

Deep learning in automated power line inspection: A review

Md. Ahasan Atick Faisal, Imene Mecheter, Yazan Qiblawey, Javier Hernandez Fernandez, Muhammad E.H. Chowdhury, Serkan Kiranyaz

Item type

Journal Contribution

Terms of use

This work is licensed under a [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

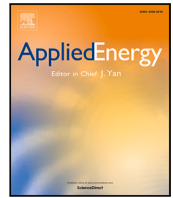
This version is available at

https://manara.qnl.qa/articles/journal_contribution/Deep_learning_in_automated_power_line_inspection_A_review/28546460/1

Access the item on Manara for more information about usage details and recommended citation.

Posted on Manara – Qatar Research Repository on

2025-02-17



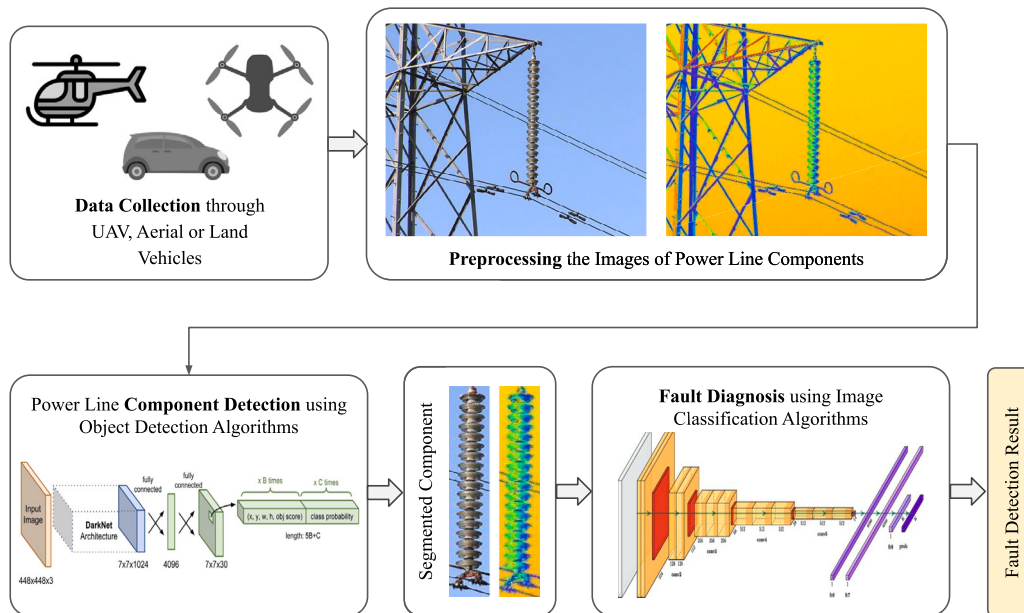
Deep learning in automated power line inspection: A review

Md. Ahasan Atick Faisal ^{a,b} , Imene Mecheter ^b , Yazan Qiblawey ^b ,
Javier Hernandez Fernandez ^b , Muhammad E.H. Chowdhury ^a , Serkan Kiranyaz ^a

^a Department of Electrical Engineering, Qatar University, Doha, 2713, Qatar

^b Iberdrola Innovation Middle East, Doha, 210177, Qatar

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Power line inspection
Fault detection
Computer vision
Deep learning

ABSTRACT

In recent years, power line maintenance has seen a paradigm shift by moving towards computer vision-powered automated inspection. The utilization of an extensive collection of videos and images has become essential for maintaining the reliability, safety, and sustainability of electricity transmission. A significant focus on applying deep learning techniques for enhancing power line inspection processes has been observed in recent research. A comprehensive review of existing studies has been conducted in this paper, to aid researchers and industries in developing improved deep learning-based systems for analyzing power line data. The conventional steps of data analysis in power line inspections have been examined, and the body of current research has been systematically categorized into two main areas: the detection of components and the diagnosis of faults. A detailed summary of the diverse methods and techniques employed in these areas has been encapsulated, providing insights into their functionality and use cases. Special attention has been given to the exploration of deep learning-based methodologies for the analysis of power line inspection data, with an exposition of

* Corresponding author at: Department of Electrical Engineering, Qatar University, Doha, 2713, Qatar.

E-mail address: atick.faisal@qu.edu.qa (M.A.A. Faisal).

<https://doi.org/10.1016/j.apenergy.2025.125507>

Received 20 May 2024; Received in revised form 8 January 2025; Accepted 7 February 2025

Available online 17 February 2025

0306-2619/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

their fundamental principles and practical applications. Moreover, a vision for future research directions has been outlined, highlighting the need for advancements such as edge-cloud collaboration, and multi-modal analysis among others. Thus, this paper serves as a comprehensive resource for researchers delving into deep learning for power line analysis, illuminating the extent of current knowledge and the potential areas for future investigation.

1. Introduction

The late 19th century saw Edison's light bulb and Tesla and Westinghouse's alternating current (AC) systems usher in the electrical age. This era not only illuminated the world but also laid the foundation for the modern power delivery system, which has become a complex network of power plants, transmission lines, and distribution networks. For many decades, this complex network relied on manual inspections, which were dangerous and limited. Technological advances improved the methods, but the recent rise of computer vision and deep learning has revolutionized how power line inspections are conducted.

A power line comprises a multitude of components, each with distinct functions, including insulators, towers, conductors, and fittings. Operating in a challenging outdoor environment, exposed to complex landforms and unpredictable weather, power line components are susceptible to frequent damage. A single faulty component, such as a conductor, or a combination of multiple damaged components, such as fittings, can trigger power outages with far-reaching consequences. These disruptions not only disrupt regional electricity supply but can also escalate to supra-regional blackouts and even catastrophic incidents, such as forest fires [1]. In California, about 10% of the state's wildfires are believed to be triggered by power lines. The severity of these fires led the California Public Utilities Commission to investigate Pacific Gas & Electric (PG&E) power line safety practices, considering drastic measures such as breaking up the utility into smaller entities for better management and accountability [2]. The resulting economic and societal costs due to power line failures can be substantial [3]. Effective power line inspection serves as the vanguard against such calamities. Its primary objective is to assess the condition of the power line components, enabling informed decisions on maintenance and replacement. A swift and accurate inspection process significantly enhances the efficiency of maintenance decision-making and, in turn, reduces the likelihood of power line failures, safeguarding the safety and reliability of the power supply to the connected load [4].

Nonetheless, power line inspection encounters a series of challenges, ranging from covering vast geographic areas to dealing with a diverse array of components and navigating complex natural environments. For decades, traditional inspection methods have relied on manual ground surveys and helicopter-assisted patrols [5]. These methods heavily depend on visual observations from humans, which involve significant costs, inherent risks, low operational efficiency, and long timeframes [6]. In recent years, the application of computer vision and deep learning technologies has helped in a transformative era for power line inspection [7]. These advanced techniques have effectively decoupled the traditional inspection process into two distinct phases: data collection and data analysis. Operators can now leverage computer vision and deep learning to automatically process images and videos captured by Unmanned Aerial Vehicles (UAVs) or other means. This transition from manual labor-intensive methods to automated inspection is driven by compelling advantages, including cost-efficiency, enhanced safety, and superior operational efficiency [7].

Today, the maintenance of power lines is being transformed by the integration of computer vision and deep learning algorithms. These technologies enable the automation of inspection processes, offering a safer, more efficient, and cost-effective method of identifying potential issues before they escalate into major failures. By leveraging high-resolution images and real-time data analysis, utility companies can now predict maintenance needs, prevent outages, and ensure the reliable delivery of electricity to consumers. However, this transition has

introduced a deluge of data. Furthermore, the conventional approach for analyzing these images and videos involves time-consuming manual efforts, which are not only costly but also fraught with potential safety hazards and may lack the necessary precision [8]. Consequently, there is an urgent demand for the development of automated methodologies to replace manual analysis [9]. Over the past years, several attempts have diligently striven to devise rapid and accurate methods for the automatic evaluation of power lines from aerial imagery or land [4]. Those attempts tried an extensive spectrum of power line components and their associated faults, primarily leveraging image processing and computer vision. Although image processing-based approaches like color [10], shape [11], or texture segmentation [12] have seen some success over the years, they are gradually being replaced by more advanced deep learning-based approaches. Moreover, a substantial portion of these endeavors are task-specific, concentrating on isolated components or specific fault types.

This review paper explores the recent works on vision-based power line inspection, casting the gaze through the lens of deep learning. The paper starts with an introduction to the foundational concepts in power line inspection, encompassing various inspection methods and data sources. Subsequently, this paper provides a brief introduction to the deep learning-based techniques applied to power line inspection. Moving forward, an extensive review of research endeavors focused on the analysis of images in power line inspection has been explored. The literature has been organized into two categories: component detection and fault diagnosis. Component detection research focuses on locating and identifying power line elements, either as a standalone task or as a crucial first step in fault analysis. Fault diagnosis studies, on the other hand, encompass both direct fault detection approaches and methods that build upon component detection to identify specific defects. This systematic division allows us to examine the unique challenges and solutions in each domain while highlighting their interconnected nature. The paper unveils the key features of each analytical approach, explores the nature of the datasets employed, and showcases representative quality analysis results, offering insight into the diverse capabilities of these methods across various applications. Finally, the paper presents a set of open research questions and uncharted territories awaiting future exploration. These questions encompass challenges related to data quality, the intricacies of small object detection, the application of deep learning in embedded systems, and the definition of robust evaluation baselines. Finally, this paper synthesizes the key findings and insights derived from this comprehensive analysis.

The structure of the rest of this paper is outlined as follows: Section 2 explores similar works and compares this review with the existing ones. Section 3 introduces the key stages of automated power line inspection. Section 4 explores various image acquisition methods and the vehicles employed for collecting power line data. In Section 5, we examine a range of imaging techniques, discussing their advantages and limitations. Section 6 presents a concise overview of available computer vision datasets pertinent to power line inspection. Section 7 highlights the leading deep learning models and architectures employed in computer vision for tasks such as object detection and classification. Following this, Section 8 offers an in-depth review of research papers focused on power line component detection. Section 9 scrutinizes literature dedicated to identifying power line faults through computer vision techniques. Section 10 provides a qualitative assessment of the reviewed articles and Section 11 addresses the prevalent challenges within this field and proposes potential areas for future research. The paper concludes with Section 12, summarizing the findings of this comprehensive review.

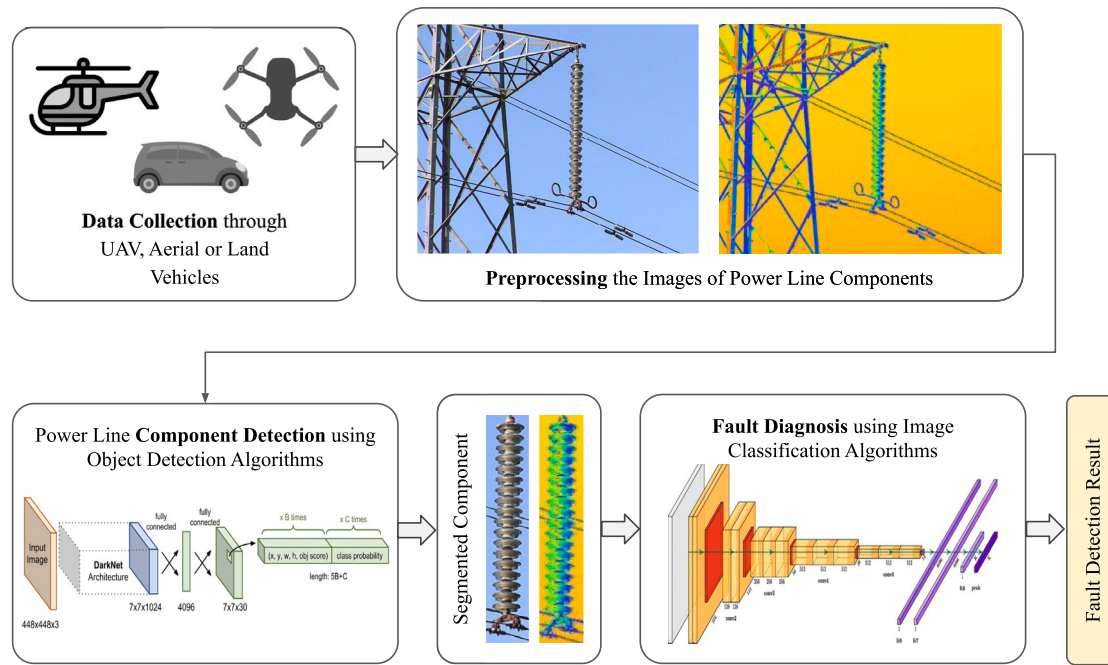


Fig. 1. Block Diagram of an automated multi-modal power line inspection system.

2. Related works

The application of computer vision and deep learning for power line inspection has garnered increasing attention in recent years, reflected in a growing body of survey papers. Chen et al. [13] and Yang et al. [14] offer general overviews of automated power line inspection techniques, including 3D reconstruction, object detection, and inspection platforms. Their work is primarily centered on the technology aspects, such as LiDAR-based reconstruction and the classification of detection techniques. While this technical perspective is valuable, the review lacks an in-depth exploration of deep learning's impact on these technologies. Sundaram et al. [15] provide an overall perspective on deep learning applications in the electrical domain, touching upon power line inspection alongside other areas. However, their review is broader in scope, covering multiple electrical applications, which dilutes the focus on power line inspection. Ruszczak et al. [16] focus specifically on the importance of specialized datasets for training deep learning models in power line inspection tasks.

Several reviews delve into the use of unmanned aerial vehicles (UAVs) for inspection. Xu et al. [17] provide a systematic summary of UAV platforms and image recognition techniques, while Foudeh et al. [18] concentrate on UAV technologies and control strategies. Nguyen et al. [4,19] explored UAV-based power line inspection techniques in the light of deep learning and proposed a concept for an autonomous UAV-based inspection system and discuss its challenges and possibilities. The above mentioned reviews are notable for their focus on UAV-based systems, but it largely centers on the hardware and system integration aspects. Although they mention deep learning, it is not the primary focus of their analysis. Finally, Liu et al. [9] present a comprehensive review of data analysis techniques, including deep learning methods, for power line inspection. Although the authors cover deep learning, the topic of the review is much broader due to the inclusion of image processing techniques and non machine learning-based approaches.

While these existing reviews offer valuable insights, this current review distinguishes itself through several key contributions. First, it provides a comprehensive and up-to-date analysis of deep learning applications specifically for power line inspection, encompassing both component detection and fault diagnosis. This focused exploration

of deep learning techniques sets it apart from previous surveys that either covered a broader range of inspection methods or touched upon power line inspection as part of a larger survey. Second, the review systematically categorizes existing research into component detection and fault diagnosis, providing a structured understanding of the field. It summarizes various methods and techniques, offering insights into their functionality and use cases. This systematic approach facilitates easier navigation and comprehension of the current state-of-the-art. Furthermore, the review places a particular emphasis on deep learning methodologies, including their fundamental principles and practical applications in power line inspection. Finally, the review outlines future research directions, highlighting areas like data quality improvement and small object detection that require further exploration. It serves as a roadmap for future research endeavors, guiding researchers towards potential breakthroughs.

In summary, this review paper contributes significantly to the field by offering a comprehensive, systematic, and deep learning-focused analysis of automated power line inspection. It builds upon existing knowledge, provides a structured understanding of current research, and charts a course for future advancements.

3. Methodology for automated power line inspection

Automated power line inspection leverages the capabilities of computer vision and deep learning to enhance the safety, reliability, and efficiency of power transmission systems. Fig. 1 shows the overview of a typical computer vision-based power line inspection process. It starts by capturing images of power line components using aerial or land vehicles sometimes using multiple imaging techniques. The images are then processed through the object detection or segmentation algorithm. Subsequently, the image classification algorithm evaluates the segmented insulators to classify them as 'good' or 'faulty', resulting in the final fault detection output. This section provides an overview of the key steps involved in this process, from data collection to fault diagnosis. It should be noted that in many cases in the literature, the component detection part has been skipped and the deep learning model has been trained for the fault detection purpose directly.

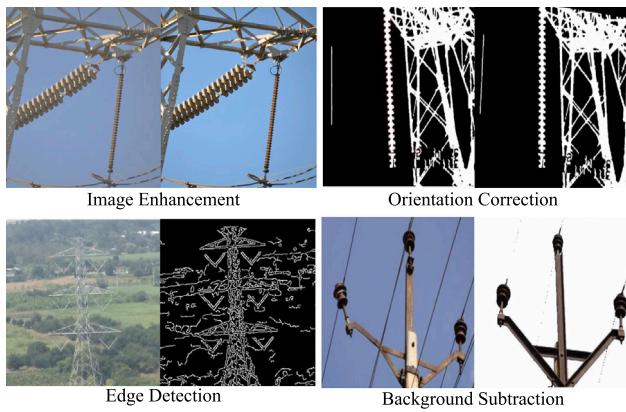


Fig. 2. Example of different image enhancement techniques used for power line inspection [9].

3.1. Data collection

The basis of automated power line inspection is high-resolution images usually captured from land or aerial vehicles. The collected images provide valuable visual data that forms the basis for subsequent analysis. The data collection process for power line inspection has been discussed in more detail in Section 4.

3.2. Preprocessing

The raw images obtained during data collection often require preprocessing to improve their suitability for analysis. Preprocessing steps may include noise reduction, image enhancement, and geometric correction to account for variations in lighting conditions and perspective distortions. These enhancements ensure that the subsequent computer vision algorithms can work effectively. Fig. 2 shows visual examples of different image processing techniques such as image enhancement, edge detection, orientation correction, and color-based background subtraction.

3.3. Component detection

Component detection plays a central role in identifying and locating power line components within the power line images. Deep learning models, such as Faster R-CNN [20], (You Only Look Once) YOLO [21], or (Single Shot Multibox Detector) SSD [22], are employed to detect a wide range of components, including insulators, suspension clamps, dampers, and conductors. These models can distinguish these components from the complex backgrounds typically encountered. Section 8 provides a comprehensive overview of the literature focusing on power line component detection.

3.4. Fault diagnosis

The final phase in automated power line inspection involves fault diagnosis, where deep learning models analyze detected components or direct image inputs to identify potential issues. Deep learning models are employed to analyze the extent and severity of faults, enabling operators to prioritize maintenance and repair efforts. These models can provide valuable insights into the overall health of the power transmission system, aiding in the prevention of outages and accidents. Section 9 provides a comprehensive overview of the literature focusing on power line fault detection.

Automated power line inspection harnesses the capabilities of computer vision and deep learning to streamline the inspection and maintenance of power transmission systems. By automating the detection and diagnosis of components and faults, these systems contribute to safer, more reliable power delivery while minimizing downtime and operational costs.

4. Image acquisition platforms and vehicles

The success of computer vision applications in the domain of transmission line maintenance largely depends on the quality and diversity of the images acquired. The choice of image acquisition techniques plays a vital role in determining the cost and labor associated with the operation. This section explores various methods employed for power line image acquisition.

4.1. Unmanned aerial vehicles

Unmanned Aerial Vehicles (UAVs), commonly known as drones, have emerged as a popular choice for image acquisition in power line inspection. UAVs equipped with high-resolution cameras can capture images and videos of transmission lines and their components from various angles and distances. They offer a unique combination of accessibility, maneuverability, and safety that is particularly well-suited to the challenges posed by transmission line maintenance. This versatility allows for comprehensive visual data collection, enabling the detection of defects and anomalies in power lines with greater accuracy [23].

4.2. Aerial vehicles

Aerial vehicles such as helicopters have been utilized for decades in power line inspection. Equipped with specialized cameras and imaging systems, they offer the advantage of covering long distances and providing a stable platform for capturing images of transmission lines. They are particularly suitable for inspecting high-voltage lines, where safety considerations are paramount [7].

4.3. Land vehicles

In situations where aerial inspection may not be feasible or cost-effective, land vehicles equipped with imaging equipment are employed. These vehicles can navigate the terrain near transmission lines, capturing images and videos of the components and their surroundings. They are especially useful for inspecting power lines in areas with limited airspace accessibility.

4.4. Fixed camera

Fixed cameras installed at strategic locations along power lines provide continuous monitoring. These cameras capture images at pre-defined intervals or when triggered by certain events. They offer a cost-effective solution for routine inspections and surveillance, although they may have limitations in terms of coverage and flexibility compared to aerial methods.

4.5. Satellite imaging

Satellite imaging technology is increasingly being explored for large-scale monitoring of transmission lines. While the resolution may not be as high as that of UAVs or helicopters, satellite imagery can provide valuable data for identifying overall trends and assessing the condition of transmission networks over vast geographic areas [24].

Table 1
Comparison of different image acquisition techniques.

Platform	Cost	Tracking ^a	Accuracy	Efficiency	Coverage	Safety	Usage
UAV	Low	Difficult	Good	Fast	Good	Safe	General purpose
Land vehicles	Low	Easy	Good	Slow	Limited	Safe	Road-side lines
Aerial vehicles	High	Easy	Bad	Fast	Good	Unsafe	Hard-to-reach areas
Fixed camera	High	Easy	Good	NA	Limited	Safe	For towers
Satellite	High	Difficult	Bad	Slow	Good	NA	Hard-to-reach areas
Inspection robots	High	Easy	Good	Slow	Limited	Safe	Detailed inspection

^a Tracking refers to following and keeping track of the power line.

4.6. Power line inspection robots

Power line inspection robots have emerged as a promising technology for automating the inspection and maintenance of transmission lines. These robots can be classified into climbing robots, flying robots (UAVs), and hybrid climbing-flying robots [25]. Climbing robots can roll along the power lines and provide high-quality inspection data, but they face challenges in obstacle avoidance and deployment onto the lines. Flying robots offer faster inspection and easier obstacle avoidance but may have limitations in terms of inspection quality. Hybrid robots aim to combine the advantages of both climbing and flying robots for more effective inspection. Research in this area focuses on improving the robots' mechanical design, power systems, control algorithms, and sensing capabilities for autonomous operation [13]. For a comprehensive review of power line inspection robots, readers are referred to [13,25,26].

Selecting the most suitable image acquisition method is crucial for obtaining high-quality data. The choice often depends on factors like cost, terrain, and the specific goals of the inspection. The choice of image acquisition method may also be influenced by weather conditions. Adverse weather, such as rain, snow, or fog, can impact the performance of aerial methods. Ground-based and fixed-camera systems may be preferred in such scenarios for their resilience to adverse weather conditions. Table 1 shows a general comparison between these techniques focusing on several different aspects such as cost, accuracy, coverage, and safety.

5. Imaging techniques

The choice of imaging technique is critical for obtaining the most relevant and accurate data. Various imaging modalities are employed, each offering unique advantages and limitations. In this section, four primary image acquisition techniques: visible light images, infrared images, UV images, and x-ray images have been explored. Furthermore, LiDAR-based geospatial mapping has also been discussed briefly.

5.1. Visible light imaging

Visible light imaging is the most widely used method and is readily available through standard cameras. It is a cost-effective choice for routine inspections. Visible light images capture fine details of power line components and surrounding vegetation and other objects, which is crucial for identifying defects and assessing structural integrity. However, visibility can be compromised in adverse weather conditions, such as fog, rain, or darkness [27,28]. While visible light images provide valuable surface-level data, they may not detect defects hidden within materials or components.

5.2. Infrared imaging (Thermography)

Infrared imaging, or thermography, detects temperature variations in power line components, highlighting issues like overheating, loose connections, and faulty insulators [29]. Faulty electrical components

often result from internal electrical defects which can lead to over-current flow and heating of the component which can be detected from its surface temperature. Infrared imaging is not reliant on visible light and can operate effectively day and night. However, Infrared imaging requires specialized cameras that can be costly and may require trained personnel. It primarily provides information about surface temperature, and its ability to penetrate materials is limited [30].

5.3. Ultraviolet (UV) imaging

UV imaging is particularly useful for detecting corona discharges — electrical discharges caused by the ionization of air surrounding high-voltage conductors — which can indicate electrical faults [31]. UV imaging is non-destructive and can reveal hidden faults without physical contact with the power lines [32]. However, UV imaging has a limited range compared to visible light or infrared imaging, which means it may not capture an entire transmission line in a single image [33]. Also, specialized UV cameras are required for this technique.

5.4. X-ray imaging

X-ray imaging can penetrate materials, providing detailed images of inner structures and components. It is effective at identifying hidden defects, such as corrosion and internal damage [34]. The use of X-ray imaging involves exposure to ionizing radiation, which requires adherence to safety protocols and limits its routine use. Also, X-ray imaging equipment is costly and requires skilled operators to acquire.

5.5. LiDAR imaging

LiDAR (Light Detection and Ranging) imaging is a cutting-edge technique increasingly used in power line inspection. It employs laser light to create high-resolution, three-dimensional representations of power line infrastructure and surrounding environments. LiDAR sensors emit pulses of laser light and measure the time taken for each pulse to bounce back after hitting an object. This data is then used to construct detailed 3D models of the power lines and their immediate surroundings. LiDAR provides precise, three-dimensional information, enabling accurate mapping of power line components and detection of even minor structural anomalies. Unlike visible light imaging, LiDAR can penetrate through mild fog, rain, and other atmospheric conditions, offering more consistent results in diverse weather scenarios. It is particularly effective in assessing vegetation encroachment and potential physical obstructions near power lines, which are crucial for maintaining line safety and preventing outages [35,36].

The choice of image acquisition technique in power line inspection depends on specific inspection goals, budget, and the expected challenges. Often, a combination of imaging modalities may be used to maximize the coverage of inspection. While visible light images provide a foundational dataset for routine inspections, advanced techniques like infrared, UV, and X-ray imaging offer deeper insights into the condition of power lines. When combined with computer vision algorithms, these techniques can significantly enhance the accuracy and efficiency of fault detection and maintenance decisions [32]. Table 2 shows a comparison between these techniques in terms of cost, lighting condition, and coverage.

Table 2
Comparison between different imaging techniques.

Modality	Cost	Lighting	Coverage	Usage
Visible	Low	Requires light	High	Structural defect, foreign object detection
IR	Moderate	Not needed	High	Internal electrical defect
UV	Moderate	Not needed	Low	Corona discharge
X-ray	High	Not needed	Low	Internal structural defect
LiDAR	High	Not needed	High	3D mapping, terrain analysis, vegetation encroachment

Table 3
Summary of some publicly available power line image datasets.

Name	Year	Type	Component	Task	No. of images	Ref	Download
Chinese Power Line Insulator Dataset (CPLID)	2018	RGB	Insulator	Segmentation, Classification	848	[37]	Link
Powerline dataset (Infrared-IR and Visible Light-VL)	2019	RGB, IR	Conductor, No conductor	Detection	8000	[40]	Link
Transmission Towers and Power Lines (TTPLA)	2020	RGB	Insulator, Tower, Conductor	Segmentation	1234	[41]	Link
Power transmission line dataset	2021	RGB	Conductor	Classification	1044	[42]	Link
(Recognizance - 2) Power lines detection	2021	RGB	Conductor	Classification	16 078	[43]	Link
STN PLAD: A dataset for multi-size Power Line Assets Detection in high-resolution UAV images	2021	RGB	Tower, Insulator, Damper	Segmentation	2409	[44]	Link
Aerial power infrastructure detection dataset	2023	RGB	Tower	Segmentation	3956	[45]	Link
RSIn-dataset: An UAV-based insulator detection aerial images dataset and benchmark	2023	RGB	Insulator	Segmentation	1887	[46]	Link

6. Publicly available datasets

Developing deep learning tools for checking power lines automatically relies on having detailed datasets with images of power line parts. These images need to be gathered and labeled by experts, which can be tricky and expensive. It often involves using UAVs or helicopters. Labeling these images is a detailed task, especially since the parts are small and spotting problems need an expert's eye. Therefore, making even a small dataset can take a lot of time and money. However, these deep-learning models need a lot of data to learn well and make accurate predictions. To deal with the lack of enough real images, some researchers have tried making artificial images by changing the background or surroundings of power line parts using computer programs [37]. The big issue is that there aren't many large datasets available for everyone to use. The datasets that are created by utility companies are not available to the public due to privacy and data protection laws and regulations. Several companies around the world are obligated to follow rules and regulations regarding data protection such General Data Protection Regulation (GDPR) [38] in Europe or California Consumer Protection Act (CCPA) [39] in California. Those laws limit the sharing of data from Utility companies. But there are a few datasets that can be found online in Table 3. For more information on this topic, readers are referred to a study by Ruszczak et al. [16] containing a comprehensive review of the power line datasets.

7. Deep learning architectures and detection paradigms

Deep learning models have significantly advanced the field of power line inspection, offering unique advantages in terms of speed, accuracy, and the ability to handle complex scenarios [47]. Models like YOLO [21,48–51], R-CNN family [20,52,53], SSD [22], and transformer architectures [54–56] have shown remarkable performance in

detecting and classifying various power line components, including insulators, dampers, pin bolts, and conductor wires [19,57–66]. These models are increasingly being deployed in hierarchical detection systems, where lightweight variants perform initial coarse screening while more sophisticated versions handle refined secondary recognition [60]. YOLO stands out for its real-time capabilities and high frame rate [67], while region-based CNNs excel in precisely localizing objects within images [68]. SSD offers a balance between speed and accuracy [69], making it particularly suitable for edge deployment, and transformer architectures, such as ViT [54], Swin Transformers [55], and DETRs [56], have demonstrated their effectiveness in capturing global context and handling complex scenes [64–66,70].

Classification algorithms, particularly those pretrained on large datasets like ImageNet [71], have also been employed for identifying faults and anomalies in power line images. ResNet [72], VGG [73], MobileNet [74], and EfficientNet [75] have shown promising results in classifying power line components as either faulty or in good condition [60,67,76–81]. The attention mechanism [82] has also gained widespread popularity in recent years, enhancing the accuracy and efficiency of object detection tasks [76,83,84].

Various computer vision tasks, such as bounding box detection, semantic segmentation, and instance segmentation have been utilized in automating the inspection of power line components. Bounding box detection is particularly useful for identifying larger components like towers, insulators, and dampers [85]. Semantic segmentation provides detailed component-wise masks [86,87], while instance segmentation excels in scenarios where components are close together or overlapping [86,87].

For a more detailed discussion on these deep learning models and computer vision tasks, please refer to Appendix A and Appendix B, respectively.

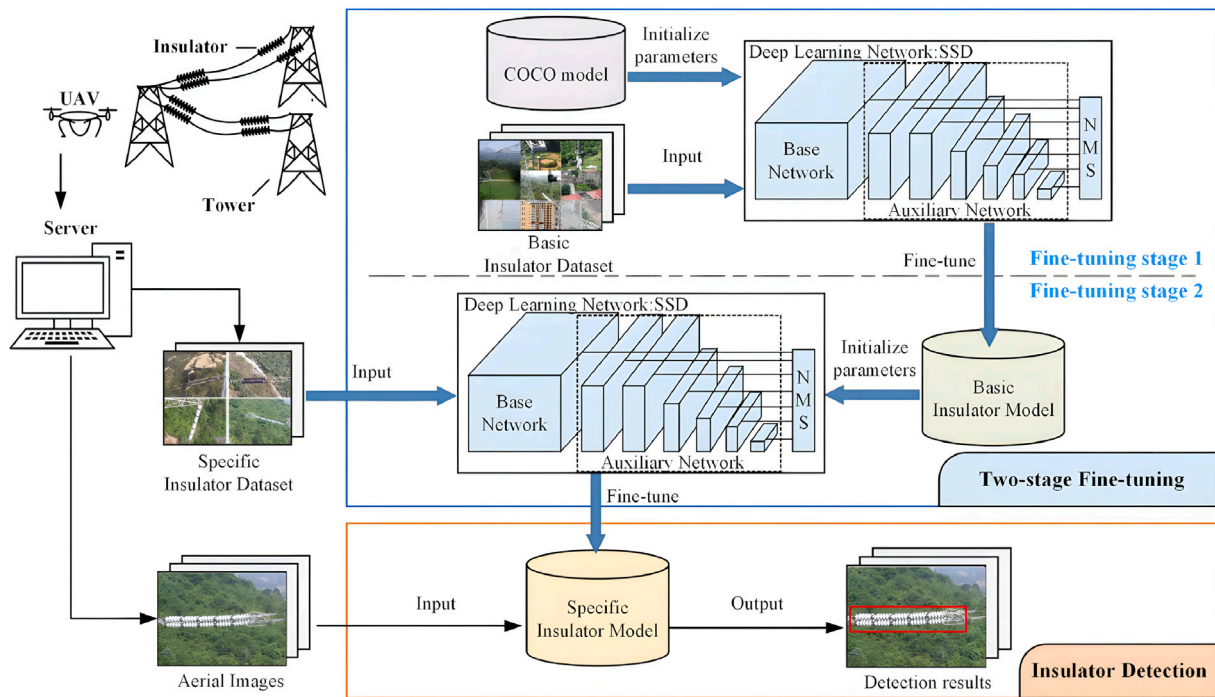


Fig. 3. Overall process of the two stage insulator detection using the SSD network [63].

8. Detection of power line components

In recent years, advancements in computer vision and deep learning have revolutionized the field of power line inspection and maintenance. Power transmission lines, a critical part of our modern infrastructure, consist of various components such as insulators, conductors, fittings, and more. The accurate and efficient detection of these components serves two crucial purposes in power line inspection: first, as a standalone task for inventory management and infrastructure mapping, and second, as a prerequisite step for fault diagnosis, where detected components are further analyzed for potential defects. This section focuses specifically on component detection methodologies, covering research works that either concentrate solely on locating and identifying power line components, or present novel component detection techniques that later serve as foundations for fault diagnosis systems. Section 9 on fault diagnosis will then examine approaches for detecting various types of faults, both in scenarios where component detection is a preliminary step and in methods that analyze faults directly from full images.

8.1. Insulator detection

Among the components of transmission lines, the insulator detection technologies are the most well-researched in the literature. These technologies offer efficient and accurate solutions to identify insulators among the other components. In a 2019 work, Miao et al. [63] proposed an efficient insulator detection method for aerial images, addressing the challenge of cluttered backgrounds. The approach utilizes the SSD network and employs a two-stage fine-tuning strategy to improve accuracy (Fig. 3). In the first stage, a basic insulator model is fine-tuned with a diverse dataset, while the second stage fine-tunes the model for specific insulator types and scenarios. This approach demonstrates that pretraining on a generalized insulator dataset can improve the model's performance when using a smaller and more specific type of insulator dataset.

The YOLO and Faster R-CNN algorithms for object detection have been widely used for insulator detection tasks. Sadykova et al. [57] proposed a cost-effective solution for automatically detecting high voltage insulators from aerial images, particularly in scenarios with uncluttered

backgrounds, varying object resolutions, and lighting conditions. The approach utilizes the YOLO network and includes data augmentation to prevent overfitting, leveraging a training dataset of 56,000 image samples after augmentation. Experimental results demonstrate the accuracy of this method in locating insulators in real-time UAV-based image data. Kang et al. [88] proposed a two-stage deep learning system for the detection of defects on the surface of high-speed railway insulators. In the first stage, a Faster R-CNN was employed to locate the insulators within images captured by the inspection vehicle. The second stage involved a novel Deep Multitask Neural Network (DMNN) that combined a Deep Material Classifier (DMC) for segmenting the insulator from the background and a Deep Denoising Autoencoder (DDAE) for detecting anomalies (defects) on the insulator surface. The DMNN was trained in a multitask learning framework, allowing the DMC and DDAE to benefit from shared convolutional layers. The authors noted that the system might struggle to detect defects that are small or have a gray value similar to the normal insulator surface.

Any kind of occlusion such as fog, smoke, etc. can be challenging for computer-vision-based insulator detection. A 2022 work by Zhang et al. [28] introduced a dataset designed for insulators and presented a benchmark model called Foggy Insulator Network (FINet), which builds upon the improved YOLOv5 framework. To enhance the dataset, the research develops and optimizes a synthetic fog algorithm, resulting in the creation of a Synthetic Foggy Insulator Dataset (SFID) containing 13,000 images. Furthermore, the YOLOv5 network is enhanced with a channel attention mechanism to form SE-YOLOv5. However, the synthetic fog generation algorithm, while effective, might not fully capture the complexity and diversity of real-world fog conditions. The evaluation of the FINet model was primarily focused on the SFID dataset, which might limit the generalizability of the results to other datasets or real-world scenarios. Table 4 summarizes the relevant literature on insulator detection.

8.2. Detection of power line fittings

Detecting power line fittings, including components like pin bolts, dampers, and suspension clamps, presents a unique set of challenges. These fittings are relatively small in size compared to the overall

Table 4
Summary of insulator detection studies in power line inspection.

Year & Ref	Component	Type of detection	Imaging platform	Dataset	Algorithm	Performance
2019 [63]	Porcelain insulator, Composite insulator	Bounding box detection	UAV	7605 RGB images	SSD	Porcelain insulator: 90.51–94.12% Composite insulator: 86.70–87.29%
2019 [57]	Glass insulator	Bounding box detection	UAV	3500 RGB images	YOLOv2	Detection accuracy: 88% Prediction time: 0.04 s
2019 [88]	Insulator	Bounding box detection	High-speed railway	18 000 RGB images	F-RCNN	mAP: 99.8%
2020 [89]	Insulator	Object presence detection	NA	4780 insulator and 13 012 background IR images	DCNN & VLAD Coding	Detection accuracy: 99.21%
2021 [90]	Insulator	Semantic segmentation	High-speed railway	800 RGB images	CDSNets	IOU: 0.94
2021 [91]	Ceramic insulator	Bounding box detection	UAV	2973 RGB images	R-CNN, SSD	Onshore mAP: 0.56–0.77; Onboard mAP: 0.24–0.27
2022 [60]	Insulator	Bounding box detection	NA	8500 insulator images	SSD, F-RCNN	Accuracy: SSD: 89%, F-RCNN: 91.6%
2022 [28]	Insulator	Bounding box detection	Synthetic	13 000 Foggy RGB images	Improved YOLOv5	mAP@0.5:0.95: 88.3%, F1 Score: 96.2%
2022 [92]	Insulator	Instance segmentation	UAV	1523 RGB images	Mask R-CNN	mAP: 87.0%
2023 [24]	Insulator	Semantic segmentation	Satellite	9900 RGB images	HRNet and OHem	F1 Score: 0.7952
2023 [93]	Insulator	Bounding box detection	UAV	1887 RGB images	YOLOv4++	mAP: 94.24%
2023 [58]	Insulator	Bounding box detection	UAV & Synthetic	CPLID dataset: 848 RGB images	YOLOv8	mAP@[0.5:0.95]: 91.25%

transmission line structure, making their detection and classification a particularly intricate task [77]. This section explores the efforts and methodologies aimed at the automatic detection of these critical power line fittings, focusing on the utilization of deep learning and computer vision techniques to address the intricacies associated with their small-scale identification. Table 5 summarizes the literature on power line fittings detection.

As mentioned earlier, one of the biggest challenges to overcome in power line fittings like bolts detection is their relatively small size compared to the other components. To tackle this, Luo et al. [77] propose a novel model (UBDDM) based on deep CNN. This model includes two modules: the ultrasmall object perception module (UOPM) for initial bolt region identification and the local bolt detection module (LBDM) for pinpointing defects within high-resolution image blocks. The authors enhance feature extraction using ResNet-50, a hybrid attention mechanism, and multiscale feature fusion. Their method simplifies data labeling requirements while maintaining end-to-end detection capabilities. While experiments demonstrate the model's superior performance in detecting bolt defects within inspection images, the model can only draw bounding boxes of some predefined dimensions.

Transmission line fittings encapsulate a wide range of components and often require a multi-task object detection model to detect all the different types. In 2022 a study done by Zhai et al. [94] proposed the Cascade Reasoning Graph Network (CRGN) which offers a novel solution to the intricate challenges posed by detecting multiple power line fittings in aerial images. CRGN uses spatial knowledge representations that capture the interrelationships among objects based on the unique characteristics of transmission line fittings. However, the model's dependency on high-resolution, close-up photographs of the fittings presents a practical challenge, as such image quality may be difficult to consistently achieve in real-world scenarios.

8.3. Conductor detection

Conductor detection using computer vision must contend with multiple complexities, including the conductor's slender profile, varying backgrounds, and potential occlusions. This subsection explores the state-of-the-art techniques and advancements in conductor detection using computer vision and deep learning, addressing the unique challenges associated with this critical task. Table 6 summarizes the literature on conductor detection.

Power lines are often a small portion of the image, leading to class imbalance issues. To address these challenges, Yang et al. [97] proposed a novel vision-based power line detection network designed to address challenges in detecting power transmission lines in complex aerial images including varying background environments, illumination conditions, and foreground-background class imbalance, where power lines occupy a small portion of the image. The proposed network utilizes an encoder-decoder architecture (similar to UNet) to create an end-to-end power line detection system. To improve segmentation precision, it incorporates an attention block to capture global contexts and emphasize target power line regions. Additionally, an attention fusion block is introduced to enhance multi-scale feature fusion and capture richer information, mitigating issues related to local contextual feature processing and information loss caused by multiple pooling operations. However, due to the complexity of the UNet-based network combined with the attention mechanism, the proposed model is not particularly suitable for real-time applications.

While deep learning, especially U-Net and its variants, has advanced pixel-level object segmentation, limitations in processing local contextual features and information loss in deep CNNs persist. To overcome these issues, in a follow-up study, Yang et al. [98] proposed a novel dual-branch residual attention network called DRA-Net. It features a dual-branch encoder with a residual CNN (RCNN) branch and a recurrent RCNN (RRCNN) branch to capture richer semantic information.

Table 5

Summary of power line fittings detection studies.

Year & Ref	Component	Type of detection	Imaging platform	Dataset	Algorithm	Performance
2021 [61]	Fittings	Bounding box detection	UAV	1455 RGB images	HK R-CNN	mAP: 59.82%–98.27%
2022 [95]	Dampers and suspension clamps	Bounding box detection	UAV	1209 RGB images	AGMNet	mAP: 95.3%
2022 [94]	Fittings	Bounding box detection	Aerial vehicle	1455 Aerial images	CRGN	mAP@0.5–0.95: 47.5%
2023 [96]	Damper	Semantic segmentation	UAV	240 RGB images	Improved GrabCut	F1 Score: 89.1–97.3%
2023 [77]	Bolts	Bounding box detection	UAV	1852 RGB images	UPOM (Based on ResNet and Attention)	Recall 0.94–1.00

Table 6

Summary of power line conductor detection studies.

Year & Ref	Component	Type of detection	Imaging platform	Dataset	Algorithm	Performance
2022 [97]	Conductor	Semantic segmentation	UAV	366 RGB images	UNet with Attention blocks	Dice: 0.957
2023 [98]	Conductor	Semantic segmentation	UAV	PLD dataset: 573 RGB images; OPL dataset: 571 RGB images	DRA-Net	mIOU: 93.19% (PLD dataset), mIOU: 96.04% (OPL dataset)

Additionally, a U-shaped noise denoising (UND) block reduces background interference, and an edge enhancement block (EEB) strengthens the network's capacity to extract useful edge features. Experimental results demonstrate DRA-Net's excellent segmentation performance, achieving a Dice coefficient of 93.26% and a mean Intersection over Union (mIoU) of 93.19% on the public power line dataset (PLD) and 96.40% Dice coefficient and 96.04% mIoU on the self-built overhead power line (OPL) dataset. However, the reported results were obtained from a relatively small training dataset of 1144 images and could be reduced when tested on a wide variety of real-world images.

8.4. Multi-component detection

A typical transmission tower consists of multiple different types of components with varying shapes and dimensions. Some studies have extended focus beyond individual component detection to multi-component detection [19,99,100]. In a 2024 study, Dong et al. [101] proposed a meta learning-based model to address the challenge of detecting key components and defects in transmission lines using aerial images, particularly when dealing with limited sample sizes for certain categories. The model incorporated a region-aware fusion (RAF) module to capture spatial relationships between support and query images, enabling effective matching and identification of objects. Additionally, cascade RAF heads were employed to progressively refine bounding box proposals and improve detection accuracy. The model was trained using a two-stage fine-tuning strategy, leveraging a larger dataset of base classes to enhance the detection of novel classes with fewer samples. Experimental results demonstrated the model's superior performance compared to traditional deep learning and few-shot object detection models. The authors acknowledged potential limitations in their work. The model's performance might be affected by variations in image quality and environmental conditions encountered in real-world power transmission line inspections. The reliance on a pre-defined set of base classes and novel classes might limit the model's adaptability to new or unexpected object categories. The authors suggested exploring the incorporation of online learning or active learning techniques to enable the model to continuously learn and adapt to new scenarios.

Self-supervised pretraining has been used effectively in other domains to tackle the challenge of minimal or non-existent annotated dataset. Liu et al. [102] introduced Tower Masking MIM (TM-MIM), a self-supervised pretraining method designed to enhance power line

component detection in aerial images, particularly in scenarios with limited labeled data. By employing a novel masking strategy that focuses on the tower–conductor region and a Siamese architecture with dual reconstruction branches, the model learns to capture discriminative features and global representations from unlabeled data. The incorporation of knowledge distillation further enhances the model's ability to retain general knowledge from pretrained models while acquiring domain-specific knowledge. The authors identified potential areas for future work. The masking strategy's reliance on tower presence in images might necessitate the development of techniques to handle images without towers. Additionally, exploring the use of varied-grained masks for different feature hierarchies could further enhance the model's ability to detect components of varying scales. The authors also suggested extending the TM-MIM approach to other object detection frameworks, such as YOLO, to broaden its applicability. Table 7 summarizes the literature on power line multi-component detection.

This section has demonstrated the remarkable progress in automating power line component detection through computer vision and deep learning techniques. The reviewed literature showcases significant achievements across various components: from insulator detection achieving mAP rates of up to 91.25% using YOLOv8, to conductor segmentation reaching Dice coefficients of 96.40% with innovative architectures like DRA-Net, and multi-component detection systems attaining mAP rates of 87.7% through self-supervised learning approaches. These advances have been enabled by architectural innovations such as attention mechanisms, multi-scale feature fusion, and hybrid networks combining CNN backbones with specialized modules. However, several domain-specific challenges persist. The detection of small-scale components like pin-bolts and dampers, which often occupy minimal pixels in aerial imagery, remains particularly challenging, with performance metrics for fitting detection typically lower than those for larger components. Additionally, real-world deployment faces obstacles such as varying imaging conditions, complex backgrounds, and class imbalance issues where critical components occupy only a small portion of the image. The limited size and diversity of publicly available datasets — with most studies utilizing fewer than 5000 images — continues to constrain model development and benchmarking. Recent promising directions include meta-learning approaches for few-shot detection, self-supervised pretraining methods like TM-MIM for leveraging unlabeled data, and specialized architectures incorporating domain knowledge about component spatial relationships. These

Table 7
Summary of multi-component detection studies in power line inspection.

Year & Ref	Component	Type of detection	Imaging platform	Dataset	Algorithm	Performance
2019 [103]	Insulator and Damper	Bounding box detection	UAV	4416 insulator & 4352 damper images	YOLOv3	Accuracy: 95.84%
2019 [100]	Tower components	Bounding box detection	NA	11 600 RGB images	F-RCNN	–
2019 [19]	Tower components	Bounding box detection	Aerial vehicle	123 151 RGB images	SSD	mAP: 0.67
2020 [99]	Tower components	Bounding box detection	UAV	City A: 2016 RGB images; City B: 3960 RGB images	Enhanced F-RCNN	City A dataset: mAP: 52.9%, City B dataset: mAP: 45.3%
2021 [80]	Insulator and Bolts	Bounding box detection	Aerial vehicle	1830 RGB images	Mask R-CNN and RetinaNet	Precision: 96.7% (Insulators), 97.9% (Bolts)
2023 [102]	Tower components	Bounding box detection	UAV	PLCD dataset: 1000 images	TM-MIM network	mAP@0.5: 87.7%
2023 [104]	Towers	Bounding box detection	Satellite	Duke university dataset: 2740 images	Improved YOLO-based network	mAP@0.5: 77.47%
2024 [101]	Tower components	Bounding box detection	UAV	9017 RGB images	Meta learning	mAP@0.5: 64.6%

challenges and emerging solutions in power line component detection will be explored more comprehensively in Section 11, along with their implications for future research and practical applications.

9. Power line fault diagnosis

Visual inspection, coupled with cutting-edge computer vision techniques has emerged as a powerful and efficient means of identifying and diagnosing a spectrum of faults. From insulator defects to conductor issues, and tower anomalies to grounding problems, the ability of computer vision to meticulously assess power line components is transforming maintenance and reliability in the electrical grid. This section delves into the application of computer vision and deep learning for the diagnosis of various faults (Fig. 4) within power lines and associated equipment, underlining its potential to enhance the resilience and performance of critical power transmission infrastructure.

9.1. Insulator fault detection

Insulators are critical components of power lines, ensuring that high-voltage currents are safely transported without coming into contact with the towers or poles that support them. However, insulators continuously endure environmental stresses and mechanical wear, which can lead to various types of faults. The following sections discuss some of the most common types of insulator faults and review significant research addressing these issues.

9.1.1. Surface defect detection

Surface defects in insulators encompass a range of anomalies that affect the outermost layer of the insulator. This category includes issues such as surface contamination, cracking, flashover marks, arcing damage, and chipping. Recent research works have made significant strides in the detection and classification of these surface defects, utilizing a variety of computer vision techniques and machine learning algorithms. Roy et al. [109] developed a deep learning framework incorporating AlexNet, VGG16, and ResNet50 models to detect sand, ash, and mud contaminations on the insulator surface. However, this approach requires images of the insulators taken at close proximity and is limited to classification and cannot detect contaminated areas. Mussina et al. [107] introduced a Fusion Convolutional Network (FCN)

for the real-time monitoring of insulators using UAVs, addressing challenges like varying resolutions and unclear surfaces. It combines a CNN with a binary ANN to form a multi-modal information fusion system that improves contamination classification accuracy on insulators to 99.76% by using image data and leakage current values. However, the proposed model was trained on a relatively small dataset of only 250 images per class and it requires close-up images of the insulator with uniform background. These limitations can affect the model's capability to generalize on real-world situations. Table 8 summarizes the literature on insulator surface faults.

9.1.2. Structural defect detection

Structural defects refer to issues that affect the internal composition and mechanical strength of the insulator. They include complete breakage, missing insulator caps, and material degradation. Table 9 summarizes the literature on insulator structural faults. The following goes over some of the notable works on structural defect detection.

In a 2022 work, [92] proposed an attention-based end-to-end framework that combines object detection and instance segmentation at the pixel level. Although trained on a relatively small dataset of 1523 images, this is one of the few works that targeted pixel-level segmentation and achieved great results. By integrating a three-branch attention structure into the backbone network, the proposed model achieved a significant improvement in detection performance, surpassing the state-of-the-art instance mask prediction models while maintaining computational efficiency. In a similar work by Wang et al. [110], the authors proposed an insulator defect detection method that leverages an improved ResNeSt network and an enhanced Region Proposal Network (RPN). The improvements to ResNeSt were aimed at refining the detection of insulator defects, particularly in aerial images where insulators might be small and have low resolution. The authors acknowledge that their method might struggle to detect certain types of defects, such as breaks on specific insulator types, due to their subtle visual characteristics. The computational efficiency of the proposed method, while improved compared to some baseline models, might still be a concern for real-time applications on resource-constrained devices.

Multi-task networks can be used to detect different types of insulator faults as separate classes. In a recent article Fu et al. [111] presented the I2D-Net, a deep learning-based method for detecting small-sized defects in overhead transmission line insulators, particularly missing caps. The

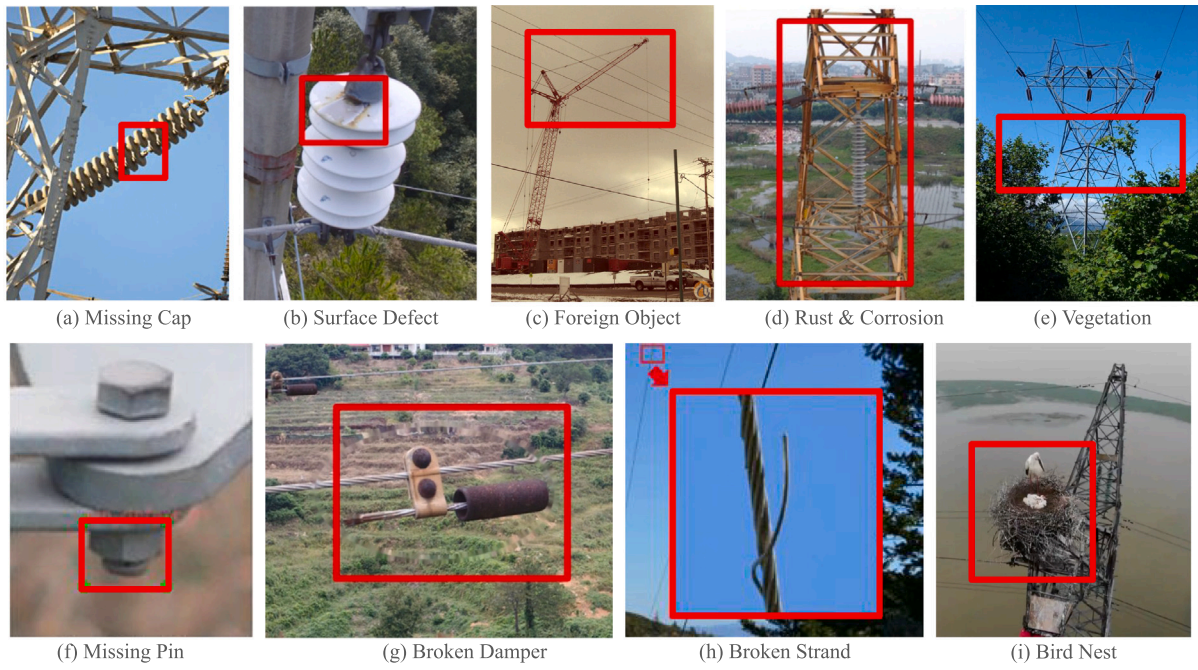


Fig. 4. Different types of power line faults [9].

Table 8

Summary of insulator surface fault detection studies.

Year & Ref	Component	Type of detection	Type of fault	Imaging platform	Dataset	Algorithm	Performance
2017 [105]	Porcelain insulators	Classification	Deterioration damage	Still camera	700 IR images	CNN	Accuracy: 93%
2019 [88]	Insulators	Classification	Surface defects	High speed railway	18 000 RGB images	DDAE (Based on CNN AutoEncoder)	F1-Score: 0.95
2019 [57]	Glass insulator	Classification	Surface defects	UAV	3500 RGB images	CNN	Accuracy: 87%
2020 [106]	Insulators	Classification	3 types of surface defects	Still camera	1240 RGB images	CNN	Accuracy: 89.5%
2020 [107]	Glass insulators	Classification	4 Types of surface defects	UAV	5000 RGB images	FCN	Accuracy: 99.76%
2021 [91]	Ceramic insulators	Bounding box detection	Insulator surface & Structural defects	UAV	2973 RGB insulator images	R-CNN, SVM, CNN, SSD Fusion	Onshore mAP: 0.56–0.77; Onboard mAP: 0.24–0.27
2022 [108]	Insulators	Classification	Kaolin defects	Still camera	600 RGB images	ANN	Accuracy: 97.50%
2023 [109]	Insulators	Classification	4 types of surface defects	Still camera	1000 RGB images	CNN	Accuracy: 97.5%

I2D-Net enhanced the Faster R-CNN object detection framework with three key components: the Three-path Feature Fusion Network (TFFN) to improve feature extraction from shallow layers, the enhanced Receptive Field Attention (RFA+) block to adapt to different-scale defects, and the Context Perception Module (CPM) to enhance defect localization. The authors acknowledged that while the I2D-Net achieved high accuracy, it came with a slight increase in inference time compared to the baseline Faster R-CNN + FPN model. Liu et al. [112] proposed another approach that utilizes a deep CNN with parallel branches to locate fault regions and estimate insulator endpoints. The method offers high accuracy and robustness, outperforming previous approaches like Faster R-CNN, SSD, and cascading networks. Zhang et al. [90] took a different approach where they used a generative network including a denoising autoencoder, discriminator, and classifier to detect defects. The method comprises two stages: first, insulator extraction using cascaded deep segmentation networks (CDSNets); second, defect detection within sampled patches using the proposed DefGAN, scoring defects based on classifier anomaly probability and denoising autoencoder

reconstruction error. Although the proposed model can detect mild deformities, it is sensitive to the noise in the image.

One of the primary causes of insulator failures is the self-explosion of caps and several research works have been done to detect these faults. Cao et al. [76] proposed an improved image augmentation method for the detection of self-detonation defects in aerial images of insulators. The method incorporated edge detection to enhance the shape features of insulators, which were then used to guide the augmentation process. The authors utilized the YOLOv3 model for insulator detection and an improved ResNet-18 model for defect classification. The proposed method was evaluated on a dataset of aerial images and showed superior performance compared to baseline models and other augmentation techniques. However, the authors acknowledged the need for further evaluation on larger and more diverse datasets, including images captured under different weather conditions. Additionally, the proposed method relied on edge detection as prior knowledge, which might not be sufficient for detecting more complex

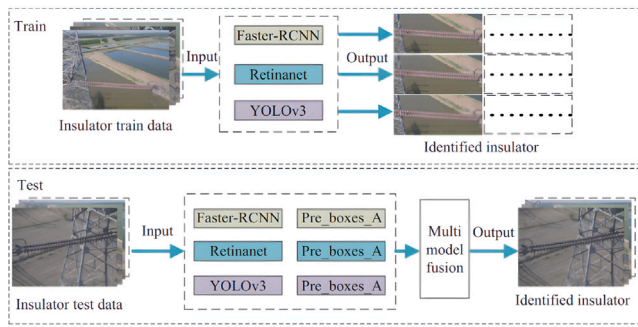


Fig. 5. Simplified diagram of multi-model fusion network for detecting insulators [60].

or subtle defects. In another work, Wei et al. [60] introduced a fault detection scheme for insulator self-explosion, leveraging edge computing and deep learning to address issues with traditional centralized cloud computing. It employs a lightweight SSD target recognition network at the edge, replacing VGG16 with MobileNets to reduce computation. In the cloud, three distinct detection algorithms (Faster-RCNN, Retinanet, YOLOv3) are used to identify insulator self-explosion areas, and a novel multimodel fusion algorithm (Fig. 5) enables overhead transmission line fault detection. Results demonstrate effective data reduction, achieving an average cloud recognition accuracy of 95.75%, with a modest increase in edge device power consumption.

9.2. Detection of conductor faults

Transmission line conductors are the lifelines of the electrical power grid, carrying electrical energy over long distances. Conductor defects, such as aging, corrosion, internal damage, and contamination, pose significant threats to the reliable and safe operation of power transmission systems. This subsection explores the challenges associated with conductor defect detection and highlights the use of computer vision technologies to enhance the efficiency and accuracy of inspections. Table 10 summarizes the literature on conductor faults.

External defects of conductor wires can be detected using visible light images. These types of defects can be due to aging, structural damages or due to foreign objects which are described in more detail in the “Detection of Foreign Objects” subsection. Aging defects of conductors are one of the most common for aluminum conductors. In a 2022 study, Yi et al. [120] introduced a novel approach to quantitatively diagnose the aging of conductor surfaces in smart high-voltage electricity grids. The model consists of an AlexNet-based deep convolution network and a specialized loss function that incorporates Gaussian label distribution. By reframing the conductor morphology aging problem as a multi-classification challenge, the model leverages a weakly labeled training dataset and a carefully designed loss function, combining entropy loss, cross-entropy loss, and Kullback–Leibler divergence loss. Comparative analysis with four commonly used CNN-based classifiers demonstrates superior performance on the collected dataset. However, the proposed model is suitable only when super-close-up high resolution conductor images are available.

Internal defects of the conductors can be challenging to detect using visible light images. To address this, Wang et al. [34] presented a novel automatic detection system for identifying internal defects in overhead aluminum conductor composite core (ACCC) transmission lines. The system utilizes an X-ray inspection robot equipped with a nondestructive testing (NDT) system to capture X-ray images of ACCC wires. The proposed system employs an anchor-free object detection model called CenterNet-NDT, which incorporates specialized modules like spatial pyramid pooling-cross stage partial convolution (SPPCSPC), polarized self-attention (PSA), and a weighted bidirectional feature pyramid network (SOFPN). Although CenterNet-NDT achieves a high mAP of 90.60% on the IN-ACCC dataset, the instrumentation and maintenance required for this method can be challenging.

9.3. Fault detection for fittings: Pin-bolts, dampers, suspension clamps

Power line fittings such as pin bolts, dampers, and suspension clamps are integral components of transmission and distribution systems. These components are tiny compared to the rest of the transmission line tower and they occupy a very small area in the images involving only a few pixels. As a result, accurately detecting faults in these components requires a high level of image processing and powerful deep-learning algorithms. Table 11 summarizes the literature on power line fittings fault detection.

Zhao et al. [123] proposed RFBD (Recognition of Bolt Faults using multilabel image recognition) that comprises two networks: VFSKnet leverages professional posterior knowledge to model relationships between bolt defect labels, while VFPKnet extracts structural features to capture fine-grained details. After combining and weighting these subnetworks, RFBD achieves label-level accuracy of 93.91% and image-level accuracy of 83.29% in detecting 5 types of bolt defects. However, the proposed solution is not end-to-end and requires properly segmented images of bolts. In a similar work, Zhang et al. [65] addressed the challenge of visually indistinguishable bolt defects in transmission lines by proposing an end-to-end detection method based on transmission line knowledge reasoning. It employs the DETR (End-to-End Object Detection with Transformers) [56] model augmented with a dilated encoder module to capture multiscale features. Additionally, a Transmission Line Image Relative Position Encoding (TL-irPE) is designed to infer bolt position knowledge. The approach includes a bolt attributes classifier and a bolt defects classifier, combining position and attributes knowledge to enhance defect detection accuracy.

One of the most prominent faults that occurs often is the missing pin of bolts. Due to its tiny nature, it can be very challenging to detect these faults. Zhao et al. [121] proposed the AVSCNet model to address the challenge of detecting pin-missing defects in bolts on transmission lines using aerial images. The model tackled the issue of varying visual shapes of bolts due to different camera angles by employing an automatic visual shape clustering method. It also incorporated feature enhancement, feature fusion, and expanded region-of-interest feature extraction to improve the detection of small-scale defects in complex aerial images. However, the model’s performance might be affected by the distributional differences in aerial images captured from diverse transmission line structures and environments.

Structural anomalies and rusting of the dampers and suspension clamps can compromise the transmission line system causing failure. Zhang et al. [95] introduced the AGMNet, an attention-guided multi-task convolutional neural network designed for the detection of power line parts in aerial images. The network incorporated a region attention mechanism to enhance the feature representation of objects, a refinable Region Proposal Network (RPN) to improve proposal quality, and a multitask learning framework to simultaneously detect power line fittings (dampers and suspension clamps), identify their rust levels, and detect abnormal conditions. However, the authors acknowledged the need for further evaluation on larger and more diverse datasets. Additionally, the distinction between different rust levels was found to be challenging, leading to less accurate rust level identification. The authors suggested further research to improve the accuracy of rust level detection.

9.4. Fault recognition on towers

Towers are the sturdy backbone that supports high-voltage conductors, ensuring the reliable delivery of electricity over long distances. Traditional methods for fault detection in transmission lines have limitations, such as susceptibility to noise and transient fluctuations. To address these challenges, Wang et al. [100] proposed a novel approach for fault zone detection that emphasizes fine-grained categorization and quality-awareness. The primary objective is to identify the most distinguishing image patches for classification. The key components of

Table 9
Summary of insulator structural fault detection studies.

Year & Ref	Component	Type of detection	Type of fault	Imaging platform	Dataset	Algorithm	Performance
2019 [113]	Glass and Porcelain insulators	Bounding box detection	Insulator structural defects	UAV	485 RGB images	SSD	Precision: 91.23%, Recall: 93.69%
2020 [114]	Insulators	Bounding box detection	Structural defects	UAV	1956 RGB images	Cascaded DNN	F1 Score: 93.3%, Speed: 135 ms/Image
2020 [110]	Insulators	Bounding box detection	Insulator structural defect	NA	2560 RGB images	Improved RetinaNet	Accuracy: 98.38%
2021 [112]	Insulators	Bounding box detection	Structural faults	Aerial vehicle	969 RGB images	Box point detector	mAP@0.5: 94.0%
2021 [115]	Insulators	Bounding box detection	Missing cap	UAV & Synthetic	CPLID dataset: 848 RGB images, Pascal VOC dataset: 5011 RGB images	Improved YOLOv3	Pascal VOC: mAP@0.75: 45.90% CPLID: mAP@0.75: 64.05%
2021 [90]	Insulators	Classification	Insulator structural defect	High speed railway catenary	800 RGB images	GAN	F1-Score: 0.95
2022 [92]	Insulators	Instance segmentation	Insulator defects	UAV	1523 RGB images	Mask R-CNN	AP: 89.4%
2022 [116]	Insulators	Bounding box detection	Bunch drop defect	UAV & Synthetic	CPLID dataset: 848 RGB images	Improved YOLOv4-ResNest	mAP: 95.63%
2022 [60]	Insulators	Bounding box detection	Self-explosion defect	UAV	8500 RGB images	F-RCNN, RetinaNet, YOLOv3 Fusion	Precision: 99.04%, Recall: 93.69%
2023 [76]	Glass insulators	Bounding box detection & Classification	Self-explosion defects	UAV	8463 RGB images	YOLOv3 & Improved ResNet-18	F1-Score: 86.25%
2023 [111]	Insulators	Bounding box detection	5 types of structural defects	UAV & Synthetic	CPLID dataset: 848 RGB images	I2D-Net (Based on F-RCNN)	mAP: 89.6%
2023 [117]	Insulator	Bounding box detection	Structural defect	Aerial vehicle	900 RGB images	PKAMNet	mAP@0.5: 95.5%
2023 [58]	Insulators	Classification	Structural defects	UAV	848 RGB images	Ps-ProtoPNet	F1-Score: 0.996
2023 [118]	Insulators	Bounding box detection	Self-explosion defect	UAV & Synthetic	CPLID dataset: 848 RGB images	GhostNet-YOLOv4	mAP: 99.50%
2024 [66]	Insulator	Bounding box detection	Structural defects	UAV	5939 RGB images	YOLO-v5 and DETR	mAP: 98%
2024 [119]	Insulator	Bounding box detection	Structural defect	UAV and Synthetic	CPLID dataset: 848 RGB images	Improved YOLO-based network	mAP@0.5: 85.8%

Table 10
Summary of conductor fault detection studies.

Year & Ref	Component	Type of detection	Type of fault	Imaging platform	Dataset	Algorithm	Performance
2019 [32]	Oil transformer and Conductor	Classification	Internal defects	Still camera	12 000 UV, IR and Visible images	Capsule network	Qualitative assessment
2021 [62]	Conductor	Bounding box detection	Vegetation	Tower mounted camera	70 RGB images	Cascaded network (Based on F-RCNN)	Accuracy: 95%
2022 [120]	Conductor	Classification	Aging defect	NA	5200 RGB images	Improved AlexNet	Avg. MAE: 3.80
2023 [34]	Conductors	Bounding box detection	4 types of composite core defects	Inspection robot	2500 X-ray images	CenterNet-NDT (Based on ResNet50)	mAP: 90.60%
2024 [36]	Conductor	Semantic segmentation	Vegetation	Land	851 LiDAR images	UNet-based network	Precision: Above 90%

their method involve extracting features from line currents using a Fast R-CNN-based image sample decomposition, with a quality module selecting the most informative regions. These extracted features are then utilized in a Support Vector Machine (SVM) [127] for fault recognition. In a similar work, Liang et al. [128] created four medium-sized datasets for training component detection and classification models. They also employed data augmentation techniques to balance the imbalanced

classes. Furthermore, the authors proposed a multi-stage component detection and classification approach using the SSD network and deep ResNet to detect small components and faults. The results demonstrate that the proposed system is both fast and accurate in detecting common faults in tower components, such as missing top caps, cracks in poles and cross arms, woodpecker damage on poles, and rot damage on cross arms. Table 12 summarizes the literature on tower faults.

Table 11
Summary of fault detection studies for power line fittings.

Year & Ref	Component	Type of detection	Type of fault	Imaging platform	Dataset	Algorithm	Performance
2020 [121]	Bolts	Bounding box detection	Missing pin defect	Tower mounted cameras	1840 RGB images	AVSCNet	mAR: 0.876
2021 [122]	Bolts	Bounding box detection	Pin defects	Aerial vehicle	1500 RGB images	Improved MTCNN	mAP: 94.76%
2022 [81]	Bolts	Bounding box detection	2 types of pin defects	UAV	482 RGB images	EfficientDet-D7	mAP: 54.3%, mAR: 63.4%
2022 [95]	Dampers and Suspension clamps	Bounding box detection	Rust defects and Abnormal conditions	UAV	1209 RGB images	AGMNet	Rust: mAP: 75.4%, Abnormal: mAP: 92.7%
2022 [123]	Bolts	Classification	5 types of bolt defects	UAV	1944 RGB images	VFPKNet	Accuracy: 83.29%
2023 [96]	Damper	Classification	Structural defects	UAV	240 Aerial images	Improved GrabCut	Accuracy: 95.76%
2023 [124]	Bolts	Bounding box detection	Structural defect	UAV	32 094 images	MARF-CCN network	mAP: 87.16%
2023 [77]	Bolts	Bounding box detection	Structural defects	UAV	1852 RGB images	UBDDM (Based on ResNet-50 & Attention)	mAP: 78.6%
2023 [125]	Fittings	Bounding box detection	Rust defect	NA	TLCF dataset: 700 RGB images	Deformable YOLOX	mAP@0.5: 0.857, mAP@0.5:0.95: 0.533
2023 [65]	Bolts	Classification	4 types of bolt defects	UAV	VIBD dataset: 8972 bolt instances in RGB images	PA-DETR (Based on ResNet50, FPN and Attention)	mAP: 81.9%
2023 [126]	Dampers	Bounding box detection	Structural defect	UAV	490 RGB images	DSA-Net	mAP@0.5: 0.935

Table 12
Summary of tower fault detection studies.

Year & Ref	Component	Type of detection	Type of fault	Imaging platform	Dataset	Algorithm	Performance
2019 [19]	Tower components	Classification	Structural defects	UAV	123 151 RGB images	ResNet	F1-Score: 77.96%
2019 [100]	Tower components	Classification	Line Faults	NA	11 600 RGB images	SVM	Accuracy: 96.73%
2020 [128]	Tower components	Bounding box detection	8 types of structural defects	UAV	WIRE_10 dataset: 8000 RGB images	F-RCNN	mAP: 91.11%
2021 [80]	Tower components	Classification	Surface and Structural defects	Helicopter	Insulators: 25 804 RGB images, Bolts: 94,619 RGB images	EfficientNetB0	ROC: 0.94 (insulator), ROC: 0.98 (Bolts)
2022 [129]	Tower components	Bounding box detection	Structural defect	UAV	1309 RGB images	Graph neural network	mAP@0.5: 89.1%
2022 [108]	Tower components	Classification	Structural defects	Still camera	240 RGB images	Inception v3	F1-Score: 84.50%
2022 [78]	Tower components	Classification	Structural defects	Still camera	240 images	Semi-ProtoPNet (Based on VGG-19)	Accuracy: 97.22%
2023 [130]	Tower components	Classification	Structural defect	Aerial vehicle	TLAD dataset: 4811 RGB images	NMHEM model	F1 Score: 80.5%
2023 [131]	Tower components	Bounding box detection	Structural defect	UAV	956 RGB images	PSTL-Net	mAP: 0.848
2024 [132]	Substation and Tower components	Classification and Bounding box detection	Structural defect	UAV and land	Substation dataset: 2000 images; UAV dataset: 5048 images; CPLID dataset 848 images	Federated learning	mAP: .990

Table 13
Summary of foreign object detection studies in power line systems.

Year & Ref	Component	Type of detection	Type of fault	Imaging platform	Dataset	Algorithm	Performance
2020 [59]	Tower, Conductor	Bounding box detection	7 types of Foreign objects	NA	926 & 2000 RGB images	YOLOv4	–
2020 [133]	Tower, Conductor	Bounding box detection	Foreign objects	Tower mounted camera	8000 RGB images	DBF-NN	mAP: 88.1%
2022 [85]	Tower	Bounding box detection	Bird nest	Aerial platform	3695 RGB images	YOLOv5	mAP: 83.4%, FPS: 85.32
2022 [67]	Tower, Conductor	Bounding box detection	4 types of Foreign objects	NA	35 000 RGB images	YOLOv3-MobileNetv2	mAP: 0.832, FPS: 60
2023 [134]	Tower, Insulator	Bounding box detection	Bird nest	UAV & Synthetic	Bird nest: 2864 images, CPLID dataset: 848 RGB images	YOLOX++	mAP: 86.8%
2023 [79]	Tower	Bounding box detection	Foreign objects	UAV & Manual	1232 RGB images	YOLOv4 with Attention	mAP: 96.71%, FPS: 45
2023 [135]	Tower, Conductor	Bounding box detection	6 types of Foreign objects	UAV	8803 RGB images	Improved YOLOv7	mAP: 92.2%, FPS: 19
2023 [136]	Conductor	Bounding box detection	Cranes	Tower mounted camera	4000 RGB images	Edge VIP (Based on YOLOv5s)	mAP: 50.60%

9.5. Detection of foreign objects

Foreign object-related anomalies and faults in power lines encompass a multitude of potential hazards that can pose significant threats to the reliable operation of electrical transmission systems. These hazards can range from inadvertent interference caused by construction equipment and vehicles to more natural occurrences like fires, bird nests, and overgrown vegetation among others. The presence of foreign objects in proximity to power line components can lead to various issues, including interruptions in electrical supply, damage to critical equipment, and heightened safety concerns for both utility personnel and the public. Table 13 summarizes the literature on foreign object faults.

Construction equipment and vehicles operating near power lines can accidentally make contact with electrical components, causing short circuits, equipment failure, and potentially even electrical fires. Similarly, fires in the vicinity of power lines can result from various sources, including wildfires, discarded cigarette butts, or even arcing caused by faulty equipment. Zhang et al. [136] presented an edge-based framework for power transmission line abnormal target detection, focusing on overcoming resource limitations and improving model performance. To mitigate the lack of labeled data, deep semi-supervised learning was introduced, which can refine the decision boundary by learning from unlabeled samples. To achieve this, the framework starts with an initial model trained on a small amount of labeled data and then updates itself using unlabeled data. In another work, Zhang et al. [59] introduced a framework that combines cloud and edge computing with deep learning techniques. Initially, a YOLOv4 model is trained in the cloud server for abnormal object detection. This trained model is then deployed to edge servers for real-time detection of abnormal objects in captured pictures. To address the limited initial data samples of only 926 images, enhancement techniques are used to increase the number of pictures, and real-time data streams are employed for incremental learning.

Bird nests, while seemingly innocuous, can also present challenges for power lines. Nests built on or near power line components can lead to electrical faults if they bridge connections or create conductive pathways. Ge et al. [85] proposed a bird's nest defect recognition method using YOLOv5, aiming to address the safety concerns posed by bird nests on power transmission towers. The method employs a YOLOv5-based architecture, comprising a backbone network, Feature Pyramid Network (FPN) [137], and YOLO head, and undergoes multiple rounds of training with a constructed bird's nest defect database and transmission line model. Results demonstrate that the YOLOv5 model

achieves an 83.40% recognition rate for bird's nests while maintaining a high FPS rate of 85.32. In a similar work, Bi et al. [134] presented a novel target detection model called YOLOX++, which is built upon the YOLOX [138] architecture to enhance the detection of abnormal targets in transmission lines. It introduces a multiscale cross-stage partial network (MS-CSPNet) to fuse multiscale features and expand the receptive field of the target, improving channel combination (Fig. 6). Depth-wise and dilated convolutions are added to the object decoupling head to capture long-range dependencies of objects in feature maps. Additionally, the alpha loss function (α -IoU) is incorporated to optimize small object localization. Experimental results demonstrate that YOLOX++ achieves detection accuracies of 86.8% for high-voltage-tower bird nests and 96.60% for power line insulators, outperforming the YOLOX model. On the PASCAL VOC dataset [139], YOLOX++ exhibits a 9.30% improvement in AP50 and a 5% improvement in APS compared to YOLOX, showcasing its enhanced robustness for small target detection.

Another common cause of power line faults is vegetation, which, if left unmanaged, can grow into power lines, potentially causing short circuits, outages, or even wildfires during dry conditions. Rong et al. [62] proposed an intelligent detection framework for real-time monitoring of vegetation encroachment on power lines. The framework utilized binocular vision sensors mounted on transmission towers to capture images, which were then processed locally. The framework employed Faster R-CNN for vegetation detection, Hough transform for power line detection, and an advanced stereovision algorithm for 3D reconstruction. The advanced stereovision algorithm incorporated calibration optimization and world coordinate system transformation to improve accuracy in large-scale scenes. The authors acknowledged that the accuracy of the proposed method might be affected by complex terrain and weather conditions, which could impact image quality and object detection. Additionally, the computational efficiency of the framework was not explicitly addressed, which could be a concern for real-time monitoring applications.

To conclude, this section has demonstrated the transformative impact of computer vision and deep learning in revolutionizing power line fault diagnosis. Through the integration of advanced neural network architectures and multi-task models, significant progress has been made across multiple domains: from detecting small insulator defects with accuracies exceeding 97%, to identifying structural anomalies in towers with mAP rates above 90%, and recognizing foreign objects with real-time processing capabilities of up to 85 FPS. However, several critical challenges persist. The scarcity of large-scale, publicly available

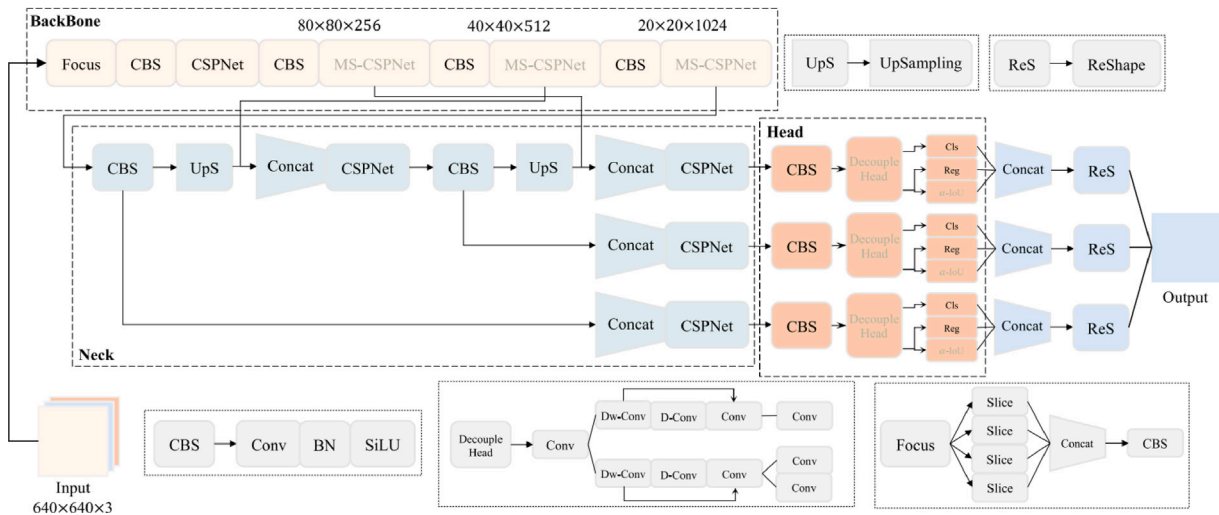


Fig. 6. The simplified network architecture of the proposed YOLOX++ network [134].

datasets remains a significant bottleneck, with most studies limited to small, proprietary datasets of under 5000 images. The detection of miniature components like pin-bolts and dampers, which often occupy only a few pixels in aerial imagery, continues to challenge even state-of-the-art models. Environmental factors such as varying lighting conditions, weather effects, and complex backgrounds further complicate reliable fault detection. To address these limitations, promising research directions include: leveraging generative adversarial networks (GANs) for synthetic data augmentation, developing semi-supervised learning approaches to maximize the utility of unlabeled data, exploring attention-based architectures specifically optimized for small object detection, and investigating multi-task learning frameworks that can simultaneously handle different types of faults. The integration of edge computing solutions, as demonstrated in several studies, also shows promise in enabling real-time fault detection while managing computational constraints. These challenges and future research directions regarding power line fault diagnosis are explored in greater detail in Section 11.

10. Discussion

In this section, we present a comprehensive overview of the research articles reviewed in this study. Additionally, we provide a qualitative assessment of these articles. Initially, a set of criteria has been established for evaluating the papers, ensuring these criteria adequately reflect the core content of each article without favoring a particular subset. This involved a meticulous review of the articles and consultation with multiple co-authors to mitigate individual biases. Following this procedure, several criteria were identified for our qualitative assessment, including the use of large datasets (sample size greater than 5000), publicly available datasets, publication of accompanying code, fault detection across multiple components, employment of image modalities beyond visible light, application of advanced image processing techniques beyond resizing and cropping, use of synthetic and augmented data, focus on small components such as bolts, localization of specific components or faults, provision of performance metrics, acknowledgment of limitations, and suitability for real-time deployment. Based on these criteria, we conducted a thorough evaluation of all selected papers, and the findings are summarized in Table 14.

Table 14 reveals several noteworthy insights regarding the articles reviewed. The scarcity of publicly available datasets is evident, with only 23% of the papers utilizing them, while the majority rely on privately generated datasets. This issue has been discussed in Section 6, “Publicly Available Datasets”. Moreover, large datasets are seldom used, with only 30% of the articles meeting our 5000-sample threshold.

To address this, numerous studies have incorporated synthetic or augmented data, with image augmentation being particularly prevalent, utilized in 67% of the studies.

The publication of source code alongside machine learning or deep learning papers is highly beneficial for enabling readers to replicate algorithms and reproduce results, provided the dataset is also available. Although not always possible, sharing the source code enhances the credibility and impact of the research. However, among the articles reviewed, only 6 out of 73 published their source code.

Detecting faults across various components presents significant challenges due to the diverse nature of components and faults, such as the noticeable size difference between power line insulators and bolts [19, 128]. Despite these challenges, 34% of the papers successfully employed multi-task learning techniques to address this issue, demonstrating effective results and highlighting further research opportunities in generalizing across more component types, faults, and environments. The detection of small objects remains particularly challenging due to their size relative to other components, often requiring sophisticated algorithms with considerable potential for advancement [80].

Unlike periodic inspections, real-time systems offer continuous surveillance, promptly identifying and addressing emergent issues. UAVs equipped with real-time algorithms can rapidly cover extensive areas, delivering immediate data to operators and aiding in the localization of faults, thus reducing the time and labor costs associated with manual inspections [59]. Real-time systems demand algorithms capable of operating at speeds typically around 30 FPS or more on low-powered edge devices [63]. Approximately 33% of the reviewed studies have tackled this challenge by developing performant algorithms suitable for real-time deployment, often relying on rapid object detection algorithms like YOLO [57] and SSD [63].

Automated power line inspection is a multi-step process where each step influences subsequent ones. For simplicity and clarity, it can be broken down into these key aspects: the component of interest, the choice of imaging platform, the size of the dataset, the type of detection, and the selected algorithm. When the process involves fault diagnosis, the type of fault should also be considered. By combining these aspects into a flow diagram, with the target component as the starting point and algorithm selection as the endpoint, we can establish a pattern for how decisions are made. With this context in mind, we carefully examined the reviewed literature to construct the diagrams in Fig. 7. Fig. 7(a) illustrates the decision-making process for automated component detection in power line infrastructure. The diagram shows that most research focused on insulators, utilized UAVs as the imaging platform, used datasets of 1000–5000 samples, and employed bounding box detection. Algorithm choices are balanced among YOLO, RCNN,

Table 14
Assessment of the reviewed literature.

Number	Authors	Year of publication	Dataset size > 5000	Dataset availability	Code availability	Multi-component	Other image modalities	Image processing	Synthetic data	Data augmentation	Small object	Fault localization	Performance metrics	Limitations	Real-time
1	Liu et al. [105]	2017	×	×	×	×	✓	×	×	✓	×	×	✓	×	×
2	Tao et al. [37]	2018	×	✓	×	×	×	✓	✓	✓	×	✓	✓	×	✓
3	Kang et al. [88]	2019	✓	×	×	×	×	×	✓	×	×	✓	✓	×	×
4	Li et al. [32]	2019	✓	×	×	✓	✓	✓	×	×	×	×	×	×	×
5	Miao et al. [63]	2019	✓	×	×	×	×	×	×	✓	×	✓	✓	×	✓
6	Chen et al. [103]	2019	✓	×	×	✓	×	✓	×	✓	×	✓	✓	×	✓
7	Jiang et al. [113]	2019	×	×	×	×	×	×	×	✓	×	✓	✓	✓	×
8	Nguyen et al. [19]	2019	✓	×	×	✓	×	×	×	✓	×	✓	✓	×	×
9	Sadykova et al. [57]	2019	×	×	×	×	×	×	×	✓	×	✓	✓	×	✓
10	Wang et al. [100]	2019	✓	×	×	✓	×	×	×	×	×	×	×	×	×
11	Sadykova et al. [57]	2019	×	×	×	×	×	×	×	×	×	×	✓	×	✓
12	Ibrahim et al. [106]	2020	×	×	×	×	×	✓	×	×	×	×	✓	×	×
13	Mussina et al. [107]	2020	✓	×	×	×	×	×	×	×	×	×	✓	×	✓
14	Wang et al. [110]	2020	×	✓	×	×	×	×	×	×	×	✓	✓	✓	✓
15	Zhang et al. [99]	2020	✓	×	×	✓	×	×	✓	×	×	✓	✓	×	×
16	Zhang et al. [59]	2020	×	×	×	✓	×	×	×	✓	×	✓	✓	×	✓
17	Zhao et al. [121]	2020	×	×	×	×	×	×	×	✓	×	✓	✓	✓	×
18	Zhao et al. [89]	2020	✓	✓	×	×	✓	×	×	×	×	×	✓	×	×
19	Zhu et al. [133]	2020	✓	×	×	✓	×	×	×	✓	×	✓	✓	✓	✓
20	Odo et al. [80]	2021	×	×	×	✓	×	×	×	✓	×	✓	✓	✓	×
21	Singh et al. [29]	2021	×	×	×	×	✓	✓	×	×	×	×	✓	×	×
22	Waleed et al. [91]	2021	×	×	×	×	×	×	×	×	×	✓	✓	✓	✓
23	Xiao et al. [122]	2021	×	×	×	×	×	×	×	✓	✓	✓	✓	×	×
24	Zhang et al. [90]	2021	×	×	×	×	×	×	✓	✓	×	✓	✓	✓	×
25	Zhai et al. [61]	2021	×	×	×	✓	×	×	×	✓	×	✓	✓	×	×
26	Rong et al. [62]	2021	×	×	×	×	×	×	×	×	×	✓	✓	×	×
27	Zhang et al. [90]	2021	×	✓	×	×	×	×	✓	✓	×	✓	✓	×	×
28	Antwi-Bekoe et al. [92]	2022	×	×	×	×	×	×	×	✓	×	✓	✓	✓	×
29	Ge et al. [85]	2022	×	×	×	×	×	×	×	✓	×	✓	✓	×	✓
30	Hao et al. [116]	2022	×	✓	×	×	×	×	✓	✓	×	✓	✓	×	✓
31	Hunag et al. [96]	2022	×	×	×	×	×	✓	×	✓	×	✓	✓	×	✓
32	Li et al. [67]	2022	✓	×	×	✓	×	×	×	✓	×	✓	✓	×	✓
33	Li et al. [81]	2022	×	×	×	×	×	×	×	×	✓	✓	✓	✓	×
34	Qiu et al. [79]	2022	×	×	×	✓	×	✓	×	✓	×	✓	✓	✓	✓
35	Stefenon et al. [108]	2022	×	✓	×	✓	×	✓	×	✓	×	×	✓	×	×
36	Stefenon et al. [78]	2022	×	✓	×	✓	×	×	×	✓	×	×	✓	✓	×
37	Wei et al. [60]	2022	✓	×	×	×	×	×	×	✓	×	✓	✓	×	×
38	Yang et al. [97]	2022	×	×	×	×	×	×	×	×	×	✓	×	×	×
39	Yi et al. [120]	2022	✓	×	×	×	×	×	×	✓	×	×	✓	×	×
40	Zhai et al. [94]	2022	×	×	×	✓	×	×	×	✓	✓	✓	✓	×	×
41	Zhang et al. [95]	2022	✓	✓	✓	×	×	×	✓	✓	×	✓	✓	×	✓
42	Zhang et al. [95]	2022	×	×	×	✓	×	×	×	×	✓	✓	✓	×	×
43	Zhao et al. [123]	2022	×	×	×	×	×	×	×	×	✓	×	✓	×	×
44	Bi et al. [134]	2023	×	✓	×	✓	×	×	✓	✓	×	✓	✓	✓	×
45	Cao et al. [76]	2023	✓	×	×	×	×	✓	×	✓	×	✓	✓	×	×
46	Dong et al. [64]	2023	×	×	×	✓	×	×	×	✓	×	✓	✓	×	×
47	Fu et al. [111]	2023	×	✓	×	×	×	×	✓	✓	×	✓	✓	×	×
48	Luo et al. [77]	2023	×	×	×	×	×	×	×	×	×	✓	✓	×	×
49	Roy et al. [109]	2023	×	×	×	×	×	✓	×	×	×	×	✓	×	✓
50	Shuang et al. [93]	2023	×	✓	×	×	×	×	×	×	×	✓	✓	✓	✓
51	Singh et al. [58]	2023	×	✓	×	×	×	×	×	✓	×	✓	✓	×	×
52	Song et al. [125]	2023	×	×	×	×	×	×	×	✓	×	✓	✓	×	✓
53	Wang et al. [34]	2023	×	×	×	×	✓	✓	×	✓	×	✓	✓	✓	×
54	Yang et al. [98]	2023	×	✓	×	×	×	×	×	×	×	✓	✓	×	×
55	Yu et al. [135]	2023	✓	×	×	✓	×	×	×	✓	×	✓	✓	✓	✓
56	Zhang et al. [136]	2023	×	×	×	×	×	×	×	✓	×	✓	✓	×	×
57	Zhang et al. [118]	2023	×	✓	×	×	×	×	✓	✓	×	✓	✓	×	✓
58	Zhang et al. [65]	2023	✓	×	×	×	×	×	×	×	✓	×	✓	×	✓
59	Zhou et al. [24]	2023	✓	×	✓	×	×	×	×	✓	×	✓	✓	×	×
60	Zhang et al. [65]	2023	✓	×	×	×	×	×	×	×	✓	×	✓	×	✓
61	Hao et al. [117]	2023	×	×	×	×	×	✓	×	✓	×	✓	✓	✓	×
62	Liu et al. [129]	2023	×	×	✓	✓	×	×	×	✓	✓	✓	✓	×	×
63	Liu et al. [130]	2023	×	×	×	✓	×	×	×	×	×	×	✓	×	×
64	Jiao et al. [124]	2023	✓	×	×	×	×	×	×	✓	✓	✓	✓	✓	×
65	Zhong et al. [132]	2024	✓	×	×	✓	×	×	✓	×	×	✓	✓	✓	×
66	Zhang et al. [126]	2024	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓

(continued on next page)

Table 14 (continued).

67	Yi et al. [131]	2024	×	✓	✓	✓	×	×	×	×	✓	✓	✓	✓	×
68	Wang et al. [119]	2024	×	✓	✓	×	×	×	✓	✓	×	✓	✓	✓	×
69	Shi et al. [104]	2024	×	✓	×	×	×	×	×	✓	✓	×	✓	✓	×
70	Liu et al. [102]	2024	×	×	✓	✓	×	×	✓	✓	✓	✓	✓	×	×
71	Jain et al. [66]	2024	✓	×	×	×	×	✓	×	✓	×	✓	✓	×	×
72	Dong et al. [101]	2024	✓	×	×	✓	×	×	×	×	✓	✓	✓	✓	×
73	Bergmann et al. [36]	2024	×	×	×	×	✓	✓	✓	×	×	✓	✓	✓	×
			30%	23%	8%	34%	8%	19%	19%	67%	30%	78%	96%	33%	33%

SSD, or custom architectures. Notably, all works targeting semantic segmentation proposed their own custom network architecture. Similarly, Fig. 7(b) reveals the decision-making process for power line fault diagnosis. Again, a large portion of the works targeted insulators, used UAVs as the imaging platform, and collected datasets of 1000–5000 samples. Component fault detection is more common in the literature than foreign object detection. While YOLO, RCNN, and SSD have been used on multiple occasions for fault detection, most works propose custom network architectures.

11. Current challenges and future directions

In the rapidly evolving field of automated power line inspection, significant advancements have been made. However, several critical challenges still persist that need to be addressed to further advance the state of the art. This section outlines these challenges and proposes potential future research directions.

11.1. Edge-cloud deployment challenges

The deployment of deep learning models for power line inspection faces significant architectural challenges in both cloud-centric and edge-based approaches. Traditional cloud computing, while offering substantial computational resources, suffers from high latency, excessive bandwidth consumption, and significant communication costs when processing the massive amount of visual data generated by inspection devices [60,140]. Edge computing attempts to address these limitations by bringing computation closer to data sources, but edge devices typically lack the computational capacity to run sophisticated deep learning models effectively [141]. This limitation is particularly critical in power line inspection, where models must detect subtle defects and anomalies with high accuracy, creating a fundamental tension between model complexity and computational efficiency.

Edge-cloud fusion architectures present a promising direction to address these challenges by combining the strengths of both approaches. This fusion has led to the emergence of effective two-stage detection systems, where lightweight models at the edge perform initial coarse screening, followed by refined secondary recognition using more sophisticated models in the cloud [60]. In this hierarchical framework, edge devices can efficiently filter and pre-process data, while cloud resources handle detailed analysis, significantly reducing data transmission while maintaining high accuracy [59,60,136]. Future research should focus on several key areas to advance this fusion approach: developing more efficient model compression techniques for edge deployment [74], improving communication protocols for edge-cloud interaction [141], creating adaptive frameworks for dynamic resource allocation [140], and investigating federated learning [142] approaches for collaborative model training. These advancements could enable more efficient and reliable power line inspection systems while maintaining the accuracy needed for critical infrastructure monitoring.

11.2. Multimodal imaging and fusion

Power line inspection research has predominantly relied on optical imaging, with our comprehensive review revealing that only around 8% of published works utilize other imaging modalities. This heavy dependence on optical imaging persists despite the known limitations

of visible spectrum cameras in various environmental conditions and their inability to detect certain types of faults. Limited studies exploring alternative modalities demonstrated the complementary capabilities of infrared and ultraviolet imaging for detecting corona discharge and heating associated with leakage current flow in composite insulators [29,32,105]. Their research highlighted how different imaging modalities can provide unique insights — with IR imaging excelling at detecting heat distribution patterns from current leakage, while UV imaging proved effective for visualizing corona discharge phenomena. However, such multimodal approaches remain vastly underutilized in power line inspection literature, despite their proven effectiveness in other domains such as medical imaging, remote sensing, and defense applications [143,144].

The future of power line inspection could benefit significantly from greater adoption of multimodal imaging approaches. Recent developments in image fusion, as outlined in the comprehensive review by Karim et al. [143], demonstrate how combining multiple imaging modalities can provide more comprehensive information about the real world than any single modality alone. While optical images excel at providing high spatial resolution and clear texture details, other modalities like infrared can detect thermal anomalies, and ultraviolet can reveal corona discharge patterns [145]. By integrating multiple modalities through advanced fusion techniques — ranging from conventional transform-based methods to emerging deep learning architectures [146,147] — future inspection systems could achieve more robust fault detection capabilities.

11.3. Lack of data availability

The scarcity of publicly available datasets remains a significant challenge in deploying deep learning for power line inspection [148]. Power line components and scenarios require vast, varied datasets for effective training. While researchers often create custom datasets [103], data protection regulations frequently prevent public sharing, as discussed in Section 6.

Several approaches show promise in addressing this challenge. Synthetic data generation using GANs [149] or Denoising Diffusion Models can create diverse, realistic power line images. For example, a recent study [150] enhanced Cycle-GAN [151] with attention mechanisms to generate insulator defect images, demonstrating significant improvements in sample quality.

Self-supervised pretraining [152] offers another solution by leveraging unlabeled data to learn useful representations before fine-tuning on smaller labeled datasets. Additionally, few-shot learning approaches [153], particularly meta-learning techniques [154], enable models to learn from limited examples. Combining these methods with transfer learning could effectively address the data scarcity challenge while maintaining robust performance in real-world applications.

11.4. Data annotation

The problem of data annotation presents a significant challenge in power line inspection, particularly given the complexity of power line components and the subtleties of potential faults. While many excellent annotation tools exist [155–157], the process remains time-consuming and labor-intensive, requiring substantial human intervention for verification. The emergence of large foundation models like SAM (Segment

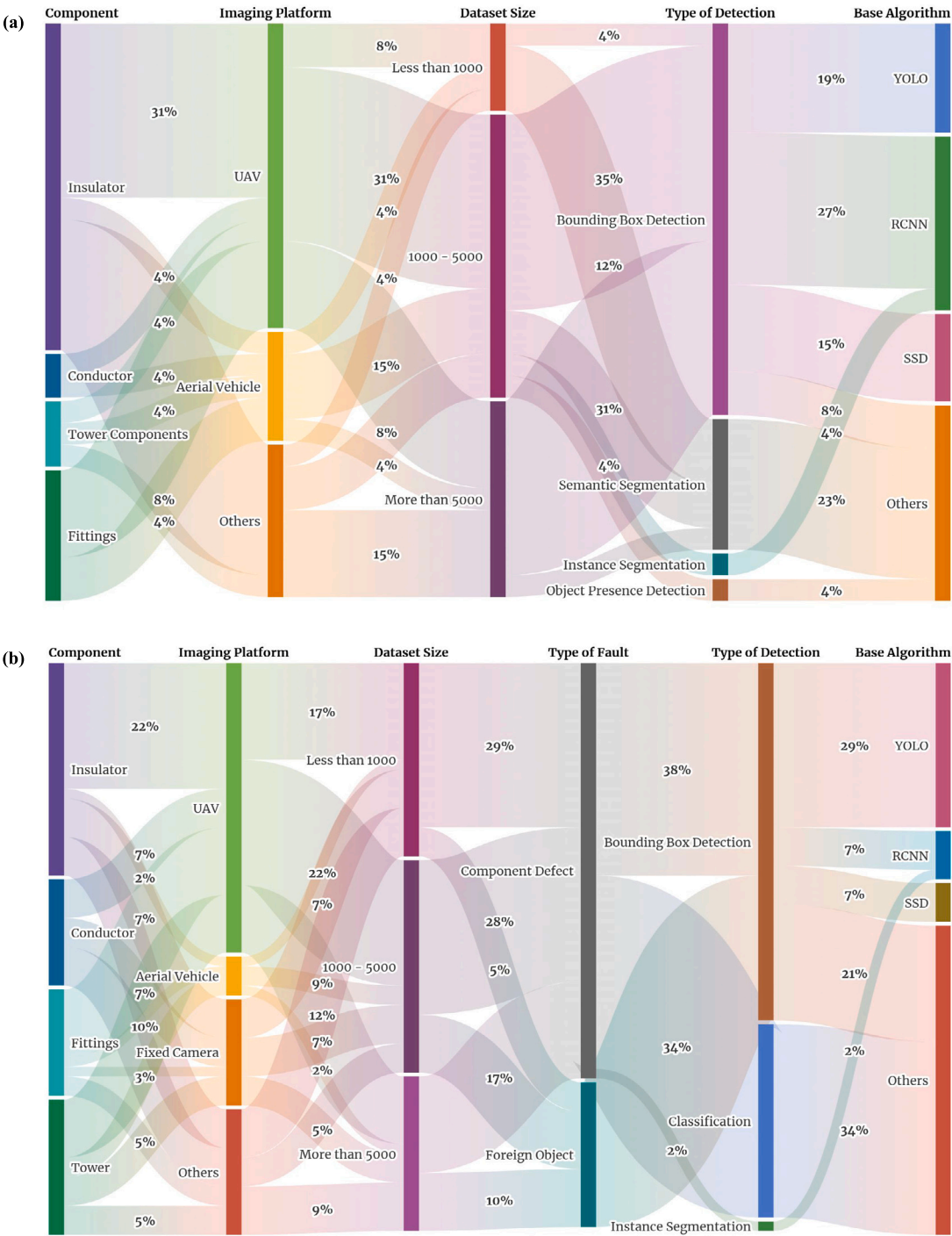


Fig. 7. (a) The inter-relation between the different aspects in power line component detection. (b) The inter-relation between the different aspects in power line fault diagnosis.

Anything Model) [158] offers new opportunities for streamlining this process through zero-shot segmentation capabilities, though these models require significant computational resources and careful adaptation to the power line domain. The integration of such models into existing annotation workflows necessitates consideration of domain-specific fine-tuning strategies to ensure reliable performance in power line inspection contexts.

Several promising approaches are being developed to address these challenges. Weakly supervised learning techniques [159] have shown potential in reducing labeling requirements, as demonstrated by Choi et al. [160] in their two-stage power line detection algorithm. The combination of these approaches with foundation models and specialized annotation tools could create more effective hybrid systems that leverage both general semantic understanding and domain-specific expertise. Additionally, self-supervised pretraining [152] can significantly reduce the dependency on labeled datasets by equipping models with a deep understanding of structural and contextual features inherent in power line images, thereby streamlining the annotation process while maintaining accuracy and reliability.

11.5. Very small components detection

Detecting small components like fittings, bolts, and fractures in power lines presents unique challenges due to their low resolution in images. These components, while critical for structural integrity, are often indistinguishable from complex backgrounds [122]. The prevalent use of UAVs introduces additional challenges: image blur from drone motion and inability to capture close-up images due to high voltage risks. Consequently, small components often occupy only a few pixels in the captured images [61].

Recent advances offer promising solutions. Transformer-based approaches [64] have shown success in reducing errors in small and occluded object detection through lightweight self-attention modules. Image super-resolution techniques [161] enhance resolution by inferring missing details and refining textures. Future research could focus on combining high-resolution imaging technology with advanced deep learning models, while improving multi-scale detection strategies and feature extraction methods for small objects [162].

11.6. Anomaly detection for unknown defects

Traditional supervised approaches often fail to identify novel power line defects absent from training data. Recent semi-supervised and unsupervised learning methods [163,164] offer solutions by learning normal patterns and detecting deviations.

Autoencoder-based methods with auxiliary anomaly localization enable end-to-end defect detection using only normal samples [165,166]. Feature embedding approaches using pre-trained networks reduce noise interference during reconstruction [167]. Memory-based methods that compare normal feature representations have improved detection of subtle anomalies [168], while combined reconstruction and discriminative training approaches enhance defect localization accuracy [169]. Future research should focus on developing efficient architectures suitable for edge deployment while maintaining detection accuracy.

12. Conclusion

This review examines the evolution of vision-based power line inspection through deep learning technologies. We explored various imaging platforms and techniques, from UAVs to X-ray imaging, analyzing their effectiveness for different inspection tasks. The review detailed how deep learning algorithms like YOLO, R-CNN, and SSD have advanced component detection and fault diagnosis, with particular success in insulator inspection. Current trends show promising

developments in edge-cloud fusion architectures and two-stage detection approaches, balancing computational efficiency with accuracy. However, challenges persist in data availability, annotation efficiency and unknown defect detection among others. Solutions are emerging through synthetic data generation, few-shot learning, and semi-automatic labeling techniques. Looking forward, the integration of edge-cloud computing, multi-modal fusion, and novel learning approaches suggests a trajectory towards more resilient and adaptive power line maintenance systems. These advancements promise to enhance the reliability and efficiency of power infrastructure inspection while reducing operational costs and safety risks.

CRediT authorship contribution statement

Md. Ahasan Atick Faisal: Writing – original draft, Formal analysis, Data curation, Conceptualization. **Imene Mecheter:** Writing – review & editing, Conceptualization. **Yazan Qiblawey:** Writing – review & editing, Conceptualization. **Javier Hernandez Fernandez:** Writing – review & editing, Project administration. **Muhammad E.H. Chowdhury:** Writing – review & editing, Validation, Supervision. **Serkan Kiranyaz:** Writing – review & editing, Validation, Supervision.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Google Gemini in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

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

Acknowledgments

This publication is supported by Iberdrola S.A. as part of its innovation department research activities. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Iberdrola Group. The open access publication cost is covered by the Qatar National Library.

Appendix A. Deep learning models for power line inspection

Historically, power line inspections relied heavily on manual labor and traditional image processing techniques, which often proved time-consuming and prone to human error [47]. With the adoption of deep learning models, a profound shift has occurred. These models have showcased their ability to learn intricate patterns and structures within images, allowing for the precise identification and localization of various power line components. This transformation is not merely theoretical; it has translated into tangible benefits for power grid operators and maintenance teams [47]. This, in turn, enhances the reliability of power transmission infrastructure and mitigates the risks of unexpected outages and associated economic and safety implications. In this section, key deep learning techniques that have found application in power line component detection and fault diagnosis have been explored. The following sections delve into their advantages and use cases, shedding light on how these techniques are pushing the boundaries of what is achievable in the field of power line maintenance.

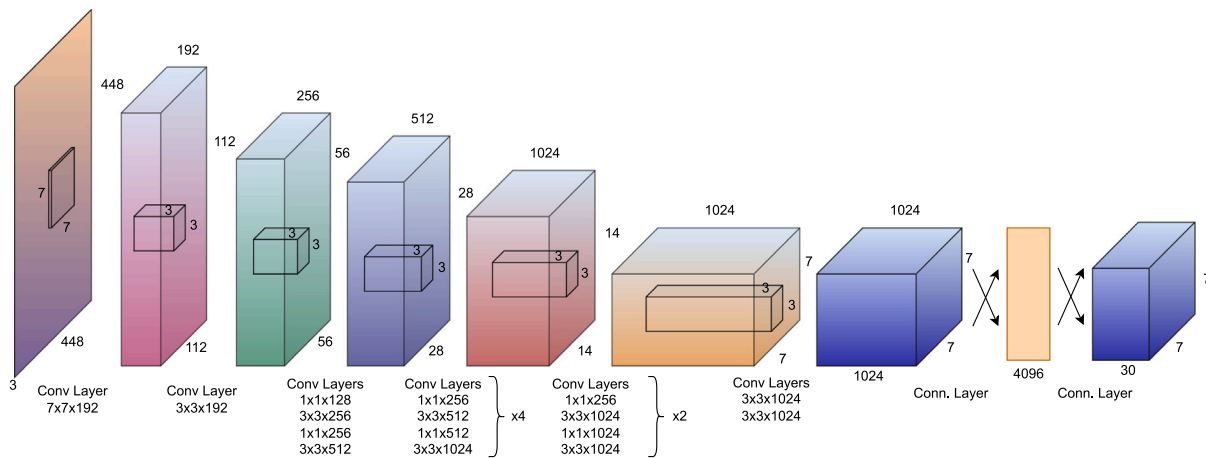


Fig. A.8. The architecture of the original YOLO network [21].

A.1. You Only Look Once (YOLO)

You Only Look Once (YOLO) [21] is a revolutionary real-time object detection system that has gained widespread recognition in computer vision applications. It stands out for its ability to swiftly process images and directly predict bounding boxes and class probabilities in a single evaluation. YOLO's efficiency and accuracy make it a compelling choice for power line component detection.

The original YOLO model (Fig. A.8) introduced the concept of end-to-end object detection in real-time. It divides an image into a grid and predicts bounding boxes and class probabilities for objects within each grid cell. YOLOv2 [48] brought improvements in accuracy and flexibility by employing anchor boxes and multi-scale detection. It was also trained on a broader dataset, allowing it to detect a wide range of objects. YOLOv3 [49] further enhanced the model’s accuracy by utilizing a three-stage detection process and the addition of more anchor boxes. YOLOv4 [50] introduced several architectural improvements, including the integration of the CSPDarknet53 backbone, PANet, and SAM block [50]. These enhancements resulted in better performance in complex scenarios and more accurate component detection. The recently proposed YOLOv8 is built on top of the previous YOLO versions and designed to be faster, and more accurate [51].

In power line inspection, YOLO is utilized to identify and classify various components, such as insulators, dampers, pin bolts, conductor wires, and fittings [57–59]. Its speed and real-time capabilities are particularly advantageous when inspecting extensive stretches of power transmission infrastructure. YOLO’s speed is one of its defining features. It operates at a significantly high frame rate, often exceeding real-time requirements [67]. This speed advantage is particularly valuable in power line inspections, where rapid assessments of extensive infrastructure can be vital. The ability to process images quickly allows for the efficient identification of components, even in cases of frequent data acquisition through aerial surveys. It’s worth noting that YOLO’s real-time performance might require a sufficiently powerful hardware setup [170], but the trade-off between accuracy and processing speed can often be optimized according to specific project requirements.

A.2. Region-based CNNs (R-CNN, Fast R-CNN, Faster R-CNN)

Region-Based Convolutional Neural Networks (CNNs) [20] represent a family of object detection models that focuses on detection accuracy while compromising on speed and complexity. This family includes R-CNN [20], Fast R-CNN [52], and Faster R-CNN [53], each building upon the other to improve efficiency and accuracy in power line component detection. R-CNN slides an image window, extracts features for each window, and then classifies and refines bounding boxes for potential objects within those windows. Fast R-CNN (Fig. A.9)

improves on R-CNN by processing the entire image at once with a single CNN to extract features, making it significantly faster. Faster R-CNN further refines the model by introducing the Region Proposal Network (RPN) [53] for generating region proposals. This innovation results in a more streamlined and faster detection process. These models have been applied in power line component detection to locate and classify insulators [60], dampers [61], pin bolts [61], conductor wires [62], and other elements.

Region-based CNNs excel in precisely localizing objects within images, making them suitable for power line component identification. Fast R-CNN and Faster R-CNN integrate region proposal and feature extraction steps, enhancing processing efficiency. However, Training and fine-tuning region-based CNNs require substantial computational resources and a large labeled dataset. These models may need hardware acceleration for real-time performance [68].

A.3. Single Shot MultiBox Detectors (SSD)

Single Shot Detectors (SSD) [22] is another object detection algorithm that combines high-speed processing with robust detection capabilities. SSD is designed for real-time object detection, eliminating the need for a separate region proposal step and streamlining all computations into a single network. SSD is employed to rapidly identify and classify various components, such as insulators [63], fittings [19], conductor wires [19], and other crucial infrastructure elements.

SSD's primary advantage is its ability to process images rapidly while keeping a high enough accuracy [69]. This characteristic is essential for real-time inspections, particularly in scenarios where assessments of extensive power line infrastructure are required. SSD's architecture (Fig. A.10) integrates all detection computations into a single network, eliminating the need for multiple stages, which simplifies implementation and results in efficient performance. While SSD offers impressive speed and efficiency, its performance may be compromised compared to models like Faster-RCNN [69]. Balancing speed and accuracy demands may require optimization for specific hardware configurations and project constraints.

A.4. Transformer architectures

Transformer architectures have revolutionized various domains in deep learning, originally emerging as a powerful tool in natural language processing. Introduced by Vaswani et al. [82], the transformer model is based on self-attention mechanisms, enabling it to capture long-range dependencies within data. Unlike traditional convolutional neural networks (CNNs), which are limited by their localized receptive fields, transformers excel in modeling global context, making them highly effective for complex tasks in computer vision, including power line inspection.

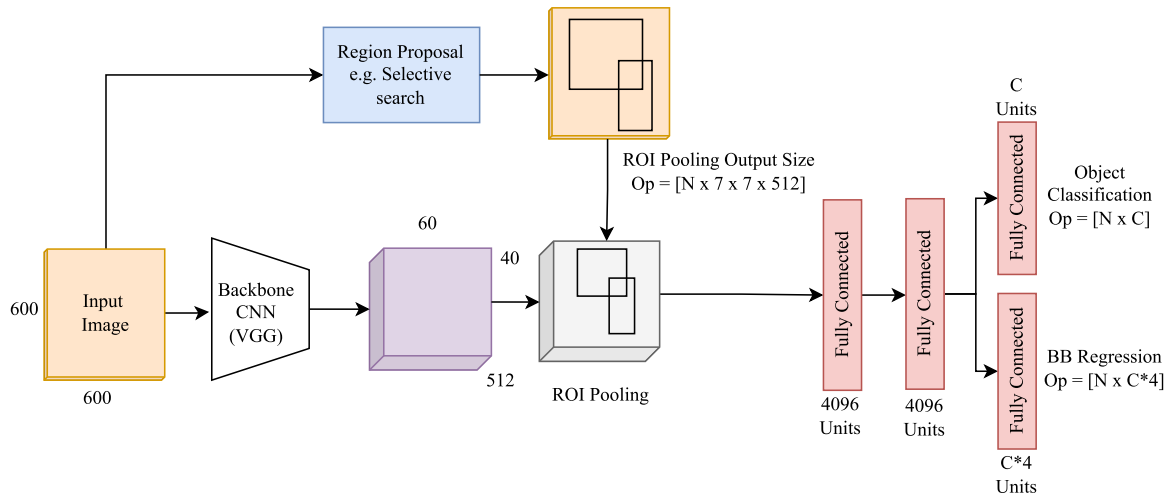


Fig. A.9. Simplified architecture of the Fast R-CNN network.

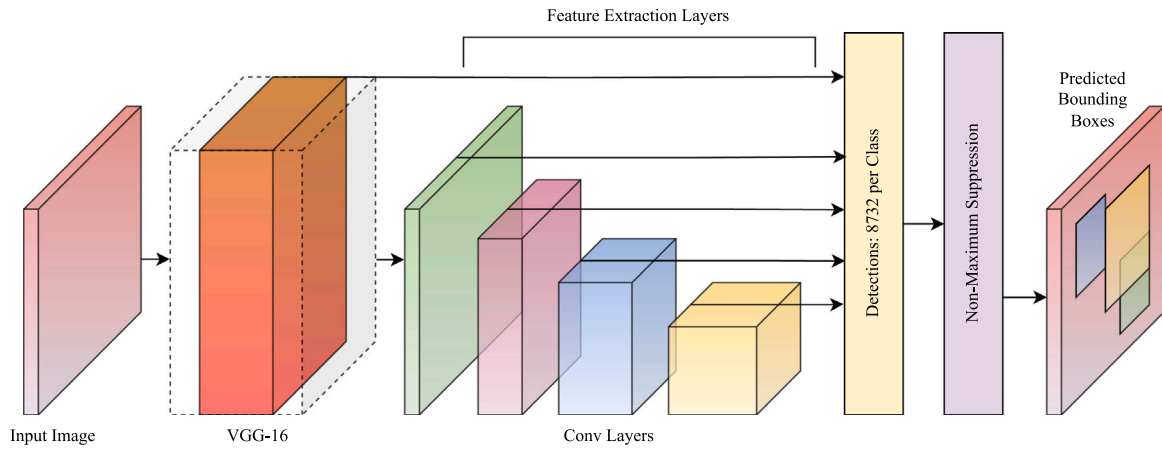


Fig. A.10. Simplified architecture of the Single Shot Multibox Detector (SSD) Network.

A.4.1. Vision Transformers (ViT)

Vision Transformers (ViT) [54] marked a paradigm shift by applying the transformer architecture directly to image data. ViTs divide an image into a sequence of patches, each treated similarly to tokens in a language model. These patches are then processed through multiple layers of self-attention, allowing the model to learn intricate relationships between different parts of the image. Although we could not find any research work on power line inspection that utilizes ViTs, they have potential in tasks requiring detailed analysis of visual data. Their ability to capture global information makes them particularly suitable for identifying subtle anomalies in power line components, such as bolt defects, micro-cracks in insulators or structural defect of conductor wires [70]. However, ViTs require a very large dataset to train on to get rid of the inductive bias [54].

A.4.2. Swin Transformers

Swin Transformers [55], or Shifted Window Transformers, build on the concept of ViTs by introducing a hierarchical structure that allows the model to operate at multiple scales. This architecture divides the image into non-overlapping windows and applies self-attention within each window. To capture cross-window information, the windows are shifted between layers, enabling the model to build a more comprehensive understanding of the image. The multi-scale feature representation of Swin Transformers is advantageous, especially in scenarios where defects or components vary in size. Swin Transformers can effectively manage high-resolution images, making them ideal for detecting and

localizing faults in expansive power transmission networks, where both small and large defects need to be identified with precision [64].

A.4.3. Detection Transformers (DETRs)

Detection Transformers (DETRs) [56] integrate the transformer architecture into object detection tasks, offering an end-to-end approach that simplifies the traditional detection pipeline. The architecture of the DETR network is shown in Fig. A.11. DETRs eliminate the need for anchor boxes and region proposals, which are common in conventional object detection models. Instead, they leverage the transformer's attention mechanism to directly predict object bounding boxes and class labels. Their ability to model complex interactions between objects within an image enhances detection accuracy, particularly in cluttered or complex scenes typical of power line infrastructure. Additionally, DETRs are robust to variations in object scale and orientation, which are common challenges in aerial imagery used for inspecting power lines [65,66].

A.5. Classification algorithms

Classification algorithms, particularly those pretrained on large datasets like ImageNet [71], have demonstrated remarkable capabilities in recognizing diverse patterns and anomalies in power line components. These algorithms excel in scenarios where pinpointing the exact location of a fault is unnecessary, and the primary goal is simply to determine whether a fault exists. When provided with an

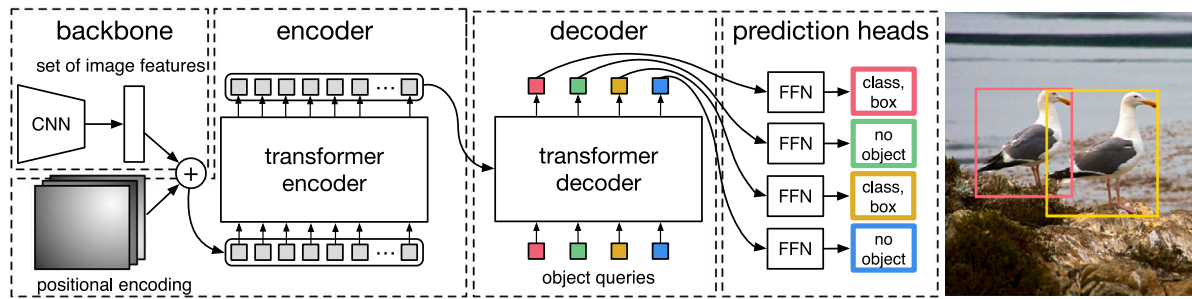


Fig. A.11. The simplified block diagram of the original DETR network [56].

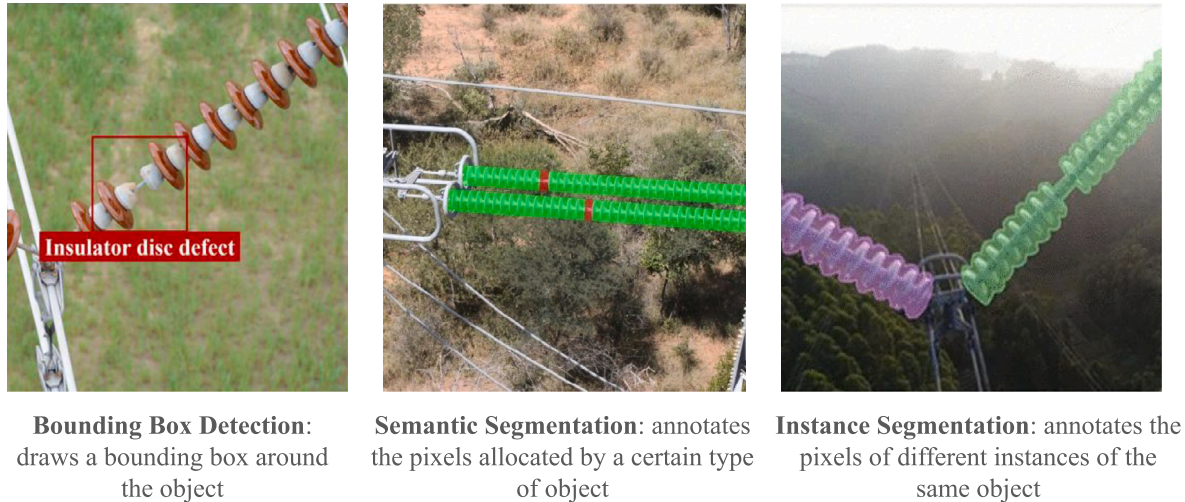


Fig. B.12. Comparison between different image segmentation techniques [86,87].

image of power line infrastructure or a specific segment of a power line component, these algorithms are capable of classifying the image as either faulty or in good condition. Key algorithms in this domain include ResNet, VGG, MobileNet and EfficientNet each bringing unique strengths to the table.

ResNet (Residual Networks) [72], a pivotal model in deep learning, introduced the concept of residual learning to ease the training of very deep networks. It employs “skip connections” to jump over some layers, effectively addressing the vanishing gradient problem. In power line inspection, ResNet’s ability to learn from a vast depth of layers makes it exceptionally good at recognizing complex patterns, crucial for identifying subtle anomalies in power lines [60,76,77].

The **VGG (Visual Geometry Group)** [73] network stands out for its simplicity and depth, with a uniform architecture that stacks convolutional layers directly on top of each other. This design, while computationally intensive, offers excellent feature extraction capabilities. For power line inspections, VGG’s depth helps in capturing intricate details necessary for accurate component classification and fault detection [78]. However, its performance may be compromised compared to other state-of-the-art models that came out in recent years [171].

MobileNet [74] architectures are designed for efficiency, making them ideal for use in mobile and edge computing scenarios. Their streamlined design, based on depthwise separable convolutions [172], allows for reduced computational load while maintaining high accuracy. In power line inspection, especially those conducted via drones or handheld devices, MobileNet’s lightweight nature enables rapid, on-site processing of images for real-time analysis [60,67,79].

EfficientNet [75] represents a new scaling method for neural networks, which uniformly scales all dimensions of depth, width, and resolution with a set of fixed scaling coefficients. This balanced scaling

results in a network that achieves state-of-the-art accuracy with a lower computational cost [171]. In power line inspection, EfficientNet can be particularly useful for processing high-resolution images effectively, allowing for detailed and accurate identification of line defects and deterioration [80,81].

The above-mentioned object detection and classification networks are often mixed with each other to design powerful networks that often have superior performance to the original networks [114]. Notably in recent years, the attention mechanism [82] and its utilization in segmentation and classification algorithms has gained widespread popularity. Its initial application was by the Google DeepMind team in 2014, where they integrated an attention module into an RNN model for image classification [54]. The potency of self-attention networks, particularly in capturing long-distance dependencies and contextual information, has made them a staple in machine vision tasks like image segmentation and classification. Cao et al. [76] innovatively used attention-guided multipath features to reconcile the contradictory needs between feature map resolution and the receptive field for high-resolution inputs. Furthermore, the introduction of the attention-RPN module by Fan et al. [83] for small sample target detection, and the combination of global attention and local restructuring by Kong et al. [84], exemplify the versatility of attention mechanisms. These adaptations enable the collection of task-oriented features across different spatial locations and scales, harnessing both global and local contexts to enhance the accuracy and efficiency of object detection.

Appendix B. Computer vision tasks in power line inspection

Various object detection techniques are employed in computer vision to automate this task, each with its strengths and application

scenarios. This section discusses key methods like bounding box detection, semantic segmentation, and instance segmentation. A visual comparison between these techniques has been shown in Fig. B.12.

B.1. Bounding Box Detection

Bounding Box Detection is a primary object detection method where a box is drawn around each object of interest in an image, marking its location and extent. It is straightforward to implement and computationally less demanding. This method is well-suited for real-time applications due to its relatively fast processing time often reaching over 80 Frames Per Second (FPS) [85]. Bounding box detection is typically used to identify and locate larger power line components such as towers, insulators, and dampers.

B.2. Semantic segmentation

Semantic segmentation involves the partitioning of an image into segments, where each pixel is classified into a predefined category. This method is capable of producing detailed component-wise masks, which are beneficial for understanding the scene at a pixel level. It provides a precise outline of the components, which is crucial for assessing their condition. Semantic segmentation allows for a comprehensive analysis of the scene by understanding the relationship between different components. This technique is particularly useful for distinguishing between different types of insulators, conductor wires, and vegetation encroachment.

B.3. Instance segmentation

Instance segmentation goes a step beyond semantic segmentation by not only separating the background from the foreground but also differentiating between individual objects of the same class. It is capable of identifying and delineating each instance of multiple objects of the same type. This method excels in scenarios where components are close together or overlapping. Instance segmentation is essential when dealing with dense power line components, such as closely spaced insulators or bundled conductor wires, to assess each component separately.

Data availability

No data was used for the research described in the article.

References

- [1] Mitchell JW. Power line failures and catastrophic wildfires under extreme weather conditions. *Eng Fail Anal* 2013;35:726–35. <http://dx.doi.org/10.1016/j.engfailanal.2013.07.006>.
- [2] The link between power lines and wildfires. 2018.
- [3] Salim NA, Othman MM, Jasni J, Musirin I, Serwan MS. Modeling and evaluating the customer interruption cost due to dynamic electrical power and energy failure. *Int J Electr Power Energy Syst* 2018;103:603–10. <http://dx.doi.org/10.1016/j.ijepes.2018.06.033>.
- [4] Nguyen VN, Jenssen R, Roverso D. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *Int J Electr Power Energy Syst* 2018;99:107–20. <http://dx.doi.org/10.1016/j.ijepes.2017.12.016>.
- [5] Liu Y, Shi J, Liu Z, Huang J, Zhou T. Two-layer routing for high-voltage powerline inspection by cooperated ground vehicle and drone. *Energies* 2019;12(7):1385. <http://dx.doi.org/10.3390/en12071385>, Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.
- [6] Matikainen L, Lehtomäki M, Ahokas E, Hyypä J, Karjalainen M, Jaakkola A, Kukko A, Heinonen T. Remote sensing methods for power line corridor surveys. *ISPRS J Photogramm Remote Sens* 2016;119:10–31. <http://dx.doi.org/10.1016/j.isprsjrs.2016.04.011>.
- [7] Yang L, Fan J, Liu Y, Li E, Peng J, Liang Z. A review on state-of-the-art power line inspection techniques. *IEEE Trans Instrum Meas* 2020;69(12):9350–65. <http://dx.doi.org/10.1109/TIM.2020.3031194>.
- [8] Martinez C, Sampedro C, Chauhan A, Collumeau JF, Campoy P. The power line inspection software (PoLIS): A versatile system for automating power line inspection. *Eng Appl Artif Intell* 2018;71:293–314. <http://dx.doi.org/10.1016/j.engappai.2018.02.008>.
- [9] Liu X, Miao X, Jiang H, Chen J. Data analysis in visual power line inspection: An in-depth review of deep learning for component detection and fault diagnosis. *Annu Rev Control* 2020;50:253–77. <http://dx.doi.org/10.1016/j.arcontrol.2020.09.002>.
- [10] Reddy MJB, B KC, Mohanta DK. Condition monitoring of 11 kV distribution system insulators incorporating complex imagery using combined DOST-SVM approach. *IEEE Trans Dielectr Electr Insul* 2013;20(2):664–74. <http://dx.doi.org/10.1109/TDEI.2013.6508770>.
- [11] Zhao Z, Liu N, Wang L. Localization of multiple insulators by orientation angle detection and binary shape prior knowledge. *IEEE Trans Dielectr Electr Insul* 2015;22(6):3421–8. <http://dx.doi.org/10.1109/TDEI.2015.004741>.
- [12] Wu Q, An J, Lin B. A texture segmentation algorithm based on PCA and global minimization active contour model for aerial insulator images. *IEEE J Sel Top Appl Earth Obs Remote Sens* 2012;5(5):1509–18. <http://dx.doi.org/10.1109/JSTARS.2012.2197672>.
- [13] Chen M, Tian Y, Xing S, Li Z, Li E, Liang Z, Guo R. Environment perception technologies for power transmission line inspection robots. *J Sens* 2021;2021(1):5559231.
- [14] Yang L, Fan J, Liu Y, Li E, Peng J, Liang Z. A review on state-of-the-art power line inspection techniques. *IEEE Trans Instrum Meas* 2020;69(12):9350–65.
- [15] Sundaram KM, Hussain A, Sanjeevikumar P, Holm-Nielsen JB, Kaliappan VK, Santhoshi BK. Deep learning for fault diagnostics in bearings, insulators, PV panels, power lines, and electric vehicle applications—The state-of-the-art approaches. *IEEE Access* 2021;9:41246–60.
- [16] Ruszcak B, Michalski P, Tomaszewski M. Overview of image datasets for deep learning applications in diagnostics of power infrastructure. *Sensors* 2023;23(16):7171.
- [17] Xu B, Zhao Y, Wang T, Chen Q. Development of power transmission line detection technology based on unmanned aerial vehicle image vision. *SN Appl Sci* 2023;5(3):72.
- [18] Foudeh HA, Luk PC-K, Whidborne JF. An advanced unmanned aerial vehicle (UAV) approach via learning-based control for overhead power line monitoring: A comprehensive review. *IEEE Access* 2021;9:130410–33.
- [19] Nguyen VN, Jenssen R, Roverso D. Intelligent monitoring and inspection of power line components powered by UAVs and deep learning. *IEEE Power Energy Technol Syst J* 2019;6(1):11–21. <http://dx.doi.org/10.1109/JPEITS.2018.2881429>.
- [20] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014. <http://dx.doi.org/10.48550/arXiv.1311.2524>.
- [21] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. 2016. <http://dx.doi.org/10.48550/arXiv.1506.02640>.
- [22] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, Berg AC. Ssd: Single shot multibox detector. In: *Computer vision—ECCV 2016: 14th European conference, Amsterdam, the Netherlands, October 11–14, 2016, proceedings, part i* 14. Springer; 2016, p. 21–37.
- [23] Zhang Y, Yuan X, Li W, Chen S. Automatic power line inspection using UAV images. *Remote Sens* 2017;9(8):824. <http://dx.doi.org/10.3390/rs9080824>, Number: 8 Publisher: Multidisciplinary Digital Publishing Institute.
- [24] Zhou F, Jin W, Zheng Z, Mou F, Li Z, Ma Y, Wei B, Huang S, Wang Q. Insulator detection for high-resolution satellite images based on deep learning. *IEEE Geosci Remote Sens Lett* 2023;20:1–5. <http://dx.doi.org/10.1109/LGRS.2023.3251372>.
- [25] Alhassan AB, 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* 2020;118:105862.
- [26] Ekren N, Karagöz Z, Şahin M. A review of line suspended inspection robots for power transmission lines. *J Electr Eng Technol* 2024;19(4):2549–83.
- [27] Zhang M, Song Z, Yang J, Gao M, Hu Y, Yuan C, Jiang Z, Cheng W. Study on the enhancement method of online monitoring image of dense fog environment with power lines in smart city. *Front Neurobotics* 2023;16.
- [28] Zhang Z-D, Zhang B, Lan Z-C, Liu H-C, Li D-Y, Pei L, Yu W-X. FInet: An insulator dataset and detection benchmark based on synthetic fog and improved YOLOv5. *IEEE Trans Instrum Meas* 2022;71:1–8. <http://dx.doi.org/10.1109/TIM.2022.3194909>.
- [29] Singh L, Alam A, Kumar KV, Kumar D, Kumar P, Jaffery ZA. Design of thermal imaging-based health condition monitoring and early fault detection technique for porcelain insulators using machine learning. *Environ Technol Innov* 2021;24:102000. <http://dx.doi.org/10.1016/j.eti.2021.102000>.
- [30] Jaffery ZA, Dubey AK. Design of early fault detection technique for electrical assets using infrared thermograms. *Int J Electr Power Energy Syst* 2014;63:753–9. <http://dx.doi.org/10.1016/j.ijepes.2014.06.049>.
- [31] Hu B, Ma L-X, Yuan S-J, Yang B. New corona ultraviolet detection system and fault location method. In: *2012 China international conference on electricity distribution*. 2012, p. 1–4. <http://dx.doi.org/10.1109/CICED.2012.6508475>.

- [32] Li Y, Yu F, Cai Q, Yuan K, Wan R, Li X, Qian M, Liu P, Guo J, Yu J, Zheng T, Yan H, Hou P, Feng Y, Wang S, Ding L. Image fusion of fault detection in power system based on deep learning. *Clust Comput* 2019;22(4):9435–43. <http://dx.doi.org/10.1007/s10586-018-2264-2>.
- [33] Zang C, He J, Xiaogen Y, Chen B, Lei H, Zhenglong J, Xinjie Z. Status and application foreground of ultraviolet technology on fault detection of power devices. In: 2008 international conference on condition monitoring and diagnosis. 2008, p. 122–5. <http://dx.doi.org/10.1109/CMD.2008.4580245>.
- [34] Wang F, Song G, Mao J, Li Y, Ji Z, Chen D, Song A. Internal defect detection of overhead aluminum conductor composite core transmission lines with an inspection robot and computer vision. *IEEE Trans Instrum Meas* 2023;72. <http://dx.doi.org/10.1109/TIM.2023.3265104>.
- [35] 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* 2021;130:106987.
- [36] Bergmann MA, Moreira LFR, Krohling B, Silveira TL, Jung CR, Tang J, Feitosa MV, Gomes RLB, Soares BN. An approach based on LiDAR and spherical images for automated vegetation inspection in urban power distribution lines. *IEEE Access* 2024.
- [37] 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* 2018;50(4):1486–98.
- [38] Voigt P, Von dem Bussche A. The eu general data protection regulation (gdpr). In: A practical guide. 1st ed.. vol. 10, Cham: Springer International Publishing; 2017, p. 10–5555, (3152676).
- [39] California consumer privacy act (CCPA). 2024, URL <https://oag.ca.gov/privacy/ccpa>.
- [40] Yetgin. Powerline image dataset (infrared-IR and visible light-VL). 2019, URL <https://data.mendeley.com/datasets/n6wrv4ry6v/8>.
- [41] Abdelfattah R, Wang X, Wang S. Ttpa: An aerial-image dataset for detection and segmentation of transmission towers and power lines. In: Proceedings of the Asian conference on computer vision. 2020, p. 601–18.
- [42] Diniz L, Santa Maria T, Pussente GA. Power transmission line dataset. 2021, <http://dx.doi.org/10.21227/t9qk-cn48>.
- [43] Diwakar T. (Recognizance - 2) power lines detection. 2021, URL <https://kaggle.com/competitions/recognizance-2>.
- [44] Vieira-e Silva ALB, de Castro Felix H, de Menezes Chaves T, Simões FPM, Teichrieb V, dos Santos MM, da Cunha Santiago H, Sgotti VAC, Neto HBDTL. STN PLAD: A dataset for multi-size power line assets detection in high-resolution UAV images. In: 2021 34th SIBGRAPI conference on graphics, patterns and images. SIBGRAPI, 2021, p. 215–22. <http://dx.doi.org/10.1109/SIBGRAPI54419.2021.00037>.
- [45] Savva A, Makrigiorgis R, Kolios P, Kyrkou C. Aerial power infrastructure detection dataset. 2023, <http://dx.doi.org/10.5281/zenodo.7781388>.
- [46] Shuang F, Han S, Li Y, Lu T. RSIn-dataset: An UAV-based insulator detection aerial images dataset and benchmark. *Drones* 2023;7(2). <http://dx.doi.org/10.3390/drones7020125>, URL <https://www.mdpi.com/2504-446X/7/2/125>.
- [47] Sundaram KM, Hussain A, Sanjeevikumar P, Holm-Nielsen JB, Kaliappan VK, Santhoshi BK. Deep learning for fault diagnostics in bearings, insulators, PV panels, power lines, and electric vehicle applications—The state-of-the-art approaches. *IEEE Access* 2021;9:41246–60. <http://dx.doi.org/10.1109/ACCESS.2021.3064360>.
- [48] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. 2016, <http://dx.doi.org/10.48550/arXiv.1612.08242>.
- [49] Redmon J, Farhadi A. YOLOv3: An incremental improvement. 2018, <http://dx.doi.org/10.48550/arXiv.1804.02767>.
- [50] Bochkovskiy A, Wang C-Y, Liao H-YM. YOLOv4: Optimal speed and accuracy of object detection. 2020, <http://dx.doi.org/10.48550/arXiv.2004.10934>.
- [51] Jocher G, Chaurasia A, Qiu J. YOLO by ultralytics. 2023.
- [52] Girshick R. Fast R-CNN. 2015, <http://dx.doi.org/10.48550/arXiv.1504.08083>.
- [53] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. 2016, <http://dx.doi.org/10.48550/arXiv.1506.01497>.
- [54] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, et al. An image is worth 16x16 words: Transformers for image recognition at scale. 2020, arXiv preprint [arXiv:2010.11929](https://arxiv.org/abs/2010.11929).
- [55] Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B. Swin transformer: Hierarchical vision transformer using shifted windows. 2021, [arXiv:2103.14030](https://arxiv.org/abs/2103.14030). URL <https://arxiv.org/abs/2103.14030>.
- [56] Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S. End-to-end object detection with transformers. In: European conference on computer vision. Springer; 2020, p. 213–29.
- [57] Sadykova D, Pernebayeva D, Bagheri M, James A. IN-YOLO: Real-time detection of outdoor high voltage insulators using UAV imaging. *IEEE Trans Power Deliv* 2019;35(3):1599–601.
- [58] Singh G, Stefenon S, Yow K-C. Interpretable visual transmission lines inspections using pseudo-prototypical part network. *Mach Vis Appl* 2023;34(3). <http://dx.doi.org/10.1007/s00138-023-01390-6>.
- [59] Zhang S, Wang J, Tong J, Zhang J, Zhang M. Cloud-edge fusion based abnormal object detection of power transmission lines using incremental learning. *IEEE Access* 2020;8:218694–701. <http://dx.doi.org/10.1109/ACCESS.2020.3037172>.
- [60] Wei B, Xie Z, Liu Y, Wen K, Deng F, Zhang P. Online monitoring method for insulator self-explosion based on edge computing and deep learning. *CSEE J Power Energy Syst* 2022;8(6):1684–96. <http://dx.doi.org/10.17775/CSEEJPES.2020.05910>.
- [61] Zhai Y, Yang X, Wang Q, Zhao Z, Zhao W. Hybrid knowledge R-CNN for transmission line multifitting detection. *IEEE Trans Instrum Meas* 2021;70:1–12. <http://dx.doi.org/10.1109/TIM.2021.3096600>.
- [62] Rong S, He L, Du L, Li Z, Yu S. Intelligent detection of vegetation encroachment of power lines with advanced stereovision. *IEEE Trans Power Deliv* 2021;36(6):3477–85. <http://dx.doi.org/10.1109/TPWRD.2020.3043433>.
- [63] 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* 2019;7:9945–56. <http://dx.doi.org/10.1109/ACCESS.2019.2891123>.
- [64] Dong K, Shen Q, Wang C, Dong Y, Liu Q, Lu Z, Lu Z. Improved swin transformer-based defect detection method for transmission line patrol inspection images. *Evol Intell* 2023. <http://dx.doi.org/10.1007/s12065-023-00837-z>.
- [65] Zhang K, Lou W, Wang J, Zhou R, Guo X, Xiao Y, Shi C, Zhao Z. PA-DETR: End-to-end visually indistinguishable bolt defects detection method based on transmission line knowledge reasoning. *IEEE Trans Instrum Meas* 2023;72:1–14. <http://dx.doi.org/10.1109/TIM.2023.3282302>.
- [66] Jain N, Bedi J, Anand A, Godara S. A transfer learning architecture to detect faulty insulators in powerlines. *IEEE Trans Power Deliv* 2024.
- [67] Li H, Liu L, Du J, Jiang F, Guo F, Hu Q, Fan L. An improved YOLOv3 for foreign objects detection of transmission lines. *IEEE Access* 2022;10:45620–8. <http://dx.doi.org/10.1109/ACCESS.2022.3170696>.
- [68] Bharati P, Pramanik A. Deep learning techniques—R-CNN to mask R-CNN: a survey. In: Computational intelligence in pattern recognition: Proceedings of CIPR 2019. Springer; 2020, p. 657–68.
- [69] Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y, Guadarrama S, et al. Speed/accuracy trade-offs for modern convolutional object detectors. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 7310–1.
- [70] Han K, Wang Y, Chen H, Chen X, Guo J, Liu Z, Tang Y, Xiao A, Xu C, Xu Y, et al. A survey on vision transformer. *IEEE Trans Pattern Anal Mach Intell* 2022;45(1):87–110.
- [71] 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. 2009, p. 248–55. <http://dx.doi.org/10.1109/CVPR.2009.5206848>.
- [72] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2015, *CoRR* [abs/1512.03385](https://arxiv.org/abs/1512.03385). [arXiv:1512.03385](https://arxiv.org/abs/1512.03385). URL <http://arxiv.org/abs/1512.03385>.
- [73] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014, arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556).
- [74] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. 2017, arXiv preprint [arXiv:1704.04861](https://arxiv.org/abs/1704.04861).
- [75] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR; 2019, p. 6105–14.
- [76] Cao Y, Xu H, Su C, Yang Q. Accurate glass insulators defect detection in power transmission grids using aerial image augmentation. *IEEE Trans Power Deliv* 2023;38(2):956–65. <http://dx.doi.org/10.1109/TPWRD.2022.3202958>.
- [77] Luo P, Wang B, Wang H, Ma F, Ma H, Wang L. An ultrasmall bolt defect detection method for transmission line inspection. *IEEE Trans Instrum Meas* 2023;72:1–12. <http://dx.doi.org/10.1109/TIM.2023.3241994>.
- [78] Stefenon SF, Singh G, Yow K-C, Cimatti A. Semi-ProtoNet deep neural network for the classification of defective power grid distribution structures. *Sensors* 2022;22(13):4859. <http://dx.doi.org/10.3390/s22134859>, Number: 13 Publisher: Multidisciplinary Digital Publishing Institute.
- [79] Qiu Z, Zhu X, Liao C, Qu W, Yu Y. A lightweight YOLOv4-EDAM model for accurate and real-time detection of foreign objects suspended on power lines. *IEEE Trans Power Deliv* 2023;38(2):1329–40. <http://dx.doi.org/10.1109/TPWRD.2022.3213598>.
- [80] Odo A, McKenna S, Flynn D, Vorstius JB. Aerial image analysis using deep learning for electrical overhead line network asset management. *IEEE Access* 2021;9:146281–95. <http://dx.doi.org/10.1109/ACCESS.2021.3123158>.
- [81] Li Y, Li Z, Liu Y, Sheng G, Jiang X. Pin bolt state identification using cascaded object detection networks. *Front Energy Res* 2022;10. <http://dx.doi.org/10.3389/fenrg.2022.813945>.
- [82] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. *Adv Neural Inf Process Syst* 2017;30.
- [83] Fan Q, Zhuo W, Tai Y. Few-shot object detection with attention-RPN and multi-relation detector. 2019, *CoRR* [abs/1908.01998](https://arxiv.org/abs/1908.01998). [arXiv:1908.01998](https://arxiv.org/abs/1908.01998). URL <http://arxiv.org/abs/1908.01998>.

- [84] Kong L, Zhu X, Wang G. Context semantics for small target detection in large-field images with two cascaded faster R-CNNs. In: Journal of physics: conference series. vol. 1069, IOP Publishing; 2018, 012138. <http://dx.doi.org/10.1088/1742-6596/1069/1/012138>.
- [85] Ge Z, Li H, Yang R, Liu H, Pei S, Jia Z, Ma Z. Bird's nest detection algorithm for transmission lines based on deep learning. In: 2022 3rd international conference on computer vision, image and deep learning & international conference on computer engineering and applications. CVIDL & ICCEA, 2022, p. 417–20. <http://dx.doi.org/10.1109/CVIDLICCEA56201.2022.9824057>.
- [86] Chen Y, Liu H, Chen J, Hu J, Zheng E. Insu-YOLO: An insulator defect detection algorithm based on multiscale feature fusion. Electronics 2023;12(15). <http://dx.doi.org/10.3390/electronics12153210>.
- [87] Oberweger M, Wendel A, Bischof H. Visual recognition and fault detection for power line insulators. In: 19th computer vision winter workshop. 2014, p. 1–8.
- [88] Kang G, Gao S, Yu L, Zhang D. Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning. IEEE Trans Instrum Meas 2019;68(8):2679–90. <http://dx.doi.org/10.1109/TIM.2018.2868490>.
- [89] Zhao Z, Qi H, Fan X, Xu G, Qi Y, Zhai Y, Zhang K. Image representation method based on relative layer entropy for insulator recognition. Entropy 2020;22(4):419.
- [90] Zhang D, Gao S, Yu L, Kang G, Wei X, Zhan D. DefGAN: Defect detection GANs with latent space pitting for high-speed railway insulator. IEEE Trans Instrum Meas 2021;70:1–10. <http://dx.doi.org/10.1109/TIM.2020.3038008>.
- [91] Waleed D, Mukhopadhyay S, Tariq U, El-Hag AH. Drone-based ceramic insulators condition monitoring. IEEE Trans Instrum Meas 2021;70:1–12. <http://dx.doi.org/10.1109/TIM.2021.3078538>.
- [92] Antwi-Bekoe E, Liu G, Ainam J-P, Sun G, Xie X. A deep learning approach for insulator instance segmentation and defect detection. Neural Comput Appl 2022;34(9):7253–69. <http://dx.doi.org/10.1007/s00521-021-06792-z>.
- [93] Shuang F, Han S, Li Y, Lu T. RSIn-dataset: An UAV-based insulator detection aerial images dataset and benchmark. Drones 2023;7(2):125. <http://dx.doi.org/10.3390/drones7020125>, Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.
- [94] Zhai Y, Wang Q, Yang X, Zhao Z, Zhao W. Multi-fitting detection on transmission line based on cascade reasoning graph network. IEEE Trans Power Deliv 2022;37(6):4858–68. <http://dx.doi.org/10.1109/TPWRD.2022.3161124>.
- [95] Zhang H, Wu L, Chen Y, Chen R, Kong S, Wang Y, Hu J, Wu J. Attention-guided multitask convolutional neural network for power line parts detection. IEEE Trans Instrum Meas 2022;71:1–13. <http://dx.doi.org/10.1109/TIM.2022.3162615>.
- [96] Huang X, Wu Y, Zhang Y, Li B. Structural defect detection technology of transmission line damper based on UAV image. IEEE Trans Instrum Meas 2023;72:1–14. <http://dx.doi.org/10.1109/TIM.2022.3228008>.
- [97] Yang L, Fan J, Xu S, Li E, Liu Y. Vision-based power line segmentation with an attention fusion network. IEEE Sens J 2022;22(8):8196–205. <http://dx.doi.org/10.1109/JSEN.2022.3157336>.
- [98] Yang L, Kong S, Deng J, Li H, Liu Y. DRA-net: A dual-branch residual attention network for pixelwise power line detection. IEEE Trans Instrum Meas 2023;72:1–13. <http://dx.doi.org/10.1109/TIM.2023.3259047>.
- [99] 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 2020;8:182105–16. <http://dx.doi.org/10.1109/ACCESS.2020.3027850>.
- [100] Wang Y, Li Q, Chen B. Image classification towards transmission line fault detection via learning deep quality-aware fine-grained categorization. J Vis Commun Image Represent 2019;64. <http://dx.doi.org/10.1016/j.jvcir.2019.102647>.
- [101] Dong C, Zhang K, Xie Z, Wang J, Guo X, Shi C, Xiao Y. Transmission line key components and defects detection based on meta-learning. IEEE Trans Instrum Meas 2024.
- [102] Liu X, Miao X, Jiang H, Chen J, Wu M, Chen Z. Tower masking mim: A self-supervised pretraining method for power line inspection. IEEE Trans Ind Inform 2023;20(1):513–23.
- [103] Chen H, He Z, Shi B, Zhong T. Research on recognition method of electrical components based on YOLO V3. IEEE Access 2019;7:157818–29. <http://dx.doi.org/10.1109/ACCESS.2019.2950053>.
- [104] Shi C, Zheng X, Zhao Z, Zhang K, Su Z, Lu Q. LSKF-YOLO: Large selective kernel feature fusion network for power tower detection in high-resolution satellite remote sensing images. IEEE Trans Geosci Remote Sens 2024.
- [105] Liu Y, Pei S, Fu W, Zhang K, Ji X, Yin Z. The discrimination method as applied to a deteriorated porcelain insulator used in transmission lines on the basis of a convolution neural network. IEEE Trans Dielectr Electr Insul 2017;24(6):3559–66. <http://dx.doi.org/10.1109/TDEL.2017.006840>.
- [106] Ibrahim A, Dalbah A, Abulsaud A, Tariq U, El-Hag A. Application of machine learning to evaluate insulator surface erosion. IEEE Trans Instrum Meas 2020;69(2):314–6. <http://dx.doi.org/10.1109/TIM.2019.2956300>.
- [107] Mussina D, Irmanova A, Jamwal PK, Bagheri M. Multi-modal data fusion using deep neural network for condition monitoring of high voltage insulator. IEEE Access 2020;8:184486–96. <http://dx.doi.org/10.1109/ACCESS.2020.3027825>.
- [108] Stefanon SF, Yow K-C, Nied A, Meyer LH. Classification of distribution power grid structures using inception v3 deep neural network. Electr Eng 2022;104(6):4557–69. <http://dx.doi.org/10.1007/s00202-022-01641-1>.
- [109] Roy SS, Paramane A, Singh J, Chatterjee S, Das AK. Accurate sensing of insulator surface contamination using customized convolutional neural network. IEEE Sens Lett 2023;7(1):1–4. <http://dx.doi.org/10.1109/LENS.2022.3232506>.
- [110] Wang S, Liu Y, Qing Y, Wang C, Lan T, Yao R. Detection of insulator defects with improved ResNet and region proposal network. IEEE Access 2020;8:184841–50. <http://dx.doi.org/10.1109/ACCESS.2020.3029857>.
- [111] Fu Q, Liu J, Zhang X, Zhang Y, Ou Y, Jiao R, Li C, Mazzanti G. A small-sized defect detection method for overhead transmission lines based on convolutional neural networks. IEEE Trans Instrum Meas 2023;72:1–12. <http://dx.doi.org/10.1109/TIM.2023.3298424>.
- [112] Liu X, Miao X, Jiang H, Chen J. Box-point detector: A diagnosis method for insulator faults in power lines using aerial images and convolutional neural networks. IEEE Trans Power Deliv 2021;36(6):3765–73. <http://dx.doi.org/10.1109/TPWRD.2020.3048935>.
- [113] 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 2019;7:61797–810. <http://dx.doi.org/10.1109/ACCESS.2019.2915985>.
- [114] 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 2020;50(4):1486–98. <http://dx.doi.org/10.1109/TSMC.2018.2871750>.
- [115] Zhang X, Zhang Y, Liu J, Zhang C, Xue X, Zhang H, Zhang W. InsuDet: A fault detection method for insulators of overhead transmission lines using convolutional neural networks. IEEE Trans Instrum Meas 2021;70:1–12. <http://dx.doi.org/10.1109/TIM.2021.3120796>.
- [116] Hao K, Chen G, Zhao L, Li Z, Liu Y, Wang C. An insulator defect detection model in aerial images based on multiscale feature pyramid network. IEEE Trans Instrum Meas 2022;71:1–12. <http://dx.doi.org/10.1109/TIM.2022.3200861>.
- [117] Hao S, An B, Ma X, Sun X, He T, Sun S. PKAMNet: a transmission line insulator parallel-gap fault detection network based on prior knowledge transfer and attention mechanism. IEEE Trans Power Deliv 2023;38(5):3387–97.
- [118] Zhang S, Qu C, Ru C, Wang X, Li Z. Multi-objects recognition and self-explosion defect detection method for insulators based on lightweight GhostNet-YOLOV4 model deployed onboard UAV. IEEE Access 2023;11:39713–25. <http://dx.doi.org/10.1109/ACCESS.2023.3268708>.
- [119] Wang Y, Song X, Feng L, Zhai Y, Zhao Z, Zhang S, Wang Q. MCI-GLA plug-in suitable for YOLO series models for transmission line insulator defect detection. IEEE Trans Instrum Meas 2024.
- [120] Yi Y, Chen Z, Wang L. Intelligent aging diagnosis of conductor in smart grid using label-distribution deep convolutional neural networks. IEEE Trans Instrum Meas 2022;71:1–8. <http://dx.doi.org/10.1109/TIM.2022.3141160>.
- [121] 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 2020;69(9):6080–91. <http://dx.doi.org/10.1109/TIM.2020.2969057>.
- [122] Xiao Y, Li Z, Zhang D, Teng L. Detection of pin defects in aerial images based on cascaded convolutional neural network. IEEE Access 2021;9:73071–82. <http://dx.doi.org/10.1109/ACCESS.2021.3079172>.
- [123] Zhao Z, Wang R, Li Y, Zhai Y, Zhao W, Zhang K. A new multilabel recognition framework for transmission lines bolt defects based on the combination of semantic knowledge and structural knowledge. IEEE Trans Instrum Meas 2022;71:1–11. <http://dx.doi.org/10.1109/TIM.2022.3200103>.
- [124] Jiao R, Fu Z, Liu Y, Zhang Y, Song Y. A defective bolt detection model with attention-based RoI fusion and cascaded classification network. IEEE Trans Instrum Meas 2023.
- [125] Song Z, Huang X, Ji C, Zhang Y. Deformable YOLOX: Detection and rust warning method of transmission line connection fittings based on image processing technology. IEEE Trans Instrum Meas 2023;72:1–21. <http://dx.doi.org/10.1109/TIM.2023.3238742>.
- [126] Zhang Y, Li B, Shang J, Huang X, Zhai P, Geng C. DSA-net: An attention-guided network for real-time defect detection of transmission line dampers applied to UAV inspections. IEEE Trans Instrum Meas 2023.
- [127] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97.
- [128] Liang H, Zuo C, Wei W. Detection and evaluation method of transmission line defects based on deep learning. IEEE Access 2020;8:38448–58. <http://dx.doi.org/10.1109/ACCESS.2020.2974798>.
- [129] Liu X, Miao X, Jiang H, Chen J, Wu M, Chen Z. Component detection for power line inspection using a graph-based relation guiding network. IEEE Trans Ind Inform 2022;19(9):9280–90.
- [130] Liu X, Miao X, Jiang H, Chen J, Chen Z. Fault diagnosis in power line inspection using normalized multihierarchy embedding matching. IEEE Trans Instrum Meas 2023;72:1–10.
- [131] Yi J, Mao J, Zhang H, Zeng K, Tao Z, Zhong H, Wang S, Wang Y. PSTL-net: A patchwise self-texture-learning network for transmission line inspection. IEEE Trans Instrum Meas 2023.
- [132] Zhong L, Liu K. Visual classification and detection of power inspection images based on federated learning. IEEE Trans Ind Appl 2024.
- [133] 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 2020;8:94065–75. <http://dx.doi.org/10.1109/ACCESS.2020.2995608>.

- [134] Bi Z, Jing L, Sun C, Shan M. YOLOX++ for transmission line abnormal target detection. *IEEE Access* 2023;11:38157–67. <http://dx.doi.org/10.1109/ACCESS.2023.3268106>.
- [135] Yu C, Liu Y, Zhang W, Zhang X, Zhang Y, Jiang X. Foreign objects identification of transmission line based on improved YOLOv7. *IEEE Access* 2023;11:51997–2008. <http://dx.doi.org/10.1109/ACCESS.2023.3277954>.
- [136] Zhang J, Wang J, Song R, Peng G, Pu T, Zhang S. An edge visual incremental perception framework based on deep semi-supervised learning for monitoring power transmission lines. *CSEE J Power Energy Syst* 2023;9(2):759–68. <http://dx.doi.org/10.17775/CSEEJPES.2022.03120>.
- [137] 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*. 2017, p. 2117–25.
- [138] Ge Z, Liu S, Wang F, Li Z, Sun J. YoloX: Exceeding yolo series in 2021. 2021, arXiv preprint arXiv:2107.08430.
- [139] Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A. The pascal visual object classes (voc) challenge. *Int J Comput Vis* 2010;88:303–38.
- [140] Wang X, Han Y, Leung VC, Niyato D, Yan X, Chen X. Convergence of edge computing and deep learning: A comprehensive survey. *IEEE Commun Surv Tutor* 2020;22(2):869–904.
- [141] Shi W, Cao J, Zhang Q, Li Y, Xu L. Edge computing: Vision and challenges. *IEEE Internet Things J* 2016;3(5):637–46.
- [142] McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA. Communication-efficient learning of deep networks from decentralized data. In: *Artificial intelligence and statistics*. PMLR; 2017, p. 1273–82.
- [143] Karim S, Tong G, Li J, Qadir A, Farooq U, Yu Y. Current advances and future perspectives of image fusion: A comprehensive review. *Inf Fusion* 2023;90:185–217.
- [144] Meher B, Agrawal S, Panda R, Abraham A. A survey on region based image fusion methods. *Inf Fusion* 2019;48:119–32.
- [145] Shen-pei Z, Xi L, Bing-chen Q, Hui H. Research on insulator fault diagnosis and remote monitoring system based on infrared images. *Procedia Comput Sci* 2017;109:1194–9.
- [146] Ma J, Yu W, Liang P, Li C, Jiang J. FusionGAN: A generative adversarial network for infrared and visible image fusion. *Inf Fusion* 2019;48:11–26.
- [147] Han S, Yang F, Yang G, Gao B, Zhang N, Wang D. Electrical equipment identification in infrared images based on ROI-selected CNN method. *Electr Power Syst Res* 2020;188:106534.
- [148] Song C, Xu W, Wang Z, Yu S, Zeng P, Ju Z. Analysis on the impact of data augmentation on target recognition for UAV-based transmission line inspection. *Complexity* 2020;2020:1–11.
- [149] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. *Commun ACM* 2020;63(11):139–44.
- [150] Ye R, Boukerche A, Yu X-S, Zhang C, Yan B, Zhou X-J. A data augmentation method for insulators based on cycle GAN. *J Electron Sci Technol* 2024;100250. <http://dx.doi.org/10.1016/j.jnlest.2024.100250>, URL <https://www.sciencedirect.com/science/article/pii/S1674862X24000181>.
- [151] Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE international conference on computer vision*. 2017, p. 2223–32.
- [152] Raina R, Battle A, Lee H, Packer B, Ng AY. Self-taught learning: transfer learning from unlabeled data. In: *Proceedings of the 24th international conference on machine learning*. 2007, p. 759–66.
- [153] Wang Y, Yao Q, Kwok JT, Ni LM. Generalizing from a few examples: A survey on few-shot learning. *ACM Comput Surv (csur)* 2020;53(3):1–34.
- [154] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. In: *International conference on machine learning*. PMLR; 2017, p. 1126–35.
- [155] Dwyer B, Nelson J, Hansen T, et al. Roboflow. 2024, URL <https://roboflow.com>.
- [156] HumanSignal. HumanSignal/Awesome-Data-Labeling: A curated list of awesome data labeling tools. 2024, URL <https://github.com/HumanSignal/awesome-data-labeling>.
- [157] Tkachenko M, Malyuk M, Holmanyuk A, Liubimov N. Label Studio: Data labeling software. 2020, URL <https://github.com/heartexlabs/label-studio>. Open source software available from <https://github.com/heartexlabs/label-studio>.
- [158] Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo W-Y, et al. Segment anything. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2023, p. 4015–26.
- [159] Zhou Z-H. A brief introduction to weakly supervised learning. *Natl Sci Rev* 2018;5(1):44–53.
- [160] Choi H, Koo G, Kim BJ, Kim SW. Weakly supervised power line detection algorithm using a recursive noisy label update with refined broken line segments. *Expert Syst Appl* 2021;165:113895.
- [161] Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, p. 4681–90.
- [162] Hu GX, Yang Z, Hu L, Huang L, Han JM, et al. Small object detection with multiscale features. *Int J Digit Multimed Broadcast* 2018;2018.
- [163] Defard T, Setkov A, Loesch A, Audigier R. Padim: a patch distribution modeling framework for anomaly detection and localization. In: *International conference on pattern recognition*. Springer; 2021, p. 475–89.
- [164] Batzner K, Heckler L, König R. Efficientad: Accurate visual anomaly detection at millisecond-level latencies. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2024, p. 128–38.
- [165] Tsai D-M, Jen P-H. Autoencoder-based anomaly detection for surface defect inspection. *Adv Eng Inform* 2021;48:101272.
- [166] Sun S, Liu Y, Hu X, Zhang W. A semisupervised autoencoder-based method for anomaly detection in cutting tools. *J Manuf Process* 2023;93:315–27.
- [167] Roth K, Pemula L, Zepeda J, Schölkopf B, Brox T, Gehler P. Towards total recall in industrial anomaly detection. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, p. 14318–28.
- [168] Yang M, Wu P, Feng H. MemSeg: A semi-supervised method for image surface defect detection using differences and commonalities. *Eng Appl Artif Intell* 2023;119:105835.
- [169] Zavrtnik V, Kristan M, Skočaj D. Draem-a discriminatively trained reconstruction embedding for surface anomaly detection. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2021, p. 8330–9.
- [170] Ultralytics. Frequently asked questions (FAQ) — docs.ultralytics.com. 2024, <https://docs.ultralytics.com/help/FAQ/>. [Accessed 18 February 2024].
- [171] Team K. Keras documentation: Keras applications. 2024, URL <https://keras.io/api/applications/>.
- [172] Chollet F. Xception: Deep learning with depthwise separable convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, p. 1251–8.