Machine Vision and Applications (2023) 34:41
https://doi.org/10.1007/s00138-023-01390-6

ORIGINAL PAPER

Interpretable visual transmission lines inspections using pseudo-prototypical part network

Gurmail Singh[1] · Stefano Frizzo Stefenon[2,3] · Kin-Choong Yow[4]

Received: 26 May 2022 / Revised: 27 February 2023 / Accepted: 8 March 2023 / Published online: 5 April 2023

Abstract
To guarantee the reliability of the electric energy supply, it is necessary that the transmission lines are operating without interruptions. To improve the identification of faults in the electrical power system, the unmanned aerial vehicle is used for inspection by recording photos. Based on computer vision, deep learning structures stand out for image classification have been an alternative to improve the identification of defects in transmission lines inspections. In this paper, the Pseudo-Prototypical Part Network (Ps-ProtoPNet) model is applied to perform the classification of missing insulators of high voltage transmission lines. To identify the position of the insulators chain and have the focus of the classification on the difference of insulators with failure, the YOLOv5 (n, s, m, l, and x), YOLOv6 (n, t, s, m, and l), YOLOv7 (std and x), and YOLOv8 (n, s, m, l, and x) are compared. The YOLOv8m is defined as the standard architecture for object detection since it has an mAP[0.5] of 0.9950 and mAP[0.5:0.95] of 0.9125. To classify the images, the Ps-ProtoPNet compares its various parts with the prototypes from all classes, and the image is classified based on the closest similarity to the prototypes class. The results show that the Ps-ProtoPNet achieves accuracy values sufficient to be applied in field inspections.

Keywords Deep learning · Power line inspection · Transmission line faults

1 Introduction

Transmission lines are responsible for connecting electricity generation plants to large consumption centers. As the generation plants are designed according to the relief characteristics, often the distance between generation and load can be considerably big, especially when the transmission system's objective is to balance the energy supply between distant regions [1]. One of the main components of transmission networks are the insulators, responsible for supporting the cables and isolating the electrical potential. To guarantee the supply of electrical energy, it is necessary to identify insulators that are faulty, since the non-insulation of the high voltage can result in leakage current and disruptive discharges that can lead to shutdowns [2]. Identifying insulators that are damaged is a difficult task, as the network is often installed in hard-to-reach places and faulty insulators can have hidden flaws.

Some specific equipment can be used to improve the fault identification in the electrical power system, as the ultrasound, a directional equipment that is used in inspections of the power network [3]. This equipment is promising for locating defects because it is directional, which facilitates the exact identification of the source of ultrasonic noise that may be related to the failure [4]. The great difficulty in using this type of equipment for electrical inspection is that it requires a lot of experience from the operator and noise in the vicinity can make it difficult to identify the fault, as it generates interference in the signal [5]. Based on these characteristics, image analysis could be a better alternative, considering that

✉ Gurmail Singh
gurmail.singh@wisc.edu

Stefano Frizzo Stefenon
sfrizzostefenon@fbk.eu

Kin-Choong Yow
kin-choong.yow@uregina.ca

1 Department of Computer Sciences, University of Wisconsin-Madison, 1210 W. Dayton Street, Madison, WI 53706-1613, USA

2 Digital Industry Center, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Trento, TN, Italy

3 Department of Mathematics, Computer Science and Physics, University of Udine, Via delle Scienze 206, 33100 Udine, UD, Italy

4 Faculty of Engineering and Applied Science, University of Regina, 3737 Wascana Pkwy, Regina, SK S4S 0A2, Canada

investments in specific equipment are not necessary and the analysis of the inspection can be performed even offline [6].

The inspection of the electrical power system has evolved over the years. It is increasingly common to apply artificial intelligence techniques to interpret the results of the inspections of the power grid [7]. In addition to the evaluation of failures using specific equipment, the analysis can also be performed regarding images of the aerial inspections. In this sense computer vision algorithms are applied to classify defects to improve the reliability of the network [8]. One of the main components that are evaluated in the inspections of the electrical power system are the insulators, which are responsible for carrying out the isolation of the electrical potential [9].

With a focus on improving the identification of defects in transmission lines, some authors have carried out promising studies using deep learning. According to Manninen et al. [10], to ensure safety and reliability in the transmission of electricity, utilities invest in improving the diagnosis of the network's infrastructure, which is usually evaluated mainly by visual inspection. In addition to the study of specific components such as insulators, there is a need to evaluate the structures that support the network, such as posts and transmission towers. The results show that from deep learning models it is possible to improve the detection of defects in the power system, which can reduce the need for visual inspection and improve the speed and accuracy of fault diagnosis.

According to Fahim et al. [11] detection of transmission line failures is a key factor in the fast restoration of the electrical power grid. Using deep learning, it is possible to extract representative features of the power system conditions that can aid failure analysis. To have a fault diagnosis using deep learning it is necessary to have a wide variety of representative fault conditions, which can be difficult considering that faults are difficult to be identified. Alternatively, an unsupervised algorithm can be applied to classify faults in transmission lines based on the capsule network. This approach can voluntarily learn the failure features and thus improve performance in processing large datasets.

The use of deep learning to identify faults extends to the evaluation of gas isolated switches, which are important components of the electric power system that are part of substations connected to transmission lines [12]. As with insulators and switches, failures can be related to partial discharges, which when not monitored can lead in disruptive failures resulting in a threat to the safe operation of the electrical power network [13].

Currently, deep learning is standing out for the identification of adverse conditions in the electrical power system [14]. From a model of a convolutional neural network (CNN) it is possible to locate short-circuit faults in transmission lines [15], disturbances caused by lightning [16], damaged structures [17], high impedance conditions [18], and network transients [19]. Among the most advanced deep learning models for computer vision, pseudo-prototypical part network (Ps-ProtoPNet) stands out due to its interpretability capacity of the adverse condition, thus facilitating the classification of images [20]. To improve the accuracy of Ps-ProtoPNet, the analysis is based on specific images of the insulators chain, the identification of the position where the insulators chain are located is performed using the *you only look once* (YOLO) architecture.

The contributions of this paper to the inspection of the electrical power grid are summarized below:

- The first contribution is the improvement in the identification of failures in transmission lines using deep learning to classify adverse conditions. From the proposed method is possible to improve the identification of failures with high accuracy before discharges occur, preventing network shutdowns.
- The second contribution is the interpretability of the model. The interpretability helps to understand why the deep learning model classifies the object in a certain way. This will help to identify if the algorithm is making a classification using the background instead of the object itself, which is a common problem for this type of assessment.
- The third contribution is related to the improvement in the model due to the combination of YOLOv8m to identify the position of the chain of insulators and the Ps-ProtoPNet model to perform the classification. This combination shows to be superior to models proposed by other authors, as presented in this paper.

Table 1 defines the symbols used in this paper.

This paper is organized as follows: Sect. 2 describes a background on transmission line inspections using aerial images and the dataset used for the analysis of the proposed method is explained. Section 3 presents the description of the proposed method for analyzing aerial inspection images of electrical power transmission lines. Section 4 evaluates the model, compares the impact of changing the base model on the improvement in the accuracy, and a comparison to other authors is presented. Finally, Sect. 5 presents the final considerations about this work.

2 Transmission lines inspection using aerial images

To maintain the availability of electricity supply, electricity utilities conduct inspections in the electric power system [21]. Inspections are typically carried out on foot or using helicopters. The objective of these evaluations is to plan maintenance on the network and thus avoid system faults

Table 1 Symbol list

Symbol	Definition
d	Distance function
d_l	Dense layer
h and w	Spatial dimensions of the prototype
k	Class of the prototype
i and j	Iteration
ℓ	Convolutional layer
$\ell(x)$	Output of the convolutional layer
m	Number of prototypes for each class
m_w	Matrix of weight
n	Length of the signal
n_c	Number of classes
p	Prototypes
p_p	Prototypical unit
x	Input image
x_s and y_t	Row vector
y	Observed value
$\hat{y}$	Output value
z	Output the convolutional layer ℓ
E	Loss function
P	Set of prototypes
$\mathcal{Z}$	Patch of z
α	Learning rate
β	Dependent variable
γ	Independent variable
η	Linear coefficient
θ	Parameter vector to be optimized
λ_1 and λ_2	Hyperparameter
ρ	Slope

such as shutdowns [22]. Although this approach is carried out by most electric utilities, it is an expensive and slow solution, given the difficulty of access on foot and the high cost of using helicopters. Currently, research is being carried out to automate electrical inspection, using artificial intelligence techniques and automated aircraft [23].

Although, there are several advantages of vision-based inspection, this approach is not widely used. As a result of technological evolution, it becomes possible to carry out inspection of transmission lines from automatic autonomous vision using an unmanned aerial vehicle (UAV) [24], an example of its use is shown in Fig. 1. The use of UAVs is an emerging alternative for inspection of transmission lines because through this there is greater work speed, flexibility, and lower cost. One of the main failures in transmission lines that can be detected by aerial images is the lack of components in an insulator chain. When a support insulator has its insulation broken, it remains supporting the network, but it

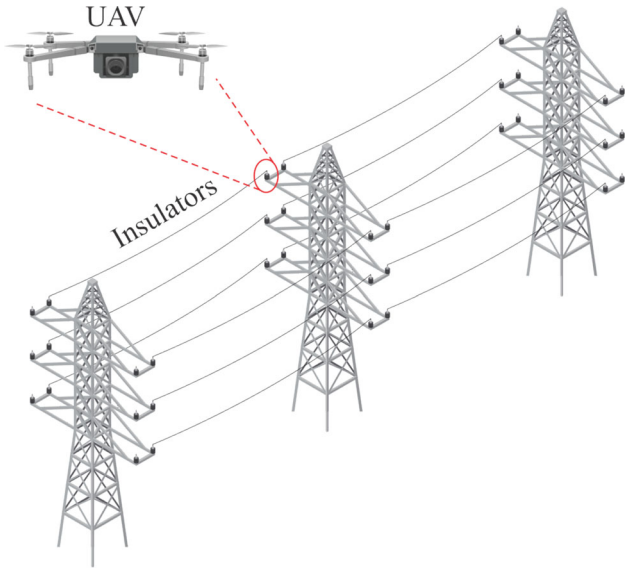

Fig. 1 Inspection of transmission lines using UAV

loses its electrical insulation capacity being more vulnerable to disruptive discharges [25].

For the inspection to be carried out accurately, it is necessary to have a tracking system that can locate the exact position and direction of the UAV in relation to the transmission line, making it possible to use commercial drones. Through a robust navigation system, the UAV can be used even when there is a communication failure with the navigation system. From an image processing algorithm, the UAV system calculates its location using a geometric approach that takes into account the transmission line structure. According to Menéndez et al. [26] one difficulty in maintaining accuracy in positioning is the increase in error in relation to the drone's movement, so the position must be calculated based on an error minimization algorithm, which can result in up to 91.44% positioning efficiency in the field. Another issue that should also be considered when planning to carry out an inspection using a UAV is the high electromagnetic field close to the grid, which can result in aircraft instability during the inspection [27].

Regarding the use of deep learning techniques to perform transmission line inspections, one difficulty is that models are generally designed to identify specific network failures [28]. According to Liang et al. [29] there are several defects that need attention so that the inspection is carried out properly. In addition to problems related to the insulator chain, there are problems in the structure, such as the presence of foreign materials near the network. Another difficulty in performing inspection using images is the large variation in the background, caused by variations in lighting or photographs taken at different angles [30]. Taking all these variables into account, the model needs to be robust to effectively identify

defects using aerial images with complex backgrounds and varying lighting.

The detection of foreign objects using deep learning in transmission line inspections is the focus of the work of Zhu et al. [31]. The evaluation proposal is based on a regression strategy with bounding boxes to locate foreign objects and their category. An effective strategy to improve the fault identification capacity is the use of public data, thus increasing the model's training capacity, allowing the analysis of patterns of different conditions to perform a pre-training and improve the generality of the application.

Advanced CNN models such as AlexNet are promising for identifying anomalies, one of the main advantages in its application is due to its characteristic of using several layers, some of which are specific for feature detection. According to Guo et al. [32] these models are superior to traditional classification methods, as they improve the image recognition rate in transmission line monitoring. Models such as the deep extreme learning machine can be applied to maximize the estimation of the transmission lines deterioration level [33].

According to Wang et al. [34], ResNeSt is a model that stands out for identifying defective insulators in transmission lines. When the insulator identification is performed from a regional proposal network (RPN), the classification achieves higher accuracy and speed values than other methods. From a fast and accurate model, it is possible to reduce the cost of maintenance of the electrical power system, since maintenance is carried out in a preventive manner, avoiding the loss of time in locating the fault during a shutdown. Models that do not consider the identification of the location of the insulator have low accuracy, as the inspection images have high resolution, and a large part of the image does not need to be analyzed. Thus, using RPN to extract features and the improved ResNeSt model to classify the failure, it is possible to achieve an accuracy rate of 98.38%, detecting up to 12.8 images per second in UAV inspections.

The RetinaNet model [35] also stands out among the algorithms based on deep learning for evaluating images of inspections carried out by UAV. From the RetinaNet model based on DenseNet as a backbone network, an improvement in precision is obtained, in addition to resulting in an algorithm that requires less computational effort, which is considerably important considering an application for online inspection [36]. The adaptive detection of domain objects is an alternative to improve the labeling problem in power transmission line inspections. As the scale of objects varies in aerial inspections because the camera is in motion, there is a difficulty in identifying specific components, for this reason, it is necessary to use a multi-scale regularization to optimize the learning of resources that help in the classification process [37].

From studies involving the use of deep learning for locating faults in transmission lines using UAVs, some factors are important to obtain a satisfactory accuracy in the classification [38]. The first point to be highlighted is the issue of insulator location, when the location is not performed properly the model tends to classify the images according to the background and not specifically in relation to the fault. Thus identifying the insulators in the aerial inspection images is important to improve the robustness of the model. Another point to be considered is that there are few faults in the power grid, so to be able to train the model it is necessary to apply segmentation techniques to expand the dataset [39].

2.1 Database

The images used in this paper are based on the work of Tao et al. [40], who developed a study on the application of CNN for fault evaluation in transmission line insulators using aerial images. The images were recorded by UAVs during inspections of the electrical power system and were provided by the *State Grid Corporation of China*. This dataset has 248 images of defective insulators and 600 images of normal insulators[1] Some images of defective insulators from this dataset are shown in Fig. 2.

As can be seen, there is a big difference in the image background in this dataset, the insulators chains are positioned in different positions, these aspects of image variation are found in the field. For this reason, it is difficult to spot the location of faulty insulators, and without interpretability a model may be incorrectly trained to recognize error patterns from the image background variations instead.

A characteristic of the background of the inspection image is its wide variation because the network is installed for long stretches. There may be variation in the background of the image with plantations, cleared areas, lakes, streams [41]. According to Lin and Liu [42] other insulating chains and transmission lines can lead to an incorrect interpretation of the algorithm. These image variations make the model need to be robust to be able to reach acceptable accuracy values in the classification.

3 Proposed method

The method proposed in this paper is divided into three steps that will be described in this section. A representative diagram of the structure of these steps is shown in Fig. 3. The first step aims to identify the position of the insulator chain, in this step the YOLO architecture is used, resulting in the coordinates where there is the insulator chain. To achieve the best performance regarding the use of the architecture, the

[1] https://github.com/InsulatorData/InsulatorDataSet.

Fig. 2 Original images of faulty insulator chains from the used dataset

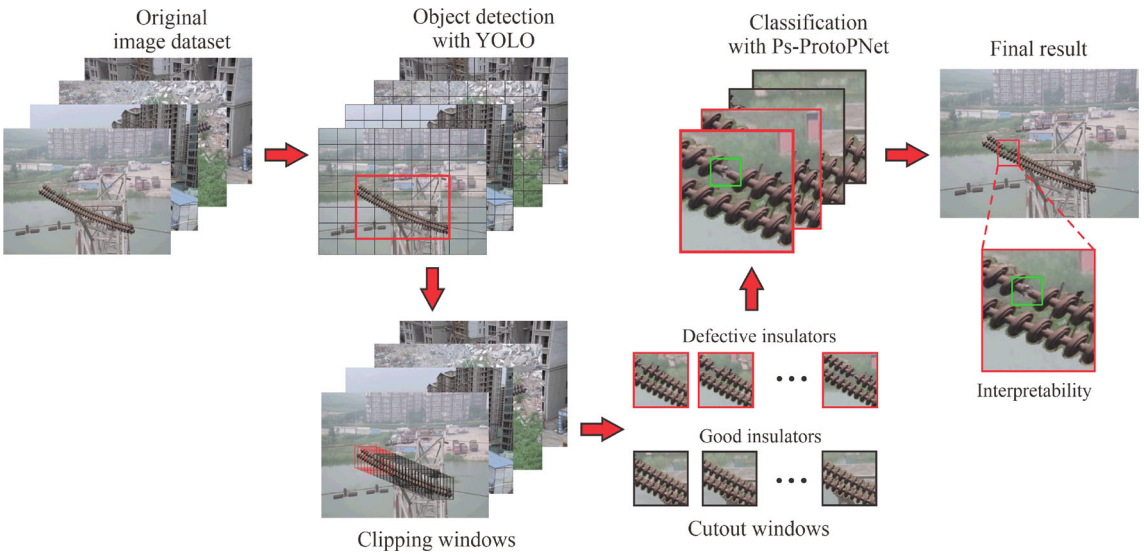

Fig. 3 Architecture of the proposed YOLO + Ps-ProtoPNet

YOLOv5, YOLOv6, YOLOv7, and YOLOv8 are compared. From the identification of the position of the insulators, windows are cropped along the chain of insulators.

The second step of the analysis is based on the failure classification, for this step the Pseudo-Prototypical Part Network model is used [20]. The result of this step are images classified with interpretability, that is, the model indicates the exact location of the failure. Based on the results of this analysis, models from other authors are compared to the proposed method.

The last step aims to present the result of fault identification on the original image, this step is carried out with the objective of making the fault analysis simpler to understand for the operator. The positioning of the image superimposed on the original is given by the coordinates previously identified through the location of the objects and image clipping.

3.1 Object detection

There are several models that can be used for object detection, highlighting the YOLO architecture, which has the advantage of its speed, considering that it simultaneously predicts the object class and location of the object in the input image. To use the YOLO model initially it is necessary to annotate the bounding boxes of the insulators chain in the images. In this paper, the LabelImg software was used to define these annotations which are the purpose of detection.

The YOLO network predicts an objectivity score for each bounding box based on logistic regression. If a previous bounding box is not assigned to a fundamental truth object, there is no loss of coordinates or class predictions [43]. Initially, the YOLOv3 proposed by Farhadi and Redmon [43], being a superior model than RetinaNet [44], feature pyramid networks (FPN) [45], R-CNN, Fast R-CNN [46], and the previous YOLO methods [47]. Nowadays, there is a wide number of versions that are being explored such as YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, YOLOv6n, YOLOv6t, YOLOv6s, YOLOv6m, YOLOv6l, YOLOv7, and YOLOv7x [48]. Ultralytics recently released the eighth version of the YOLO architecture [49], with the YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x available for object detection.

The difference among the variations of YOLOv6, YOLOv7, and YOLOv8 are also relied on Depth Multiple and Width Multiple, such as the YOLOv5. In YOLOv6 proposed by Li et al. [50], the model requires two runs to be evaluated, since the training is separated from the model evaluation, besides having less flexibility regarding the training configuration for analysis. The YOLOv8 [49] and the YOLOv7 proposed by Wang et al. [51] are similar to the YOLOv5 previously introduced by Ultralytics [52].

In this paper, after performing the annotations with the LabelImg, the training for object detection was performed. The YOLOv5 (n, s, m, l, and x), YOLOv6 (n, t, s, m, and l), YOLOv7 (std and x), and YOLOv8 (n, s, m, l, and x) are evaluated for this task. From the object detection result, the position of the insulator chain was identified in all images.

From the definition of the position of the insulator chain, the images are cropped. The windows cutout is performed based on the inclination of the insulator chain. The slope is calculated by the reduced straight-line equation ($\beta = \rho\gamma + \eta$), where γ and β are respectively the independent and dependent variables; ρ is the slope, and η is the linear coefficient. Considering that the beginning of the insulators chain $A(\gamma_1, \beta_1)$ and its end $B(\gamma_2, \beta_2)$ are known, the slope of the cropped made, is given by:

$$\rho = \frac{\beta_2 - \beta_1}{\gamma_2 - \gamma_1}. \tag{1}$$

Based on the slope, 25 cutouts are made in windows, which correspond to the separation of approximately one insulator in the chain, then 1073 images of defective insulators and 2791 images of normal insulators are obtained. After argumentation, the training has 32,677 images of defective insulators and 85,649 images of normal insulators.[2] All the images were resized to the dimensions 224×224 as required by the base models.

3.2 Pseudo-prototypical part network

The existing models for the detection of defective insulators lack the interpretability of the reasoning process of their predictions. In this paper, the Ps-ProtoPNet is presented for the detection of defective insulators [20]. The Ps-ProtoPNet is an interpretable deep learning model that is used over the dataset of the insulators in different conditions. Ps-ProtoPNet is related to ProtoPNet [53], Gen-ProtoPNet [54], and NP-Proto-PNet [55], however, there are differences in the model structure. To use Ps-ProtoPNet in this paper, it was necessary to include an additional convolution layer to ensure interpretability and adapting the model to the format of the used images, as will be explained in detail in this section.

A prototype structure represents a patch of an image. To perform classification, the ProtoPNet compares the different parts of the learned prototypes of images from all classes to the test image. The classification is based on the weighted combination of similarity scores of the model [53]. To calculate the scores between learned prototypes (square spatial dimensions of 1×1) and parts of the test image, the ProtoPNet and NP-ProtoPNet use $L2$ distance function, while Gen-ProtoPNet uses a generalized version of $L2$.

The Ps-ProtoPNet does not do the last-layer convex optimization to maintain the impact of the negative reasoning

[2] https://github.com/SFStefenon/InsulatorsTransmission.

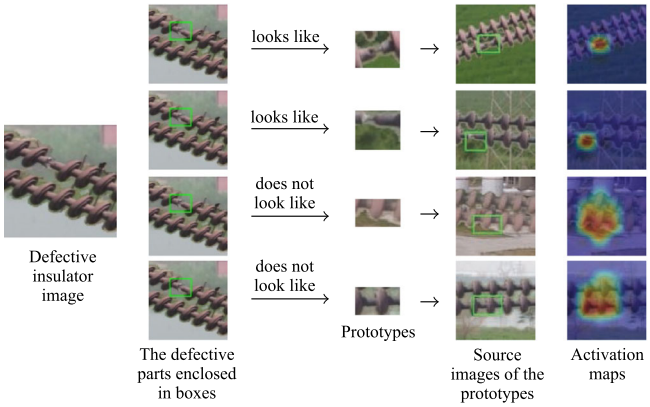

Fig. 4 From a defective insulator, Ps-ProtoPNet identifies the parts of the image that look like the prototypical parts, and the parts that do not look like any part of the images of normal insulators

process on image ratings. The non-optimization of the last layer helps to preserve the negative connection of the logits with incorrect classes. In the applied Ps-ProtoPNet a generalized version of the $L2$ distance function is used, and the prototypes can have any kind of spatial dimension. Comparing parts of the image with the learned prototypes is an integral part of how the method reasons about new images [20].

A new image is compared with the learned prototypes of all classes and is sorted to the class whose prototypes are most similar to the parts of the image. There are two classes of images: Defective and Normal. A defective insulator is distinguished from the normal insulator based on the greater similarity of parts of the image with the prototype of the Defective class. Figure 4 presents how this evaluation is performed.

Several perspectives have emerged to explain CNN, including *posthoc* interpretability analysis. A neural network with analysis *posthoc* is interpreted following classifications made by the model. Activation maximization [56–58], deconvolution [59], and saliency visualization [60–63] are examples of *posthoc* approach. However, these techniques do not solve the reasoning process with transparency.

Another approach to clarifying the model's reasoning process is attention-based interpretability, which includes class activation maps. The model is intended to point out the parts of a test image that are its focus of attention [64–66]. These approaches do not point to prototypes that are similar to parts of the tested image. Li et al. [67] propose a model that uses full-size image prototypes to find similarity scores. An important improvement was made by Chen et al. [53] with the development of the ProtoPNet model.

3.2.1 Ps-ProtoPNet novelty

ProtoPNet classifies the image based on a weighted combination of the similarity scores. A fixed number of prototypes are selected for each class. The ProtoPNet calculates the Euclidean distance (Eq. 2) of each prototype from each latent patch of the test image that has spatial dimensions equal to 1×1. These distances are inverted, and a maximum of the result is called the prototype similarity score. For each image, only one similarity index for each prototype is obtained. In the dense layer of the model, the similarity scores are multiplied with the weights to calculate the logits. In the training process, the model performs convex optimization of the last layer to make certain weights zero [53].

$$d_{st}^2 = (x_s - y_t)(x_s - y_t)', \tag{2}$$

where x_s and y_t are the row vectors.

With larger prototypes spatial dimensions than 1×1, Ps-ProtoPNet uses the negative weights to make negative connection between similarity scores and incorrect classes. Besides the positive reasoning process, Ps-ProtoPNet does not do last-layer convex optimization to maintain the impact of the negative reasoning process on image classification, while the ProtoPNet emphasizes the positive reasoning process. Not optimizing the last layer considerably reduces training time.

3.2.2 Ps-ProtoPNet architecture

The Ps-ProtoPNet is constructed over baseline models, which in this paper are based on ImageNet [68]. For a definition of the best structure the VGG-13 [69], VGG-16 [70], VGG-19 [71], ResNet-34 [72], ResNet-50 [73], and ResNet-152 [74] baselines were evaluated. Figure 5 presents the Ps-ProtoPNet architecture, the model has convolution layers which are followed by two additional layers 1×1 (the convolution layers

are ℓ), and the convolution layers are followed by a generalized [75] convolution layer p_p of prototypical parts and a dense layer d_l with matrix of weight m_w. In this structure the dense layer does not have bias [76]. The ReLU activation function is used for the additional convolution layers.

For an input image x, $\ell(x)$ is the output of the convolutional layers ℓ, and the shape of $\ell(x)$ is $512 \times 7 \times 7$. Considering that $P^k = \{p_l^k\}_{l=1}^{m'}$ is the set of prototypes of a class k and $P = \{P^k\}_{k=1}^{n_c}$ is the set of prototypes of all classes, then m' is the number of prototypes for each class and n_c is the total number of classes.

In this paper, $m' = 10$ and $n_c = 2$, and the hyperparameter $m' = 10$ are chosen randomly. The shape of the prototypes is $512 \times h \times w$, where $1 \times 1 < h \times w < 7 \times 7$, that is, h and w are not simultaneously equal to 1 or 7. Therefore, every prototype could be considered as a representation of some prototypical part of some insulator image.

3.2.3 Ps-ProtoPNet training

The generalized version d of the distance function $L2$ (Euclidean distance) is used to find distance between prototypes and latent patched of the input images. For an image x, the shape of $\ell(x)$ is $512 \times 7 \times 7$, where the depth of $\ell(x)$ is 512 and the spatial dimensions of $\ell(x)$ are 7×7. Given a prototype p of the shape $512 \times h \times w$, where $1 \leq h, w \leq 7$, and h and w are not simultaneously equal to 1 or 7. The output z of the convolutional layers ℓ has $(8-h)(8-w)$ patches of dimensions $h \times w$. Therefore, the square of the distance $d(\mathcal{Z}_{ij}, p)$ between the prototype p and (i, j) patch $\mathcal{Z}_{ij}$ of z is given by:

$$d^2(\mathcal{Z}_{ij}, p) = \sum_{l=1}^{h} \sum_{m=1}^{w} \sum_{k=1}^{512} \|z_{(i+l-1)(j+m-1)k} - p_{lmk}\|_2^2. \quad (3)$$

Considering the prototypes of spatial dimension 1×1, where $h = w = 1$, the square of the distance is:

$$d^2(\mathcal{Z}_{ij}, p) = \sum_{k=1}^{512} \|z_{ijk} - p_{11k}\|_2^2, \quad (4)$$

which is the square of the Euclidean distance between p and a patch of z, where $p_{11k} \simeq p_k$. The prototypical unit p_p is calculated by:

$$p_p(z) = \max_{1 \leq i \leq 8-h,\ 1 \leq j \leq 8-w} \log\left(\frac{d^2(\mathcal{Z}_{ij}, p) + 1}{d^2(\mathcal{Z}_{ij}, p) + \epsilon}\right), \quad (5)$$

that is,

$$p_p(z) = \max_{\mathcal{Z} \in \text{patches}(z)} \log\left(\frac{d^2(\mathcal{Z}, p) + 1}{d^2(\mathcal{Z}, p) + \epsilon}\right). \quad (6)$$

where p is more similar to x if the inverse of the distance between the patch is smaller.

3.2.4 Optimization of layers

To optimize the layers before the dense layer, considering that $X = \{x_1 \ldots x_n\}$ and $Y = \{y_1 \ldots y_n\}$ are sets of images and corresponding labels respectively, given $D = \{(x_i, y_i) : x_i \in X, y_i \in Y\}$, the objective function is:

$$\min_{P, \ell_{\text{conv}}} \frac{1}{n} \sum_{i=1}^{n} \text{CrosEnt}(h \circ p_p \circ \ell(x_i), y_i) + \lambda_1 \text{ClstCst} + \lambda_2 \text{SepCst}, \quad (7)$$

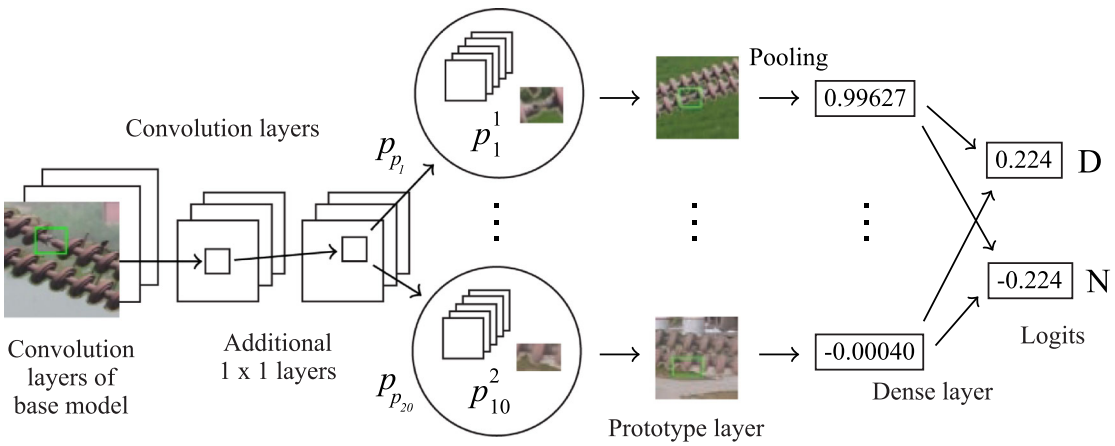

Fig. 5 Ps-ProtoPNet architecture

where λ_1 and λ_2 are hyperparameters. The cluster cost (ClstCst) and the separation cost (SepCst) are:

$$\text{ClstCst} = \frac{1}{n}\sum_{i=1}^{n} \min_{j:p_j \in P_{y_i}} \min_{\mathcal{Z} \in \text{patches}(\ell(x_i))} d^2(\mathcal{Z}, p_j), \quad (8)$$

$$\text{SepCst} = -\frac{1}{n}\sum_{i=1}^{n} \min_{j:p_j \notin P_{y_i}} \min_{\mathcal{Z} \in \text{patches}(\ell(x_i))} d^2(\mathcal{Z}, p_j). \quad (9)$$

The decrease of the ClstCst leads to the clustering of prototypes around their respective classes. However, decreasing the SepCst keeps prototypes away from their incorrect classes [53], the drop in cross entropy leads to better rankings. Whereas m_w is the weight matrix for the last layer, $m_w^{(i,j)}$ is the weight assigned to the connection between the prototype similarity score jth and logit of ith class, for a class k is included $m_w^{(i,j)} = 1$ for all j with $p_j^i \in P^i$, and for all $p_j^k \notin P^i$ with $k \neq i$, $m_w^{(k,j)}$ equal to -1.

Since the distance function is not negative, the optimization of all layers except the last layer with the stochastic gradient descent (SGD) optimizer (Eq. 10) can help the Ps-ProtoPNet learn the latent space. The SGD update the parameters to minimize the loss function by performing steps in each iteration toward the negative loss gradient,

$$\theta_{i+1} = \theta_i - \alpha \nabla E(\theta_i) \quad (10)$$

where α is the learning rate, θ is the parameter vector to be minimized, and $E(\theta)$ is the loss function [77].

The Ps-ProtoPNet pushes the prototypes in the patches from the output $\ell(x)$ of an image x that has the smallest distances from the prototypes, and then, the Ps-ProtoPNet perform the following update:

$$p_j^k \longleftarrow \arg \min_{\{\mathcal{Z}:\mathcal{Z} \in \text{patches}(\ell(x_i)) \; \forall i \text{ s.t. } y_i = k\}} d(\mathcal{Z}, p_j^k).$$

The prototype layer gets updated prototypical parts that are closer to their respective classes. The image patch that is most similar to p is used for the visualization. The prototype activation value must be at least 94th percentile of the activation values of p_p [53].

3.2.5 Final presentation

The final presentation of the result is carried out to improve the operator's understanding. In this step, the results of the interpretability of fault classification are presented on the original images. To facilitate the understanding of the image, the classification result is superimposed exactly on the place where it was cutout, the coordinates for the superposition are defined according to the position of records of the images previously taken.

4 Results and discussion

In this section, the results of applying the proposed method will be presented and discussed. Initially, the results of the training process will be presented, followed by the testing results. At the end of this section a comparison with other authors for this assessment will be discussed. For the evaluation of the model the *accuracy*, *precision*, *recall*, and *F-measure* were used.

For the insulators' classification task, the experiments were performed in the *Lambda Labs Cluster* of the University of Regina, Canada. For these evaluations, the requirements were allocated: a graphic processing unit (GPU) NVIDIA Quadro RTX 5000 and 256 GB of random access memory (RAM). The algorithm presented in this paper was developed in Python. In this section, the best comparative performance results are highlighted in bold.

4.1 Identifying the insulator chain

Since the YOLO model has several variations, the versions of YOLOv5, YOLOv6, YOLOv7, and YOLOv8 are compared in Table 2. The testing time presented in this table represents the time required to test an image after the model has been trained. Since the task is detection, the mean Average Precision (mAP) is considered to determine the version of the model that has better performance. For the object detection task, the Google Colab platform is used with a GPU Tesla K80. The models that had the best mAP@[0.5], had results around 0.995, there was a greater difference in performance when analyzing mAP@[0.5:0.95]. In this case, YOLOv8m had the best performance results for object detection, with an mAP@[0.5:0.95] of 0.91251, therefore, in this paper, this structure is defined as the standard for the object detection task.

The time required to perform the training is considerably longer than the time to test the model, this makes the application of CNNs promising for industrial problems since the model can be previously trained using high computational power and the application (in the testing phase) can be accomplished using low computational resources. For the fastest model to complete training (YOLOv6n), the time needed to train the model was 6.03×10^5 longer than the time needed to test (post-process per image). For all evaluated models, the testing time for object detection was less than 3 milliseconds per image.

From the use of YOLOv8m the chain (string) of insulators were identified, and from these identifications, the coordinates for the image clipping were defined. The YOLOv8m creates a bounding box with the position in which the insulator chain was found (see Fig. 6), besides the confidence in its identification, which can be used as a threshold value for this result to be considered.

Table 2 Evaluation of the YOLO version for detection of insulator chains

Model	F1-score	mAP		Training time (s)	Testing time (ms)
		[0.5]	[0.5:0.95]		
YOLOv5n	0.99945	0.99500	0.77151	1472.86	**0.93**
YOLOv5s	0.99907	0.99500	0.78086	1310.11	0.97
YOLOv5m	0.99944	0.99500	0.83422	2299.05	1.25
YOLOv5l	0.99920	0.99500	0.87596	4874.22	1.55
YOLOv5x	0.99945	0.99500	0.84479	6989.85	2.15
YOLOv6n	0.98311	0.97592	0.83302	**1212.29**	2.01
YOLOv6t	0.98312	0.97404	0.82781	1257.91	2.11
YOLOv6s	**0.99999**	0.99514	0.82289	1747.19	2.01
YOLOv6m	0.92033	0.95645	0.65296	2685.73	2.18
YOLOv6l	0.99983	**0.99516**	0.84969	3095.16	2.20
YOLOv7 (std)	0.98626	0.99437	0.75218	4126.18	1.23
YOLOv7x	0.99319	0.99478	0.83837	5196.21	1.62
YOLOv8n	0.99960	0.99500	0.87728	1810.84	1.63
YOLOv8s	0.99958	0.99500	0.89678	1918.96	2.12
YOLOv8m	0.99939	0.99500	**0.91251**	2371.29	1.51
YOLOv8l	0.99954	0.99500	0.90683	2617.43	1.57
YOLOv8x	0.99954	0.99500	0.90683	3079.68	2.74

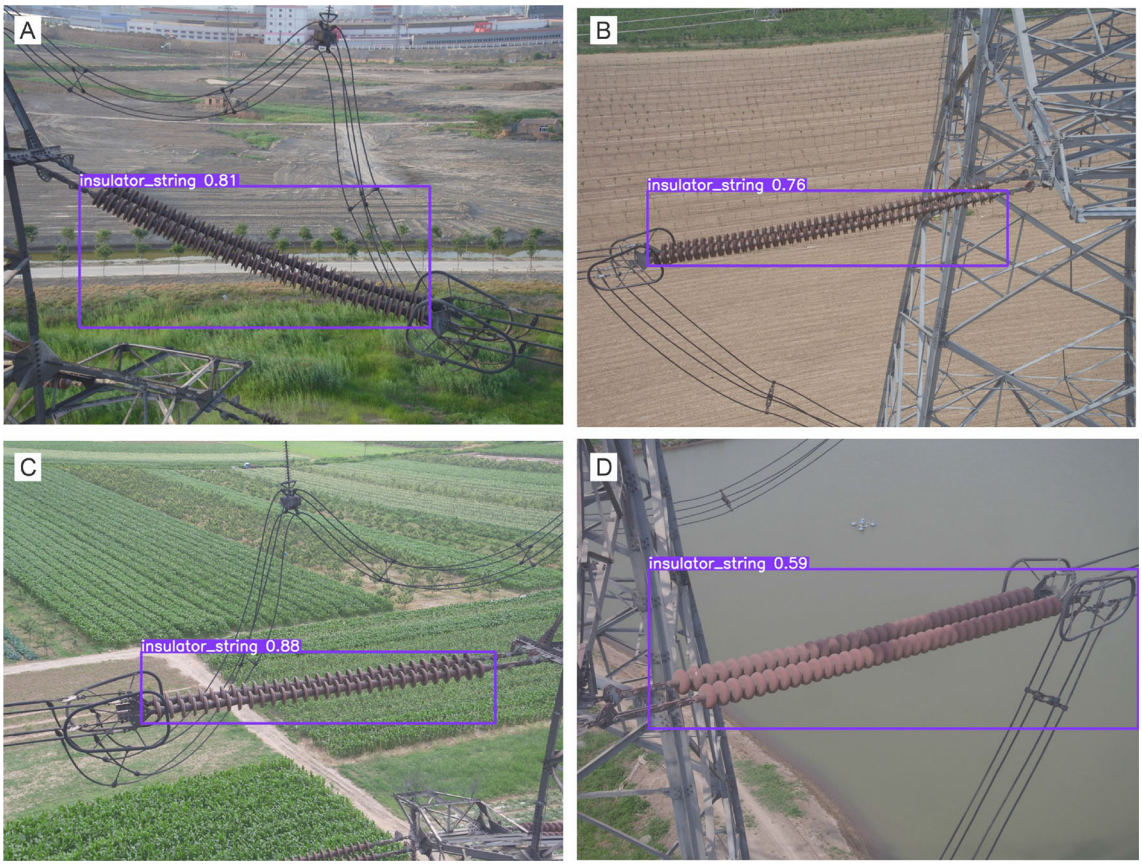

Fig. 6 Insulators chain detection using YOLOv8m model

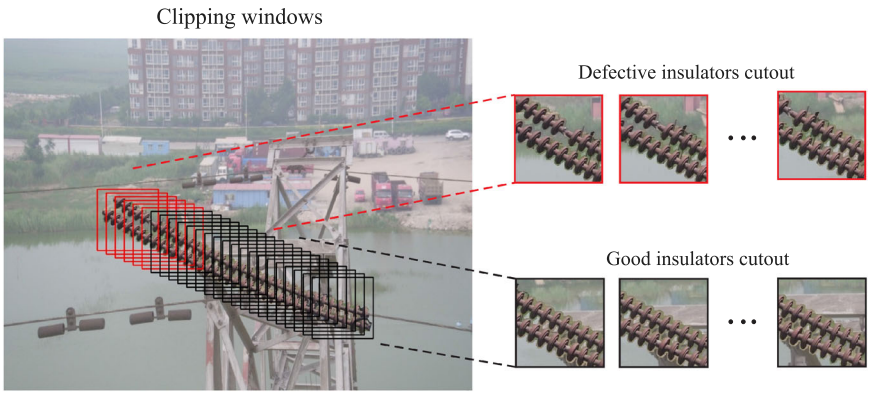

Fig. 7 Clipping windows made in the insulator chain

It has been observed that the model is able to detect the chain of insulators in different positions, from different angles and approximations. In some results, the edges of the insulator chain extrapolate the limits of the rectangle defined as the analysis margin, this is not a problem, since in the considered rectangle there are only two extreme edges for every 25 cutouts made.

There is variety of backgrounds in the considered database, besides the position in which the photograph was registered is different. In Fig. 6A there is a large variation in the background of the image possibly because there is a region below the transmission line that is not planted with crops (only presence of vegetation). In Figs. 6B and 6C, there is a cultivated area of different conditions, and in Fig. 6 there is a river/lake in the background.

In addition to these four images presented, there are images in locations with higher population density, with a larger number of buildings. These image differences make the problem complex, and for this reason, the use of combined techniques (YOLOv8m + Ps-ProtoPNet) as proposed in this paper is a promising alternative to classify the problem.

The authors Chen et al. [78, 79] use the YOLOv3 structure to perform the fault classification, in this paper YOLOv8m was applied specifically to locate the insulator chain, the classification is performed by the Ps-ProtoPNet model. Equation (1) defines the place where the image clipping starts and the insulator chain inclination, which defines the direction of the cutouts windows.

The clipping windows made in relation to one of the images are shown in Fig. 7. The images representing faulty insulators are highlighted with a red outline, in this case, they correspond to only 6 cropping windows. As there are more insulators in good condition than failing insulators there is an imbalance between the classes, as mentioned in the Sect. 2.1. The procedure was performed for all images in the used database.

From the clipping windows, smaller images cutouts are obtained that correspond to the parts of the grid that have faults and others that are in good condition. These images are used to carry out the training and classification of the condition that must be located during the inspection of the electrical power system.

4.2 Model classification

In the network training process, the interpretability of the fault occurs using the Ps-ProtoPNet, in which the model specifically defines the critical points to be analyzed. Interpretability is one of the most important factors in an approach using deep learning. When there is no fault interpretability, the model can result in high accuracy by performing the classification through the background of the image, which makes the model unsuitable for practical applications. As can be observed the test image in the first column of Fig. 8 belongs to the defective class.

In the second column, the image has some green rectangles, these patches provide the highest similarity score to be compared to the corresponding prototypes in the third column. In the third column, the prototypes are taken from the corresponding source images in the fourth column. The rectangles in the source image indicate the patches from where the corresponding prototypes are taken.

The fifth column has prototypes similarity scores, and the sixth column has the weights. The seventh column is obtained by multiplying the similarity scores with the corresponding class connections. The logit for the first class result of 2.244 is the sum of the entries in the seventh column. Equivalently, the logit for the second class is obtained by multiplying the second row of the weight matrix with the respective similarity score matrix.

Ps-ProtoPNet calculates the similarity scores between the input image and the prototypical parts $p_1^1 - p_{10}^1$ and $p_1^2 - p_{10}^2$. The similarity score of the prototype p_1^1 is 0.9922, which is greater than the similarity scores of p_{10}^2 been 0.00040. The Ps-ProtoPNet model tracks the spatial relationship of the convolutional output and the prototypical parts and upsamples

the parts to the input image size to point the patches on the source images that correspond to the prototypes.

4.3 Models assessment

The confusion matrices, presented in Fig. 9, allows visualization of an algorithm's performance. In this assessment, the VGG-13, VGG-16, VGG-19, ResNet-34, ResNet-50, and ResNet-152 are evaluated as the base models to define the best structure of the network. Each row of the confusion matrix represents instances in an actual class, while each column represents instances in a predicted class. Based on true positives, true negatives, false positives, and false negatives of the defective class, the performance results are obtained.

The results based on the confusion matrices are presented in Table 3. As can be seen the Ps-ProtoPNet has

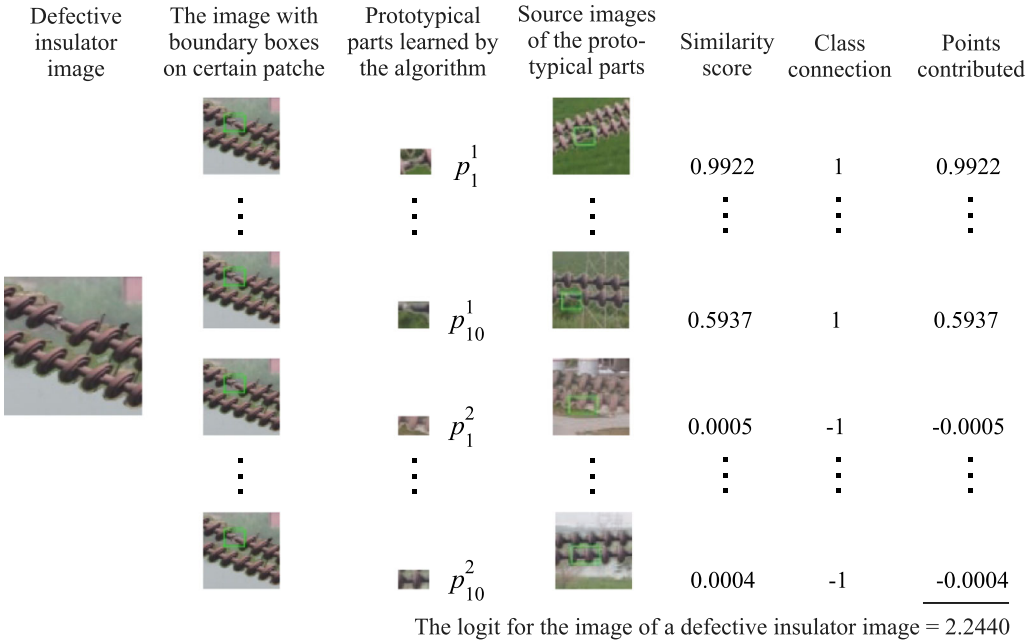

Fig. 8 The explanation of the classification process of the model

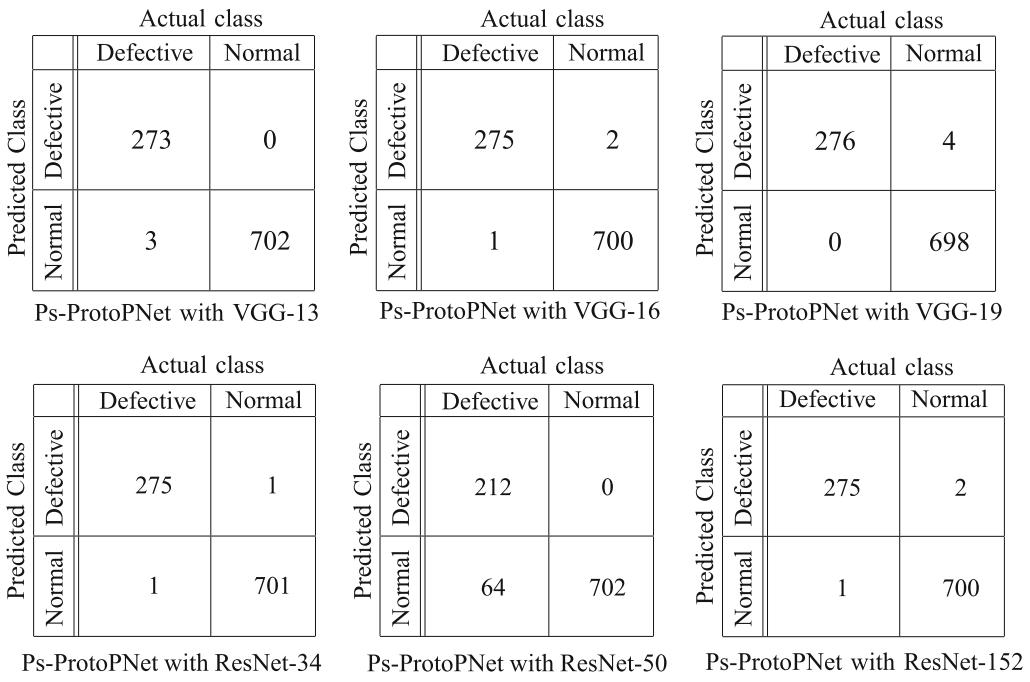

Fig. 9 Confusion matrices of Ps-ProtoPNet

Table 3 Comparison of the performances of the models with different base model

Model/method	Accuracy	Precision	Recall	F1-score
VGG-13	99.69	**0.9999**	0.9891	0.9945
VGG-16	99.69	0.9927	0.9963	0.9945
VGG-19	99.59	0.9857	**0.9999**	0.9928
ResNet-34	**99.79**	0.9964	0.9964	**0.9964**
ResNet-50	99.48	0.9963	0.9855	0.9908
ResNet-152	99.69	0.9927	0.9963	0.9945

high accuracy, precision, recall, and F1-score, in addition, it has interpretability, as mentioned above, thus being a robust model for the analysis in question. The best accuracy found for the classification of a missing insulator in transmission lines was using the Ps-ProtoPNet model with the ResNet34 as a base model, reaching the value of 99.79 % accuracy and an F1-score of 0.9964. This structure is defined as standard and is used for further comparisons.

4.4 Comparison against other methods

The paper of Tao et al. [40] based on a deep convolutional neural network reached the precision of 0.91 for the defect detection. Based on the same dataset, the Ps-ProtoPNet using ResNet-34 as the base model presented in this paper reached the value of 0.9964 for the precision and recall based on the same dataset. This shows that the combination of high-performance models generates results that are superior to conventional computer vision approaches based on deep learning.

Using the YOLOv5, Feng et al. [80] reached the accuracy of 86.8%, using the Ps-ProtoPNet the accuracy was 99.79%. Using the YOLO v3, Chen et al. [78] reached the accuracy of 95.45%. Using the YOLOv3 based on a DarkNet backbone [79] reached the precision of 0.837 and with their YOLOv3 modified version the precision was 0.926. Longjun et al. [81] presented a research where other types of faults are evaluated as vibration damper, missing vibration damper, broken strand, and foreign body. The classification model results in less accuracy and recall, resulting in a F1-score of 0.843, being lower than other models based on convolutional neural network. Using a model based on advanced techniques image classification Li et al. [82] achieved an precision of 0.919 and recall of 0.957 for defect recognition.

A visual comparison of the best results of each author is presented in Table 4.

A promising work for fault detection in transmission insulators with an accuracy greater than 97% was presented by Sampedro et al. [84]. The analysis show that the combination of high-performance models result in a more robust algorithm, considering that it was proposed to use the Up-Net network within a generative adversarial network framework. Wang et al. [34] had one of the best results for the problem analyzed, with a accuracy rate of 98.38%, found using the advanced models ResNeSt and region proposal network. Many papers have been carried out using deep learning models for the classification of failures in transmission insulators, the best results were found combining techniques. However, the Pseudo-Prototypical model was superior to the compared models.

Some models were not compared in this analysis because they did not present the F1-score or the variations required in the model. The work of Sampedro et al. [84] do not present comparative variations in the model, as they focus on analyzing the variations in relation to the type of failure. Chen et al. [78], Vigneshwaran et al. [85], and Deng [86] only use accuracy to compare model variations.

Table 4 Comparison with other authors

Author	Model/method	Backbone/base	F1-score
Tao et al. [40]	ILN	VGG-16	0.934
	ILN	ResNet-50	0.926
	DNN	ResNet-101	0.933
Chen et al. [79]	RetinaNet	ResNet-101	0.845
	YOLOv3	DarkNet-53	0.836
	M-YOLOv3	Dense-DarkNet	0.938
Li et al. [82]	Faster R-CNN+FPN	ResNet-101	0.917
		ResNeXt-101	0.929
		ResNeXt-101(OHEM)	0.944
Wen et al. [83]	ERCN	ResNet50	0.872
	CME-CNN	ResNet50	0.906
	CME-CNN	MoblieNetV2	0.889
Our method	YOLOv8m + Ps-ProtoPNet	ResNet-34	**0.996**

Fig. 10 Final result of the application of YOLO8m and Ps-ProtoPNet

The final results of this comparative analysis made it clear that the proposed method based on the combination between YOLO8m and Ps-ProtoPNet is a superior model for analyzing the problem in question, observing that the F1-score result was higher than all the analyzes presented by other authors. The final result of the identification of failures in inspection images is shown in Fig. 10.

Liu et al. [87] does not present variations in the base model for a comparison of their proposed algorithm, the variations YOLO-tiny, YOLOv2, YOLOv3, and MTI-YOLO are applied with from the DarkNet framework for all comparisons. The F1-score calculations were performed for the papers that presented the precision and recall.

5 Conclusion

The use of artificial intelligence techniques shows promise for identifying faults in the electrical power system. From a model that has a high accuracy it is possible to automate the analysis of images captured during field inspections. Using unmanned vehicles provides greater speed in recording the conditions of the electrical network, thus obtaining an efficient approach for analyzing transmission networks.

Interpretability in a deep learning model is considerably important for specific fault classifications, such as the identification of missing insulators in transmission lines. When there is no interpretability, the network can have high accuracy based on image background characteristics, which results in inefficiency for field analysis, since variations in environmental conditions result in changes in image background and luminosity.

The accuracy results of the Ps-ProtoPNet were suitable for a real application from images recorded by UAVs, considering the accuracy result of 99.79%, obtained from the Ps-ProtoPNet using the ResNet-34 as the base model. The non-optimization of the last layer and the use of prototypes with rectangular spatial dimensions and square spatial dimensions greater than 1×1 helped the model to improve its performance over NP-ProtoPNet, Gen-ProtoPNet and ProtoPNet.

The future challenge for this approach is the online evaluation of insulator conditions. Through an embedded system, it would be possible to have a greater speed in inspection with teams to check the conditions of the network after defining the critical places where there are failures. Thus, it would be possible to identify the exact location and reason for the

failure when a shutdown occurs and then have greater agility to mobilize a field team to correct the failure.

Acknowledgements We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number DDG-2020-00034. Cette recherche a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), numéro de référence DDG-2020-00034.

References

1. Aboshady, F.M.: Modified distance protection for transmission line with hexagonal phase-shifting transformer. Int. J. Electr. Power Energy Syst. **134**, 107379 (2022). https://doi.org/10.1016/j.ijepes.2021.107379
2. Wang, B., Dong, M., Ren, M., Wu, Z., Guo, C., Zhuang, T., Pischler, O., Xie, J.: Automatic fault diagnosis of infrared insulator images based on image instance segmentation and temperature analysis. IEEE Trans. Instrum. Meas. **69**(8), 5345–5355 (2020). https://doi.org/10.1109/TIM.2020.2965635
3. Stefenon, S.F., Ribeiro, M.H.D.M., Nied, A., Mariani, V.C., Coelho, L.D.S., Leithardt, V.R.Q., Silva, L.A., Seman, L.O.: Hybrid wavelet stacking ensemble model for insulators contamination forecasting. IEEE Access **9**, 66387–66397 (2021). https://doi.org/10.1109/ACCESS.2021.3076410
4. Wang, H., Cheng, L., Liao, R., Zhang, S., Yang, L.: Nonlinear ultrasonic nondestructive detection and modelling of kissing defects in high voltage composite insulators. IEEE Trans. Dielectr. Electr. Insul. **27**(3), 924–931 (2020). https://doi.org/10.1109/TDEI.2019.008523
5. Stefenon, S.F., Bruns, R., Sartori, A., Meyer, L.H., Ovejero, R.G., Leithardt, V.R.Q.: Analysis of the ultrasonic signal in polymeric contaminated insulators through ensemble learning methods. IEEE Access **10**, 33980–33991 (2022). https://doi.org/10.1109/ACCESS.2022.3161506
6. Jiang, H., Qiu, X., Chen, J., Liu, X., Miao, X., Zhuang, S.: Insulator fault detection in aerial images based on ensemble learning with multi-level perception. IEEE Access **7**, 61797–61810 (2019). https://doi.org/10.1109/ACCESS.2019.2915985
7. Alhassan, A.B., Zhang, X., Shen, H., Xu, H.: Power transmission line inspection robots: a review, trends and challenges for future research. Int. J. Electr. Power Energy Syst. **118**, 105862 (2020). https://doi.org/10.1016/j.ijepes.2020.105862
8. Samadi, M., Seifi, H., Haghifam, M.-R.: Midterm system level maintenance scheduling of transmission equipment using inspection based model. Int. J. Electr. Power Energy Syst. **110**, 467–476 (2019). https://doi.org/10.1016/j.ijepes.2019.03.050
9. Stefenon, S.F., Seman, L.O., Sopelsa Neto, N.F., Meyer, L.H., Nied, A., Yow, K.-C.: Echo state network applied for classification of medium voltage insulators. Int. J. Electr. Power Energy Syst. **134**, 107336 (2022). https://doi.org/10.1016/j.ijepes.2021.107336
10. Manninen, H., Ramlal, C.J., 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. **128**, 106726 (2021). https://doi.org/10.1016/j.ijepes.2020.106726
11. Fahim, S.R., Sarker, S.K., Muyeen, S.M., Das, S.K., Kamwa, I.: A deep learning based intelligent approach in detection and classification of transmission line faults. Int. J. Electrical Power Energy Syst. **133**, 107102 (2021). https://doi.org/10.1016/j.ijepes.2021.107102
12. Wang, Y., Yan, J., Yang, Z., Zhao, Y., Liu, T.: Optimizing GIS partial discharge pattern recognition in the ubiquitous power internet of things context: a MixNet deep learning model. Int. J. Electr. Power Energy Syst. **125**, 106484 (2021). https://doi.org/10.1016/j.ijepes.2020.106484
13. Branco, N.W., Cavalca, M.S.M., Stefenon, S.F., Leithardt, V.R.Q.: Wavelet LSTM for fault forecasting in electrical power grids. Sensors **22**(21), 8323 (2022). https://doi.org/10.3390/s22218323
14. Stefenon, S.F., Singh, G., Yow, K.-C., Cimatti, A.: Semi-ProtoPNet deep neural network for the classification of defective power grid distribution structures. Sensors **22**(13), 4859 (2022). https://doi.org/10.3390/s22134859
15. Teimourzadeh, H., Moradzadeh, A., Shoaran, M., Mohammadi-Ivatloo, B., Razzaghi, R.: High impedance single-phase faults diagnosis in transmission lines via deep reinforcement learning of transfer functions. IEEE Access **9**, 15796–15809 (2021). https://doi.org/10.1109/ACCESS.2021.3051411
16. Luo, G., Hei, J., Yao, C., He, J., Li, M.: An end-to-end transient recognition method for VSC-HVDC based on deep belief network. J. Mod. Power Syst. Clean Energy **8**(6), 1070–1079 (2020). https://doi.org/10.35833/MPCE.2020.000190
17. Manninen, H., Ramlal, C.J., 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. **128**, 106726 (2021). https://doi.org/10.1016/j.ijepes.2020.106726
18. Wang, S., Dehghanian, P.: On the use of artificial intelligence for high impedance fault detection and electrical safety. IEEE Trans. Ind. Appl. **56**(6), 7208–7216 (2020). https://doi.org/10.1109/TIA.2020.3017698
19. Dabbaghjamanesh, M., Moeini, A., Hatziargyriou, N.D., Zhang, J.: Deep learning-based real-time switching of hybrid AC/DC transmission networks. IEEE Trans. Smart Grid **12**(3), 2331–2342 (2021). https://doi.org/10.1109/TSG.2020.3041853
20. Singh, G., Yow, K.-C.: Object or background: an interpretable deep learning model for Covid-19 detection from CT-scan images. Diagnostics **11**(9), 1732 (2021). https://doi.org/10.3390/diagnostics11091732
21. Stefenon, S.F., Corso, M.P., Nied, A., Perez, F.L., Yow, K.-C., Gonzalez, G.V., Leithardt, V.R.Q.: Classification of insulators using neural network based on computer vision. IET Gener. Transm. Distrib. **16**(6), 1096–1107 (2021). https://doi.org/10.1049/gtd2.12353
22. Rocha, P.H.V., Costa, E.G., Serres, A.R., Xavier, G.V.R., Peixoto, J.E.B., Lins, R.L.: Inspection in overhead insulators through the analysis of the irradiated RF spectrum. Int. J. Electr. Power Energy Syst. **113**, 355–361 (2019). https://doi.org/10.1016/j.ijepes.2019.05.060
23. Nguyen, V.N., 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. **99**, 107–120 (2018). https://doi.org/10.1016/j.ijepes.2017.12.016
24. Hui, X., Bian, J., Zhao, X., Tan, M.: Vision-based autonomous navigation approach for unmanned aerial vehicle transmission-line inspection. Int. J. Adv. Rob. Syst. **15**(1), 1–15 (2018). https://doi.org/10.1177/1729881417752821
25. Zhao, Z., Qi, H., Qi, Y., Zhang, K., Zhai, Y., Zhao, W.: Detection method based on automatic visual shape clustering for pin-missing defect in transmission lines. IEEE Trans. Instrum. Meas. **69**(9), 6080–6091 (2020). https://doi.org/10.1109/TIM.2020.2969057
26. Menéndez, O., Pérez, M., Auat Cheein, F.: Visual-based positioning of aerial maintenance platforms on overhead transmission lines. Appl. Sci. **9**(1), 165 (2019). https://doi.org/10.3390/app9010165
27. da Silva, M.F., Honório, L.M., Marcato, A.L.M., Vidal, V.F., Santos, M.F.: Unmanned aerial vehicle for transmission line inspection using an extended Kalman filter with colored electromagnetic interference. ISA Trans. **100**, 322–333 (2020). https://doi.org/10.1016/j.isatra.2019.11.007
28. Stefenon, S.F., Yow, K.-C., Nied, A., Meyer, L.H.: Classification of distribution power grid structures using inception v3 deep neural

network. Electr. Eng. **104**, 4557–4569 (2022). https://doi.org/10.1007/s00202-022-01641-1
29. Liang, H., Zuo, C., Wei, W.: Detection and evaluation method of transmission line defects based on deep learning. IEEE Access **8**, 38448–38458 (2020). https://doi.org/10.1109/ACCESS.2020.2974798
30. Miao, X., Liu, X., Chen, J., Zhuang, S., Fan, J., Jiang, H.: Insulator detection in aerial images for transmission line inspection using single shot multibox detector. IEEE Access **7**, 9945–9956 (2019). https://doi.org/10.1109/ACCESS.2019.2891123
31. Zhu, J., Guo, Y., Yue, F., Yuan, H., Yang, A., Wang, X., Rong, M.: A deep learning method to detect foreign objects for inspecting power transmission lines. IEEE Access **8**, 94065–94075 (2020). https://doi.org/10.1109/ACCESS.2020.2995608
32. Guo, Y., Pang, Z., Du, J., Jiang, F., Hu, Q.: An improved AlexNet for power edge transmission line anomaly detection. IEEE Access **8**, 97830–97838 (2020). https://doi.org/10.1109/ACCESS.2020.2995910
33. Maeda, K., Takahashi, S., Ogawa, T., Haseyama, M.: Estimation of deterioration levels of transmission towers via deep learning maximizing canonical correlation between heterogeneous features. IEEE J. Sel. Top. Signal Process. **12**(4), 633–644 (2018). https://doi.org/10.1109/JSTSP.2018.2849593
34. Wang, S., Liu, Y., Qing, Y., Wang, C., Lan, T., Yao, R.: Detection of insulator defects with improved ResNeSt and region proposal network. IEEE Access **8**, 184841–184850 (2020). https://doi.org/10.1109/ACCESS.2020.3029857
35. Wang, Y., Wang, C., Zhang, H., Dong, Y., Wei, S.: Automatic ship detection based on RetinaNet using multi-resolution Gaofen-3 imagery. Remote Sens. **11**(5), 531 (2019). https://doi.org/10.3390/rs11050531
36. Liu, J., Jia, R., Li, W., Ma, F., Abdullah, H.M., Ma, H., Mohamed, M.A.: High precision detection algorithm based on improved RetinaNet for defect recognition of transmission lines. Energy Rep. **6**, 2430–2440 (2020). https://doi.org/10.1016/j.egyr.2020.09.002
37. Zhang, P., Zhang, Z., Hao, Y., Zhou, Z., Luo, B., Wang, T.: Multi-scale feature enhanced domain adaptive object detection for power transmission line inspection. IEEE Access **8**, 182105–182116 (2020). https://doi.org/10.1109/ACCESS.2020.3027850
38. Kim, S., Kim, D., Jeong, S., Ham, J.-W., Lee, J.-K., Oh, K.-Y.: Fault diagnosis of power transmission lines using a UAV-mounted smart inspection system. IEEE Access **8**, 149999–150009 (2020). https://doi.org/10.1109/ACCESS.2020.3016213
39. Stefenon, S.F., Freire, R.Z., Meyer, L.H., Corso, M.P., Sartori, A., Nied, A., Klaar, A.C.R., Yow, K.-C.: Fault detection in insulators based on ultrasonic signal processing using a hybrid deep learning technique. IET Sci. Meas. Technol. **14**(10), 953–961 (2020). https://doi.org/10.1049/iet-smt.2020.0083
40. Tao, X., Zhang, D., Wang, Z., Liu, X., Zhang, H., Xu, D.: Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. IEEE Trans. Syst. Man Cybern. Syst. **50**(4), 1486–1498 (2020). https://doi.org/10.1109/TSMC.2018.2871750
41. Guan, H., Sun, X., Su, Y., Hu, T., Wang, H., Wang, H., Peng, C., Guo, Q.: UAV-lidar aids automatic intelligent powerline inspection. Int. J. Electr. Power Energy Syst. **130**, 106987 (2021). https://doi.org/10.1016/j.ijepes.2021.106987
42. Lin, T., Liu, X.: An intelligent recognition system for insulator string defects based on dimension correction and optimized faster R-CNN. Electr. Eng. **103**(1), 541–549 (2021). https://doi.org/10.1007/s00202-020-01099-z
43. Farhadi, A., Redmon, J.: Yolov3: an incremental improvement. In: Computer Vision and Pattern Recognition, pp. 1804–02767. Springer, Berlin (2018)
44. Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. **42**(2), 318–327 (2020). https://doi.org/10.1109/TPAMI.2018.2858826
45. Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2117–2125 (2017)
46. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 770–778. IEEE, Las Vegas (2016). https://doi.org/10.1109/CVPR.2016.90
47. Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517–6525 (2017). https://doi.org/10.1109/CVPR.2017.690
48. Souza, B.J., Stefenon, S.F., Singh, G., Freire, R.Z.: Hybrid-yolo for classification of insulators defects in transmission lines based on UAV. Int. J. Electr. Power Energy Syst. **148**, 108982 (2023). https://doi.org/10.1016/j.ijepes.2023.108982
49. Ultralytics, G.: YOLOv8 in PyTorch. https://github.com/ultralytics/ultralytics (2023)
50. Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., Ke, Z., Li, Q., Cheng, M., Nie, W., Li, Y., Zhang, B., Liang, Y., Zhou, L., Xu, X., Chu, X., Wei, X., Wei, X.: YOLOv6: a single-stage object detection framework for industrial applications. arXiv (2022). https://doi.org/10.48550/ARXIV.2209.02976
51. Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y.M.: YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv (2022). https://doi.org/10.48550/ARXIV.2207.02696
52. Ultralytics, G.: YOLOv5 in PyTorch. https://github.com/ultralytics/yolov5 (2022)
53. Chen, C., Li, O., Tao, C., Barnett, A.J., Su, J., Rudin, C.: This looks like that: deep learning for interpretable image recognition. arXiv preprint arXiv:1806.10574 **5**, 1–12 (2018)
54. Singh, G., Yow, K.-C.: An interpretable deep learning model for Covid-19 detection with chest X-ray images. IEEE Access **9**, 85198–85208 (2021). https://doi.org/10.1109/ACCESS.2021.3087583
55. Singh, G., Yow, K.-C.: These do not look like those: an interpretable deep learning model for image recognition. IEEE Access **9**, 41482–41493 (2021). https://doi.org/10.1109/ACCESS.2021.3064838
56. Erhan, D., Bengio, Y., Courville, A., Vincent, P.: Visualizing higher-layer features of a deep network. Univ. Montreal **1341**(3), 1–13 (2009)
57. Hinton, G.E.: A Practical Guide to Training Restricted Boltzmann Machines, pp. 599–619. Springer, Berlin (2012). https://doi.org/10.1007/978-3-642-35289-8_32
58. Lee, H., Grosse, R., Ranganath, R., Ng, A.Y.: Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In: Proceedings of the 26th Annual International Conference on Machine Learning. ICML09, pp. 609–616. Association for Computing Machinery, New York (2009). https://doi.org/10.1145/1553374.1553453
59. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Computer Vision (ECCV 2014), pp. 818–833. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_53
60. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep inside convolutional networks: visualising image classification models and saliency maps. In: Workshop at International Conference on Learning Representations, pp. 1–8 (2014)
61. Sundararajan, M., Taly, A., Yan, Q.: Axiomatic attribution for deep networks. arXiv (2017). https://doi.org/10.48550/ARXIV.1703.01365
62. Smilkov, D., Thorat, N., Kim, B., Viégas, F., Wattenberg, M.: Smoothgrad: removing noise by adding noise. arXiv preprint arXiv:1706.03825 **1**, 1–10 (2017)

63. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 618–626. IEEE, Venice (2017). https://doi.org/10.1109/ICCV.2017.74
64. Zheng, H., Fu, J., Mei, T., Luo, J.: Learning multi-attention convolutional neural network for fine-grained image recognition. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 5219–5227. IEEE, Venice (2017). https://doi.org/10.1109/ICCV.2017.557
65. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Learning deep features for discriminative localization. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2921–2929. IEEE, Las Vegas (2016). https://doi.org/10.1109/CVPR.2016.319
66. Zhang, N., Donahue, J., Girshick, R., Darrell, T.: Part-based R-CNNs for fine-grained category detection. In: Computer Vision (ECCV2014), pp. 834–849. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_54
67. Li, O., Liu, H., Chen, C., Rudin, C.: Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, pp. 3530–3537 (2018)
68. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE, Miami (2009). https://doi.org/10.1109/CVPR.2009.5206848
69. Prasad, K.N.R.S.V., D'souza, K.B., Bhargava, V.K.: A downscaled faster-RCNN framework for signal detection and time-frequency localization in wideband RF systems. IEEE Trans. Wirel. Commun. **19**(7), 4847–4862 (2020). https://doi.org/10.1109/TWC.2020.2987990
70. Jahangeer, G.S.B., Rajkumar, T.D.: Early detection of breast cancer using hybrid of series network and VGG-16. Multimed. Tools Appl. **80**(5), 7853–7886 (2021). https://doi.org/10.1007/s11042-020-09914-2
71. Murali, S., Deepu, R., Shivamurthy, R., et al.: ResNet-50 vs VGG-19 vs training from scratch: a comparative analysis of the segmentation and classification of pneumonia from chest X-ray images. Glob. Transit. Proc. (2021). https://doi.org/10.1016/j.gltp.2021.08.027
72. Gao, M., Qi, D., Mu, H., Chen, J.: A transfer residual neural network based on ResNet-34 for detection of wood knot defects. Forests **12**(2), 212 (2021). https://doi.org/10.3390/f12020212
73. Fulton, L.V., Dolezel, D., Harrop, J., Yan, Y., Fulton, C.P.: Classification of Alzheimer's disease with and without imagery using gradient boosted machines and ResNet-50. Brain Sci. **9**(9), 212 (2019). https://doi.org/10.3390/brainsci9090212
74. Pan, T.-S., Huang, H.-C., Lee, J.-C., Chen, C.-H.: Multi-scale ResNet for real-time underwater object detection. In: Signal, Image and Video Processing, pp. 1–9 (2020). https://doi.org/10.1007/s11760-020-01818-w
75. Ghiasi-Shirazi, K.: Generalizing the convolution operator in convolutional neural networks. Neural Process. Lett. **50**(3), 2627–2646 (2019). https://doi.org/10.1007/s11063-019-10043-7
76. Nalaie, K., Ghiasi-Shirazi, K., Akbarzadeh-T, M.-R.: Efficient implementation of a generalized convolutional neural networks based on weighted euclidean distance. In: 2017 7th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 211–216. IEEE, Mashhad (2017). https://doi.org/10.1109/ICCKE.2017.8167877
77. Stefenon, S.F., Kasburg, C., Nied, A., Klaar, A.C.R., Ferreira, F.C.S., Branco, N.W.: Hybrid deep learning for power generation forecasting in active solar trackers. IET Gener. Transm. Distrib. **14**(23), 5667–5674 (2020). https://doi.org/10.1049/iet-gtd.2020.0814
78. Chen, H., He, Z., Shi, B., Zhong, T.: Research on recognition method of electrical components based on yolo v3. IEEE Access **7**, 157818–157829 (2019). https://doi.org/10.1109/ACCESS.2019.2950053
79. Chen, Z., Xiao, Y., Zhou, Y., Li, Z., Liu, Y.: Insulator recognition method for distribution network overhead transmission lines based on modified yolov3. In: 2020 Chinese Automation Congress (CAC), pp. 2815–2820 (2020). https://doi.org/10.1109/CAC51589.2020.9327352
80. Feng, Z., Guo, L., Huang, D., Li, R.: Electrical insulator defects detection method based on yolov5. In: 2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS), Suzhou, China, pp. 979–984 (2021). https://doi.org/10.1109/DDCLS52934.2021.9455519
81. Wang, L., Chen, Z., Hua, D., Zheng, Z.: Semantic segmentation of transmission lines and their accessories based on UAV-taken images. IEEE Access **7**, 80829–80839 (2019). https://doi.org/10.1109/ACCESS.2019.2923024
82. Li, X., Su, H., Liu, G.: Insulator defect recognition based on global detection and local segmentation. IEEE Access **8**, 59934–59946 (2020). https://doi.org/10.1109/ACCESS.2020.2982288
83. Wen, Q., Luo, Z., Chen, R., Yang, Y., Li, G.: Deep learning approaches on defect detection in high resolution aerial images of insulators. Sensors **21**(4), 1033 (2021). https://doi.org/10.3390/s21041033
84. Sampedro, C., Rodriguez-Vazquez, J., Rodriguez-Ramos, A., Carrio, A., Campoy, P.: Deep learning-based system for automatic recognition and diagnosis of electrical insulator strings. IEEE Access **7**, 101283–101308 (2019). https://doi.org/10.1109/ACCESS.2019.2931144
85. Vigneshwaran, B., Maheswari, R.V., Kalaivani, L., Shanmuganathan, V., Rho, S., Kadry, S., Lee, M.Y.: Recognition of pollution layer location in 11 kv polymer insulators used in smart power grid using dual-input VGG convolutional neural network. Energy Rep. (2021). https://doi.org/10.1016/j.egyr.2020.12.044
86. Deng, C.: The method of insulator defect recognition based on group theory. IEEE Access **9**, 96380–96389 (2021). https://doi.org/10.1109/ACCESS.2021.3095382
87. Liu, C., Wu, Y., Liu, J., Han, J.: MTI-YOLO: a light-weight and real-time deep neural network for insulator detection in complex aerial images. Energies **14**(5), 1426 (2021). https://doi.org/10.3390/en14051426

Gurmail Singh received his Ph.D. in Algebra from the University of Regina, Canada, in 2015. Currently, he is working with the Department of Computer Sciences at the University of Wisconsin-Madison, USA.

Stefano Frizzo Stefenon received his Ph.D. in Electrical Engineering from the State University of Santa Catarina, Brazil in 2021. He currently works at Fondazione Bruno Kessler, Trento, Italy. Condition monitoring and predictive maintenance of complex industrial systems: Model-based reasoning meets data science.

Kin-Choong Yow received his Ph.D. from Cambridge University, UK in 1998. He joined the University of Regina (UofR) in 2018, where he is presently a Professor in the Faculty of Engineering and Applied Science, Canada.