MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SIENCE AND TECHNOLOGY

Trinh Hien Anh

RESEARCH ON DEEP LEARNING-BASED TECHNIQUES FOR DETECTING AND IDENTIFYING ANOMALIES IN HIGH-VOLTAGE POWER GRID EQUIPMENT FROM UAV IMAGES

SUMMARY OF DISSERTATION ON INFORMATION SYSTEM

Code: 9 48 01 04

Ha Noi- 2026

The dissertation is completed at: Graduate University of Science and Technology, Vietnam Academy of Science and Technology

Supervisors:

Supervisor 1: Dr. Nguyen Thi Thanh Tan, Department of Information Technology, Electric Power University

Supervisor 2: Prof. Dr. Ngo Quoc Tao, Institute of Information Technology

Referee 1: …

Referee 2: …

Referee 3: ….

The dissertation will be examined by Examination Board of Graduate University of Science and Technology, Vietnam Academy of Science and Technology at……………………….. (time, date, year…)

This dissertation can be found at:

1) Graduate University of Science and Technology Library
2) National Library of Vietnam

INTRODUCTION

1.Research motivation and significance

With the increasing demand for electricity, ensuring the safe and reliable operation of high-voltage transmission grids is essential for national energy security. Insulator strings and other grid equipment are exposed to harsh environmental conditions, leading to defects such as cracks, broken or missing insulator discs, material degradation, and contamination. These defects can cause equipment failures and power interruptions.

UAVs, high-resolution cameras, and thermal imaging have improved the efficiency of power grid inspection. However, the large volume of acquired data makes manual analysis time-consuming and costly, while small defects remain difficult to detect and often require expert intervention.

Deep learning and computer vision provide promising solutions, but UAV-based inspection remains challenging due to small and multi-scale objects, complex backgrounds, varying environmental conditions, occlusion, limited defective samples, and the need for efficient inference on resource-constrained edge platforms.

Therefore, this dissertation investigates “Deep learning techniques for detecting and identifying equipment anomalies in high-voltage power grids from UAV-acquired images. The research aims to improve inspection efficiency, enable early defect detection, support predictive maintenance, promote digital transformation in the power sector, and enhance the reliability of the national power system.

2. Research objectives

To propose deep learning models that improve the effectiveness of detecting and identifying abnormalities in power transmission grid equipment from UAV-acquired images under challenging conditions, including small objects, complex backgrounds, and imbalanced data.

Specific Objectives

To achieve the aforementioned overall objective, this dissertation focuses on the following specific objectives:

• Investigate deep learning-based object detection models for UAV-acquired images;

• Propose a model for detecting and identifying abnormalities in insulator strings;

• Investigate data imbalance handling techniques to improve training effectiveness;

• Develop and experimentally evaluate the proposed system on real-world UAV datasets.

3. Research content

To achieve the proposed objectives, the following research tasks are addressed:

• Investigate the current state of power transmission grid equipment inspection and monitoring;

• Analyze deep learning models for object detection from UAV-acquired images;

• Construct and evaluate a UAV-acquired dataset of power grid equipment;

• Propose solutions to improve the effectiveness of equipment detection and anomaly identification, including:

– A model for detecting equipment on power transmission grids;

– A model for detecting abnormalities in insulator strings;

– Data augmentation and balancing techniques.

4. Research objectives, scope, and methodology

Research Subjects

In accordance with the objectives of the dissertation, the research subjects include:

• Deep learning models for object detection from UAV-acquired images;

• Insulator strings and common types of abnormalities occurring in power transmission grids;

• Deep learning-based object detection and anomaly detection models;

• Data imbalance handling strategies in computer vision.

Research Scope

The dissertation focuses on the detection and identification of abnormalities in high-voltage power grid equipment from optical images acquired by UAVs in real-world environments. The study does not extend to long-term time-series forecasting or power system simulation.

Research Methodology

The dissertation combines theoretical research with experimental investigations.

The proposed models are implemented, developed, and trained. Their performance is evaluated using real-world datasets, and the results are compared with state-of-the-art methods.

5. Research contributions

The dissertation makes the following scientific and practical contributions:

• Development of an improved HVE-YOLO model to enhance the effectiveness of detecting power transmission grid equipment from UAV-acquired images;

• Development of an integrated deep learning framework for detecting and classifying abnormalities in insulator strings under challenging conditions involving small objects and complex backgrounds;

• Development of an Attention-Aware Prompting-based data augmentation strategy to address data imbalance during model training.

The dissertation also conducts comprehensive experimental evaluations, including comparisons with stateof-the-art models, ablation studies, and qualitative analyses. These experiments demonstrate the effectiveness, stability, and practical applicability of the proposed methods to the inspection and monitoring of power transmission grid equipment using UAV-acquired images.

6. Dissertation structure

The dissertation consists of an introduction, four main chapters, and a conclusion. The introduction presents the research objectives, scope, methodology, and contributions. Chapter 1 provides an overview of high-voltage power grid inspection using UAV-acquired images, analyzes existing approaches, and identifies research gaps. Chapter 2 proposes the HVE-YOLO model for detecting power grid equipment from UAV-acquired images. Chapter 3 investigates the FGS model for detecting and segmenting abnormalities in insulator strings. Chapter 4 proposes AAPID within the Few-Shot Class-Incremental Learning (FSCIL) framework to identify novel defect classes with a limited number of training samples while mitigating catastrophic forgetting. Finally, the conclusion summarizes the research findings, contributions, and directions for future research.

CHAPTER 1. INTRODUCTION TO THE PROBLEM OF EQUIPMENT ANOMALY DETECTION AND IDENTIFICATION IN POWER GRIDS USING UAV-ACQUIRED IMAGES

1.1. Problem Introduction

Power systems, particularly high-voltage and extra-high-voltage transmission grids, play a critical role in ensuring energy security and stable power system operation. Traditional inspection methods remain limited in terms of cost, safety, and accessibility; therefore, the integration of unmanned aerial vehicles (UAVs) with computer vision and deep learning represents a promising approach for automating inspection and monitoring activities.

However, UAV images often contain small objects captured from varying viewpoints and distances, with complex backgrounds, while data on abnormal conditions remain scarce. These factors pose substantial challenges to model training and deployment.

Accordingly, this thesis focuses on three main problems:

• Detection of transmission-grid equipment from UAV images;

• Identification of anomalies on insulator strings under complex environmental conditions;

• Improvement of model learning capability when abnormal data are scarce.

The ultimate goal is to contribute to the development of an intelligent and automated power grid inspection system that supports asset management and predictive maintenance, in line with the digital transformation of the power sector.

1.1.1. Equipment Detection in High-Voltage Transmission Grids

The development of UAVs combined with computer vision and deep learning has opened an effective pathway toward automating the inspection and management of transmission grids. UAV-acquired images and videos can be used to detect, recognize, and assess equipment conditions, thereby contributing to the construction of digital databases and supporting the detection of anomalies, damage, and potential failures.

Nevertheless, UAV imagery presents several challenges, including small objects, varying viewpoints and distances, complex backgrounds, and the need for fast processing. Therefore, deep learning models must be developed to accurately detect small equipment, adapt to diverse imaging conditions, and satisfy real-time processing requirements, thereby providing a foundation for intelligent power grid inspection systems.

1.1.2. Anomaly Detection on Insulator Strings

Insulator strings in high-voltage transmission systems serve to electrically isolate conductors from transmission towers while also supporting mechanical loads, thereby directly affecting the safety and reliability of the power system.

These abnormalities are typically small, localized, and irregularly distributed, making early detection challenging when relying solely on manual inspection.

1.1.3. Characteristics and Challenges of UAV Data for Power Grid Inspection

Characteristics of UAV Images for Power Grid Inspection

UAV imagery for power grid inspection has several distinctive characteristics: (i) high spatial resolution and the need for multi-scale feature extraction; (ii) significant variations in viewing angles, geometry, and perspective; (iii) many target objects are small and slender; (iv) complex and diverse backgrounds; and (v) varying lighting and weather conditions. These characteristics increase the difficulty of detecting power grid equipment and anomalies, while also requiring deep learning architectures capable of preserving spatial information, enhancing multi-scale features, and adapting effectively to real-world inspection environments.

1.2. Related works

Current studies on equipment and anomaly detection in power grids primarily employ deep learning and computer vision and can be broadly categorized into three main directions: (1) Conventional images: high image quality and clearly visible objects, but limited observation coverage and insufficient representation of real-world inspection conditions; (2) UAV images: wide observation coverage and suitability for field inspection, but challenges arise from small objects, complex backgrounds, varying viewpoints and illumination, and data imbalance and (3) Synthetic data generation: intended to supplement scarce abnormal samples, balance data distributions, and improve generalization, particularly for long-tailed data distributions.

Overall, although these directions have achieved encouraging results, gaps remain in the integration of UAV data, small-object detection, and abnormal-data generation. These gaps motivate the investigation of deep learning and data generation methods to improve the detection, identification, and diagnosis of anomalies on insulator strings.

1.2.1. Equipment and Anomaly Detection from Conventional Images

Current object detection methods have mainly evolved along three directions: two-stage models, one-stage models, and Transformer-based methods. Two-stage models first generate region proposals and then perform classification and bounding-box regression, whereas one-stage models directly predict object locations and classes, offering advantages in inference speed. Transformer-based methods exploit attention mechanisms to better model relationships among features and improve detection performance.

1.2.2. Equipment and Anomaly Detection from UAV Images

Studies using UAV images for power grid inspection mainly address three challenges: detecting small objects, handling complex backgrounds, and ensuring real-time inference on resource-constrained computing platforms. The main approaches include two-stage models, one-stage models, Transformer architectures, and hybrid models. However, variations in viewing angles, object sizes, illumination conditions, and background environments remain major obstacles to maintaining detection accuracy and generalization capability.

1.3. Synthetic Data Generation Methods for Addressing Anomaly Sample Scarcity and Data Imbalance

Power grid inspection datasets often exhibit severe class imbalance, with normal samples accounting for a large proportion while abnormal samples are rare, in some cases accounting for less than 1%. This imbalance can cause models to favor the majority class, resulting in low Recall and F1-score for abnormal cases. Synthetic data generation methods have therefore been investigated to supplement rare samples, balance data distributions, and improve the detection of damage types that are difficult to collect in real-world conditions.

1.3.1. Research Gaps

The review of existing studies reveals three major research gaps: (i) limited capability for detecting small objects in complex UAV image backgrounds; (ii) the lack of effective solutions for severe data imbalance; and (iii) insufficient exploitation of synthetic data and limited integration of power-grid domain knowledge.

1.3.2. Research Questions

These three gaps lead to the following three main research questions: **RQ1**: How can the detection capability for small equipment be improved in UAV images with complex backgrounds and varying viewpoints? **RQ2**: How can small, localized anomalies on insulator strings be accurately identified when the boundary between normal and abnormal conditions is unclear? and **RQ3**: How can data imbalance be addressed and the quality and utility of synthetic data be improved by integrating power-grid domain knowledge into generative models?

1.3.3. Research Contents and Main Contributions of the Thesis

To answer the three research questions, this thesis focuses on three main research directions: (i) proposing the **HVE-YOLO** model to improve the detection of small equipment in UAV images; (ii) developing the **FGS** framework, which combines detection and segmentation to identify detailed anomalies on insulator strings; and (iii) proposing the **AAPID** method for learning and diagnosing anomalies under limited-data conditions while exploiting synthetic data and domain knowledge.

1.4. Chapter Conclusion

Chapter 1 has presented the theoretical background, an overview of existing methods, and the characteristics of power transmission grid inspection from UAV images. The core challenges, including small objects, complex backgrounds, viewpoint variations, and data imbalance, have been analyzed. Based on this analysis, three research gaps were identified, leading to three corresponding solutions: **HVE-YOLO** for equipment detection, **FGS** for anomaly identification on insulator strings, and **AAPID** for incremental learning under limited-data conditions. The proposed methods and experimental results are presented in the following chapters.

CHAPTER 2. DETECTION OF EQUIPMENT ON HIGH-VOLTAGE POWER GRIDS FROM UAV-ACQUIRED IMAGES

This chapter focuses on detecting high-voltage power grid equipment from UAV images to support power system inspection, monitoring, and maintenance. The problem presents several challenges due to small objects, slender shapes, complex backgrounds, and varying illumination conditions, which can degrade the performance of models such as Faster R-CNN and YOLO.

To address these challenges, this study proposes HVE-YOLO, an improved model based on YOLOv11, specifically designed to detect six core equipment classes on high-voltage power grids. The model is evaluated on data collected at the CT6 site.

2.1. Related Studies on Power Grid Equipment Detection

2.1.1. Studies on Power Grid Equipment Detection from UAV Images

High-voltage transmission grids play a vital role in maintaining national energy security and power system stability. However, because grid components operate outdoors, they are continuously exposed to adverse weather conditions, increasing the risk of damage and cascading failures. Periodic inspection is therefore essential to ensure operational safety, comply with technical standards, and optimize maintenance costs. In recent years, unmanned aerial vehicle (UAV) technology has demonstrated significant effectiveness in monitoring and collecting visual data of high-voltage grid components [1, 2, 3].

2.1.2. Power Grid Equipment Detection Using One-Stage Models

With their real-time object detection capability, high accuracy, and relatively efficient computational cost, YOLO variants (CT1, CT4) have been widely applied to the recognition of insulator strings, line fittings, abnormal thermal hotspots, and various types of damage to electrical equipment. Recent studies show that YOLO not only improves the automation of inspection tasks but also helps reduce risks to operating personnel and enhance power system reliability.

YOLO11 Backbone Architecture

Model Limitations

Although YOLOv11 achieves substantial improvements in accuracy and inference speed for real-time object detection, it still exhibits several limitations when applied to equipment detection on high-voltage transmission grids. One important challenge is the recognition of small objects such as locking pins, spacers, vibration dampers, or small cracks on insulator strings.

Another notable limitation of YOLOv11 is its strong dependence on high-quality, fully annotated training datasets.

2.2. HVE-YOLO Model (High Voltage Equipment Detection)

The proposed HVE-YOLO is designed to improve high-voltage transmission grid equipment detection from UAV images while retaining the Backbone–Neck–Head architecture and anchor-free prediction mechanism of YOLO11. The model is enhanced with three main components: (i) SPD-Conv, which replaces stride-based downsampling to preserve spatial information and improve small-object detection; (ii) CBAM after multi-scale feature fusion to reduce the influence of background noise and enhance informative features; and (iii) Wise-IoU v3 to improve bounding-box regression quality, particularly for difficult samples.

2.2.1. HVE-YOLO Model

The HVE-YOLO model [CT6] is developed to improve the detection and recognition of high-voltage transmission grid equipment from UAV images. The model can detect six basic equipment classes: *monopole, steel pole, electrical conductor, vibration damper, glass insulator string, and silicon insulator string*. These are all important components for monitoring, operating, and maintaining transmission grids.

2.2.2. Overall Architecture of HVE-YOLO

The overall architecture is designed to optimize small-object detection and improve model robustness under complex environmental conditions. The HVE-YOLO architecture is illustrated in Figure 2.1 and consists of three main components: Backbone, Neck, and Head.

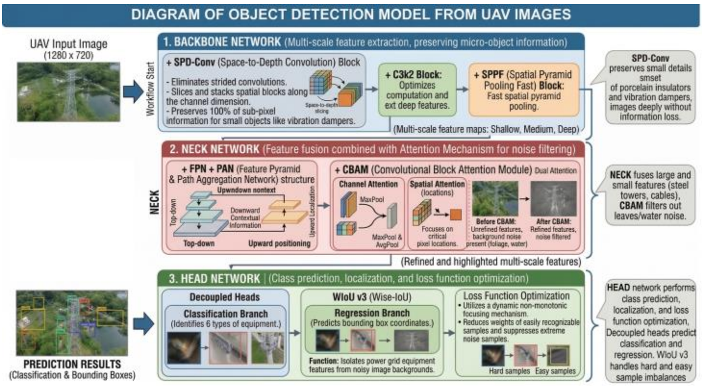

Figure 2.1: HVE-YOLO Architecture

The input image acquired by the UAV is first fed into the Backbone to extract hierarchical features. These features are then aggregated through the Neck using the Feature Pyramid Network (FPN) and Path Aggregation Network (PAN), and further refined by CBAM attention. Finally, the Detection Head simultaneously performs three tasks: object classification, bounding-box regression, and objectness probability estimation.

2.2.3. Improvements in HVE-YOLO

Backbone Improvements for Enhanced Small-Object Detection

The HVE-YOLO backbone is enhanced with SPD-Conv, C3k2, and SPPF to preserve spatial information, improve feature representation, and capture multi-scale contextual information. SPD-Conv reduces information loss during downsampling, making it particularly effective for detecting small and slender objects in UAV imagery. C3k2 enhances feature learning and improves discrimination between targets and complex backgrounds, while SPPF expands the receptive field and aggregates multi-scale information with low computational cost.

The combination of these three components improves the detection of small, occluded, and low-contrast objects, while maintaining reasonable computational complexity and real-time inference capability, making HVE-YOLO suitable for embedded GPUs and UAV-based Edge AI platforms.

CBAM Attention

HVE-YOLO employs FPN and PAN to fuse multi-scale features, combining high-level semantic information with detailed spatial features to improve the detection of devices with diverse sizes, particularly small objects. The model further integrates CBAM, consisting of Channel Attention and Spatial Attention, to focus on informative features and relevant spatial locations while reducing the influence of complex backgrounds.

The combination of FPN, PAN, and CBAM enhances HVE-YOLO's ability to detect small, occluded, and low-contrast objects, while maintaining a lightweight architecture and computational cost suitable for real-time inference.

Improvement of the Localization Loss Function

To address the challenges of detecting small electrical equipment, HVE-YOLO adopts Wise-IoU version 3 (WIoU-v3) as the localization loss function. WIoU-v3 employs a dynamic non-monotonic focusing mechanism, in which the contribution of each sample is adaptively adjusted through a gradient gain factor. This mechanism reduces the influence of excessively easy or difficult samples while focusing optimization on samples with appropriate difficulty levels and greater potential to provide useful learning information.

The total loss function of HVE-YOLO is defined as follows:

$\mathcal{L} = \lambda_1 \mathcal{L}_{WIoU-v3} + \lambda_2 \mathcal{L}_{cls} + \lambda_3 \mathcal{L}_{obj}$ (2.1)

in which $\mathcal{L}_{WIoU-v3}$ is localization loss, $\mathcal{L}_{cls}$ is classification loss, $\mathcal{L}_{obj}$ is objectness loss $\lambda_1, \lambda_2, \lambda_3$ are the balancing coefficients.

$WIoU - v3$ is formulated as $\mathcal{L}_{WIoU-v3} = \text{r}.\mathcal{L}_{WIoU-v1}$ (2.2)

with the focusing factor $r = \frac{\beta}{\delta.\alpha^{(\beta-\delta)}}$ $\beta = \frac{\mathcal{L}_{IoU}}{\overline{\mathcal{L}_{IoU}}}$ (2.3)

where the term represents the outlier degree of the sample, $\overline{\mathcal{L}_{IoU}}$ denotes the moving average of the IoU loss. The parameters and are hyperparameters that control the dynamic focusing mechanism.

Unlike the conventional mechanism that monotonically increases the weight according to sample difficulty, WIoU-v3 employs a non-monotonic focusing mechanism to reduce the influence of noisy or excessively difficult samples while prioritizing samples with an appropriate learning quality. Consequently, HVE-YOLO improves bounding-box regression for small and occluded electrical equipment or objects observed from complex viewing angles, thereby enhancing the model's accuracy and robustness in UAV-based power grid inspection environments.

2.2.4. HVE-YOLO Training Procedure

The development and training of the HVE-YOLO model consists of three main stages: (i) data preparation, (ii) model development and training, and (iii) model evaluation. These stages are organized in an integrated workflow to ensure high detection accuracy and good generalization when the model is deployed in real-world environments.

Data Preparation

The data were collected using unmanned aerial vehicles (UAVs) during field inspections of high-voltage transmission lines, following the procedure illustrated in Figure 2.2.

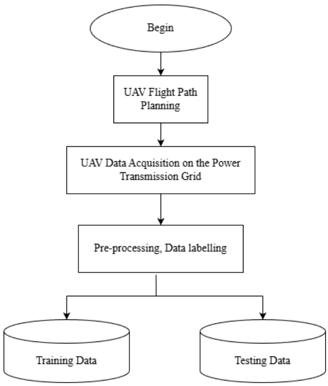

Figure 2.2: UAV-Based Data Collection Procedure for High-Voltage Power Grids

The data were divided into training and validation sets at a ratio of 80%/20% using a stratified sampling method, ensuring a relatively balanced distribution across the equipment classes. In addition, consecutively acquired images from the same location were carefully examined to minimize data leakage between the two sets. This procedure helps ensure that the evaluation results objectively reflect the generalization capability of HVE-YOLO on data not used during training.

Model Construction and Training

HVE-YOLO was initialized with pretrained weights to reduce convergence time and improve its capability to detect small-scale equipment in complex backgrounds.

The training process is described in Algorithm 1.

Algorithm 1: HVE-YOLO Training and Inference Procedure

Input: Training set $T = \{(Ik, Gk)\}M\ k$=1, learning rate η, number of epochs E

Output: Optimized parameters $\Theta*$ and detection results Dk

1: Initialize model parameters Θ;

2. For epoch = 1 to E:

2.1. For *each data batch* $\mathscr{B}$:

F_b ← Backbone(I_k).

F_n ← Neck(F_b).

(P_b, P_c, P_o) ← Head(F_n).

$\mathscr{L}$_box ← $\mathscr{L}$_WIoU-v3(P_b, $\mathcal{G}_k$).

$\mathscr{L}$_cls ← $\mathscr{L}$_cls(P_c, $\mathcal{G}_k$).

$\mathscr{L}$_obj ← $\mathscr{L}$_obj(P_o, $\mathcal{G}_k$).

$\mathscr{L}$_$\mathscr{B}$ ← $\lambda_1\mathscr{L}$_box + $\lambda_2\mathscr{L}$_cls + $\lambda_3\mathscr{L}$_obj.

Θ ← Θ − η∇_Θ $\mathscr{L}$_$\mathscr{B}$.

3. Θ* ← Θ.

4. Inference Stage:

5. For *each test image* I_k:

F_b ← Backbone(I_k).

F_n ← Neck(F_b).

(P_b, P_c, P_o) ← Head(F_n).

S ← P_c ⊙ P_o.

$\mathcal{D}_k$ ← NMS(P_b, S).

6. Return {Θ*, $\mathcal{D}_k$}.

Model Evaluation

HVE-YOLO was evaluated on the validation set using Precision, Recall, F1-score, AP, mAP, and FPS, combined with a qualitative analysis of its ability to detect small, occluded objects and objects in complex backgrounds. The model achieved an mAP@0.5 of 0.972, demonstrating high detection performance under the conditions of the studied dataset.

However, these results cannot directly confirm the model's generalization capability to 220 kV and 500 kV transmission lines, as greater UAV-to-target distances may reduce the apparent size of objects in images and increase detection difficulty. Maintaining performance depends on image quality, spatial resolution, UAV stability, and the preservation of fine-grained visual features.

Future work should expand the dataset to cover multiple voltage levels and observation distances, while investigating super-resolution techniques, extremely small object detection, and multimodal sensor fusion to improve the reliability of long-range UAV-based inspection.

2.3. Experiments

Nghiên cứu xây dựng môi trường thử nghiệm nhằm đánh giá khả năng triển khai mô hình trên UAV, với mô hình được huấn luyện bằng Python 3 và PyTorch, sau đó triển khai trên NVIDIA Jetson Orin Nano và tối ưu bằng TensorRT. Mô hình được chuyển đổi sang FP16 và INT8 kết hợp PTQ để đánh giá độ chính xác, tốc độ suy luận, mức sử dụng tài nguyên và khả năng triển khai Edge AI.

2.3.1. Test Dataset

The dataset was collected using a DJI Mavic 2 Pro UAV on 110 kV power transmission lines under various environmental conditions. It contains 69,955 instances across six classes, including conductors, silicon insulator strings, glass insulator strings, steel lattice towers, monopoles, and vibration dampers. Of these, 53,784 instances were used for training and 16,171 for testing, with annotations stored in the YOLO format.

2.3.2. Experimental Scenarios

HVE-YOLO was evaluated through five groups of experiments to comprehensively assess its capabilities:

(1) Overall accuracy: The model was trained using an 80/20% split, with an input size of 640 × 640 pixels, AdamW optimizer, a learning rate of 10^{-3}, a batch size of 16, and 300 epochs. Performance was evaluated using Precision, Recall, and mAP@0.5:0.95.

(2) Small object detection: The evaluation focused on conductors, insulator strings, and vibration dampers, particularly objects occupying less than 5% of the image area, as well as cases involving occlusion and complex viewing angles.

(3) Environmental robustness: The model was evaluated under varying lighting, weather, background, viewing-angle, and image-quality conditions. Performance degradation was assessed based on mAP, Recall, and F1-score.

(4) Real-time performance: HVE-YOLO was deployed on an NVIDIA Jetson Orin Nano and evaluated using Intensive Test, Parallel Processing Test, and Stress Test scenarios. The evaluation considered FPS, inference time, resource utilization, temperature, and power consumption.

(5) Model comparison: HVE-YOLO was compared with HVE-Mask R-CNN and different YOLO variants in terms of detection accuracy, small object detection capability, inference speed, resource requirements, and real-time deployment capability

2.3.3. Experimental Results

The experimental results show that HVE-YOLO outperformed HVE-Mask R-CNN on most evaluation criteria. Specifically, HVE-YOLO achieved an mAP@0.5 of 0.972, a maximum F1-score of 0.950, and a Recall of 0.990, as presented in Table 2.1

Table 2.1. Performance Comparison between HVE-YOLO and HVE-Mask R-CNN

Metric	HVE-YOLO	HVE-Mask R-CNN
Architecture	One-stage	Two-stage
mAP@0.5	0,972	0,855
Maximum F1-score	0,950	0,851
Overall Recall	0,990	0,906
FPS – Intensive Test	162,9	75,8
FPS – Stress Test	165,5	82,3
FPS – Parallel Processing	95,6	98,6

2.4. Chapter 2 Conclusion

Chapter 2 developed the HVE-YOLO model for detecting six types of equipment on 110 kV transmission lines from UAV imagery. The model incorporates multi-scale feature representation, an additional P2 feature layer, and an attention mechanism to enhance the detection of small objects in complex backgrounds.

Experimental results showed that HVE-YOLO achieved an mAP@0.5 of 0.972 and a maximum speed of 165.5 FPS, meeting the requirements for real-time detection while maintaining stable performance under variations in viewing angle, illumination, observation distance, and image background. The insulator string localization accuracy exceeding 96% provides a foundation for subsequent equipment condition analysis.

These results provide the basis for Chapter 3, which further investigates the detection and identification of anomalies in insulator strings from UAV imagery.

CHAPTER 3. ANOMALY DETECTION ON INSULATOR STRINGS FROM UAV-ACQUIRED IMAGES

Chapter 2 proposed HVE-YOLO for detecting high-voltage power grid equipment from UAV imagery, achieving an mAP@0.5 of 97% and a maximum speed of 165.5 FPS, thereby meeting the requirements for real-time inspection. However, equipment detection is only the first step; further analysis of equipment conditions and anomaly detection is required, particularly for insulator strings, where defects are often small, diverse, and associated with limited data.

Therefore, Chapter 3 focuses on developing the FGS (FPN-GELAN-

Diffusion Segmentation) model, which integrates multi-scale feature representation with a diffusion mechanism to enhance anomaly detection and segmentation on insulator strings under limited-data and complex environmental conditions.

3.1. Related Studies on Anomaly Detection in Insulator Strings

3.1.1. Types of Anomalies in Insulator Strings

3.1.2. Studies on Anomaly Detection in Insulator Strings from UAV Images

In UAV imagery, anomalies in insulator strings are often small, exhibit diverse morphologies, and can easily be occluded by complex backgrounds, making feature extraction and detection challenging. YOLO-based models are widely adopted due to their high processing speed and ability to detect small objects. Various versions have exploited FPN and PAN architectures and, more recently, incorporated DETR or RT-DETR to enhance feature representation. However, DETR-based models generally require large amounts of training data, whereas YOLOv9 improves feature representation through the GELAN architecture, contributing to more effective anomaly detection in UAV imagery.

YOLOv9 Backbone

Advantages and Limitations of YOLOv9 for UAV-Based Insulator String Inspection

Advantages: High processing speed, enhanced capability for detecting small objects, and strong learning and adaptation capacity under variations in viewing angle, object scale, and illumination conditions.

Limitations: Requires a relatively large amount of training data and has limited capability in recognizing previously unseen fault types.

3.2. Proposed FGS (FPN–GELAN–Diffusion Segmentation) Model

3.2.1. FGS (FPN–GELAN–Diffusion Segmentation)

In UAV imagery, anomaly detection on insulator strings is challenging due to the small size of defects, diverse defect morphologies, complex image backgrounds, and varying acquisition conditions. Factors such as UAV vibration, viewing angle and distance, weather, illumination, and optical differences between insulator types can reduce the stability of image features. Meanwhile, downsampling operations may cause the loss of fine details associated with small anomalies, such as cracks, peeling, or localized contamination.

To address these limitations, this thesis proposes FGS (FPN–GELAN–Diffusion Segmentation), which integrates multi-scale feature fusion, GELAN, and a diffusion model to preserve fine-grained information, enhance the detection of small anomalies, and achieve accurate anomaly segmentation under complex UAV imaging conditions.

3.2.2. Overall Architecture

The FGS model consists of three main components: the multi-scale FPN feature network, the GELAN backbone, and the Diffusion Model

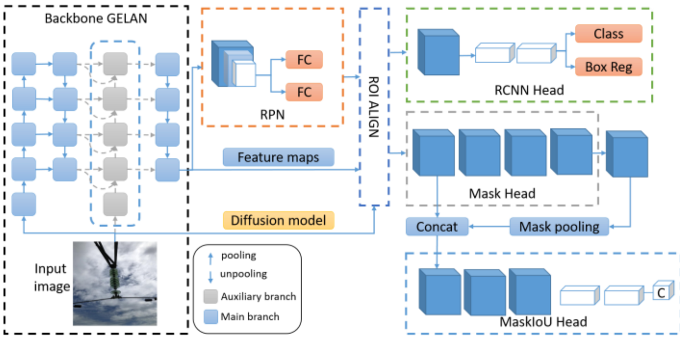

Figure 3.1: Overall architecture of the FGS model

Figure 3.1 illustrates the overall structure of the FGS system. The architecture is designed as a unified integration of feature extraction, region proposal, representation enhancement, and anomaly segmentation stages.

3.2.3. Improvements in the FGS Model

GELAN Backbone

The GELAN backbone enhances multi-level feature aggregation through multi-branch connections and feature reuse, improving the detection of small, complex anomalies while maintaining computational efficiency. The integrated SE attention mechanism adaptively weights feature channels, helping the model focus on anomaly-relevant regions.

Diffusion Model

The Diffusion Model enhances feature representation by distinguishing meaningful insulator structures from random noise, improving robustness to complex backgrounds, illumination variations, and image noise in UAV imagery.

RoI and Prediction Branches

The RoI Align layer combines features from the RPN, GELAN backbone, and Diffusion Model to generate more discriminative RoI representations for subsequent prediction and segmentation.

Post-processing

The RCNN Head, Mask Head, and MaskIoU Head generate object locations, class labels, and segmentation masks. A post-processing stage then removes duplicate or low-confidence predictions and refines the final results for improved accuracy and stability.

3.2.4. FGS Model Training Procedure

The FGS training procedure is designed as an end-to-end process consisting of three stages: (i) data preparation; (ii) training configuration and execution; and (iii) model evaluation.

Data Preparation

Training Configuration and Procedure The FGS model is trained following a five-step procedure. First, the model components, including GELAN-SE-DWConv, Diffusion/DiT, FPN, RPN, RoIAlign, RCNN Head, Mask Head, and MaskIoU Head, are initialized and trained from scratch..

Model Evaluation

3.3. Experiments

3.3.1. Experimental Setup

All experiments were conducted using Python 3.10 on an Intel Core i9-13900K CPU, 64 GB RAM, and an NVIDIA RTX 4090 GPU (24 GB VRAM). Models were trained with 640×640 input images for 50 epochs using the AdamX optimizer and an initial learning rate of 5×10^{-4}. A linear warm-up was applied for the first five epochs, followed by cosine decay, while data augmentation was disabled during the final 10 epochs to reduce over-reliance on augmented data.

3.3.2. Data Preparation

The study uses two datasets, VNElectric and Insulator Defect Detection, to evaluate the effectiveness and generalization capability of the proposed method. VNElectric consists of 1,596 UAV images collected from 110 kV power transmission lines in northern Vietnam, covering diverse viewing angles, environmental conditions, and complex backgrounds. The dataset was split at a ratio of 3:1:1 for training, validation, and testing, with data augmentation techniques applied to improve model generalization.

3.3.3. Experimental Results

Experimental results on the VNElectric dataset show that FGS achieved a Precision of 92.3%, Recall of 91.4%, and mAP@50 of 93.7%, representing improvements of 1.7, 1.1, and 0.9 percentage points, respectively, over YOLOv9-E. These results demonstrate that combining the GELAN backbone, multi-scale FPN feature fusion, and the Diffusion Model improves the detection of small anomalies in complex UAV imagery.

The ablation study further confirms the complementary contributions of the individual components. FPN had the greatest impact, with its removal reducing mAP@50 by 5.6 percentage points, followed by GELAN and the Diffusion Model, which caused decreases of 3.9 and 1.7 percentage points, respectively. Qualitative results also show that FGS maintains relatively stable detection performance under complex backgrounds, small and occluded objects, and uneven illumination, demonstrating its effectiveness on real-world UAV imagery.

3.4. Chapter 3 Conclusion

Chapter 3 proposed FGS (FPN-GELAN-Diffusion Segmentation) for anomaly detection and segmentation in insulator strings from UAV imagery. By integrating GELAN, FPN, and a Diffusion Transformer, FGS improves multi-scale feature representation and the detection of small, low-contrast anomalies in complex backgrounds.

On VNElectric, FGS achieved 92.3% Precision and 93.7% mAP@50, outperforming YOLOv9-E. However, its high computational cost and parameter count require further optimization for Edge AI deployment and improved generalization.

CHAPTER 4. INCREMENTAL LEARNING WITH LIMITED DATA FOR INSULATOR STRING ANOMALY DIAGNOSIS

Chapter 4 addresses limited anomaly data in UAV-based power grid inspection. AAPID is proposed within the FSCIL framework, integrating invariant and adaptive prompts with Information Bottleneck Loss and Anchor Loss to learn new fault classes from few samples while retaining previous knowledge.

Experiments on CIFAR100 and VPI-Fault-2500 demonstrate its effectiveness and potential for recognizing new anomaly types in insulator strings.

4.1. Related Studies on Data Generation for Deep Learning Model Training

4.1.1. Traditional Data Generation Methods

Deep learning performance strongly depends on data quality and quantity, whereas electrical equipment fault data are often scarce and imbalanced. Traditional augmentation methods remain limited because they mainly transform existing data. Therefore, GANs, VAEs, and Diffusion Models are used to generate diverse and realistic data, supplement rare fault samples, and improve model generalization when real-world data are limited.

4.1.2. Data Generation and Learning Methods for Anomaly Detection under Limited Data

Under limited anomaly data, Prompt Learning combined with Few-Shot and Incremental Learning enables models to learn new classes from few samples while reducing catastrophic forgetting. However, existing methods are not well suited to small, low-contrast insulator anomalies and newly emerging faults in complex UAV imagery. Therefore, AAPID is proposed to preserve prior knowledge, learn new fault classes efficiently, and improve feature discrimination in the FSCIL setting.

4.2. AAPID (Attention-aware Adaptive Prompting for Insulator Diagnosis)

Traditional deep learning models can achieve high performance in detecting and identifying anomalies on insulator strings, but they still depend on scarce and imbalanced fault data and must cope with newly emerging fault classes. To address these challenges, the thesis adopts Few-Shot Class-Incremental Learning (FSCIL) and proposes AAPID to learn new fault classes from a limited number of samples, preserve previously learned knowledge, and mitigate catastrophic forgetting, thereby improving system adaptability.

4.2.1. Architecture overall

AAPID is proposed within the FSCIL framework and is based on the Vision Transformer (ViT), with the objective of learning new anomaly classes from few samples while preserving previously acquired knowledge and mitigating catastrophic forgetting.

The model freezes the pretrained ViT encoder and updates only the prompt parameters and adaptive prompt generator, thereby reducing overfitting and training cost. The architecture consists of four main components: a frozen ViT, Task-Invariant Prompt (TIP), a Prompt Encoder, and a loss function combining Information Bottleneck Loss and Anchor Loss, thereby improving adaptability and discrimination among new fault classes.

4.2.2. Attention-based Adaptive Prompting Mechanism Frozen Vision Transformer Encoder

AAPID uses a pretrained Vision Transformer (ViT) as its backbone and freezes all backbone parameters during FSCIL. The input image is divided into patches, converted into patch embeddings, augmented with positional information, and processed through Transformer blocks to extract features.

Freezing the backbone reduces overfitting, preserves learned knowledge, decreases the number of parameters that need to be optimized, and improves training stability in the few-shot setting. However, the frozen backbone cannot directly adapt to new classes; therefore, the prompting mechanism plays the primary role in adaptation and provides the foundation for AAPID's adaptive prompting system.

Task-Invariant Prompt

The Task-Invariant Prompt (TIP) in AAPID is a learnable set of prompts used to provide general knowledge to the model without modifying the backbone parameters. TIP aims to learn features shared across all classes, such as the overall shape, color, geometric structure, and surface characteristics of insulator strings.

TIP consists of M tokens, all initialized with the same value. This design encourages the prompt to learn features shared across tasks rather than memorizing class-specific characteristics.

TIP is then concatenated with the image patch tokens: $Z = [PT\ IP; X]$

The prompt tokens and image tokens jointly participate in the self-attention process within the Transformer. Thus, TIP serves as a task-invariant source of background knowledge that supports adaptation to new classes in the FSCIL setting.

Adaptive Prompt Generator

Although Task-Invariant Prompt (TIP) can store shared knowledge, it is insufficient to distinguish different fault types. Therefore, AAPID introduces a Task-Specific Prompt (TSP) mechanism to generate adaptive prompts for each input image.

Unlike methods such as L2P and DualPrompt, AAPID does not learn a fixed prompt for each class. Instead, the model uses a Prompt Encoder to generate prompts directly from image features. Let the input image feature be represented by $F = f(X)$. The Prompt Encoder performs the mapping $PTSP = G(F)$.

where $G(\cdot)$ denotes the Prompt Encoder and $PTSP$ denotes the adaptive prompt generated from the input data.

TSP can encode characteristics specific to each fault type, such as surface cracks, rim chips, body breakage, and severe surface contamination. By generating adaptive prompts according to the input features, AAPID can distinguish different fault types and effectively learn new classes even when the number of training samples is extremely limited.

4.2.3. Prompt Representation Learning with Information Bottleneck Information Bottleneck Constraint

To improve the generalization capability of the Prompt Encoder, AAPID applies the *Information Bottleneck* (IB) principle to retain information relevant to class labels while removing redundant information from the input data. The IB objective is expressed as min $I(P; X) - \gamma I(P; y)$.

where $I(P; X)$ and $I(P; y)$ denote the mutual information between the prompt and the input data and between the prompt and the class label, respectively, while γ is a balancing coefficient.

Since direct computation of mutual information is difficult, AAPID uses an approximation based on a variational lower bound: $LIB = \mathrm{E}\ [-\ \log P(y|x)] + \beta KLDKL\ (P(p|x)|r(p))$ where the first term is the crossentropy loss, which ensures that the prompt retains information necessary for classification; the second term is the Kullback–Leibler (KL) divergence, which acts as a regularizer by limiting the amount of instance-specific information encoded by the prompt. With *r*(*p*) commonly assumed to follow a Gaussian distribution, this mechanism encourages the prompt to focus on discriminative features and reduces dependence on noisy or sample-specific features. Consequently, the Prompt Encoder can generate highly generalizable prompts suitable for few-shot learning and adaptation to new classes.

4.2.4. Anchor Loss and Feature Space Optimization

Although Information Bottleneck encourages the prompt to focus on semantically rich information, overlap may still occur among classes in the feature space. To improve representation discriminability, AAPID additionally employs *Anchor Loss*. For each class *k*, a representative sample is selected as the anchor, and the corresponding feature representation is defined as $\hat{c}_k = f_{\theta,p(\hat{x}_k)}$

Anchor Loss is constructed based on the cosine similarity between the feature of sample x_k and its anchor:

$$\mathcal{L}_{anchor} = 1 - \frac{\hat{c} * k^T f * \theta, p(x_k)}{|\hat{c}| * 2|f * \theta, p(x_k)|_2}$$

This mechanism encourages samples from the same class to have representations close to their corresponding anchor, thereby forming compact feature clusters and improving class separability. This is particularly useful in the few-shot setting, where the number of training samples per class is very limited..

4.2.5. Objective Function and Model Analysis

Overall Objective Function

AAPID jointly optimizes the semantic representation capability of the prompts and the discriminability of the feature space through the following objective: $\mathcal{L} = \mathcal{L}_{IB} + \lambda \mathcal{L}_{anchor}$ where λ is the balancing coefficient between Information Bottleneck Loss and Anchor Loss. This combination forms a dual optimization mechanism: $\mathcal{L}_{IB}$removes redundant information while preserving features relevant to class labels, whereas $\mathcal{L}_{anchor}$ improves the compactness and separability of feature representations. Consequently, the model maintains effective classification performance and better adaptability when only a very limited number of samples are available for new classes.

Architecture Evaluation

AAPID combines three approaches: a pretrained Vision Transformer, Prompt Learning, and Few-Shot Class-Incremental Learning. A key feature of the architecture is the Information Bottleneck-based Prompt Encoder, which generates adaptive prompts and transfers knowledge from previously learned classes to new classes with limited samples. In addition, Anchor Loss improves feature-space discriminability and helps mitigate catastrophic forgetting during incremental learning. This combination makes AAPID suitable for electrical equipment fault recognition, where new fault types may continuously emerge while only a small amount of training data is available.

4.3. Experiments

The experiments are conducted to comprehensively evaluate AAPID for Few-Shot Class-Incremental Learning (FSCIL) in insulator fault recognition from UAV images. Three main aspects are considered: the ability to learn new classes from a limited number of samples, the ability to preserve knowledge of previously learned classes, and the ability to form a highly discriminative feature space. Experiments are conducted on both the standard CIFAR100 benchmark and the domain-specific VPI-Fault-2500 dataset to jointly assess methodological effectiveness and practical applicability**.**

4.3.1. Experimental Environment

The model is implemented using PyTorch 2.0 with CUDA, with a ViT-B/16 pretrained on ImageNet-1K serving as the backbone. Input images are normalized according to ImageNet statistics and resized to 224 ×224 pixels. Basic data augmentation operations are applied during training to improve generalization.

The base-learning stage is trained for 20 epochs using the SGD optimizer with an initial learning rate of 0.01 and a batch size of 48. The prompt length is set to 3 for CIFAR100 and 5 for VPI-Fault-2500. The coefficient λ in the overall loss function is selected using a validation set. For incremental sessions, the model uses only 5 samples per new class, reflecting a realistic few-shot setting.

4.3.2. Experimental Datasets

The experiments are evaluated on **two datasets**: CIFAR100 and VPI-Fault-2500. **CIFAR100** is used as the standard FSCIL benchmark with 60 base classes and 8 incremental sessions under the 5-way 5-shot setting; images are resized to 224 × 224 for ViT-B/16. **VPI-Fault-2500** contains 2,500 insulator-string images belonging to 5 classes, where *Normal*, *SC*, and *RC* are base classes, while *BB* and *SSC* are introduced in the incremental session under the **2-way 5-shot** setting. AAPID uses only **5 images per new fault class** and does not reuse base-class data, thereby simulating recognition of newly emerging fault types under limited-data conditions.

4.3.3. Experimental Results

The experiments are designed to answer the following three research questions: • **RQ1:** Does AAPID improve performance compared with existing CIL and FSCIL methods? • **RQ2:** What is the contribution of each AAPID component to the overall performance? • **RQ3:** Can AAPID form a highly discriminative feature space between base and new classes?

Evaluation Metrics

Model performance is evaluated using three main metrics:

- Average Accuracy (*Aavg*): the average accuracy across all classes and learning sessions, reflecting the overall model performance.
- Performance Dropping (PD): the accuracy degradation of the base classes after incremental learning defined as $PD = A_0 - A_T$ where A_0, A_T denote the accuracy on the base classes after the first and final sessions, respectively.
- Harmonic Accuracy (HAcc): easures the balance between retaining base classes and learning new classes::$H_{Acc} = \frac{2A_0A_n}{A_0+A_n}$ where A_0 ,A_n is the average accuracy on the new classes..

Overall Evaluation Results

AAPID is compared with three groups of methods (1) The Conventional CIL methods; (2) the specialized FSCIL methods and (3) the prompt-based incremental learning methods

The results in Table 4.1 and Table 4.2 show that AAPID achieves the highest performance on both datasets.

Table 4.1: Performance comparison on the CIFAR100 dataset (Setting: 60+8×5-way 5-shot)

Method	*Aavg*(%)↑	PD(%)↓	HAcc(%)↑
iCaRL	78.4	27.2	57.5
Foster	73.9	35.9	11.0
CEC	81.0	19.0	64.1
FACT	78.6	21.7	55.5
TEEN	87.3	8.8	81.2
L2P	71.1	37.0	0.0
DualP	70.6	36.9	0.1
CodaP	72.0	37.4	0.0
DualP+	81.1	8.5	75.3
AAPID (proposed)	**89.5**	**4.8**	**86.1**

Table 4.2: Performance comparison on the VPI-Fault-2500 dataset (Setting: 3+2-way 5-shot)

Method	*Aavg*(%)↑	PD(%)↓	HAcc(%)↑
iCaRL	85.1	15.4	78.2
Foster	82.6	20.1	65.7
TEEN	91.5	6.2	88.3
DualP+	88.9	8.0	85.1
AAPID (proposed)	**94.2**	**3.1**	**92.5**

Nghiên cứu loại bỏ – Ablation Study (RQ2)

To evaluate the contribution of each component, the modules in AAPID are removed one at a time. The results in Table 4.3 show that every ablated variant reduces *Aavg* on both CIFAR100 and VPI-Fault-2500, confirming the complementary roles of the model components

Method	CIFAR100	VPI-Fault-2500
AAPID (full)	**89.5**	**94.2**
Without Task-Invariant Prompts (TIP)	88.9	93.6
Without Task-Specific Prompts (TSP) (TSP)	87.8	92.1
Without Anchor Loss	88.7	93.4
Without IB Loss	88.2	92.9

Feature Space (RQ3)

To evaluate the model's ability to learn a discriminative feature space, the features extracted from the Query Set of VPI-Fault-2500 were visualized using t-SNE, as illustrated in Fig. 4.2.

Figure 4.2: t-SNE visualization of the ablation study results on the VPI-Fault-2500 dataset.

4.4. Chapter 4 Conclusion

Chapter 4 proposed AAPID for insulator fault recognition under limited data for new classes within the Few-Shot Class-Incremental Learning (FSCIL) framework. The method employs a pretrained Vision Transformer with adaptive prompting, incorporating Task-Invariant Prompts, Task-result Prompts, Information Bottleneck Loss, and Anchor Loss.

CONCLUSIONS AND RECOMMENDATIONS

The thesis investigates and develops deep learning techniques for UAV-based monitoring of power transmission lines, focusing on three main tasks: equipment detection, anomaly detection and segmentation in insulator strings, and few-shot incremental learning for newly emerging fault types.

First, the proposed HVE-YOLO model detects six groups of power transmission equipment, achieving an mAP@0.5:0.95 of 88.5%, while reaching

45.2 FPS, 11.4 GFLOPs, and 9.8 W on an NVIDIA Jetson Orin Nano, demonstrating its potential for edge deployment.

Second, the FGS model for anomaly detection in insulator strings achieved a Precision of 92.3%, Recall of 91.4%, and mAP@0.5 of 93.7% on the VNelectric dataset. Ablation experiments confirmed the contributions of GELAN, RPN, and Diffusion components to equipment detection, classification, and anomaly segmentation in insulator strings.

Third, AAPID for Few-Shot Class-Incremental Learning enabling the model to learn new fault classes from very limited samples while retaining previously acquired knowledge. With only five samples per new class, AAPID achieved an of 89.5%, of 4.8%, and of 86.1% on CIFAR100, and 94.2%, 3.1%, and 92.5%, respectively, on VPI-Fault-2500.

Limitations

However, several limitations remain: (i) model performance depends on image quality and environmental conditions; (ii) some models, particularly FGS, have high computational costs; (iii) the number of newly emerging fault types evaluated in the FSCIL setting remains limited; and (iv) AAPID still depends on the representational capability of the pretrained ViT and has not been fully optimized for resource-constrained hardware.

Future Research Directions

Future research will focus on five main directions (1)Multimodal data integration: Combining optical images with thermal images, LiDAR, and sensor data to improve the detection of anomalies that are difficult to observe from conventional imagery; (2) Real-time model optimization: Investigating pruning, quantization, knowledge distillation, low-rank adaptation, and lightweight architectures to reduce computational costs, particularly for large models such as FGS;(3) Early fault prediction: Combining temporal image sequences, thermal imagery, and sensor data with LSTM, Transformer, or other deep sequential models to monitor degradation trends and predict potential failures;(4) Improved generalization: Expanding datasets and equipment types while investigating transfer learning, domain adaptation, and continual learning to enable adaptation across different regions, operating conditions, and power systems ; and (5)Intelligent decision-support systems: Integrating equipment detection, anomaly recognition, and incremental learning with UAV imagery, thermal images, LiDAR, SCADA data, and expert knowledge to assess risks and support inspection and maintenance.

Long-Term Direction is to develop an intelligent power-grid inspection system based on UAVs and artificial intelligence, integrating multimodal data, deep learning, and continual learning to enhance automation, adaptability, and reliability in power-grid inspection, operation, and maintenance.

LIST OF THE PUBLICATIONS RELATED TO THE DISSERTATION

CT1. **Trịnh Hiền Anh**, Nguyễn Thị Thanh Tân, Bùi Xuân Hùng, Đoàn Ngọc Duy, "*Phát hiện thiết bị trên lưới truyền tải điện cao thế 110kV sử dụng kỹ thuật học sâu*". Hội thảo Quốc gia lần thứ XXV: Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông, Hà Nội, 8–9/12/2022, pp.174–179.

CT2. **Anh Trinh Hien,** A. D. Tran, D. C. Viet, Q. D. T. Thuy and Q. N. Huu, "Insulator defect detection based on feature pyramid network and diffusion model". Signal, Image and Video Processing, 2025, Vol. 19, Issue 5, p.372 (SCIE-Q3).

CT3. **Anh Trinh Hien**, Tao Ngo Quoc, Thanh-Tan Nguyen Thi, "AttentionAware Prompting for Learning New Faults on Insulators with Limited Data". Journal of Information Hiding and Multimedia Signal Processing, Vol. 16, No. 4, December 2025, pp.1199–1208 (SCOPUS-Q4).

CT4. **Trịnh Hiền Anh**, Nguyễn Thị Thanh Tân, "Một phương pháp hiệu quả để phát hiện đường dây, cột điện và các thiết bị trên lưới truyền tải điện cao thế 110kV bằng kỹ thuật học sâu". Tạp chí Khoa học và Công nghệ Năng lượng – Trường Đại học Điện lực, Số 32, 2023, ISSN: 1859-4557, pp.1–11.

CT5. **Trịnh Hiền Anh**, Nguyễn Thị Thanh Tân, "Một kỹ thuật tăng cường chất lượng ảnh UAV thu nhận trên hệ thống truyền tải điện cao thế 110kV ". Hội thảo Quốc gia lần thứ XXVI: Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông, Bắc Ninh, 5–6/10/2023, pp.206–209.

CT6. Thi-Thanh-Tan Nguyen, **Anh Trinh Hien**, Tao Ngo Quoc, "HVE-YOLO: A multi-scale neural network integrated for high-voltage power grid equipment recognition", Journal of Information Hiding and Multimedia Signal Processing, Accepted