53]. Another method that can be used for interpreting CNNs is by using saliency maps. These maps illustrate the parts of the input image that have the most influence on the final forecast. By examining the saliency maps, it is possible to discern which portions of the input image the network is concentrating on while making a prediction [54].

The use of class activation maps (CAM) with deep learning approaches to have explainable artificial intelligence (XAI) and YOLO models is becoming popular in several fields in addition to the power system analysis, the next subsections cover other applications of these approaches highlighting its achievements.

2.1. CAMs for other fields

In [55] the interpretation of defects based on CAM for risk management is evaluated. They explore deep learning models for the recognition of structural defects in existing bridge heritage. Traditional inspection methods, while effective, are labor-intensive and time-consuming. The proposed models are a promising tool to automate defect recognition processes and improve the efficiency of bridge maintenance.

Jiang et al. Jiang et al. [56] explored the CAMs based on a novel approach called LayerCAM. The LayerCAM extends the concept of CAMs by introducing hierarchy to the localization process. LayerCAM harnesses the hierarchical structure within CNNs to generate multi-layered localization maps, offering a more fine-grained and interpretable understanding of object presence.

Marvasti-Zadeh et al. [57] presented the class activation mapping for tree crown detection (Crown-CAM), a methodology that combines deep learning with interpretability techniques to provide transparent and visually intuitive explanations for tree crown detection in aerial images. Crown-CAM leverages CNNs for tree crown detection while simultaneously generating CAMs that highlight the regions contributing to the network's decision. The Crown-CAM introduces a hierarchical interpretation scheme tailored to tree crown detection. It hierarchically segments the crown regions, providing insights into the structural characteristics of individual trees and their distribution within the image.

Yu et al. [58] applied an XAI model for fault diagnosis of rotating machinery, they combined a residual neural network using 6 layers (ResNet06) to two CAMs approaches, these being the gradient-based class activation map (Grad-CAM) and Eigen-CAM. These CAM techniques generate heatmaps that highlight the regions of input data contributing most to the fault diagnosis, thereby enabling engineers and operators to understand and trust the diagnostic results.

2.2. YOLO for other fields

According to Zheng et al. [59], the detection of failures in wind turbines is crucial for ensuring the safety and optimal performance of these energy systems. Based on the YOLOv5, they proposed an automated detection of surface cracks on the wind turbine blades. The proposed algorithm combines a deep learning model with enhancements tailored for the task of crack detection on curved and irregular surfaces.

YOLO has been used for a large range of applications, including the detection or identification of defects of steel surface [60], uneaten feed pellets [61], medical face masks [62], and traffic signs [63]. As presented by Li et al. [47], even considering infrared images this method shows promising results, proving that it is a promising framework to be used in various fields of research.

3. Object detection method

CNNs are neural networks commonly employed in image, navigation [64], remote sensing [65], and video processing jobs [66]. The fundamental component of a CNN is the convolutional layer, which extracts features from the input image using a collection of filters [67]. The mathematical operations of a convolutional layer are represented by the convolution operation, which is defined as follows:

$$(f * g)(n) = \sum_{m=-\infty}^{\infty} f(m)g(n-m) \tag{1}$$

where f and g are the input and filter respectively, and the output is represented by $(f * g)(n)$.

The YOLO detects objects in real-time. The design is made up of two primary components: a CNN for feature extraction and a fully linked layer for object detection [68]. The CNN is made up of several convolutional and max-pooling layers that extract information from the input image [69]. The CNN's final feature mappings are fed into a fully connected layer, which generates a collection of bounding boxes and class probabilities.

3.1. YOLO architecture

The object identification layer, which uses a multi-task loss function to predict bounding boxes and class probabilities, is the fundamental mathematical component of YOLO [70]. The architecture had a lot of improvements since its first release: YOLOv1 is the 2015 introduction of the first version of YOLO. It employs a single CNN to extract characteristics from the input image and to predict the bounding boxes and class labels for items inside the image. Real-time object detection was possible with YOLOv1, but accuracy and the capacity to recognize small items were limited. Introduced in 2016, YOLOv2 enhanced the original YOLO design by incorporating anchor boxes, a technique that enables the network to forecast bounding boxes at various scales and aspect ratios.

YOLOv2 also featured the Darknet-19 CNN architecture [71], which enhanced the accuracy of object detection. Released in 2018, YOLOv3 introduced a number of new features and improvements to the architecture, including a new CNN architecture known as Darknet-53, a technique known as feature pyramid networks to improve the detection of small objects, and a new method known as non-maximum suppression to improve the object detection.

YOLOv4 incorporates a number of new features, including a new backbone network, spatial pyramid pooling and path aggregation network for improved performance, and Mosaic data augmentation for improved generalization. Cross mini-batch normalization is a technique proposed by YOLOv4 to improve stability. YOLOv5 was announced in 2021, and it improves upon YOLOv4 by introducing new approaches, such as EfficientNet-Lite backbones and optimized multi-scale training and inference, which increase the accuracy and speed of object identification [72].

YOLOv5 employs a new grid-based method that allows it to recognize objects of varying sizes with greater precision. The major advantage of the latest versions YOLOv5, YOLOv6, and YOLOv7 is their flexibility in modifying the backbone size (based on the variation of the depth and width multiple), which is a promising option when the classification task has many nonlinearities [73]. The YOLO model is increasingly being improved in order to perform a more accurate fault identification, and combined with other methods to improve its capabilities [51].

Presently, the sum of localization loss, classification loss, and confidence loss constitutes the loss function of YOLO. Localization loss is the difference between the ground truth bounding box and the expected bounding box for grid cells including or without an object. The classification loss is the difference between the expected probability and actual probability of a class for grid cells containing an object. Formally, this is given by:

$$\begin{aligned} L = \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \Big[1_{ij}^{obj} \Big((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \Big) \\ + \lambda_{noobj} 1_{ij}^{noobj} \Big((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (w_i - \hat{w}_i)^2 + (h_i - \hat{hi})^2 \Big) \Big] \\ + \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} \sum_{c \in classes} \big(p_i(c) - \hat{p}_i(c) \big)^2 \end{aligned} \tag{2}$$

where (x_i, y_i, w_i, h_i) and $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$ are the coordinates and dimensions of the ground truth bounding box and the predicted bounding box, 1_{ij}^{obj} is an indicator function equal to 1 if the object is present in the $i, j-th$ grid cell and bounding box, and 0 otherwise, 1_{ij}^{noobj} is an indicator function equal to 1 if no object is present in the $i, j-th$ grid cell and bounding box, and 0 otherwise, S^2 is the number of grid cells in the feature map.

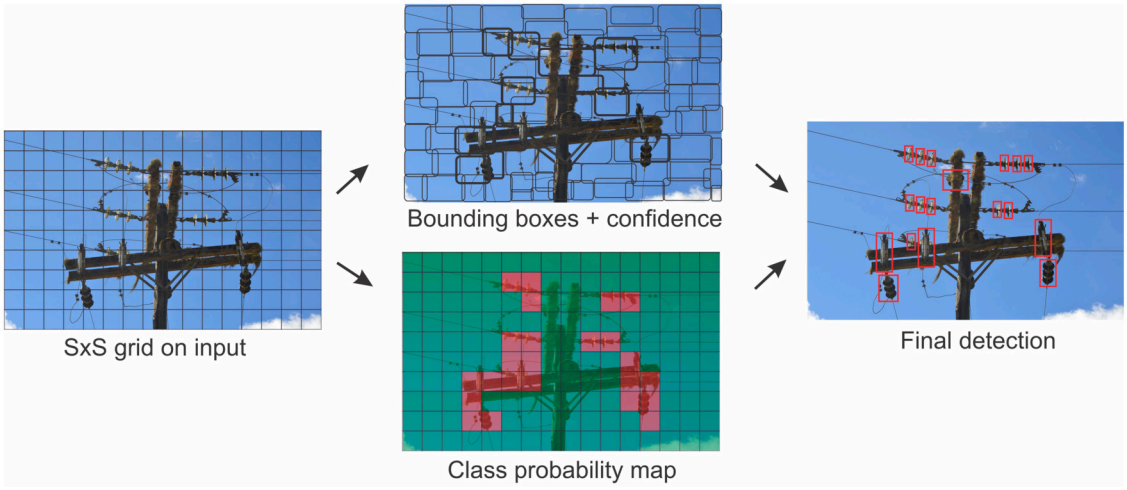

Fig. 1. The detection principle of YOLO method.

B is the number of bounding boxes per grid cell, $p_i(c)$ is the predicted probability of class c for the $i, j-th$ grid cell and bounding box, $\hat{p}_i(c)$ is the ground truth probability of class c for the $i, j-th$ grid cell and bounding box, λ_{coord} and λ_{noobj} are hyperparameters that control the relative importance of the localization and classification losses.

Let $P(class_i|x, y, w, h)$, $P(obj|x, y, w, h)$, and $P(x, y, w, h|obj)$ be the predicted probabilities for the classes, objectness and the bounding box parameters respectively:

$$P(class_i|x, y, w, h) = \frac{e^{(b_{i,j}+C_{i,j})}}{\sum_{c=0}^{C-1} e^{(b_{c,j}+C_{c,j})}} \tag{3}$$

$$P(obj|x, y, w, h) = \sigma(t_{i,j}) \tag{4}$$

$$P(x, y, w, h|obj) = \frac{1}{w_{obj} h_{obj}} \tag{5}$$

where x, y is the center of the bounding box, w, h is the width and height. During the training process, the loss function compares these predicted probabilities with the ground truth probabilities and adjusts the parameters of the model accordingly to minimize the loss.

The YOLO employs anchor boxes, which are pre-defined bounding boxes with varying aspect ratios [74]. The final bounding box is determined by comparing the predicted bounding box to these anchor boxes to see which one is the closest match to the ground truth box. This allows the network to detect objects with varying sizes and aspect ratios [73].

The anchor boxes are determined during training. The confidence of the predictions is given by the product of the probability of classes $P(class_i|object)$, the probability that the cell contains an object $P(obj|x, y, w, h)$, and the intersection over union (IoU) between the predicted bounding box and the ground truth bounding box, as follows:

$$P(class_i|object) \cdot P(object) \cdot IoU_{pred}^{truth} = P(class_i) \cdot IoU_{pred}^{truth} \tag{6}$$

where IoU_{pred}^{truth} is the IoU between the predicted and the ground truth bounding boxes. The working principle of the YOLO model is shown in Fig. 1.

3.2. Transfer learning

Transfer learning is an approach in which a pre-trained model, trained on a large dataset, is used as a starting point to solve a related but different task. The idea behind transfer learning is to leverage the knowledge learned by the pre-trained model on the original task and transfer it to the new task with the hope that it will improve the performance of the model on the new task [75].

$$\mathcal{L}(D_{new}, \theta) = \mathcal{L}(D_{new}, \theta_{pre-trained}) + \lambda \mathcal{L}(D_{pre-trained}, \theta_{pre-trained}) \tag{7}$$

where D_{new} is the new dataset, θ are the model parameters, $\theta_{pre-trained}$ are the pre-trained model parameters, $\mathcal{L}(D, \theta)$ is the loss function of the model on dataset D with parameters θ, and λ is a hyperparameter that controls the trade-off between the new task loss and the pre-trained task loss. In this work, pre-trained models from the fifth-generation YOLO released by Ultralytics [76] were used.

3.3. Eigen-CAM

Eigen-CAM aims to show the parts of an image that are significant for a certain class prediction provided by a CNN. Eigen-CAM's fundamental concept is to use the eigenvectors of the final convolutional layer's feature maps to determine the weights necessary to build the heatmap [77]. The eigenvectors are computed using the covariance matrix of the feature maps, which represents the relationships between various feature maps. Eigen-CAM is calculated as:

$$L_c = \sum_i \alpha_i^c A_i^c \tag{8}$$

where L_c is the output class score, A_i^c are the feature maps of the final convolutional layer, and α_i^c are the weights that are computed by taking the dot product between the eigenvectors of the feature maps A_i^c and the output class score L_c. The final heatmap is then obtained by taking the weighted sum of the feature maps A_i^c using the weights α_i^c.

Eigen-CAM can be considered superior to Grad-CAM because it is more resistant to changes in the feature maps and less susceptible to noise. By using the eigenvectors of the feature maps from the final convolutional layer to compute the weights that are used to build the heatmap, the approach may be used to show the portions of an image that are significant for a certain class prediction generated by a CNN. This method is less susceptible to noise and more resistant to changes in the feature maps than Grad-CAM [78]. The optimization of the hyperparameters of the CNN models makes them promising and therefore can be used in various applications related to fault identification in the power electrical system [79].

3.4. Hypertuning

A genetic algorithm (Algorithm 1) is an optimization approach that can be used to determine the optimal hyperparameters for machine

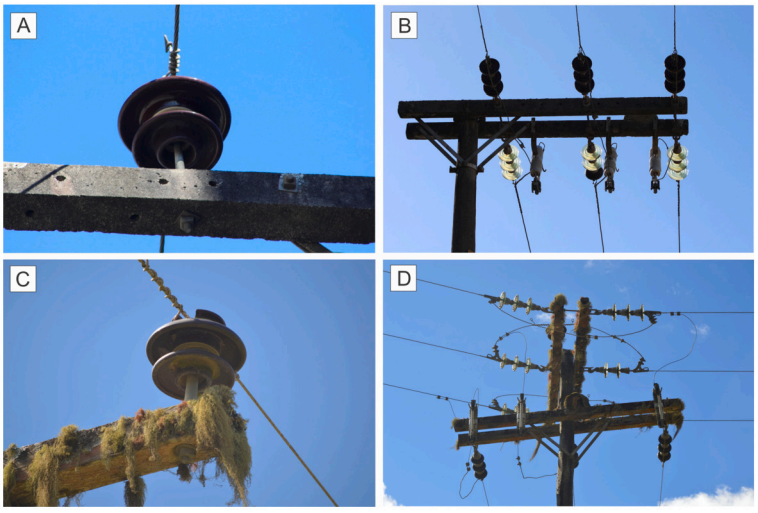

Fig. 2. Field inspection pictures of the dataset: A) insulator in good condition; B) structure in good condition; C) insulator that requires maintenance; D) structure that requires maintenance.

learning models such as YOLO. The main idea is to generate a population of diverse sets of hyperparameters and then evolve the population over numerous generations using a set of genetic operators (such as crossover and mutation). The objective is to identify the optimal hyperparameter values that optimize a performance metric, such as the mean average precision (mAP), which was considered in this paper.

Algorithm 1: Genetic algorithm.

```
Input  : Initial population P, number of generations g, crossover
         probability p_c, mutation probability p_m
Output: Optimal solution
for i ← 1 to g do
    for j ← 1 to |P| do
        if rand() < p_c then
            offspring_j ← crossover(select(P), select(P))
        if rand() < p_m then
            offspring_j ← mutation(offspring_j)
    P ← selection(offspring, P)
return best(P)
```

3.5. Dataset

The images used in this paper were registered during inspections of the electrical power distribution system in the state of Santa Catarina, in the southern region of Brazil. The distribution power grids are of the 13.8 kV class, which use pillar-type and mainly porcelain pin profile grid support insulators, and porcelain shell and suspension composite insulators.

The dataset has 240 images, of which 120 are photographs taken from structures in good condition, and the other 120 photographs are from inspections where maintenance was required. Considering the following conditions of structures that need maintenance: Contaminated or broken insulators; contaminated crosshead or fuse switch; foreign objects near the structure (birds' nests on the crosshead or fungi strongly attached). An example of images in different conditions of the used dataset is shown in Fig. 2.

The pictures taken during the inspection are in high resolution with mostly structures with contamination (Fig. 2D) and in good condition (Fig. 2B), some of the dataset images are specifically of insulators (Fig. 2A and Fig. 2C), which were zoomed in for a more detailed record of the component's condition. As contamination increases conductivity, maintenance should be carried out when there is a high concentration of contaminants on supporting structures of the grid or insulating components. For future comparisons, the dataset is available.[1]

3.6. Experiment

The experiments were performed in a cluster, in which the following requirements have been allocated: an NVIDIA Tesla V100S 32GB Graphics Card using 64GB of random access memory (RAM). The YOLOv5 (n, s, m, l, and x) [76], detectron (masked R-CNN) [80], and YOLOv7 (standard and x) [81] are evaluated. To avoid overfitting, the early stop criterion was used, in which training is ended when there is no improvement in the model. For practical purposes, a maximum value of 500 epochs was adopted. The precision, recall, and F1_score were used to evaluate the classification task, given by:

$$\text{precision} = \frac{tp}{tp + fp}, \tag{9}$$

$$\text{recall} = \frac{tp}{tp + fn}, \tag{10}$$

$$\text{F1\_score} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}}, \tag{11}$$

where tp is the true positive, fp is the false positive, and fn is the false negative.

For object detection, the mAP is considered, given by:

$$\text{mAP} = \frac{1}{n}\sum_{k=1}^{n}\left(\sum_{\eta}(\text{recall}_{\eta} - \text{recall}_{\eta-1})\text{precision}_{\eta}\right)_k \tag{12}$$

where n is the number of classes, k is the corresponding class, and η is the nth threshold.

4. Results and discussion

In this session the results of the application of the proposed approach will be presented, initially, the adequate version of YOLOv5 will be evaluated to detect components of the electrical distribution network, considering 2 classes: insulating components in good condition; and insulating components that need maintenance (with contamination, near strange objects, or broken). The best results of each evaluation are highlighted in bold.

[1] https://github.com/SFStefenon/InspectionDataSet.

Table 1
Object detection evaluation using YOLOv5 versions.

Model	Batch Size	Mem. (GB)	Precision	Recall	F1_score	mAP	
						[0.5]	[0.5:0.95]
YOLOv5n	4	0.59	0.82669	**0.92391**	**0.87260**	0.90134	0.82005
	8	1.08	0.85600	0.86957	0.86273	0.90017	0.76861
	16	1.89	0.81350	0.90217	0.85554	0.89150	0.80990
YOLOv5s	4	1.38	0.82568	0.86957	0.84706	0.90058	0.82619
	8	2.34	0.84309	0.87625	0.85935	0.88750	0.80582
	16	3.76	0.80449	0.86957	0.83576	0.86915	0.79712
YOLOv5m	4	2.37	0.80725	0.90217	0.85207	0.90246	0.83453
	8	3.83	0.80756	0.86957	0.83742	0.90344	0.85084
	16	6.41	0.79185	0.92391	0.85280	0.89035	0.84243
YOLOv5l	4	3.62	**0.82828**	0.89253	0.85921	**0.91969**	**0.88160**
	8	6.10	0.84028	0.89508	0.86681	0.88968	0.85256
	16	9.96	0.82548	0.88175	0.85269	0.87036	0.83472
YOLOv5x	4	5.22	0.81324	0.88043	0.84550	0.90210	0.84432
	8	8.99	0.81441	0.86957	0.84109	0.86916	0.82929
	16	14.2	0.81820	0.92391	0.86785	0.88657	0.84701

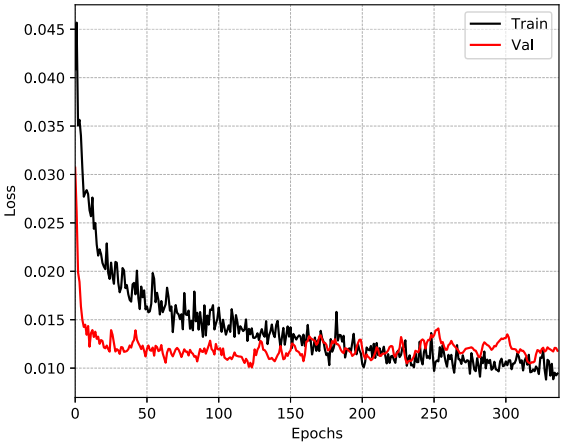

Fig. 3. Loss function evaluation.

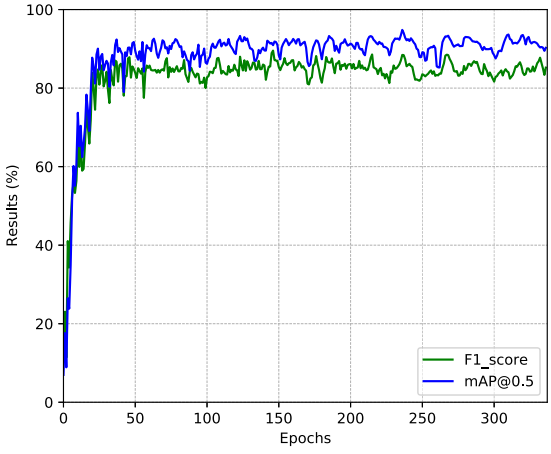

Fig. 4. F1_score and mAP@0.5 over the epochs.

The model training was considered completed when it reached the early stop criterion regarding the loss function, Fig. 3 shows an example of an experiment where model overfitting started after 170 epochs, however, the model stabilized only after 320 epochs when the validation loss (val) function became distant from the training loss (train) function and there was no improvement in the model.

4.1. Definition of the architecture

To define the framework for solving the classification task of the conditions of the power grid components the YOLOv5 was considered as the base model and its versions will be evaluated. In Table 1 the results of the evaluation of YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x are presented according to the batch size variation.

The first observation that can be made in this experiment, is that the models that use more parameters require more memory to run, whereas the smaller the batch size, the less memory capacity is required. Therefore when a benchmark is performed in which there is a limitation in processing power it is possible to reduce the batch size to be able to perform the analysis, disregarding the time needed for training.

Regarding the batch size for the YOLOv5 versions, in some cases there was a higher mAP using a smaller batch size value, however, it is not possible to affirm that the number of images loaded for training has a direct influence on the model performance, so it is necessary to perform analyses varying this parameter to define a better configuration. Considering that increasing the batch size did not bring better results and that it required more computational power, experiments were performed with a maximum value of 16. In a preliminary test to compute the YOLOv5x model (the largest of the YOLOv5 versions) with a batch size equal to 32, it is necessary to have 27.5 GB of GPU available.

In relation to mAP@0.5 and mAP@0.5:0.95, the best-performing model was YOLOv5l which did not present the best F1_score, however, there was the highest precision in the comparison between models of this class, for this reason, this model was considered more suitable in this evaluation. Moreover, the difference for the best F1_score was only 1.34%, being the value of 0.85921 acceptable for a preliminary analysis. An example of the performance of these metrics during the training is shown in Fig. 4.

It took about 30 hours to perform each experiment for each variation of the model and batch size. It was observed that there was no direct relationship between the size of the model or the batch size with the speed of convergence, the main reason that some models converged faster happened because the early stop was reached, as the weights are initialized randomly, the results have no meaning in terms of the model configuration (convergence time evaluation), for this reason, they were not presented.

4.2. Parameters optimization

Considering that the YOLOv5l had the best results in an initial evaluation (for the mAP), a hypertuning of its hyperparameters was performed to obtain an optimized structure. The result of applying

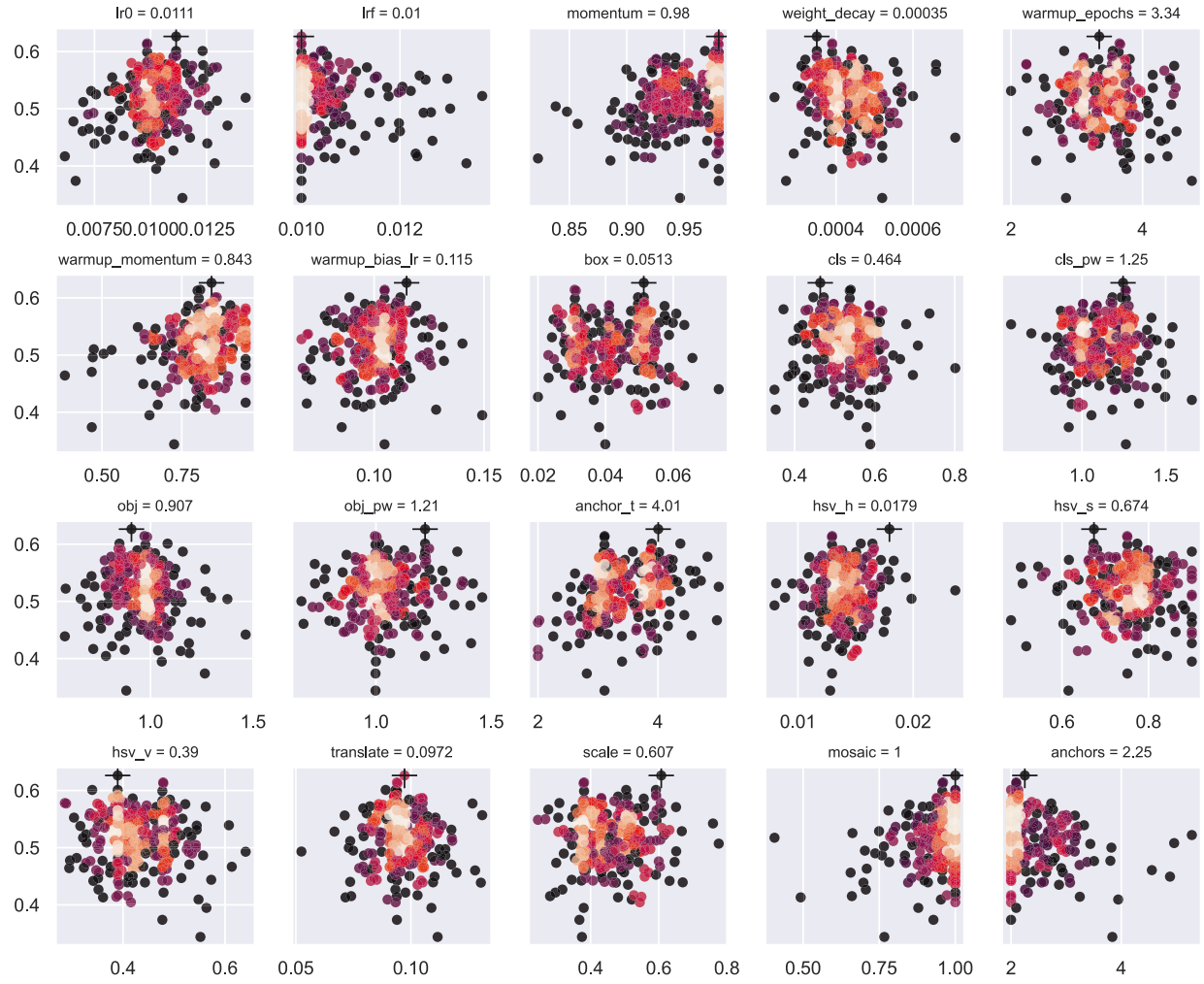

Fig. 5. Parameter sensibility.

the genetic algorithm for hyperparameter optimization, regarding its sensibility, is shown in Fig. 5 representing fitness (y-axis) versus hyperparameter values (x-axis), while the correlation between the parameters can be observed in Fig. 6.

The resulting best value for their use is presented for each component in Table 2 and was used as the default value for the hypertuning. Further, for a complete analysis, optimizers, batch sizes, and patience thresholds are evaluated. Fig. 7 shows how the metrics change regarding the sensibility during hypertuning. It was possible to observe that SGD was consistently the best optimizer for this given problem.

The procedure to evaluate the optimization of the hyperparameters performs several experiments and compares them according to the variation of the hyperparameters to define their best performance, this interactive process requires great computational effort, requiring 48 hours to compute the model considering 10 epochs per evolution.

4.3. Interpretability evaluation

Activation maps can be used to comprehend how CNNs make their predictions by revealing which features of the input image the network is focusing on. For instance, if an activation map displays high values in a particular region of the image, this indicates that the network is utilizing information from that region to make its prediction. By analyzing the activation maps, researchers can better comprehend how the network extracts the features from the input image and use them to make predictions.

Table 2
Best parameters resulting from the hypertuning procedure.

Parameter	Value
lr0	0.01114
lrf	0.01
momentum	0.98
weight_decay	0.00035
warmup_epochs	3.3407
warmup_momentum	0.84312
warmup_bias_lr	0.11476
box	0.05128
cls	0.4638
cls_pw	1.2452
obj	0.9074
obj_pw	1.2133
iou_t	0.2
anchor_t	4.0051
hsv_h	0.01794
hsv_s	0.67387
hsv_v	0.38951
translate	0.09722
scale	0.60727
mosaic	1.0
anchors	2.2496

In addition, activation maps can assist in identifying regions of the image where the model is not paying attention, allowing for a more interpretable explanation of the model's predictions. In addition, it can be

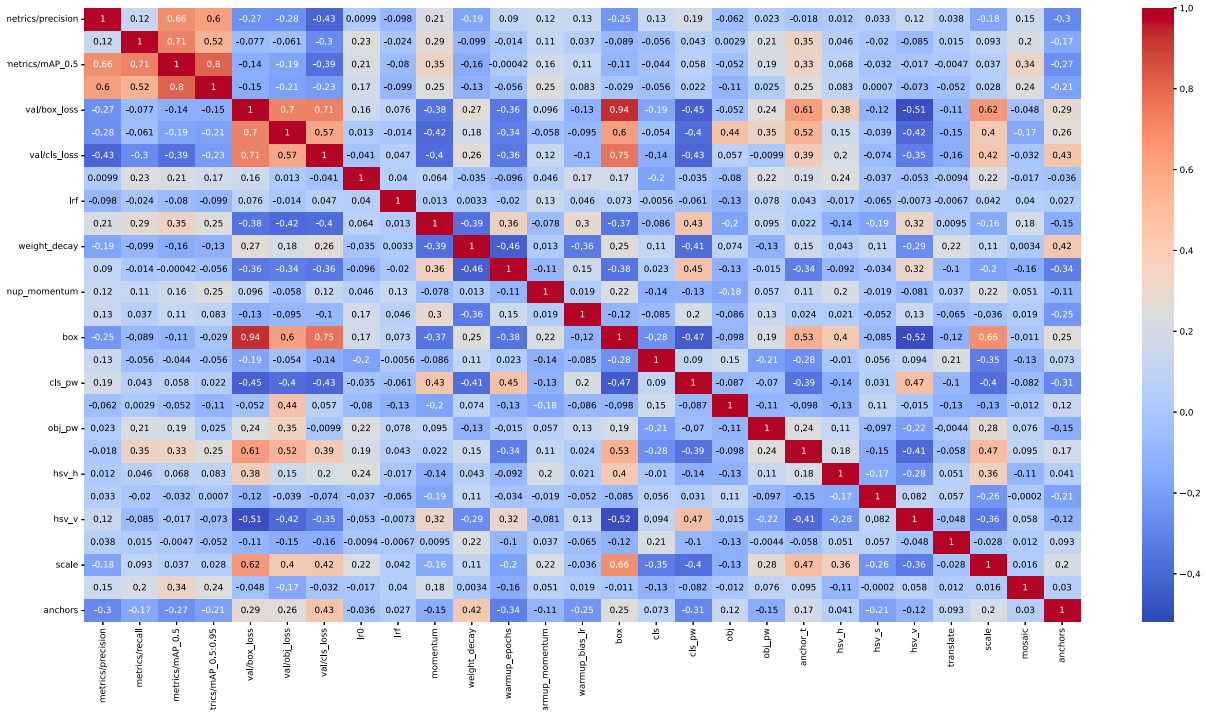

Fig. 6. Parameter sensibility correlation matrix.

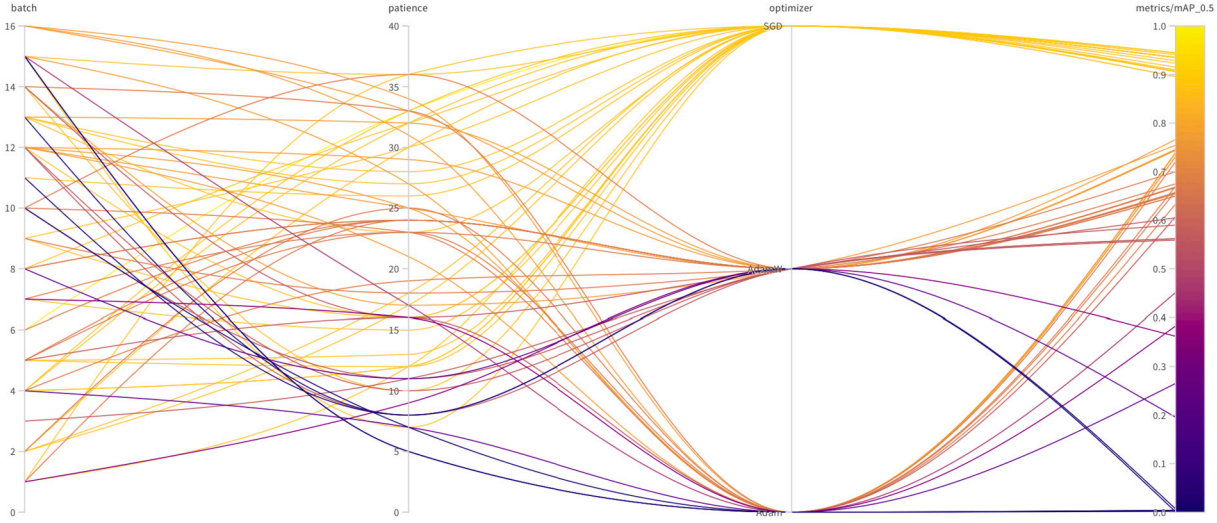

Fig. 7. Parameter sensibility for different optimizers, batch size, and patience thresholds.

used to identify the model's failure cases by identifying regions where the model is not paying attention, which could result in incorrect predictions. Eigen-CAM applied to four images (2 containing defective and 2 containing normal insulators) are presented in Fig. 8.

4.4. Benchmark

The benchmarking will be presented by comparing the proposed Hypertuned-YOLO to YOLOv5 (standard), YOLOv7, and detectron (masked R-CNN), the results of this evaluation are presented in Table 3. In version 7 of the YOLO model, two variations are available: YOLOv7 (std) and YOLOv7x. The variations of these versions were not as influenced by the early stop, thus the models required more time to converge and complete the training, each experiment took around 45 to 50 hours, and as with the other models the difference in simulation time was not significant due to the variation in batch size.

The results of the model based on Masked R-CNN were inferior in all metrics considered, thus the YOLO versions were a better alternative for the objectives proposed in this paper. The model that had its parameters optimized had an improvement in mAP@0.5 and was the second best in F1_score, showing to be promising for the application that was evaluated.

5. Final remarks

This study focused on deploying object detection algorithms, specifically a Hypertuned-YOLO model, to assess the condition of insulating components in electrical power distribution systems in Santa Catarina, Brazil. Utilizing a dataset of 240 inspection photographs taken from 13.8 kV-class power grids, the research examined the model's ability to distinguish between insulating components in good condition and those in need of repair or replacement.

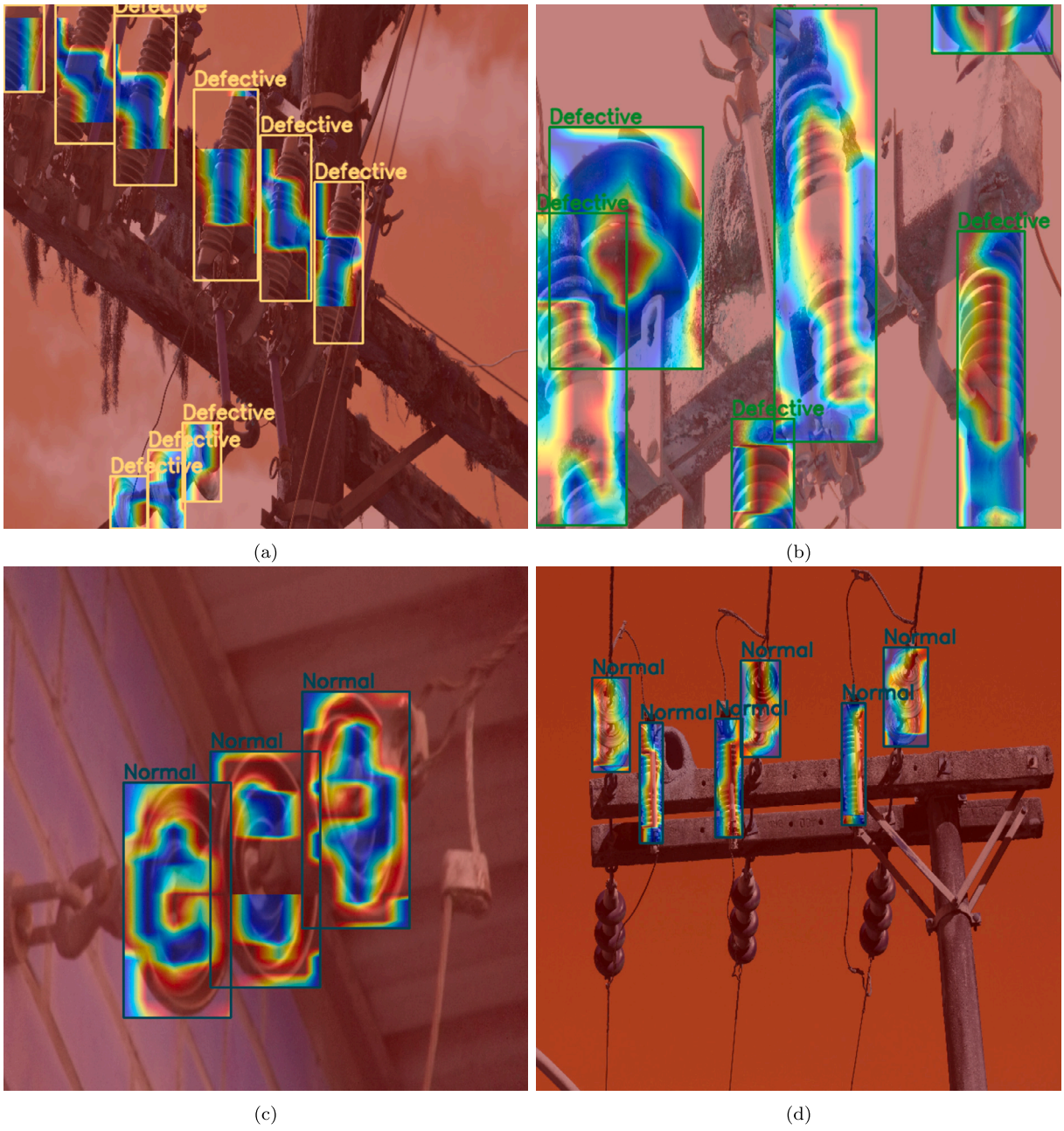

Fig. 8. Activation maps obtained with Eigen-CAM applied to YOLO5 (best model): A) Fuse switches and shell-type insulators with contamination; B) Fuse switches and pin-type insulators with contamination; C) Shell-type insulators in normal condition; D) Fuse switches and shell-type insulators in normal condition.

Table 3
Benchmark for object detection.

Model	Precision	Recall	F1_score	mAP		Time (h)
				[0.5]	[0.5:0.95]	
Masked R-CNN	0.81912	0.76964	0.79361	0.87508	0.76964	7.42
YOLOv5 (best)	0.82828	**0.89253**	0.85921	0.91969	**0.88160**	20.74
YOLOv7 (std)	**0.87006**	0.86663	**0.86834**	0.89976	0.85079	47.10
YOLOv7 (x)	0.82942	0.90075	0.86361	0.88617	0.84558	48.71
Our Method	0.84411	0.89130	0.86706	**0.92241**	0.80201	4.85

One notable aspect of this research was incorporating a genetic algorithm to optimize hyperparameters of the YOLO model. The optimization process successfully enhanced several performance metrics and reduced computational requirements. Specifically, the optimized model exhibited faster convergence rates during the training phase, outperforming standard YOLOv5 by 4.3 times, YOLOv7 by 9.7 times, and YOLOv7x by 10 times.

This efficiency in identifying faults is advantageous for real-world applications where frequent model updates and retraining are necessary. Evaluation criteria included precision, recall, F1-score, and

mAP@0.5. The performance results were promising, demonstrating that the method is suitable for practical applications in inspecting electrical power distribution systems. Eigen-CAM was integrated into the model to offer an explainable framework, facilitating better understanding and interpretation for end-users or system operators.

The findings of this study support the viability of using an optimized object detection model for reliable and efficient monitoring of electrical power distribution networks. The model's performance was comparable to, in some evaluation metrics, exceeded that of existing object detection models. This research thereby contributes to the ongoing efforts to improve the accuracy and efficiency of power grid inspections.

For future studies, a larger dataset could enable more complex modeling approaches, such as multi-classification, which would be beneficial for identifying a range of conditions affecting insulating components. It could also extend the model's capability to differentiate among other electrical infrastructure elements like fuse switches. Such advancements would further aid operators in conducting more comprehensive and effective inspections.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by "Caracterización, análisis e intervención en la prevención de riesgos laborales en entornos de trabajo tradicionales mediante la aplicación de tecnologías disruptivas". De la Consejeria de Empleo e Industria, under project 2022/00384/001. This research was funded (in part) by the Portuguese FCT program, Center of Technology and Systems (CTS), UIDB/00066/2020/UIDP/00066/2020.

References

[1] Seman LO, Stefenon SF, Mariani VC, dos Santos Coelho L. Ensemble learning methods using the Hodrick–Prescott filter for fault forecasting in insulators of the electrical power grids. Int J Electr Power Energy Syst 2023;152:109269. https://doi.org/10.1016/j.ijepes.2023.109269.

[2] Branco NW, Cavalca MSM, Stefenon SF, Leithardt VRQ. Wavelet LSTM for fault forecasting in electrical power grids. Sensors 2022;22(21):8323. https://doi.org/10.3390/s22218323.

[3] El-Refaie ESM, Abd Elrahman M, Mohamed MK. Electric field distribution of optimized composite insulator profiles under different pollution conditions. Ain Shams Eng J 2018;9(4):1349–56. https://doi.org/10.1016/j.asej.2016.08.012.

[4] Sopelsa Neto NF, Stefenon SF, Meyer LH, Ovejero RG, Leithardt VRQ. Fault prediction based on leakage current in contaminated insulators using enhanced time series forecasting models. Sensors 2022;22(16):6121. https://doi.org/10.3390/s22166121.

[5] Nguyen TP, Yeh CT, Cho MY, Chang CL, Chen MJ. Convolutional neural network bidirectional long short-term memory to online classify the distribution insulator leakage currents. Electr Power Syst Res 2022;208:107923. https://doi.org/10.1016/j.epsr.2022.107923.

[6] Liu Y, Li F, Guan Q, Zhao Y, Yan S. Power equipment fault diagnosis method based on energy spectrogram and deep learning. Sensors 2022;22(19):7330. https://doi.org/10.3390/s22197330.

[7] Liu Y, Pei S, Fu W, Zhang K, Ji X, Yin Z. The discrimination method as applied to a deteriorated porcelain insulator used in transmission lines on the basis of a convolution neural network. IEEE Trans Dielectr Electr Insul 2017;24(6):3559–66. https://doi.org/10.1109/TDEI.2017.006840.

[8] Singh G, Yow KC. Object or background: An interpretable deep learning model for COVID-19 detection from CT-Scan images. Diagnostics 2021;11(9):1732. https://doi.org/10.3390/diagnostics11091732.

[9] Corso MP, Stefenon SF, Singh G, Matsuo MV, Perez FL, Leithardt VRQ. Evaluation of visible contamination on power grid insulators using convolutional neural networks. Electrical Engineering 2023;105:3881–94. https://doi.org/10.1007/s00202-023-01915-2.

[10] Singh G. Think positive: An interpretable neural network for image recognition. Neural Netw 2022;151:178–89. https://doi.org/10.1016/j.neunet.2022.03.034.

[11] Stefenon SF, Corso MP, Nied A, Perez FL, Yow KC, Gonzalez GV, et al. Classification of insulators using neural network based on computer vision. IET Gener Transm Distrib 2021;16(6):1096–107. https://doi.org/10.1049/gtd2.12353.

[12] Liu X, Lin Y, Jiang H, Miao X, Chen J. Slippage fault diagnosis of dampers for transmission lines based on faster R-CNN and distance constraint. Electr Power Syst Res 2021;199:107449. https://doi.org/10.1016/j.epsr.2021.107449.

[13] Jiang P, Ergu D, Liu F, Cai Y, Ma B. A review of YOLO algorithm developments. Proc Comput Sci 2022;199:1066–73. https://doi.org/10.1016/j.procs.2022.01.135.

[14] Han S, Yang F, Yang G, Gao B, Zhang N, Wang D. Electrical equipment identification in infrared images based on roi-selected CNN method. Electr Power Syst Res 2020;188:106534. https://doi.org/10.1016/j.epsr.2020.106534.

[15] Salem AA, Abd-Rahman R, Ishak MTB, Lau KY, Abdul-Malek Z, Al-ameri S, et al. Influence of contamination distribution in characterizing the flashover phenomenon on outdoor insulator. Ain Shams Eng J 2023:102249. https://doi.org/10.1016/j.asej.2023.102249.

[16] Corso MP, Perez FL, Stefenon SF, Yow KC, Ovejero RG, Leithardt VRQ. Classification of contaminated insulators using k-nearest neighbors based on computer vision. Computers 2021;10(9):112. https://doi.org/10.3390/computers10090112.

[17] Cao B, Shen S, Li Z, Yang D, Wang L. Insulator contamination monitoring based on its hygroscopicity under unsaturated humidity. IEEE Trans Instrum Meas 2022;71:1–8. https://doi.org/10.1109/TIM.2022.3196450.

[18] Stefenon SF, Seman LO, Pavan BA, Ovejero RG, Leithardt VRQ. Optimal design of electrical power distribution grid spacers using finite element method. IET Gener Transm Distrib 2022;16(9):1865–76. https://doi.org/10.1049/gtd2.12425.

[19] Medeiros A, Sartori A, Stefenon SF, Meyer LH, Nied A. Comparison of artificial intelligence techniques to failure prediction in contaminated insulators based on leakage current. J Intell Fuzzy Syst 2022;42(4):3285–98. https://doi.org/10.3233/JIFS-211126.

[20] Klaar ACR, Stefenon SF, Seman LO, Mariani VC, Coelho LdS. Optimized EWT-Seq2Seq-LSTM with attention mechanism to insulators fault prediction. Sensors 2023;23(6):3202. https://doi.org/10.3390/s23063202.

[21] Stefenon SF, Seman LO, Sopelsa Neto NF, Meyer LH, Mariani VC, Coelho LdS. Group method of data handling using Christiano–Fitzgerald random walk filter for insulator fault prediction. Sensors 2023;23(13):6118. https://doi.org/10.3390/s23136118.

[22] Stefenon SF, Oliveira JR, Coelho AS, Meyer LH. Diagnostic of insulators of conventional grid through labview analysis of fft signal generated from ultrasound detector. IEEE Lat Am Trans 2017;15(5):884–9. https://doi.org/10.1109/TLA.2017.7910202.

[23] Stefenon SF, Americo JP, Meyer LH, Grebogi RB, Nied A. Analysis of the electric field in porcelain pin-type insulators via finite elements software. IEEE Lat Am Trans 2018;16(10):2505–12. https://doi.org/10.1109/TLA.2018.8795129.

[24] Stefenon SF, Furtado Neto CS, Coelho TS, Nied A, Yamaguchi CK, Yow KC. Particle swarm optimization for design of insulators of distribution power system based on finite element method. Electr Eng 2022;104(2):615–22. https://doi.org/10.1007/s00202-021-01332-3.

[25] Manninen H, Ramlal CJ, Singh A, Kilter J, Landsberg M. Multi-stage deep learning networks for automated assessment of electricity transmission infrastructure using fly-by images. Electr Power Syst Res 2022;209:107948. https://doi.org/10.1016/j.epsr.2022.107948.

[26] Nguyen VN, Jenssen R, Roverso D. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. Int J Electr Power Energy Syst 2018;99:107–20. https://doi.org/10.1016/j.ijepes.2017.12.016.

[27] Alsadik B, Remondino F, Nex F. Simulating a hybrid acquisition system for UAV platforms. Drones 2022;6(11):314. https://doi.org/10.3390/drones6110314.

[28] Sadykova D, Pernebayeva D, Bagheri M, James A. IN-YOLO: Real-time detection of outdoor high voltage insulators using UAV imaging. IEEE Trans Power Deliv 2020;35(3):1599–601. https://doi.org/10.1109/TPWRD.2019.2944741.

[29] Tan L, Lv X, Lian X, Wang G. YOLOv4_Drone: UAV image target detection based on an improved YOLOv4 algorithm. Comput Electr Eng 2021;93:107261. https://doi.org/10.1016/j.compeleceng.2021.107261.

[30] Stefenon SF, Yow KC, Nied A, Meyer LH. Classification of distribution power grid structures using inception v3 deep neural network. Electr Eng 2022;104:4557–69. https://doi.org/10.1007/s00202-022-01641-1.

[31] Ahmed M, Mohanta J, Sanyal A. Inspection and identification of transmission line insulator breakdown based on deep learning using aerial images. Electr Power Syst Res 2022;211:108199. https://doi.org/10.1016/j.epsr.2022.108199.

[32] Stefenon SF, Bruns R, Sartori A, Meyer LH, Ovejero RG, Leithardt VRQ. Analysis of the ultrasonic signal in polymeric contaminated insulators through ensemble learning methods. IEEE Access 2022;10:33980–91. https://doi.org/10.1109/ACCESS.2022.3161506.

[33] Hao K, Chen G, Zhao L, Li Z, Liu Y, Wang C. An insulator defect detection model in aerial images based on multiscale feature pyramid network. IEEE Trans Instrum Meas 2022;71:1–12. https://doi.org/10.1109/TIM.2022.3200861.

[34] Lei X, Sui Z. Intelligent fault detection of high voltage line based on the faster R-CNN. Measurement 2019;138:379–85. https://doi.org/10.1016/j.measurement.2019.01.072.

[35] Zhao Z, Zhen Z, Zhang L, Qi Y, Kong Y, Zhang K. Insulator detection method in inspection image based on improved faster R-CNN. Energies 2019;12(7):1204. https://doi.org/10.3390/en12071204.

[36] Grilli E, Daniele A, Bassier M, Remondino F, Serafini L. Knowledge enhanced neural networks for point cloud semantic segmentation. Remote Sens 2023;15(10):2590. https://doi.org/10.3390/rs15102590.

[37] Ben Atitallah A, Said Y, Ben Atitallah MA, Albekairi M, Kaaniche K, Alanazi TM, et al. Embedded implementation of an obstacle detection system for blind and visually impaired persons' assistance navigation. Comput Electr Eng 2023;108:108714. https://doi.org/10.1016/j.compeleceng.2023.108714.
[38] Li H, Ge S, Gao C, Gao H. Few-shot object detection via high-and-low resolution representation. Comput Electr Eng 2022;104:108438. https://doi.org/10.1016/j.compeleceng.2022.108438.
[39] Cai J, Makita Y, Zheng Y, Takahashi S, Hao W, Nakatoh Y. Single shot multibox detector for honeybee detection. Comput Electr Eng 2022;104:108465. https://doi.org/10.1016/j.compeleceng.2022.108465.
[40] Li M, Wang H, Wan Z. Surface defect detection of steel strips based on improved YOLOv4. Comput Electr Eng 2022;102:108208. https://doi.org/10.1016/j.compeleceng.2022.108208.
[41] Alghamdi S, Algethami A, Tan T. Vehicle-camel collisions in Saudi Arabia: application of single and multi-stage deep learning object detectors. Ain Shams Eng J 2023:102328. https://doi.org/10.1016/j.asej.2023.102328.
[42] Ji SJ, Ling QH, Han F. An improved algorithm for small object detection based on YOLO v4 and multi-scale contextual information. Comput Electr Eng 2023;105:108490. https://doi.org/10.1016/j.compeleceng.2022.108490.
[43] Qian X, Wang X, Yang S, Lei J. LFF-YOLO: A YOLO algorithm with lightweight feature fusion network for multi-scale defect detection. IEEE Access 2022;10:130339–49. https://doi.org/10.1109/ACCESS.2022.3227205.
[44] Stefenon SF, Singh G, Souza BJ, Freire RZ, Yow KC. Optimized hybrid YOLOu-Quasi-ProtoPNet for insulators classification. IET Gener Transm Distrib 2023;17(15):3501–11. https://doi.org/10.1049/gtd2.12886.
[45] Wu W, Li Q. Machine vision inspection of electrical connectors based on improved YOLO v3. IEEE Access 2020;8:166184–96. https://doi.org/10.1109/ACCESS.2020.3022405.
[46] Zhao S, Liu B, Chi Z, Li T, Li S. Characteristics based fire detection system under the effect of electric fields with improved YOLO-v4 and ViBe. IEEE Access 2022;10:81899–909. https://doi.org/10.1109/ACCESS.2022.3190867.
[47] Li S, Li Y, Li Y, Li M, Xu X. YOLO-FIRI: Improved YOLOv5 for infrared image object detection. IEEE Access 2021;9:141861–75. https://doi.org/10.1109/ACCESS.2021.3120870.
[48] Aburaed N, Alsaad M, Mansoori SA, Al-Ahmad H. A study on the autonomous detection of impact craters. In: El Gayar N, Trentin E, Ravanelli M, Abbas H, editors. Artificial neural networks in pattern recognition. Cham: Springer International Publishing; 2023. p. 181–94.
[49] Wang Y, Wang H, Xin Z. Efficient detection model of steel strip surface defects based on YOLO-v7. IEEE Access 2022;10:133936–44. https://doi.org/10.1109/ACCESS.2022.3230894.
[50] Liu L, Liang J, Wang J, Hu P, Wan L, Zheng Q. An improved YOLOv5-based approach to soybean phenotype information perception. Comput Electr Eng 2023;106:108582. https://doi.org/10.1016/j.compeleceng.2023.108582.
[51] Singh G, Stefenon SF, Yow KC. Interpretable visual transmission lines inspections using pseudo-prototypical part network. Mach Vis Appl 2023;34(3):41. https://doi.org/10.1007/s00138-023-01390-6.
[52] Singh G, Yow KC. These do not look like those: An interpretable deep learning model for image recognition. IEEE Access 2021;9:41482–93. https://doi.org/10.1109/ACCESS.2021.3064838.
[53] Hohman F, Park H, Robinson C, Polo Chau DH. Summit: Scaling deep learning interpretability by visualizing activation and attribution summarizations. IEEE Trans Vis Comput Graph 2020;26(1):1096–106. https://doi.org/10.1109/TVCG.2019.2934659.
[54] Zhao L, Zeng Y, Liu P, Su X. Band selection with the explanatory gradient saliency maps of convolutional neural networks. IEEE Geosci Remote Sens Lett 2020;17(12):2105–9. https://doi.org/10.1109/LGRS.2020.3012140.
[55] Cardellicchio A, Ruggieri S, Nettis A, Renò V, Uva G. Physical interpretation of machine learning-based recognition of defects for the risk management of existing bridge heritage. Eng Fail Anal 2023;149:107237. https://doi.org/10.1016/j.engfailanal.2023.107237.
[56] Jiang PT, Zhang CB, Hou Q, Cheng MM, Wei Y. Layercam: Exploring hierarchical class activation maps for localization. IEEE Trans Image Process 2021;30:5875–88. https://doi.org/10.1109/TIP.2021.3089943.
[57] Marvasti-Zadeh SM, Goodsman D, Ray N, Erbilgin N. CROWN-CAM: Interpretable visual explanations for tree crown detection in aerial images. IEEE Geosci Remote Sens Lett 2023;20:1–5. https://doi.org/10.1109/LGRS.2023.3271649.
[58] Yu S, Wang M, Pang S, Song L, Qiao S. Intelligent fault diagnosis and visual interpretability of rotating machinery based on residual neural network. Measurement 2022;196:111228. https://doi.org/10.1016/j.measurement.2022.111228.
[59] Zheng Y, Liu Y, Wei T, Jiang D, Wang M. Wind turbine blades surface crack-detection algorithm based on improved YOLO-v5 model. J Electron Imaging 2023;32(3):033012. https://doi.org/10.1117/1.JEI.32.3.033012.
[60] Guo Z, Wang C, Yang G, Huang Z, Li G. MSFT-YOLO: Improved YOLOv5 based on transformer for detecting defects of steel surface. Sensors 2022;22(9):3467. https://doi.org/10.3390/s22093467.
[61] Hu X, Liu Y, Zhao Z, Liu J, Yang X, Sun C, et al. Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-v4 network. Comput Electron Agric 2021;185:106135. https://doi.org/10.1016/j.compag.2021.106135.
[62] Loey M, Manogaran G, Taha MHN, Khalifa NEM. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustain Cities Soc 2021;65:102600. https://doi.org/10.1016/j.scs.2020.102600.
[63] Dewi C, Chen RC, Liu YT, Jiang X, Hartomo KD. YOLO v4 for advanced traffic sign recognition with synthetic training data generated by various GAN. IEEE Access 2021;9:97228–42. https://doi.org/10.1109/ACCESS.2021.3094201.
[64] dos Santos GH, Seman LO, Bezerra EA, Leithardt VRQ, Mendes AS, Stefenon SF. Static attitude determination using convolutional neural networks. Sensors 2021;21(19):6419. https://doi.org/10.3390/s21196419.
[65] Kattenborn T, Leitloff J, Schiefer F, Hinz S. Review on convolutional neural networks (CNN) in vegetation remote sensing. ISPRS J Photogramm Remote Sens 2021;173:24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010.
[66] Lu S, Wang B, Wang H, Chen L, Linjian M, Zhang X. A real-time object detection algorithm for video. Comput Electr Eng 2019;77:398–408. https://doi.org/10.1016/j.compeleceng.2019.05.009.
[67] Borré A, Seman LO, Camponogara E, Stefenon SF, Mariani VC, Coelho LS. Machine fault detection using a hybrid CNN-LSTM attention-based model. Sensors 2023;23(9):4512. https://doi.org/10.3390/s23094512.
[68] Huang J, Chen J, Wang H. A lightweight and efficient one-stage detection framework. Comput Electr Eng 2023;105:108520. https://doi.org/10.1016/j.compeleceng.2022.108520.
[69] Li M, Zhang Z, Lei L, Wang X, Guo X. Agricultural greenhouses detection in high-resolution satellite images based on convolutional neural networks: Comparison of faster R-CNN, YOLO v3 and SSD. Sensors 2020;20(17):4938. https://doi.org/10.3390/s20174938.
[70] Chen H, He Z, Shi B, Zhong T. Research on recognition method of electrical components based on YOLO V3. IEEE Access 2019;7:157818–29. https://doi.org/10.1109/ACCESS.2019.2950053.
[71] Manninen H, Ramlal CJ, Singh A, Rocke S, Kilter J, Landsberg M. Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques. Int J Electr Power Energy Syst 2021;128:106726. https://doi.org/10.1016/j.ijepes.2020.106726.
[72] Deng F, Xie Z, Mao W, Li B, Shan Y, Wei B, et al. Research on edge intelligent recognition method oriented to transmission line insulator fault detection. Int J Electr Power Energy Syst 2022;139:108054. https://doi.org/10.1016/j.ijepes.2022.108054.
[73] Souza BJ, Stefenon SF, Singh G, Freire RZ. Hybrid-YOLO for classification of insulators defects in transmission lines based on UAV. Int J Electr Power Energy Syst 2023;148:108982. https://doi.org/10.1016/j.ijepes.2023.108982.
[74] Alrowais F, Asiri MM, Alabdan R, Marzouk R, Hilal AM, alkhayyat A, et al. Hybrid leader based optimization with deep learning driven weed detection on Internet of things enabled smart agriculture environment. Comput Electr Eng 2022;104:108411. https://doi.org/10.1016/j.compeleceng.2022.108411.
[75] Hua J, Zeng L, Li G, Ju Z. Learning for a robot: Deep reinforcement learning, imitation learning, transfer learning. Sensors 2021;21(4):1278. https://doi.org/10.3390/s21041278.
[76] Ultralytics G. YOLOv5 in PyTorch. Available from: https://github.com/ultralytics/yolov5, 2022.
[77] Bany Muhammad M, Yeasin M. Eigen-CAM: Visual explanations for deep convolutional neural networks. SN Computer Science 2021;2:47. https://doi.org/10.1007/s42979-021-00449-3.
[78] Rahman AN, Andriana D, Machbub C. Comparison between Grad-CAM and Eigen-CAM on YOLOv5 detection model. In: 2022 international symposium on electronics and smart devices (ISESD), vol. 1. Bandung, Indonesia: IEEE; 2022. p. 1–5.
[79] Uckol HI, Ilhan S, Ozdemir A. Workmanship defect classification in medium voltage cable terminations with convolutional neural network. Electr Power Syst Res 2021;194:107105. https://doi.org/10.1016/j.epsr.2021.107105.
[80] Wu Y, Kirillov A, Massa F, Lo WY, Girshick R. Detectron2. Available from: https://github.com/facebookresearch/detectron2, 2019.
[81] Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. https://doi.org/10.48550/ARXIV.2207.02696, 2022.