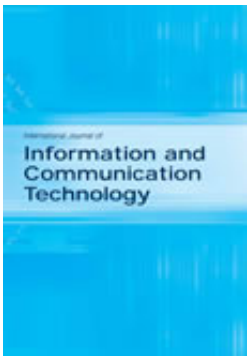


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

A fine-tuned YOLOv11-based insulator icing detection algorithm for intelligent inspections of power systems

Hai Huang, Xun Zhang, Dianli Chen, Yong Du, Shenli Wang, Xiaohua Liu, Quan Fang, Yuhang Xia

DOI: 10.1504/IJICT.2025.10075001

Article History:

Received:	18 January 2025
Last revised:	19 March 2025
Accepted:	19 March 2025
Published online:	05 January 2026

A fine-tuned YOLOv11-based insulator icing detection algorithm for intelligent inspections of power systems

Hai Huang, Xun Zhang, Dianli Chen, Yong Du, Shenli Wang, Xiaohua Liu and Quan Fang

State Grid Hubei Extra High Voltage Company,
Wuhan, 430050, China
and
Hubei Super-Energic Electric Power Co., Ltd.,
Wuhan, 430050, China
Email: 2922254268@qq.com
Email: 448544771@qq.com
Email: 2822254268@qq.com
Email: 2077678160@qq.com
Email: 2124695301@qq.com
Email: 995885579@qq.com
Email: 3925897955@qq.com

Yuhang Xia*

Hubei Key Lab of Micro-Nanoelectronic Materials and Devices,
Faculty of Microelectronics,
Hubei University,
Wuhan, 430062,
Shanghai, 200240, China
Email: hubxiay@stu.hubu.edu.cn
*Corresponding author

Abstract: Ice accumulation on insulators can lead to electrical breakdown, equipment damage, and line outages, making timely and accurate detection essential for maintaining the safe and stable operation of power systems. This paper proposes an ice accretion detection method for insulators based on You Only Look Once version 11 (YOLOv11), integrating image processing and deep learning techniques to achieve automated detection. A self-built dataset was used to fine-tune YOLOv11, enhancing the model's accuracy and robustness in complex environments. Compared to its predecessors, YOLOv11 features an improved backbone network for more efficient feature extraction, advanced attention mechanisms for enhanced focus on critical regions, and an anchor-free detection head that reduces computational complexity while maintaining high precision. Multi-scale feature fusion ensures the accurate detection of ice accretion of various sizes, while dynamic label assignment optimises alignment between predictions and ground truth. Experimental results demonstrate that the fine-tuned YOLOv11-based algorithm achieves high mean average precision (mAP) and F1-scores on the test set, indicating

robust detection performance. The proposed method not only enhances detection efficiency but also reduces labour costs, making it well-suited for large-scale power line monitoring.

Keywords: ice accumulation; insulator icing detection; YOLOv11; ice accretion detection.

Reference to this paper should be made as follows: Huang, H., Zhang, X., Chen, D., Du, Y., Wang, S., Liu, X., Fang, Q. and Xia, Y. (2025) 'A fine-tuned YOLOv11-based insulator icing detection algorithm for intelligent inspections of power systems', *Int. J. Information and Communication Technology*, Vol. 26, No. 48, pp.1–22.

Biographical notes: Hai Huang centres his research on live working technology for transmission lines and intelligent operation and maintenance technology for transmission lines. He has in-depth expertise in designing safe and efficient live working schemes, developing specialised tools for live maintenance, and exploring the integration of intelligent technologies (such as UAVs and robotics) into live operation scenarios. His research contributes to minimising power outages during maintenance, enhancing the safety of on-site operations, and promoting the intelligent upgrading of live working in power systems.

Xun Zhang is a Senior Engineer with a focus on intelligent power system design and key technology research for smart grids. His work involves the overall design of intelligent power systems, including the planning of grid structures, the integration of renewable energy resources, and the development of smart grid dispatch and control platforms. He also conducts in-depth research on core technologies of smart grids, such as energy storage integration, demand response, and grid stability control, aiming to build a more flexible, efficient, and sustainable modern power system.

Dianli Chen is a Senior Engineer in the power industry, and specialises in transmission line live working technology and intelligent operation and maintenance technology for transmission lines. His research focuses on solving practical challenges in live working, such as adapting to complex terrain and extreme weather conditions, and optimising the performance of live working equipment. He also explores the application of AI and digital twin technologies in transmission line inspection and maintenance, striving to realise the automation and intelligence of live operation and maintenance processes.

Yong Du is a Senior Engineer dedicated to key technology research for smart grids and intelligent power system design. His research interests cover the development of smart grid communication networks, the application of big data analytics in power system management, and the design of microgrid systems for distributed energy integration. He works to promote the innovation and industrialisation of key smart grid technologies, supporting the transformation of traditional power grids into intelligent, green, and user-centric energy systems.

Shenli Wang focuses on the research and application of power transmission line operation and maintenance technologies. His core research directions include transmission line operation and maintenance and intelligent operation and maintenance technology for transmission lines, where he has dedicated efforts to optimising routine inspection processes, developing intelligent monitoring systems, and improving the reliability and efficiency of transmission line operations. His work aims to integrate advanced sensing and data analysis technologies into traditional maintenance workflows, providing technical support for the stable operation of power grids.

Xiaohua Liu focuses on transmission line operation and maintenance and intelligent operation and maintenance technology for transmission lines with extensive experience as a Senior Engineer. His research involves the development of condition-based maintenance strategies for transmission lines, the improvement of fault diagnosis accuracy using advanced sensor technologies, and the construction of digital management platforms for transmission line assets. His work aims to reduce maintenance costs, extend the service life of transmission line equipment, and ensure the safe and stable operation of the power transmission network.

Quan Fang is a Senior Engineer specialising in the intelligent transformation of distribution networks and the integration of energy internet technologies. His research interests cover the development of intelligent dispatching algorithms for distribution networks, the collaborative control of distributed energy storage and distribution grids, and the exploration of user-side interaction mechanisms in energy internet systems. He works to advance the flexible and efficient operation of distribution networks, promoting the integration of distributed energy resources into the main grid and supporting the construction of a reliable, economical, and user-friendly intelligent distribution system.

Yuhang Xia is currently a graduate student at Hubei University, with a research focus on deep learning, particularly in object detection and image segmentation. He is dedicated to advancing the precision and efficiency of image processing through cutting-edge computer vision technologies, contributing to the progress of related fields.

1 Introduction

In electrical power systems, insulators serve as a crucial component of transmission lines, fulfilling dual roles of mechanical connection and electrical insulation. The primary function of insulators is to ensure electrical isolation of power lines, preventing current leakage through line supports or other ground facilities. However, due to their constant exposure to outdoor environments, insulators are subjected to the erosion of natural conditions such as frost, rain, snow, and sand, especially in cold regions where ice accretion is particularly severe. Ma et al. (2021) found that during the cold season, the accumulation of snow and frost not only increases the weight and mechanical load on the insulators but also triggers the degradation of their electrical performance, significantly reducing their electrical insulation strength, thereby increasing the risk of power system failures.

Ice accretion poses a severe threat to power systems. Firstly, snow and ice adhering to the surface of insulators can alter their surface morphology and electrical properties, leading to electrical breakdown and equipment damage. Secondly, ice accumulation may cause insulators to fracture or detach, resulting in line outages. Additionally, Wang et al. (2023b) proposed that the weight of accumulated ice can also lead to line breaks or tower inclination issues. Therefore, timely and effective detection of ice accretion on insulators is crucial for the safe and stable operation of power systems.

Traditional methods for detecting ice accretion on insulators primarily rely on manual inspections and ground checks. Han et al. (2019) proposed while these methods can identify issues to some extent, they are inefficient, costly, and not well-suited for

large-scale line inspections. With the rapid advancement of drone technology and computer vision, automated detection technologies based on image processing have become a focus of research. By deploying high-definition cameras on drones or other elevated platforms to capture extensive high-quality images, and employing image processing and machine learning algorithms to automatically detect ice accretion on insulators, the efficiency and accuracy of inspections can be significantly enhanced.

Among the many object detection algorithms, you only look once (YOLO) stands out as an efficient and fast real-time object detection algorithm. It has been widely applied to various computer vision tasks because of its ability to simultaneously perform object classification and localisation. YOLO transforms the object detection problem into a regression problem, significantly improving the speed and accuracy of object detection. According to Bharati and Pramanik (2020), YOLO achieves a speed performance of approximately 21 to 155 fps, making it one of the fastest object detection algorithms available. Zhang et al. (2024) believes that although YOLO has performed quite well on some standard datasets, it still faces numerous challenges in practical applications, especially for specific tasks such as ice accretion detection on insulators. These challenges include the varying manifestations of ice under different weather conditions and the unevenness of ice accretion levels.

In comparison to other object detection algorithms, faster R-CNN has also been widely used in the field of object detection. Li (2021) analysed the performance of faster R-CNN models based on different pre-training models and conducted a comprehensive evaluation of the performance of faster R-CNN. The experimental results showed the accuracy and detection speed of R-CNN, fast R-CNN, and faster R-CNN based on three different datasets. Faster R-CNN significantly improves the overall performance by adding a region proposal network (RPN), especially in terms of detection speed. However, the application of different pre-training models will result in a great difference in the performance of faster R-CNN. This analysis provides a benchmark for evaluating the performance of object detection algorithms, including YOLO, in various applications.

Furthermore, Girshick (2015) introduced fast R-CNN, which builds on the original R-CNN by introducing a more efficient training and detection process. Fast R-CNN uses a single-stage training process that jointly optimises classification and bounding box regression, significantly reducing the computational cost and improving detection accuracy. This work laid the foundation for subsequent advancements in object detection, including Faster R-CNN and YOLO.

Therefore, this paper aims to explore the feasibility and methods of ice accretion detection on insulators based on the YOLOv11 algorithm. Specifically, we use a self-built dataset to fine-tune YOLOv11 to enhance the model's accuracy and robustness in complex environments. By training YOLOv11 specifically, it hopes to achieve efficient and accurate ice detection, providing effective technical support for the operation and maintenance of power systems. Compared to traditional detection methods, automated detection based on YOLOv11 can not only significantly improve detection efficiency but also reduce labour costs to some extent, and better cope with the monitoring tasks of large-scale lines.

The contributions of this study are mainly reflected in the following aspects:

1 A method for ice accretion detection on insulators based on the YOLOv11 algorithm is proposed, combining image processing and deep learning technologies to achieve automation in ice detection.

2 A self-built dataset is used to fine-tune YOLOv11, overcoming detection difficulties in complex environments such as low temperatures, snow, and ice.

3 The effectiveness and robustness of this method are verified through experiments, achieving good detection results.

Through the research presented in this paper, we aim to provide an efficient, accurate, and scalable ice accretion detection solution for the intelligent operation and maintenance of power systems, thereby promoting the development of power equipment inspection technology.

2 Related work

In the field of insulator ice accretion detection, as the demand for power equipment inspection continues to grow, the associated detection methods are constantly evolving and improving. Liu et al. (2020b) proposed early research methods primarily relied on manual patrols and traditional image processing techniques, which suffered from low efficiency and poor accuracy, and have gradually been replaced by newer technologies. Here are some typical research advancements:

2.1 Ice accretion detection based on image processing

In the early stages, ice accretion detection on insulators mainly depended on image processing techniques. By acquiring static images or video frames and combining basic image processing algorithms, researchers were able to perform simple target recognition and analysis. Pernebayeva et al. (2019) proposed a method based on image smoothing and threshold transformation involving filtering the image to remove noise and then extracting the outline of the insulator Hao et al. (2022) proposed threshold transformation distinguished between ice and snow areas and non-ice areas by setting a pixel value threshold, thereby identifying the iced insulators. Irene et al. (2009) and Solangi (2018) believe these methods were simple and computationally light, but their recognition effects were easily affected in complex environments, especially when it was difficult to provide accurate results in low-contrast or high-noise backgrounds.

2.2 Traditional feature extraction methods

With the development of computer vision technology, researchers began to introduce more complex feature extraction methods. For example, Mikolajczyk and Schmid (2005) proposed a feature extraction technique based on the GLOH descriptor, which can effectively capture edge and texture information in images for target recognition and classification. These methods extracted local features of insulators from images and then used machine learning algorithms for classification to determine whether there was ice accretion (Liu et al., 2020a). Additionally, the GVF snake model was proposed as an image segmentation method based on active contour models. This method accurately fitted the contours of insulators to further analyse whether they were iced. The GVF snake method could effectively detect irregular insulator contours in images, especially suitable for images with complex shapes or unclear edges (Zhang, 2022). Despite

performing well in experimental settings, these methods still face significant challenges in complex outdoor environments, particularly under strong lighting, fog, or snowy weather, especially when dealing with background interference in images (Nusantika et al., 2024).

2.3 Detection methods based on 3D reconstruction

In recent years, with the continuous advancement of computer vision and stereo vision technology, researchers have begun to explore ice accretion detection methods for insulators based on 3D reconstruction. By employing binocular vision or multi-view imaging techniques, it is possible to reconstruct a 3D point cloud model of the insulator. Based on these 3D data, researchers can more accurately calculate the thickness of the ice accretion and thereby assess the impact of the ice on the insulator (Liu et al., 2023). For example, Wang et al. (2023a) used binocular vision technology to match two images of the same target object taken from different angles, reconstruct a three-dimensional point cloud model, and extract surface morphology information of the insulator from it. These methods have a significant advantage in terms of accuracy, especially when calculating the thickness of ice accretion, providing more precise results than 2D image processing (Guo and Hu, 2017). However, 3D reconstruction methods have high requirements for image acquisition, typically requiring multi-perspective shooting data, and also have stringent demands on image quality and camera calibration (Marek et al., 2021). Moreover, these methods have high computational complexity and poor real-time performance, which limits their application in large-scale, high-frequency inspection tasks (Gonçalves et al., 2022).

2.4 Deep learning methods

In recent years, the rapid development of deep learning technology has provided new solutions for insulator ice accretion detection. Deep learning, especially the application of convolutional neural networks (CNNs) and object detection algorithms, has become the mainstream method in the field of image recognition. The YOLO series of algorithms, as an efficient object detection method, has achieved excellent results in various object detection tasks due to its fast detection speed and high accuracy. In the detection of ice accretion on insulators, the YOLO model can simultaneously perform target detection and classification in a single forward propagation process, thus achieving rapid localisation and identification of iced insulators (Qiu et al., 2022). Researchers have fine-tuned the YOLO model on self-built datasets, significantly enhancing the detection accuracy in complex environments. For example, Chen (2024) applied YOLOv11 to target detection in high-resolution images, and further improved the recognition effect under different meteorological conditions through data enhancement, transfer learning and other means.

Compared to traditional image processing and 3D reconstruction methods, deep learning-based object detection methods show strong robustness and flexibility in complex environments. Particularly, the YOLO algorithm can provide rapid responses when processing real-time images, meeting the needs of large-scale inspections (Liu et al., 2023). Additionally, the YOLO model can be combined with automated equipment such as drones to achieve automated inspections over a wide area, reducing labour costs

and improving the safety and efficiency of power system operations (Nguyen et al., 2018).

3 Method

In this study, we propose an ice detection method for insulators based on the YOLOv11 model. YOLO is an efficient object detection algorithm that can simultaneously complete object classification and localisation tasks through a single forward pass. To enable YOLOv11 to accurately detect whether insulators are iced, we used our own collected dataset of insulator images and fine-tuned the YOLOv11 model. This method, through refined training, enables the YOLOv11 model to accurately identify and locate iced insulators under adverse weather conditions.

YOLO is a CNN-based, end-to-end object detection algorithm. The YOLO model revolutionised object detection algorithms by introducing a unified neural network architecture that can simultaneously handle bounding box regression and object classification tasks. This integrated approach marks a significant departure from traditional two-stage detection methods, as the YOLO model offers end-to-end training capabilities through a fully differentiable design.

The main feature of YOLOv11 is transforming the object detection task into a regression problem, directly predicting bounding boxes and class probabilities from the input image. Its model structure mainly includes the backbone, neck, and head. The backbone acts as a feature extractor, using CNNs to convert raw image data into multi-scale feature maps. Secondly, the neck serves as an intermediate processing stage, primarily used for aggregating and enhancing feature representations across different scales. The Head acts as the prediction structure, generating the final output based on the refined feature maps.

In YOLOv11, the model requires input images to have a fixed size, typically 640 × 640 pixels. If the provided image does not conform to this size standard, the system will automatically resize the image to the required dimensions, ensuring consistency in all inputs. The model's output is a tensor with dimensions (S, S, $B \times (5 + C)$), where $S \times S$ refers to the number of grids into which the model divides the input image. The entire image is split into $S \times S$ grid cells, with each cell responsible for predicting objects within its coverage area.

B represents the number of bounding boxes predicted by each grid cell. To improve detection accuracy, YOLO allows each cell to predict multiple bounding boxes (B of them), to better accommodate objects of different shapes and sizes that may appear within that cell.

The number 5 refers to the five parameters associated with each bounding box: four parameters define the position and size of the bounding box (the coordinates of the centre point x, y, as well as the width w and height h, all of which use relative coordinates with respect to the grid cell), and the fifth parameter is the confidence score, which reflects the system's belief in the presence of an object within the bounding box and is measured by the intersection over union (IoU) between the bounding box and the actual object.

For each bounding box, the model outputs the following information: four positional parameters – x, y, w, h, which correspond to the relative coordinates of the centre point of the bounding box and the relative width and height of the bounding box. The confidence

score represents the estimated probability that the bounding box contains an object, calculated based on the IoU between the bounding box and the true label.

Additionally, *C* denotes the number of classes, meaning that each bounding box also includes a vector of length *C* representing the probability distribution across various classes, thereby determining the most likely object class within the bounding box. In your description, you specifically mentioned the prediction of insulators, which implies that the model has been trained to recognise these specific objects.

Additionally, *C* represents the number of classes, meaning that each bounding box also includes a vector of length *C* that represents the probability distribution across various classes, thereby determining the most likely object class within the bounding box. In your description, you specifically mentioned the prediction of insulators, which implies that the model has been trained to recognise this particular type of object.

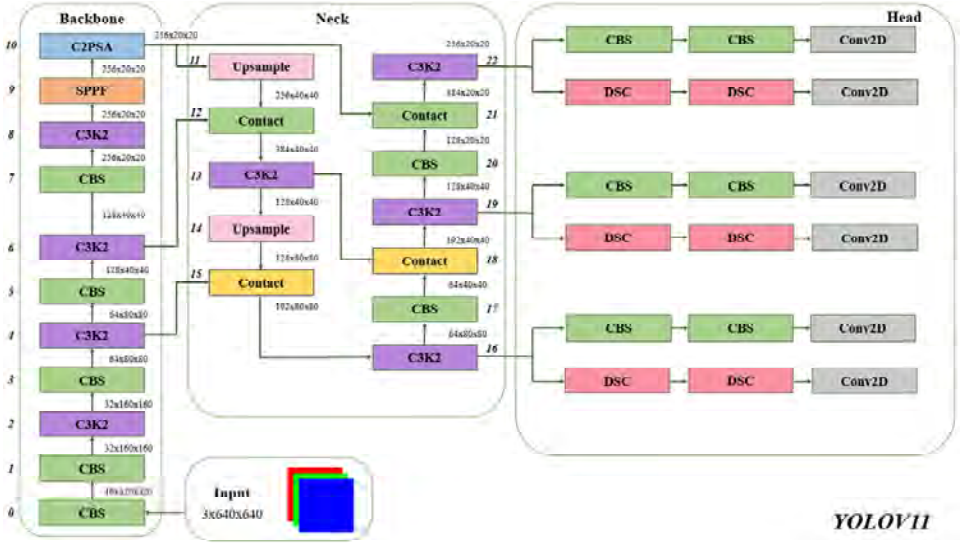
$$Confidence = P(Object) * IoU \tag{1}$$

C represents the number of classes, with each bounding box predicting a class probability distribution, i.e., the probability that the bounding box belongs to each class.

YOLOv11 processes input images through a grid-based approach, where each grid cell is responsible for predicting objects that fall within its region. Each grid cell predicts multiple bounding boxes and predicts the IoU for each bounding box with the ground truth object, as well as the class confidence.

The YOLO series has seen multiple iterations such as YOLOv5, YOLOv8, and so on. YOLOv11 is the latest iteration in the YOLO series, building upon the foundation of YOLOv1. It represents a significant leap in real-time object detection technology. This new version introduces substantial enhancements in both architecture and training methods, pushing the boundaries of accuracy, speed, and efficiency.

Figure 1 YOLOv11 detection framework (see online version for colours)



3.1 Model initialisation

3.1.1 Backbone

The backbone is a crucial component of the YOLO architecture, responsible for extracting features from input images at multiple scales. This network primarily involves stacking convolutional layers and specialised blocks to extract image features from the given images.

The YOLOv11 model used in this study maintains a structure similar to its predecessors in the Backbone section, mainly utilising initial convolutional layers to downsample the image. These layers form the foundation of the feature extraction process, gradually reducing spatial dimensions while increasing the number of channels. A notable improvement in YOLOv11 is the introduction of the C3k2 block, which replaces the C2f block used in previous versions. The C3k2 block is a more computationally efficient implementation of the cross-stage partial (CSP) bottleneck. It employs two smaller convolutions instead of a large convolutional kernel. The 'k2' in C3k2 indicates the smaller kernel size, which helps to speed up processing while maintaining performance.

The YOLOv11 model we used inherits the spatial pyramid pooling – fast (SPPF) block from previous versions and introduces a cross-stage partial with spatial attention (C2PSA) module based on it. The C2PSA module, as a key innovation, significantly enhances the model's ability to focus on important areas in the feature map. By integrating a spatial attention mechanism, the C2PSA module allows YOLOv11 to more accurately focus on key areas in the image, thereby improving detection efficiency and accuracy. Additionally, this module ensures that the model can focus on specific regions of interest by performing spatial pooling operations on features, thus optimising the detection of objects of different scales and positions.

3.1.2 Loss function

The loss function is primarily defined and calculated within the backbone, and its main purpose in the main network is to guide the learning process of the model. It measures the difference between the model's predicted values and the actual labels, and adjusts the model parameters to minimise this difference through backpropagation algorithms. The YOLO loss function consists of three parts: localisation loss (bounding box loss), confidence loss, and classification loss.

Localisation loss typically uses mean squared error (MSE) to calculate the difference between the predicted bounding boxes and the true bounding boxes. For each bounding box, its loss can be represented as:

$$L_{loc} = \sum_{i=0}^{B} 1_i^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + \left(\sqrt{\omega_i} - \sqrt{\hat{\omega}_i}\right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i}\right)^2 \right] \quad (2)$$

In the context, x_i, y_i, ω_i, h_i represent the centre coordinates and the width and height of the predicted bounding box, respectively. $\hat{x}_i$, $\hat{y}_i$, $\hat{\omega}_i$, $\hat{h}_i$ are the corresponding parameters of the true bounding box. 1_i^{obj} is the indicator function, which signifies whether there is a detection target within that grid cell.

The confidence loss calculates the difference between the predicted box's confidence and the true target's confidence, typically using MSE:

$$L_{conf} = \sum_{i=0}^{B} 1_i^{obj} \left(C_i - \hat{C}_i \right)^2 \tag{3}$$

where C_i is the confidence of the predicted box, representing the probability that the box contains an object, and $\hat{C}_i$ is the confidence of the true box.

The classification loss calculates the difference between the predicted class and the true class, often using cross-entropy loss:

$$L_{cls} = -\sum_{C=0}^{C} p_c \log\left(\hat{p}_c \right) \tag{4}$$

where $\hat{p}_c$ is the probability of the true class (in this study, it refers to the probability of whether the insulator is iced), and $\hat{p}_c$ is the predicted class probability.

The final total loss is the weighted sum of the three losses:

$$L_{total} = \lambda_{loc} L_{loc} + \lambda_{conf} L_{conf} + \lambda_{cls} L_{cls} \tag{5}$$

where L_{total} is the localisation loss, which calculates the difference between the predicted box's coordinates and size and the true box's coordinates and size.

3.1.3 Neck

The neck layer in object detection models is situated between the backbone and the Head, serving to enhance, fuse, and optimise features from different levels (scales). This layer enables the model to capture target information of varying sizes and complexities. By employing architectures like feature pyramid networks (FPN) and path aggregation network (PANet), the neck layer increases the model's flexibility and accuracy, particularly in complex environments, and significantly improves the detection capability for smaller objects.

In the detection task for insulators, the size variation of insulators in images due to distance can be substantial, making it challenging for a single-scale feature map to handle such data. To address this issue, multi-scale feature fusion techniques are introduced in the neck part of the model, primarily by fusing feature maps from different levels (depths), allowing the network to maintain good perceptual abilities across different scales.

YOLOv11 replaced the original C2f block in the neck section with the C3k2 block. The C3k2 block is designed to enhance speed and efficiency, thereby improving the overall performance of the feature aggregation process. After upsampling and concatenation, YOLOv11's neck section utilises this improved block to further increase speed and performance. YOLOv11 also strengthened the spatial attention mechanism, particularly through the C2PSA module. This attention mechanism enables the model to focus on key areas of the image, potentially improving detection accuracy, especially when dealing with small objects or partially occluded objects.

3.1.4 Head

The head component is the final part of the network, responsible for generating the final output results from the features extracted and fused through the backbone and neck. The primary function of the head is to predict objects, i.e., to output the class, bounding box position, and confidence for each candidate region (or grid cell).

The head predicts the class for each bounding box. For multiple bounding boxes per grid cell, YOLO predicts the probability of each bounding box belonging to different classes, meaning the head outputs a probability distribution of whether it is an insulator. Subsequently, this component also performs positional regression for each predicted bounding box, outputting the relative position coordinates and predicting the confidence for each box. Confidence indicates the likelihood that the system believes the predicted box contains an object, and it is also related to the IoU between the predicted box and the true box.

3.2 Weight initialisation

Initialising weights is an important step before training a neural network. Good initialisation can accelerate training and improve model performance. YOLO can use two methods for weight initialisation:

1 Loading pre-trained weights: If using a pre-trained model, you can directly load the pre-trained weights, skipping the training from scratch.

2 Random weight initialisation: If training the model from scratch, YOLO will initialise each layer of the network through random initialisation.

This article adopts the method of loading a pre-trained model, where yolov11m.pt is a file that contains both the network structure and the trained weights. After loading, all the model's weights are initialised to the pre-trained state. This model has been trained on large-scale datasets such as COCO and ImageNet. Using a pre-trained model can significantly speed up the training process and provide better initial performance.

3.3 Input process

To enhance the model's generalisation and accuracy, we performed standardisation on the images before processing them, ensuring that the input image size is unified to 640 × 640, and normalising the pixel values to scale the pixel values to the [0, 1] range. In addition to standardisation, we also carried out data augmentation, including random cropping, rotation, translation, scaling, and colour jittering operations. This helps to increase the diversity of the data and improve the model's robustness to different environments.

3.4 Optimiser

Optimisers are one of the core components used to train neural networks. They are responsible for adjusting the network weights through gradient descent algorithms to minimise the loss function. An optimiser typically consists of the following parts: the chosen optimisation algorithm (such as Adam, SGD, etc.), hyperparameter settings like

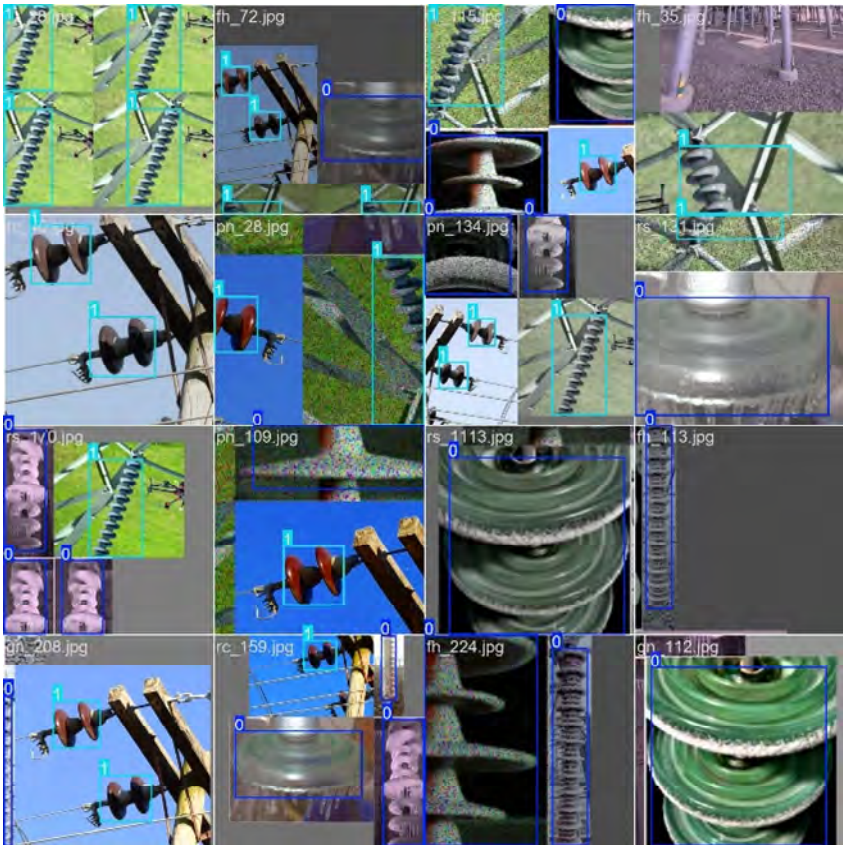
learning rate and momentum, and how to update the network weights. During fine-tuning, the following training strategies were used:

Optimiser: the Adam optimiser was adopted, which can adaptively adjust the learning rate based on the historical gradients of each parameter, thereby effectively accelerating training. Adam has good convergence speed and stability. The Adam optimiser adjusts the learning rate for parameter updates by calculating the first moment (mean) and second moment (variance) of the gradients. The parameter update formula is as follows:

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{m_t}{\sqrt{v_t} + \in} \tag{6}$$

where α is the learning rate (commonly set to 0.001), ϵ is a small constant to prevent division by zero errors, and m_t and v_t are the first and second moments of the gradients, respectively.

Figure 2 Dataset (label 1: insulator without ice; label 0: insulator with ice) (see online version for colours)



After fine-tuning, relevant hyperparameters for model training such as learning rate, batch size, and epochs were set. After training, the model was tested on the test set to evaluate its precision, recall, F1-score, and other metrics across different categories, to

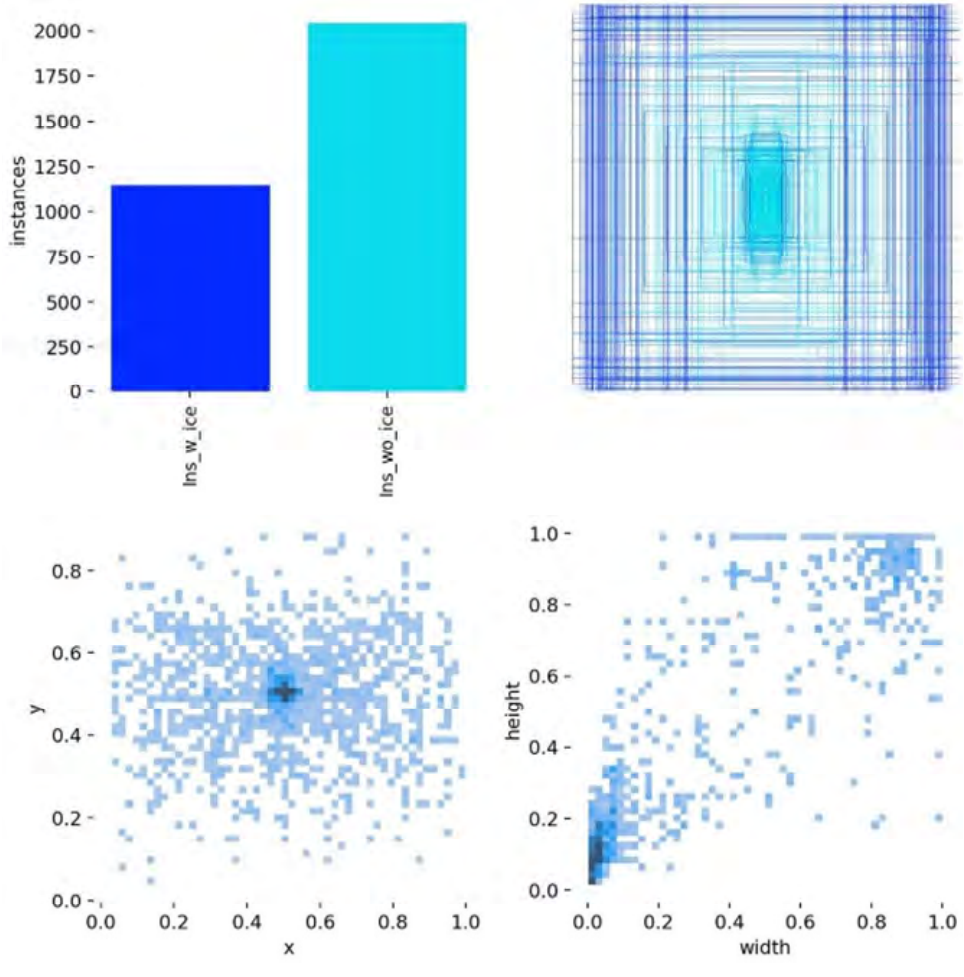
verify its effectiveness in practical applications. The evaluation results indicate that the fine-tuned YOLO model performs excellently in detecting iced insulators, achieving high accuracy in object detection under various environmental conditions.

4 Result

4.1 Dataset

There collected a dataset of insulator images captured under various environmental conditions, such as clear skies, rain and snow, and low temperatures. The dataset includes images of multiple categories, including insulators without ice and insulators with varying degrees of ice coverage. Each image in the dataset is annotated with the location of the insulator (through bounding boxes) and labels indicating the presence of ice.

Figure 3 Dataset (information of labelled samples) (see online version for colours)



For each image, the LabelImg tool had been applied to annotate the insulators, defining two category labels: no ice and ice. Each object is annotated as a rectangular bounding box, and each bounding box is associated with a category label. The annotated data includes the image path, coordinates (x_{min}, y_{min}, x_{max}, y_{max}) for each object, and the category of that object. We divided the organised dataset into two parts: a training set (accounting for 80% of the total data) and a validation set (accounting for 20% of the total data).

4.2 Train

The total number of training epochs is set to 100. The training set consists of 2,493 images, while the validation set contains 634 images. Each batch contains 36 images, and the program updates the weights using the average gradient of 36 samples each time. The choice of 36 as the batch size is to balance training speed and memory usage. A smaller batch size makes training more stable but less computationally efficient; a larger batch size aids in faster training but requires more GPU memory.

Each batch consists of 36 images, and the model performs forward and backward propagation on each batch, then updates the network parameters. During training, the network iterates over this data multiple times until the training is completed or stopped early.

To prevent the model from overfitting during training, we employed an early stopping strategy. Early stopping is a method that decides whether to stop training based on the performance changes on the validation set. We set the patience to 100, which means that if there is no significant improvement in the performance metrics on the validation set for 100 consecutive batches, the training will be stopped early.

Monitoring a key performance metric on the validation set (such as loss or accuracy) requires setting an early stopping strategy. If there is no significant improvement in performance within a specified 'patience period', it indicates that the model has stabilised and may not be able to improve further, and the risk of overfitting begins to increase. Therefore, stopping the training prevents the program from overfitting and saves computational resources.

4.3 Training results

Having completed 889 batches of training, we have acquired the optimal weight parameters, which are the result of fine-tuning on this particular training dataset. These optimal weights effectively capture the characteristics and patterns present in the training data, thus improving the model's reliability for object detection tasks. The training process included numerous iterative optimisations, refining the model through backpropagation and minimising the loss function, ultimately enabling the model to precisely recognise both the positions and categories of targets within images.

From Figure 4, it can be observed that the mAP50(B) and mAP50-95(B) metrics consistently increase as the number of iterations grows, indicating a continuous improvement in the model's detection capability. mAP50(B) represents the mean average precision (mAP) at an IoU threshold of 0.5. This metric exhibits a rapid increase during training and eventually approaches 1.0, suggesting that the model achieves high detection accuracy under a relatively lenient IoU criterion. In contrast, mAP50-95(B) is computed as the mAP across multiple IoU thresholds ranging from 0.5 to 0.95. While its overall

trend closely follows that of mAP50(B), the final value is slightly lower, reflecting a decline in detection accuracy under stricter IoU conditions. This trend aligns with common patterns observed in object detection training, demonstrating that the model maintains robust performance across various IoU thresholds. Furthermore, as training progresses, the model's performance gradually converges, indicating an improvement in its generalisation capability.

Figure 4 Training results (see online version for colours)

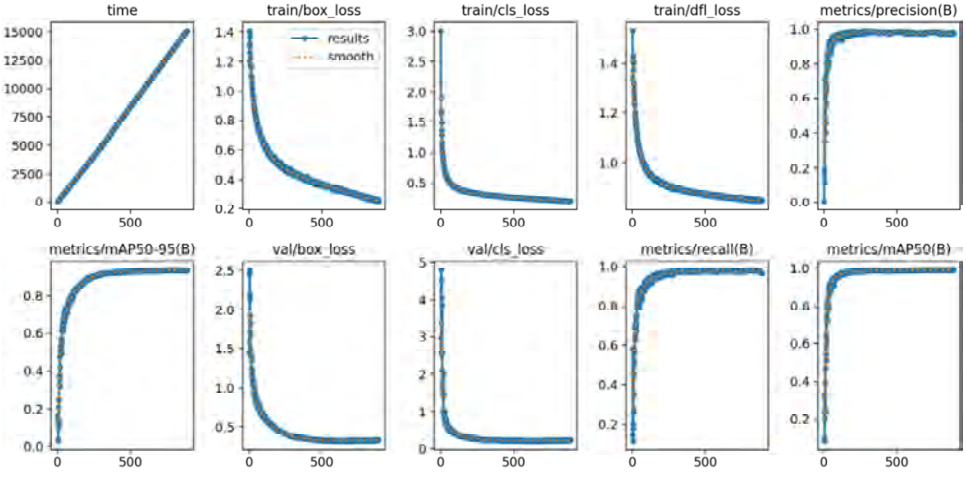
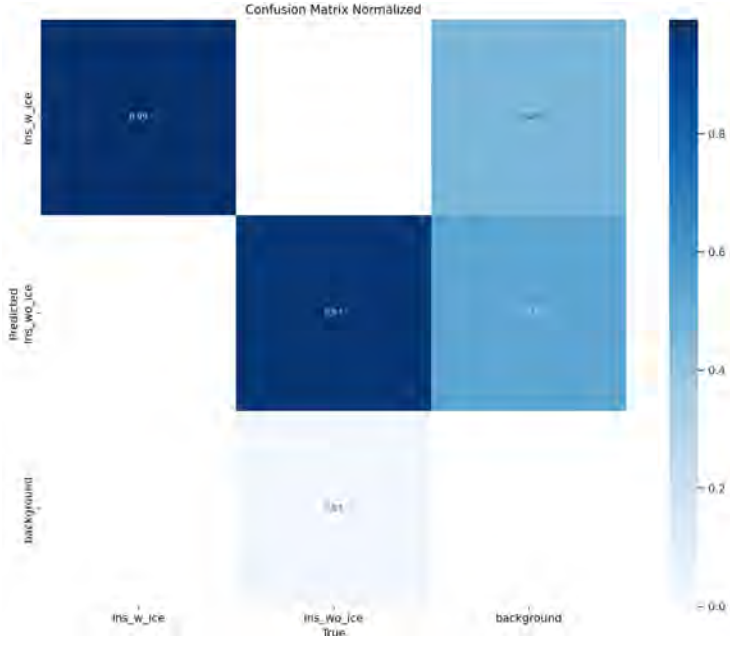


Figure 5 Confusion matrix normalised (see online version for colours)



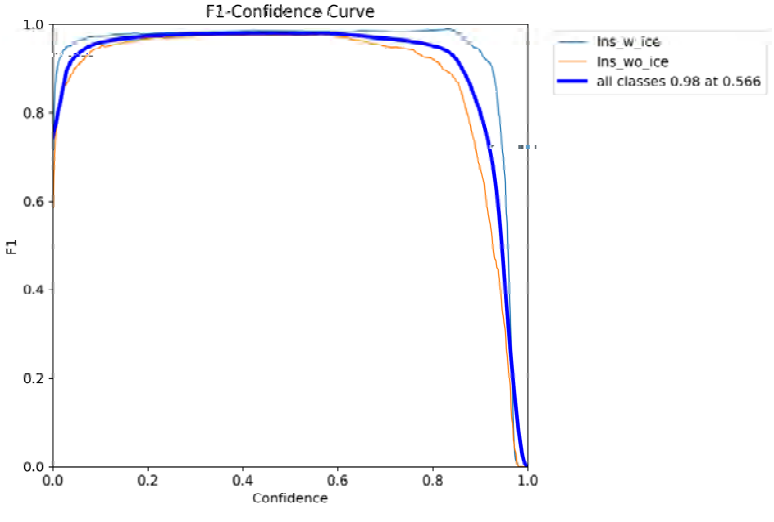
Going forward, these optimal weight parameters can be loaded into a standard YOLOv11 model, leveraging these parameters to conduct inference on new datasets. This means that even when using different datasets or scenarios, as long as these optimal weights are loaded, the YOLOv11 model should still deliver detection results comparable to those achieved during our training process.

4.3.1 Confusion matrix

A confusion matrix is a tool for assessing the performance of a classification model. It provides a more detailed evaluation of model performance than a single accuracy score by summarising the relationships between the model's predicted outcomes and the actual labels. Each row of the confusion matrix corresponds to the established categories (in this case, whether the insulators are icy), and each column corresponds to the predicted categories. The confusion matrix allows us to closely examine the model's classification performance, including both correct predictions and instances of misclassification.

Based on the results from the confusion matrix, the accuracy of predicting whether insulators are iced or not iced both exceed 97%. Although the model exhibits a certain level of misclassification when distinguishing the background, it generally demonstrates high accuracy in predicting the icing condition of insulators. For the next steps in optimising the model, consideration can be given to reducing the misidentification rate of the background to enhance its recognition capability in real-world scenarios.

Figure 6 F1-confidence curve (see online version for colours)



4.3.2 F1-score curve

The F1 curve evaluates a model's classification performance by assessing changes in precision and recall. The F1-score is the harmonic mean of precision and recall, designed to provide a balanced measure of a model's performance on these two indicators. The F1 curve shows the change trends of the F1-score at different thresholds, highlighting the model's capabilities. In particular, when handling classification issues and encountering

class imbalance, the F1 curve provides a more accurate assessment method than relying solely on accuracy.

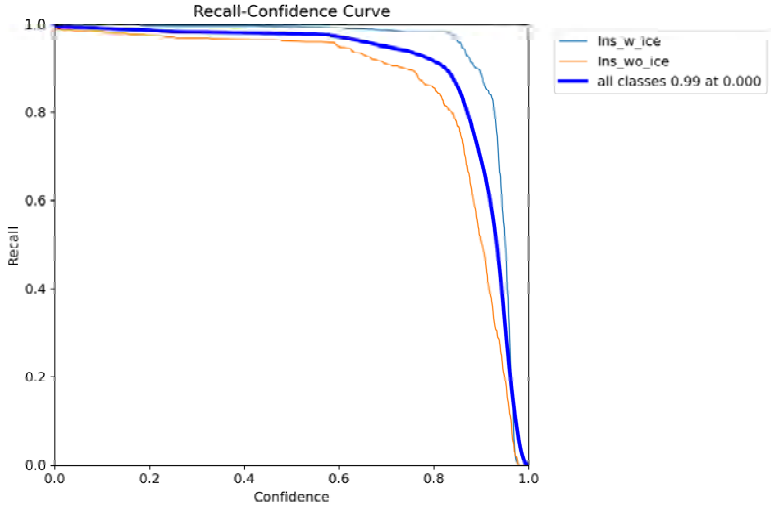
The x-axis (confidence) indicates the confidence threshold of the classification model, which is the degree of certainty the model has in its predictions. The threshold spans from 0 to 1, showcasing the model's performance at varying levels of confidence. The y-axis (F1) represents the F1-score, a metric that integrates precision and recall to evaluate the performance of a classification model. A higher F1-score signifies superior balance between precision and recall.

From the graph, it can be observed that the overall curve remains stable within the confidence range of approximately 0.1 to 0.8, with an F1-score close to 0.98, indicating that the model achieves high overall predictive performance across all categories. When the confidence exceeds 0.8, the F1-score starts to drop rapidly, possibly due to some correct predictions being missed at higher confidence levels. The highest F1-score is 0.98, corresponding to a confidence of 0.566.

4.3.3 Recall-confidence curve

- Horizontal axis (confidence): Represents the confidence threshold of the model's predictions, ranging from 0 to 1. The higher the confidence, the more certain the model is about the correctness of its predictions.
- Vertical axis (recall): Represents recall, which is the proportion of correctly identified targets out of all actual targets. Higher recall indicates a lower missed detection rate by the model.

Figure 7 Recall-confidence curve (see online version for colours)



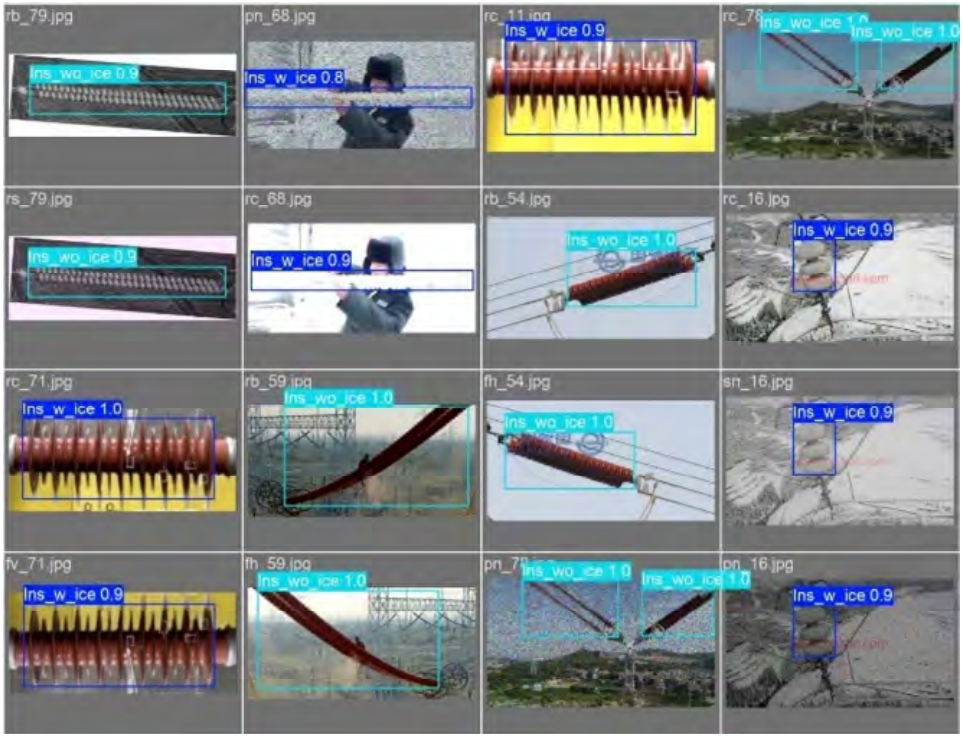
The curve shows that the recall performance of the model for the Ins_w_ice and Ins_wo_ice categories is very close, indicating that the model has a balanced detection capability for these two categories, without significant bias. At low confidence levels (close to 0), the recall is close to 1.0, suggesting that the model can almost completely detect the targets under a low threshold, but this may introduce more false positives

(affecting precision). At high confidence levels (close to 1.0), the recall drops rapidly, indicating that the model's predictions become more cautious, and some targets fail to be detected (missed detections).

4.3.4 Predict result

From the prediction result diagram figures, it can be observed that the model is capable of effectively detecting the majority of iced insulators across various scenarios, demonstrating its ability in identifying iced insulators. However, there are still cases of undetected iced insulators or misclassifications. These false detections and omissions indicate that the model requires further improvement in handling details and boundary conditions. Nevertheless, the overall detection performance of the model suggests it has the potential for practical application.

Figure 8 Predict result (samples) (see online version for colours)



To further enhance the model's accuracy and reliability, various strategies can be employed to optimise its performance. For instance, supplementing the dataset with more diverse samples can improve the model's generalisation ability, especially for cases involving weak coverage, partial occlusion, and complex backgrounds. Additionally, targeted adjustments to the model's parameters, such as optimising hyperparameter settings, improving the loss function, and enhancing data pre-processing techniques, can all contribute to achieving higher detection accuracy across different scenarios. With continuous optimisation and refinement, the model is expected to meet the standards required for practical applications.

Figure 9 Predict result (in actual substation scenarios) (see online version for colours)

5 Conclusions

This paper proposes an insulator icing detection method based on the YOLO deep learning algorithm, aiming to quickly and accurately identify whether insulators on transmission lines are iced. Insulator icing is one of the primary factors leading to transmission line failures. Traditional detection methods often suffer from low efficiency, high costs, and significant manual intervention. By introducing object detection technology, the proposed method effectively addresses these challenges.

In terms of method design, this study selects the YOLO series model as the detection framework and enhances the model's ability to recognise iced and ice-free insulators through data augmentation and feature optimisation. The model training process involves the use of a large-scale annotated dataset and employs optimisation strategies combining cross-entropy loss and IoU loss to balance detection precision and recall. Regarding the

model structure, improvements were made to the neck and head components of YOLO, introducing the C2PSA module to enhance spatial attention. This enables the model to focus more on critical features in the iced areas, thereby improving detection performance for small objects and in complex backgrounds.

Experimental results demonstrate that the proposed method achieves high levels of mAP and F1-score on the test set, indicating robust detection performance. Notably, the detection precision for the 'iced insulator' category outperforms that of the 'ice-free insulator' category.

In summary, this research indicates that the YOLOv11-based insulator icing detection method is efficient and reliable, providing technical support for intelligent inspections in power systems. Future work will focus on enhancing the model's ability to detect iced insulators, further optimising its real-time performance and robustness to meet the demands of more complex application scenarios.

Declarations

Author contributions: Conceptualisation, H.H., X.Z. and D.C.; methodology, H.H. and Y.D.; software, X.Z. and S.W.; validation, H.H., D.C. and X.L.; formal analysis, D.C., Y.D. and Q.F.; investigation, H.H., D.C. and X.L.; resources, X.Z., Y.D. and Q.F.; data curation, H.H., Y.D. and S.W.; writing – original draft preparation, H.H., X.Z. and X.L.; writing – review and editing, D.C., Y.D., S.W. and X.L.; visualisation, X.Z., Y.D., X.L. and Q.F.; supervision, H.H.; project administration, H.H. and S.W.; funding acquisition, X.Z. and Q.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Hubei Province International Scientific Research Cooperation Project (2023EHA035).

Data availability statement: The data presented in this study are available in the article.

Conflicts of interest: Authors were employed by the companies State Grid Hubei Extra High Voltage Company and Hubei Super-Energic Electric Power Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Bharati, P. and Pramanik, A. (2020) 'Deep learning techniques – R-CNN to mask R-CNN: a survey', in *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019*, pp.657–668.

Chen, L. (2024) 'Defect identification of electricity transmission line insulators based on the improved lightweight network model with computer vision assistance', *Heliyon*, Vol. 10, No. 10, pp.e30405–e30405, https://doi.org/10.1016/j.heliyon.2024.e30405.

Girshick, R. (2015) 'Fast R-CNN', in *Proceedings of the IEEE International Conference on Computer Vision*, pp.1440–1448.

Gonçalves, R.S., Agostini, G.S., Reinaldo, Homma, R.Z., Edgardo, D., Trautmann, P.V. and Clasen, B.C. (2022) 'Inspection of power line insulators', *Advances in Computational Intelligence and Robotics Book Series*, pp.462–491, https://doi.org/10.4018/978-1-7998-8686-0.ch018.

Guo, Q. and Hu, X. (2017) 'Power line icing monitoring method using binocular stereo vision', *2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, IEEE.

Han, J., Yang, Z., Zhang, Q., Chen, C., Li, H., Lai, S., Hu, G., Xu, C., Xu, H., Wang, D. and Chen, R. (2019) 'A method of insulator faults detection in aerial images for high-voltage transmission lines inspection', *Applied Sciences*, Vol. 9, No. 10, p.2009, https://doi.org/10.3390/app9102009.

Hao, Y., Liang, W., Yang, L., He, J. and Wu, J. (2022) 'Methods of image recognition of overhead power line insulators and ice types based on deep weakly-supervised and transfer learning', *IET Generation, Transmission & Distribution*, Vol. 16, No. 11, pp.2140–2153, https://doi.org/10.1049/gtd2.12428.

Irene, U.S., Berlijn, S.M. and Fahlström, A. (2009) 'Automatic surveillance and analysis of snow and ice coverage on electrical insulators of power transmission lines', *Lecture Notes in Computer Science*, pp.368–379, https://doi.org/10.1007/978-3-642-02345-3_36.

Li, W. (2021) 'Analysis of object detection performance based on faster R-CNN', in *Journal of Physics: Conference Series*, IOP Publishing, March, Vol. 1827, No. 1, p.12085.

Liu, J., Hu, M., Dong, J. and Lu, X. (2023) 'Summary of insulator defect detection based on deep learning', *Electric Power Systems Research*, Vol. 224, pp.109688–109688, https://doi.org/10.1016/j.epsr.2023.109688.

Liu, X., Miao, X., Jiang, H. and Chen, J. (2020a) 'Data analysis in visual power line inspection: an in-depth review of deep learning for component detection and fault diagnosis', *Annual Reviews in Control*, https://doi.org/10.1016/j.arcontrol.2020.09.002.

Liu, Y., Li, Q., Farzaneh, M. and Du, B.X. (2020b) 'Image characteristic extraction of ice-covered outdoor insulator for monitoring icing degree', *Energies*, Vol. 13, No. 20, pp.5305–5305, https://doi.org/10.3390/en13205305.

Ma, F., Wang, B., Li, M., Dong, X., Mao, Y., Zhou, Y. and Ma, H. (2021) 'Edge intelligent perception method for power grid icing condition based on multi-scale feature fusion target detection and model quantization', *Frontiers in Energy Research*, Vol. 9, https://doi.org/10.3389/fenrg.2021.754335.

Marek, S., Marek, H. and Otcenasova, A. (2021) 'Advanced power line diagnostics using point cloud data – possible applications and limits', *Remote Sensing*, Vol. 13, No. 10, pp.1880–1880, https://doi.org/10.3390/rs13101880.

Mikolajczyk, K. and Schmid, C. (2005) 'A performance evaluation of local descriptors', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 10, pp.1615–1630, https://doi.org/10.1109/tpami.2005.188.

Nguyen, V.N., Jenssen, R. and Roverso, D. (2018) 'Automatic autonomous vision-based power line inspection: a review of current status and the potential role of deep learning', *International Journal of Electrical Power & Energy Systems*, Vol. 99, pp.107–120, https://doi.org/10.1016/j.ijepes.2017.12.016.

Nusantika, N.R., Xiao, J. and Hu, X. (2024) 'Precision ice detection on power transmission lines: a novel approach with multi-scale retinex and advanced morphological edge detection monitoring', *Journal of Imaging*, Vol. 10, No. 11, pp.287–287, https://doi.org/10.3390/jimaging10110287.

Pernebayeva, D., Irmanova, A., Sadykova, D., Bagheri, M. and James, A. (2019) 'High voltage outdoor insulator surface condition evaluation using aerial insulator images', *High Voltage*, Vol. 4, No. 3, pp.178–185, https://doi.org/10.1049/hve.2019.0079.

Qiu, Z., Zhu, X., Liao, C., Shi, D. and Qu, W. (2022) 'Detection of transmission line insulator defects based on an improved lightweight YOLOv4 model', *Applied Sciences*, Vol. 12, No. 3, pp.1207–1207, https://doi.org/10.3390/app12031207.

Solangi, A.R. (2018) *Icing Effects on Power Lines and Anti-Icing and De-Icing Methods*, Master's thesis, UiT The Arctic University of Norway [online] https://hdl.handle.net/10037/14198. (accessed 15 October 2025).

Wang, D., Yue, J., Li, J., Xu, Z., Zhao, W. and Zhu, R. (2023a) 'Research on sag monitoring of ice-accreted transmission line arcs based on stereovision technique', *Electric Power Systems Research*, Vol. 225, p.109794, DOI: 10.1016/j.epsr.2023.109794, ISSN: 0378-7796.

Wang, Q., Fan, Z., Luan, Z. et al. (2023b) 'Insulator abnormal condition detection from small data samples', *Sensors*, Vol. 23, No. 18, p.7967.

Zhang, Q. (2022) 'Application of image processing in ice-structure interaction', *Encyclopedia of Ocean Engineering*, pp.23–42, https://doi.org/10.1007/978-981-10-6946-8_131.

Zhang, Q., Zhang, J., Li, Y., Zhu, C. and Wang, G. (2024) 'IL-YOLO: an efficient detection algorithm for insulator defects in complex backgrounds of transmission lines', *IEEE Access*, Vol. 12, pp.14532–14546, https://doi.org/10.1109/access.2024.3358205.