RESEARCH ARTICLE

Target Recognition of Multi Source Machine Vision Pan Tilt Integrated Inspection Robot for Power Inspection

MINJIE ZHU, BINGYE ZHANG , CHEN ZHOU, HONGLIANG ZOU, AND XUEYAN WANG

State Grid Zhejiang Electric Power Supply Company Taizhou Branch, Taizhou 317700, China

Corresponding author: Bingye Zhang (zhangbingye198@163.com)

ABSTRACT With the continuous changes in socio-economic needs, traditional methods of power facility inspection can no longer meet practical needs due to their low efficiency and lack of scalability. In response to this challenge, this study delves into the integrated motion control technology of inspection robots equipped with gimbal mechanisms, aiming to improve the convenience and efficiency of dynamic data collection. A customized multi-source heterogeneous visual detection and recognition model based on the YOLOv3 framework has been proposed, and simultaneously using path aggregation networks to enhance information processing capacity by fusing multi-scale features. Experimental analysis shows that as the robot's movement speed increases, the error rate correspondingly increases, indicating the direction of optimization. In the target recognition experiment, the proposed model achieved an average accuracy of 94.26% in visible light images and 68.05% in infrared images. In addition, Sub_ The YOLO algorithm demonstrates a fast detection speed of 30 frames per second, with an average accuracy of over 80%, marking an important progress in real-time object detection applications. In the linear motion test, the relative error of the robot's motion accuracy was 0.33% at a speed of 500 millimeters per second. However, when the speed was increased to 1200 millimeters per second, the error increased to 2.45%, indicating a significant increase in slip. This indicates that the linear motion accuracy of the robot is acceptable at low to medium speeds, but the accuracy decreases significantly at high speeds. Overall, the research results confirm the synergistic effect of integrated motion control between inspection robots and gimbals, as well as Sub_ The superiority of YOLO in target recognition has improved the ability to use wheeled robots for electrical inspections, bringing substantial technological progress to the field of autonomous inspection.

INDEX TERMS Integrated motion control, multi-source heterogeneous detection, Sub_YOLO, autonomous power inspection, path aggregation network.

I. INTRODUCTION

In the current era of Industry 4.0, the power system, as an indispensable and important component of modern society, its operational efficiency and safety are of crucial significance for social operation. With the progress of social economy, the complexity and scale of power system facilities are constantly increasing. The traditional inspection methods for power facilities can no longer meet the demand. This is

The associate editor coordinating the review of this manuscript and approving it for publication was Yizhang Jiang .

mainly manifested in low efficiency of manual inspections, high risks, and insufficient detection of potential hazards [1], [2], [3]. Therefore, how to effectively conduct power facility inspections to ensure the stable operation of the power system is a major challenge facing the current power industry. In this context, the research on power inspection robots based on advanced machine vision technology has emerged. Among them, the multi-source machine vision pan tilt integrated inspection robot for power inspection has led the new trend of modern power inspection due to its stability and practicality [4], [5], [6]. It can integrate multiple sources of

information. The machine vision technology is applied to complete efficient target recognition. While ensuring inspection quality, it improves inspection efficiency, greatly reduces the labor intensity and safety hazards of inspection personnel, and provides strong technical support for the safe and stable operation of the power system. However, the challenges faced by the multi-source machine vision pan tilt integrated inspection robot are still considerable.For example, how to improve the accuracy of target recognition. How to better handle complex information from different sources. How to optimize the operability of robots. Based on these issues, the research on multi-source machine vision pan tilt integrated inspection robots needs to be further deepened [7], [8], [9]. Therefore, the integrated motion control technology of power inspection robots and pan tiltis analyzed to provide a more effective and safe new mode of power facility inspection. The first part introduces the research purpose through recent research. The second part designs a motion control and target recognition strategy for a multi-source machine vision integrated inspection robot. The third part verifies the effectiveness of the strategy through simulation experiments. The fourth part draws conclusions.

With the development of social economy and the increasing complexity of the power system, the current power facility inspection method relying on manual operation can no longer meet the growing demand for inspection efficiency and accuracy. Traditional methods not only have low efficiency, but also pose potential safety risks in high voltage working environments. The main motivation of this study is to improve work efficiency while avoiding security risks.

This study provides an advanced integrated control, multi-source heterogeneous visual processing, and information flow system, which improves the automation level of power facility inspections and enhances the performance of robot vision systems in real-world work environments.

The contributions are 1) to provide a more efficient and improved monitoring algorithm for environments with small targets and high occlusion, which enables intelligent detection of power facilities to maintain monitoring efficiency in more complex environments; 2) to combine different levels of information features, which can resultin a more practical improvement method for multi-source heterogeneous machine vision technology, and further enhances the accuracy of object detection; and 3) to combine multi-scale feature integration technology and information flow technology, which can promote the development of object detection technology and verify the application value of this technology combination.

There are still certain limitations in the practical application of existing technologies. Firstly, existing machine vision systems mainly face the problem of difficulty in accurately identifying targets in complex environments, especially in multi-source information processing. Ensuring effective synchronization and integration of information from different types of sensors remains a technical challenge. Secondly,

image capture and processing in special environments (such as harsh weather) are challenges faced by existing technologies, and harsh environments often require higher robustness of machine vision systems. Finally, from the perspective of operability and ease of use, existing inspection robot systems require highly skilled operators for monitoring and control, and designing a more intuitive and automated operating interface is a potential requirement. The multi-source machine vision gimbal integrated inspection robot proposed in this study has certain gaps compared to existing technologies in the above aspects. The research attempts to improve the accuracy of target recognition and adaptability to complex environments by integrating machine vision information from different sources. Enhance the autonomy of inspection robots and the ease of use of operating interfaces to reduce reliance on manual operations.

The first section of the study proposes the research objectives, and the second section designs an integrated motion control scheme and a multi-source heterogeneous machine vision target recognition model. The third section analyzes the performance of the model and compares it with the benchmark model. Finally, the fourth section draws research conclusions

II. RELATED WORK

At present, many scholars both domestically and internationally have conducted research on hardware improvements and technological upgrades of power inspection robots. Zhonglin et al. studied a lightweight nuclear power plant inspection robot. Static analysis is conducted on the key parts of the robot. The stiffness and strength of the mechanical structure are determined to meet the requirements of lightweight design. With the help of modal analysis, it has been proven that there is no resonance during motor operation. A hierarchical control system is established on LabVIEW to simulate walking control, proving that adaptive fuzzy PID control has better performance than conventional PID control. Accurate control of lifting mechanism position positioning is achieved using the S-type acceleration and deceleration algorithm [10]. Liu et al. proposed a new type of motion balance adjustment controller for power line inspection robots. Sliding mode control is applied to complete motion balance adjustment work. A sliding surface is designed using the Ackermann formula. A universal type-2 fuzzy system is applied to improve the anti-interference performance. It enhances the robot's anti-interference ability while achieving motion balance control [11]. Li et al. proposed a patrol robot equipped with retractable double links in crawling performance analysis of a power line inspection robot equipped with a retractable double link controller. Through static analysis and dynamic evaluation indicators, it is shown that the new inspection robot has stronger performance in the crawling process. The dynamic model demonstrates that this new robot has a smaller driving torque when crawling. It can better cope with different obstacles [12]. Huang et al. proposed a new

ranging method for the robustness of obstacle distance measurement under changing lighting conditions. By integrating image enhancement and monocular vision, it can adapt to various illumination conditions along the line. According to the findings, the error of static distance measurement is less than 3%. The error of dynamic distance measurement is less than 5%. This method has made significant contributions to the sustainable development of inspection robots [13]. Wang et al. studied a new type of cable crawling robot. The robot can control the ground station wirelessly. It is powered by a built-in lithium battery. A static model of the robot during obstacle negotiation is established. The relationship between driving force, resistance, and obstacle height is analyzed to determine the robot's obstacle negotiation ability. The experiment has verified that the robot can climb any inclined cable. It can overcome obstacles with a height of 2.42mm and a load of 5kg [1].

The robot target recognition in power inspection has also received widespread attention. Liu et al. explored the issue of using radar and camera information fusion for target recognition and tracking of self driving vehicles under adverse weather conditions. The fusion scheme is based on a framework with radar as the main hardware and cameras as auxiliary hardware. The Markov distance is applied to match the observed values of the target sequence. The test results of sensor data collected by actual vehicles indicate that the fusion algorithm of radar and camera has better environmental perception performance than a single sensor in harsh weather. It can effectively reduce the missed detection rate of self driving on environmental perception in adverse weather conditions [14]. Shi et al. proposed the synthetic aperture radar (SAR) image target recognition algorithm based on attention mechanism. The model extracts information in two stages. Firstly, the noise is reduced. Then mixed attention is added to gradually extract advanced features. On the MSTAR dataset, compared to advanced algorithms, it can significantly reduce the parameters and complexity [15]. Du et al. proposed a practical deceptive interference method based on fragile position perception for adversarial attacks. It is used for radar high-resolution range profile (HRRP) target recognition. It learns fragile distance units from HRRP samples. Several interference pulses with specific amplitudes are injected into these distance units. This interference signal is easy to generate. It can achieve high prediction of incorrect target categories [16]. Wang et al. proposed a dense capsule network for SAR automatic target recognition. Their main contributions are twofold. Firstly, the original convolutional layer is replaced with dense blocks to increase the depth of the network. Secondly, the fully connected layer is replaced with a deconvolution layer to construct a reconstructed network. On the MSTAR dataset, dense capsule network outperforms other methods on limited data [17].

Diwan et al. compared the performance and application of single-stage and two-stage detectors, and pointed out the advantages and disadvantages of YOLO algorithm and its

evolutionary version in terms of speed and accuracy. The YOLO algorithm has significant advantages in computational speed and slightly inferior accuracy compared to traditional two-stage object detectors [18]. Liu et al. proposed a new framework called Image Adaptive YOLO (IA-YOLO), which improves YOLOv3 in an end-to-end manner to adapt to various climate conditions through a differentiated image processing module and a small CNN network [19]. In order to improve the target detection rate and reduce false alarms in infrared images, Li et al. optimized the YOLOv5 algorithm and designed an improved network called YOLO-FIRI. By expanding the modules in the feature extraction network, introducing an improved attention mechanism, and adding multi-scale detection, the detection accuracy of small targets has been significantly improved, and further performance improvements have been achieved on multiple datasets compared to existing optimal methods [20].

In the relevant research and commercial applications, substation inspection robots have made significant progress in both theory and practice. Among them, autonomous wheeled robots have become the main development direction due to the high transmission efficiency, flexible movement, and simple control methods. However, there are still some issues with these robots in motion control accuracy, navigation and positioning performance, and flexibility in scheduling inspection tasks. Further optimization and improvement are needed in robustness and intelligence. Therefore, an integrated motion control technology for dynamic acquisition of inspection data is proposed to improve the performance of power inspection robots and meet practical application requirements.

III. TARGET RECOGNITION OF MULTI-SOURCE MACHINE VISION PAN TILT INTEGRATED INSPECTION ROBOT FOR POWER INSPECTION

The integrated motion control technology of inspection robots and pan tilt is studied. Kinematic analysis is conducted on the robot and the platform. An integrated kinematic model of the robot and platform is established, achieving visual servo control for dynamic acquisition of image data. At the same time, multi-source heterogeneous machine vision detection and object recognition technologies are studied. YOLOv3 is used to construct a universal detection model for multi-source heterogeneous vision. Integrating path aggregation networks and deep separable convolutions, high-performance multi-source heterogeneous visual detection and target recognition are achieved. A software system suitable for robot inspection is developed based on the design of wheeled autonomous mobile robots and platform hardware.

The main challenge in establishing an accurate kinematic model for inspection robots is to interpret and analyze their motion states in complex electrical environments. For example, robots are affected by non holonomic constraints, control system dynamics, environmental uncertainty, and other issues. To overcome these issues, this study adopted
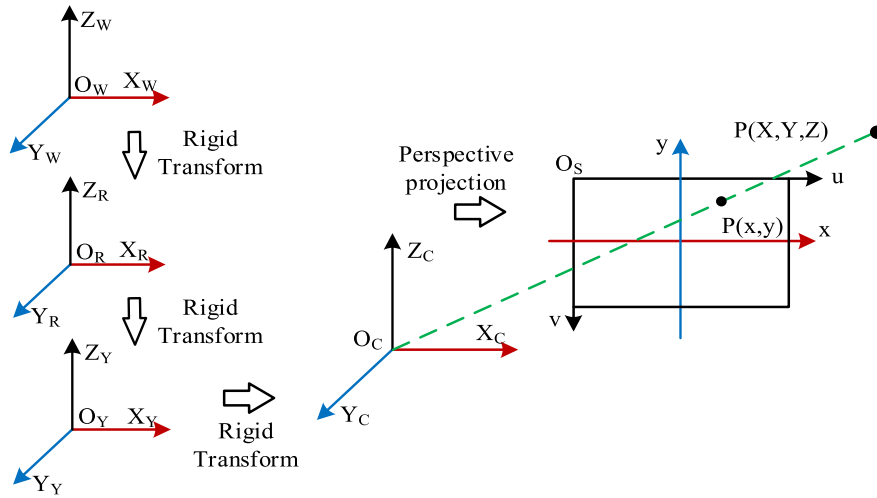
FIGURE 1. Robot inspection coordinate system.

the differential drive assumption to simplify the model into a planar three degree of freedom motion model, where the robot can translate along the X and Y axes and rotate around the Z axis. In addition, the study simplifies nonlinear models into linear models and adopts methods suitable for linear systems when designing control algorithms. This approach greatly improves the feasibility and practicality of the model, enabling effective robot positioning and path planning under limited computing resources.

A. INTEGRATED MOTION CONTROL TECHNOLOGY OF POWER INSPECTION ROBOT AND PAN TILT

In three-dimensional space, the position of any point can be represented by the position vector of 3×1. It also indicates which coordinate system the point is located on. For example, in the $\{A\}$ coordinate system, the position vector of P can be represented by $^A P$. A represents the reference coordinate system $\{A\}$. To determine the pose changes, a Cartesian coordinate system $\{B\}$ can be established that is fixedly connected to object K. It is used to describe the pose of object K relative to coordinate system $\{A\}$. The principal axis vector of coordinate system $\{B\}$ is used to construct a rotation matrix of 3×3. This matrix is called the rotation matrix relative to the reference coordinate system $\{A\}$. Before conducting kinematic modeling of the substation robot inspection system, the associated coordinate system is introduced. The relationship between each coordinate system is shown in Figure 1.

The camera model is applied to describe the mapping relationship between spatial 3D points and planar 2D pixel points. The most commonly used camera model is the monocular pinhole camera model. It is based on the principle of small hole imaging, describing a set of linear transformation relationships. The model includes two parts, internal and external parameters. The internal parameters include information such as focal length f and the optical center coordinate (cx, cy) of the image. The external parameters represent the pose

transformation relationship between the world coordinate system $\{W\}$ and the camera coordinate system $\{C\}$. Through these internal and external parameters, the camera model can map the 3D point P in space to the 2D pixel P' on the physical imaging plane. In practical applications, due to factors such as camera manufacturing process and internal lens assembly, nonlinear distortion can occur in the image imaging process. This distortion is mainly divided into radial distortion and tangential distortion. To accurately describe the correspondence between the target and the camera image, camera calibration is necessary to determine the linear and nonlinear parameters of the camera.

Kinematic modeling is conducted on the inspection robot. The inspection robot has non holonomic constraints. Moreover, establishing an accurate kinematic model is relatively difficult. The study adopts differential drive. According to the assumption, it is simplified as a planar three degree of freedom motion model, which involves translation along the X and Y directions and deflection around the Z axis. Figure 2 illustrates the kinematic model of the inspection robot.

The coordinate of the robot relative to the world coordinate system $\{W\}$ is (x, y). The yaw angle is θ. b is the general value of the left and right wheel spacing. b is the radius of the wheel. v_i is the linear speed of each wheel. v and ω are the movement speed and angular velocity of the inspection robot, respectively. R is the instantaneous turning radius of the inspection robot. The state variable $q = [x\ y\ \theta]^T$ represents the pose of the inspection robot. The robot pose is described in equation (1).

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ \omega \end{bmatrix} \quad (1)$$

The substation inspection robot moves from (X_t, Y_t, θ_t) to $(X_{t+1}, Y_{t+1}, \theta_{t+1})$ during the time interval t. Its motion trajectory is a small arc. $\Delta\theta = \theta_{t+1} - \theta_t$. If the position of the inspection robot at t is known as (X_t, Y_t, θ_t), then the

(a) Robot motion model

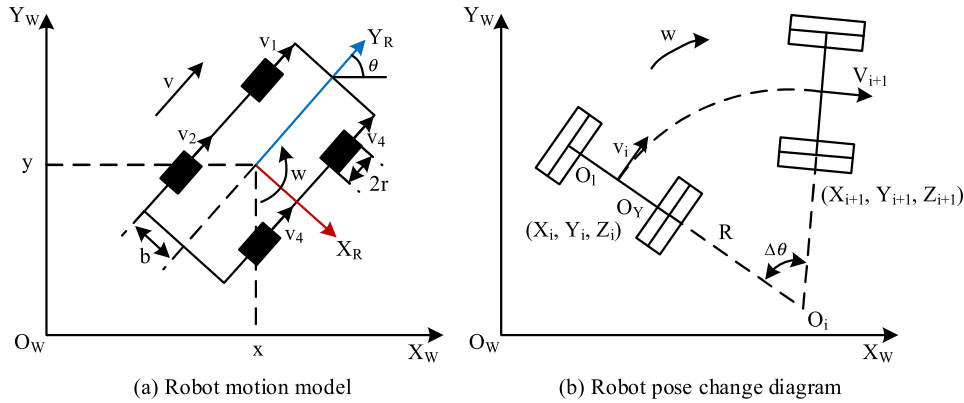(b) Robot pose change diagram

FIGURE 2. Robot motion model and pose changes.

position information $(X_{t+1}, Y_{t+1}, \theta_{t+1})$ at $t + 1$ is known as equation (2).

$$
\begin{cases}
X_{t+1} = X_t + \Delta x_t \cos \theta_t - \Delta y_t \sin \theta_t \\
Y_{t+1} = Y_t + \Delta x_t \sin \theta_t - \Delta y_t \cos \theta_t \\
\theta_{t+1} = \theta_t + \omega \Delta t
\end{cases} \tag{2}
$$

Pan tilt is a device used to support fixed detection devices, widely used in scenes that require moving object detection and tracking. As a two degree of freedom robotic arm, the pan tilt can be kinematic described using the D-H method. The D-H method uses four parameters $(a_{i-1}, \alpha_{i-1}, d_i, \theta_i)$ to represent the transformation relationship between adjacent connecting rods. a_{i-1} and α_{i-1} represent the distance and angle between the joint axes at both ends of the connecting rod. d_i and θ_i represent the positional relationship between adjacent connecting rods. The pose transformation matrix A is commonly used to describe the pose transformation relationship between two linkage coordinate systems. If A_1 and A_2 are the pose matrices of the corresponding coordinate systems of different mechanisms, then the pose of the second mechanism in the base coordinate system {0} is displayed in equation (3).

$$
T_2 = A_1 A_2 \tag{3}
$$

The D-H is used to establish the coordinate system of the robotic arm link. Firstly, a coordinate system is established based on a few point coordinate system. The i-th coordinate system is $\{i\}$. The coordinate axis is determined. The expected motion plane of the z_i axis of the rotating joint is perpendicular. The expected motion direction of the z_i axis of the sliding joint coincides. The x_i-axis is perpendicular to the z-axis of the two connecting rods. The y-axis is determined according to the right-hand rule. Then the D-H parameters are determined. The coordinate system $\{O_{i-1}\}$ is rotated θ_i around the z_{i-1}-axis and translated d_i along the axis, so that the two connecting rods' x-axis are collinear and coincident. The coordinate system $\{O_{i-1}\}$ shifts x_{i-1} along the x_{i-1} axis and rotates α_{i-1} around the axis, so that the two coordinate systems completely coincide. The transformation matrix

is multiplied sequentially to obtain the pose transformation matrix between the linkage coordinate systems $\{i-1\}$ and $\{i\}$. According to this method, a two degree of freedom linkage coordinate system is established for the pan tilt. The origin of the pan tilt coordinate system $\{Y\}$ is in the center of the pan tilt horizontal rotation base. The origin of the pitch rotation joint is fixed at the corresponding geometric center. The joint axis is taken as the Z-axis. The end load is the camera coordinate system $\{C\}$. The two degree of freedom pan tilt is shown in Figure 3.

The Jacobian matrix describes the differential motion relationship of feature points. If the robotic arm has N joint angles, then the Jacobian matrix is a $6 \times N$ dimensional matrix. According to the principle of velocity transfer matrix, the Jacobian matrix is divided into two parts: translation and rotation, as shown in equation (4).

$$
dX_i = J(q) \cdot dp_i \tag{4}
$$

In equation (4), p is the position matrix $p = [x \ y \ z]^T$ of any point on the i-th coordinate system. The translational Jacobian matrix in the base coordinate system is shown in equation (5).

$$
J_1 = \begin{bmatrix} \dfrac{\partial p}{\partial \theta_1} & \dfrac{\partial p}{\partial \theta_2} & \cdots & \dfrac{\partial p}{\partial \theta_i} & 0 & \cdots & 0 \end{bmatrix} \tag{5}
$$

The Jacobian matrix of the rotating part is the differential of the Euler angle with respect to the rotation angles of each joint, as shown in equation (6).

$$
J_2 = \begin{bmatrix} {}^0_1 r_z & {}^0_2 r_z & \cdots & {}^0_N r_z & 0 & \cdots & 0 \end{bmatrix} \tag{6}
$$

The Jacobian matrix is shown in equation (7).

$$
J = \begin{bmatrix} \dfrac{\partial p}{\partial \theta_1} & \dfrac{\partial p}{\partial \theta_2} & \cdots & \dfrac{\partial p}{\partial \theta_i} & 0 & \cdots & 0 \\ {}^0_1 r_z & {}^0_2 r_z & \cdots & {}^0_N r_z & 0 & \cdots & 0 \end{bmatrix} \tag{7}
$$

Considering the motion constraints of the inspection robot, the overall Jacobian matrix of the substation robot inspection is shown in equation (8).

$$
J_{all} = \begin{bmatrix} O_{3 \times N} & v_1 \\ O_{3 \times N} & v_2 \end{bmatrix} \tag{8}
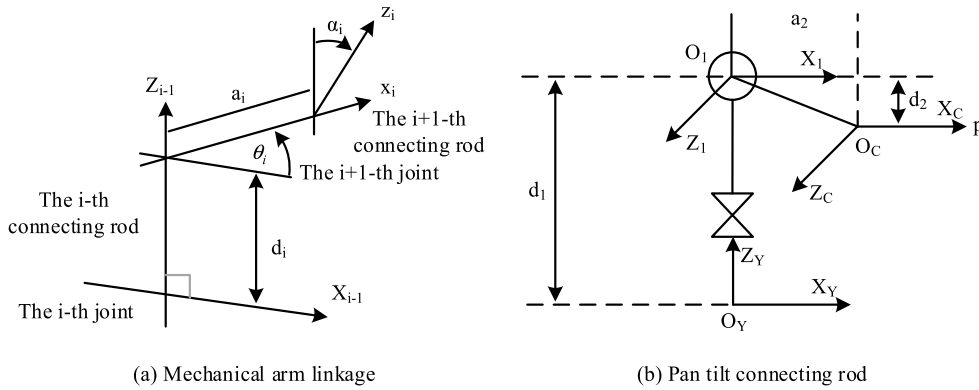$$

(a) Mechanical arm linkage   (b) Pan tilt connecting rod

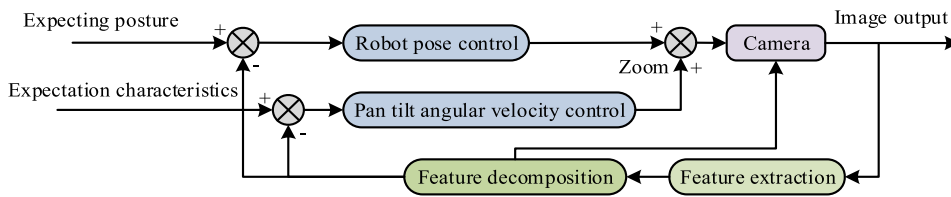**FIGURE 3.** Connecting rod coordinate system.

FIGURE 4. Motion control block diagram for dynamic acquisition of inspection data.

The pose of the target point and the robot in the $\{W\}$ is represented by vectors $^{W}P_0$ and $^{W}P_1$. The origin of $\{C\}$ is $^{R}P_2$ in robot coordinate system $\{R\}$. The pose of the target point in the $\{C\}$ is $^{C}P_3$. From this, the geometric relationship equation between each vector can be obtained, as shown in equation (9).

$$^{W}P_0 = {}^{W}P_1 + {}^{W}P_2 + {}^{W}P_3 = {}^{W}P_1 + {}_{R}^{W}R\,{}^{R}P_2 + {}_{R}^{W}R\,{}_{C}^{R}R\,{}^{C}P_3 \tag{9}$$

In equation (9), $_{R}^{W}R$ is the pose transformation matrix from $\{R\}$ to $\{W\}$. $_{C}^{R}R$ is the pose transformation matrix from $\{C\}$ to $\{R\}$.

An integrated motion control technology for dynamic acquisition of inspection data is proposed. By combining visual features with robot motion, the challenges of large-scale scanning and image acquisition are solved. This technology includes a robot visual servo system consisting of a mobile robot, a pan head, and a camera, and a control strategy consisting of two inner and outer loops. The outer ring is responsible for image acquisition and feature extraction. The inner ring is used for the positioning and scanning motion of the robot. Figure 4 displays the motion control block diagram.

At present, visual servo control systems often use PID algorithm. Its principle is simple and easy to implement. The feedback information of the image loop is the two-dimensional pixel deviation of the target, which is adjusted to converge to 0. However, when the deviation is small, it is prone to regulatory oscillations. To solve this problem, a PID control with dead zone is studied. A threshold is set to eliminate the controller's jitter in the near steady state.

When the target is within the circular threshold range of the camera's view center, the motor is controlled to maintain the camera's current pose. The threshold is set to 10 pixel values.

PID control with dead zone plays a regulatory role in visual servo systems. In the case of small deviations, traditional PID algorithms are prone to causing control signal jitter, which affects the stability of the system. To solve this problem, the concept of dead zone is introduced, which means that when the deviation between the target and the center of the field of view is within a certain range, the system does not respond and maintains its current state. This significantly reduces frequent adjustments caused by minor positional changes, avoiding excessive system response. In the application of this study, the dead zone threshold was set to 10 pixel values. When the center point of the prediction box aligns with the center point of the field of view box within this range, the robot will stop moving until it detects the next target tracking. Under the action of PID control with dead zone, the robot can smoothly and accurately complete inspection tasks, and still maintain high detection ability when encountering small range target position changes.

When the target position deviates from the center of the view, the camera position is adjusted to align the center point of the prediction box with the center point of the view box, as shown in equation (10).

$$\begin{cases} X_c = \dfrac{W}{2}, Y_c = \dfrac{H}{2} \\ x_c = w_c + \dfrac{w}{2}, y_c = h_c + \dfrac{h}{2} \end{cases} \tag{10}$$

In equation (10), the size of the captured image is $W \times H$. The center point is (X_c, Y_c). The prediction box size is $w \times h$. The

coordinate is (x_c, y_c). The coordinate of the top left corner of the prediction box is (w_c, h_c). According to the relationship between the proportion of the current prediction box and the view area and the ideal proportion, the zoom factor of the camera is adjusted, as shown in equation (11).

$$p = \sqrt{q \cdot S_2 / S_1} \tag{11}$$

In equation (11), S_1 is the area of the instrument prediction box. S_2 is the area of the captured image. q is the ratio of the predicted box area to the view area under ideal conditions.

The integrated motion control technology proposed in this study utilizes advanced visual servo systems to effectively solve the challenges of scanning and image acquisition in large-scale power facility inspections. The core of the technology lies in a dual loop control strategy, with the outer loop responsible for automatic image acquisition and feature extraction, capturing high-quality images by dynamically adjusting camera parameters. The inner loop controls the precise positioning and scanning motion of the robot, ensuring coverage of all inspection points by precisely controlling the movement of the gimbal and robot. This method not only improves inspection efficiency and coverage, but also reduces human errors and missed detection during the inspection process through automatic calibration and optimization of the image acquisition process.

B. MULTI SOURCE HETEROGENEOUS MACHINE VISION DETECTION AND OBJECT RECOGNITION TECHNOLOGY

Visual detection and object recognition locate the position of substation equipment targets in the image and determine the category of the targets by processing the collected images. However, single source visual inspection cannot meet the detection needs of various equipment characteristics in substations. There are various types of substation equipment, and there are significant differences within the categories. Therefore, traditional detection algorithms require customized research and development for each category. When the size and position of the device image change, the parameters need to be readjusted. Its adaptive ability is poor. Therefore, multi-source heterogeneous machine vision detection and object recognition technologies are analyzed.

The research is based on the YOLOv3 network framework for improvement. A multi-source heterogeneous visual detection and recognition model Sub_YOLO is designed. The YOLO network framework is a typical single stage algorithm with the characteristics of fast computation and high-precision detection, making it very suitable for substation equipment detection. The YOLOv3 network has three parts, a backbone network layer, a feature fusion layer, and a universal detection layer. The backbone network layer adopts deep convolutional neural networks to aggregate multiple features on images of different dimensions. The feature fusion layer integrates multiple feature layers generated by the backbone network layer, integrating features from different levels. The universal detection layer classifies and regresses the fused image features. Different strategies are used to generate

prediction boxes and label categories. The general architecture of YOLOv3 is shown in Figure 5.

The backbone network layer adopts the Darknet53 structure. Each residual block consists of two convolutional layers and a skip connection, which can alleviate the network degradation problem caused by deep convolution. The feature fusion layer is located in the middle, middle, and bottom layers of the backbone network, extracting three feature layers for object detection. After processing the feature layers, they are fused with other feature layers through deconvolution and output to the detection layer. K-means clustering is used to form anchor boxes. The feature fusion layer outputs feature layers for multi-scale prediction, obtaining regression results for classification and position.

The Intersection Over Union (IOU) of the prediction box represents the overlap degree between the prediction box generated by the detection model and the actual box. The ratio of the intersection and union of two rectangular boxes is shown in equation (12).

$$IOU = \frac{Pre \cap GT}{Pre \cup GT} \tag{12}$$

In equation (12), *Pre* is the prediction box. *GT* is the real box. When the loss is 1, *IOU* is the difference in the intersection and union ratio between the predicted box *A* and the actual box *B*, as shown in equation (13).

$$L_{IOU} = 1 - IOU(A, B) \tag{13}$$

From a geometric perspective, *IOU* can measure the overlap degree between two rectangular boxes by calculating the ratio of the intersection area to the union area. However, *IOU* cannot fully reflect the inclusion and complete non overlapping situations in the loss function. When the size of the predicted box is different from the actual box, the center point is misaligned, and the distance relationship does not match, *IOU* cannot be accurately reflected. A series of improved loss function methods are proposed to address the non overlapping and inclusion relationships between prediction boxes and real boxes. Among them, CIOU (Complete IoU) considers factors such as the distance, overlap rate, and scale between the target box and the prediction box. The penalty factors are used to describe the aspect ratio relationship between the two, thereby providing more stable regression prediction results. The calculation of CIOU is shown in equation (14).

$$CIOU = IOU - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \tag{14}$$

In equation (14), $\rho^2(b, b^{gt})$ is the Euclidean distance between the predicted box and the real box. c stands for the diagonal distance between the smallest rectangle containing the predicted box and the actual box. The definitions of α and v are shown in equation (15).

$$\begin{cases} \alpha = \dfrac{v}{1 - IOU + v} \\ v = \dfrac{4}{\pi^2}\left(\arctan\dfrac{w^{gt}}{h^{gt}} - \arctan\dfrac{w}{h}\right)^2 \end{cases} \tag{15}$$
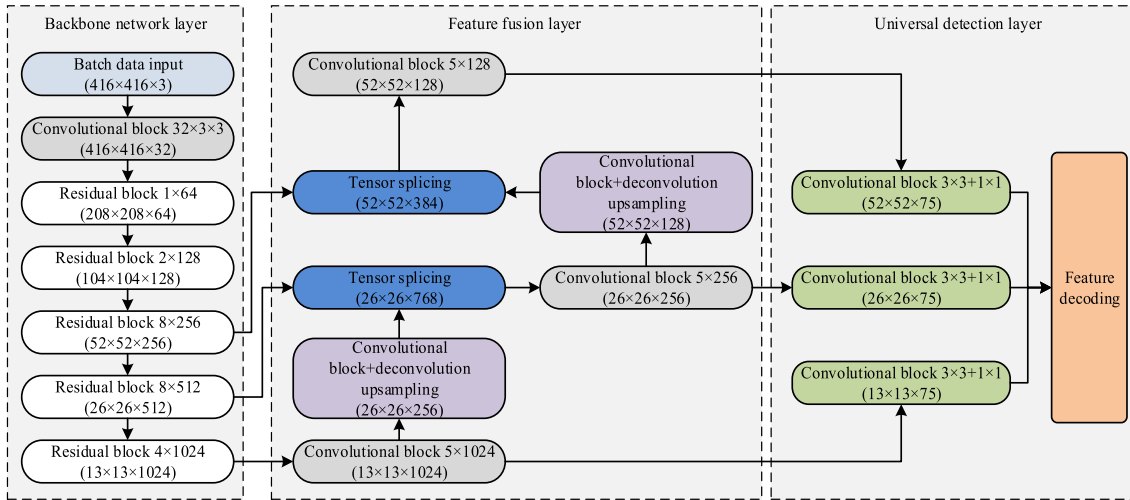
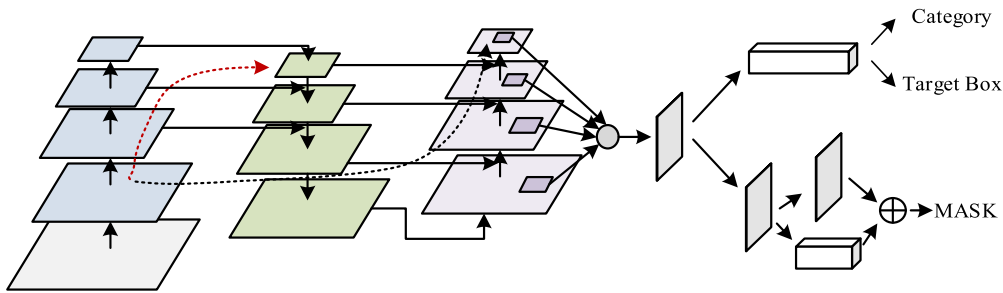**FIGURE 5.** YOLOv3 network framework structure.

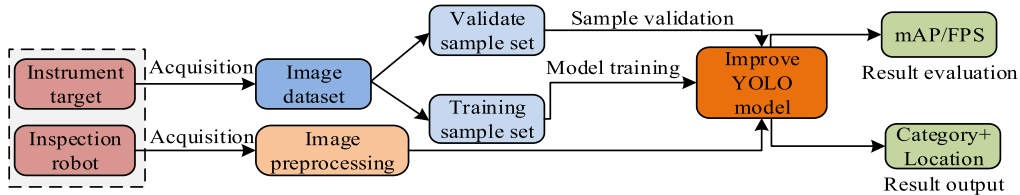**FIGURE 6.** PANet network architecture.

FIGURE 7. Target detection and recognition process.

In equation (15), *a* and *v* introduce aspect ratios to optimize the difference in width and height between the predicted box and the actual box. The corresponding loss function is shown in equation (16).

$$L_{CIOU} = 1 - IOU + \frac{\rho^2 \left(b, b^{gt} \right)}{c^2} + \alpha v \qquad (16)$$

A series of methods are adopted in target detection to improve the performance. The network structure of PANet is shown in Figure 6.

Firstly, data augmentation and fusion paths are performed to enhance the data during training. Subsequently, a multi-source heterogeneous sensor detection and recognition network for internal substations is designed. Finally, the path aggregation network PANet is introduced to accelerate

information flow and integrate features from different levels to more accurately locate pixels. PANet adopts a bottom-up path enhancement structure, which takes the previous layer feature mapping as input and processes it. Then it integrates with the same layer features of FPN to enhance the propagation of underlying features. At the same time, adaptive feature pooling is used to aggregate candidate regions and all features, avoiding arbitrary allocation of candidate regions. Finally, the fully connected layer integrates dual branches for more accurate pixel segmentation and improves detection performance.

The network model and training process are implemented using PyTorch and trained on the VoC2007 dataset. PyTorch is a simple and widely used framework for building and training neural networks, which can quickly build and debug

different detection frameworks. During the training process, transfer learning methods are adopted. Starting from zero to train the model requires a lot of time and data. However, the image dataset of substation equipment is relatively small, which cannot fully support the training of the network framework. Therefore, optimization can be achieved through transfer learning. Transfer learning utilizes the already trained old model as a pre-trained model, freezes some layers for fine-tuning, and generates a new model suitable for substation equipment detection. The training process is shown in Figure 7.

A software system suitable for substation robot inspection is developed based on wheeled autonomous mobile robots and pan tilt hardware. This software system runs on the Ubuntu 16.04 LTS. The robot operating system ROS is used for development, mainly using C++ and Python programming languages. ROS is an operating system suitable for robot development, providing hardware abstraction, low-level device control, message synchronization, and other functions. It has high reusability and rich community ecology, which can improve the efficiency of robot development. Based on the ROS framework, the substation robot inspection system software consists of a system management module, a positioning and navigation module, a visualization module, and a visual detection and tracking module. In order to meet the needs of real-time image processing, the study sets the input image to 416×416 pixels. The network uses 3×3 and $1 \times$ A 1-size kernel with a step size of 1 or 2. The activation function of the network is Leaky ReLU, which increases the network's ability to process nonlinear features. The flowchart of the substation robot inspection software is shown in Figure 8.

In the proposed model, the loss function is a key factor in the optimization process, used to fine tune network parameters to improve object detection performance. Our loss function consists of three parts: localization loss, confidence loss, and classification loss, each part tailored to the different task characteristics of the detection task. The localization loss is responsible for narrowing the difference between the predicted box and the true box, and the study uses CIOU loss as a localization loss tool. The confidence loss is used to distinguish between foreground objects and background, ensuring that the model only assigns high confidence to areas where the target exists. Classification loss is used to detect inconsistencies between predicted class labels and actual labels.

In the selection of hyperparameters, the study sets the initial learning rate to 0.001 to ensure fast convergence during early training. When the performance on the validation set does not significantly improve within several consecutive epochs, the learning rate is halved to prevent oscillations and the model is fine tuned for better performance. Considering both model complexity and memory limitations, the batch size is set to 32. Find a balance between avoiding over-fitting and ensuring computational efficiency. In addition, to suppress over-fitting of the model during training, a smaller weight attenuation coefficient was introduced, set to 1e-4. The study chose the Adam optimizer as the main optimizer, with a momentum value set to 0.9 to prevent getting stuck in local minima.

IV. EXPERIMENTAL ANALYSIS OF TARGET RECOGNITION FOR MULTI-SOURCE MACHINE VISION PAN TILT INTEGRATED INSPECTION ROBOT

The developed wheeled mobile robot system for power inspection is used for integrated motion control experiments of inspection robots and pan tilt. Machine vision detection and target recognition experiments are conducted to compare the performance.

A. INTEGRATED MOTION CONTROL EXPERIMENT OF INSPECTION ROBOT AND PAN TILT

The experimental platform is based on the powerful Ubuntu 16.04 LTS operating system and is equipped with a wide range of Robot Operating System (ROS) middleware. ROS supports multi-sensor integration, data fusion, and advanced task execution, ensuring a fast and efficient development cycle. In terms of dataset, a dataset containing 300 categories of high-definition visible light images and 150 categories of high-resolution infrared images was constructed, covering the situations of various power equipment under different working conditions. The dataset includes 12000 visible light images and 5000 infrared images, all of which have been adjusted to a uniform resolution of 1920×1080 pixels to ensure that the images input into our visual system are standardized.

The key to conducting experiments on wheeled mobile robot systems is to verify the applicability and effectiveness of the algorithm and control technology in this study under practical physical constraints, such as non holonomic constraints, friction, and robot dynamics. Especially when considering linear motion and free rotation radius, this experiment can verify whether the robot can move along the expected path, further confirming the accuracy of target recognition in the robot's motion state. In addition, the experiment also emphasized the importance of dealing with uncertainty in actual motion processes, such as the model's ability to handle changes in road conditions and battery power consumption in special environments.

The robot system designed in the research combines machine vision technology with platform motion technology. According to the motion model designed in the research, the robot can move from three angles: X-axis, Y-axis, and Z-axis, with the Z-axis being a rotational motion. These three modes of motion essentially provide possibilities for robot obstacle avoidance. In linear motion, the robot moves along a predetermined path, and machine vision recognition matches the identified target with the predetermined target during this process. During the rotation process, multiple angles can be taken for shooting. The detection of the power of both can effectively verify the power performance of the robot in
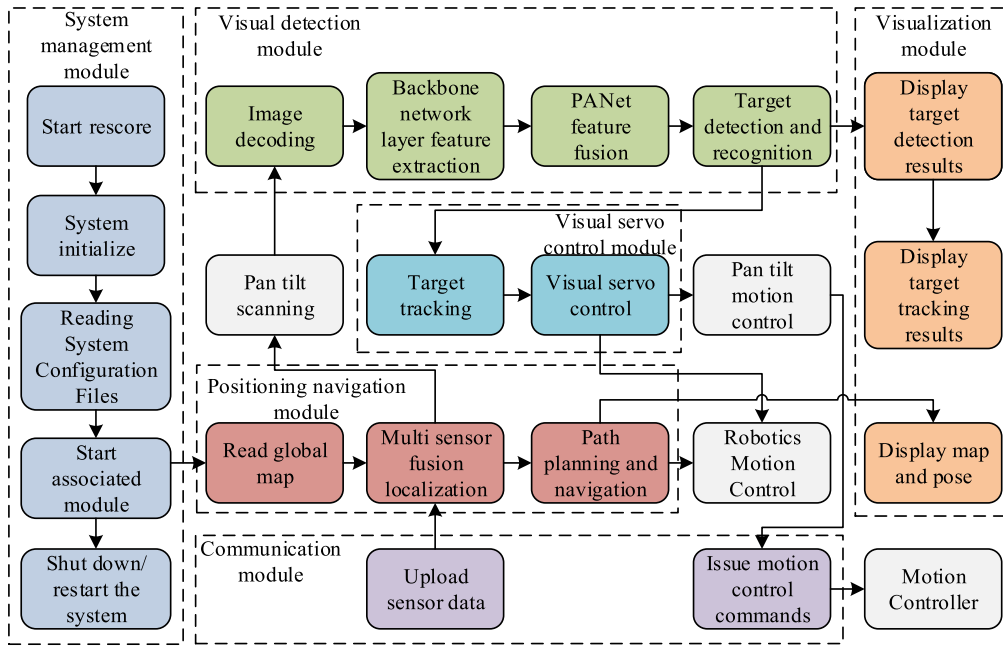
FIGURE 8. Electrical inspection robot inspection software diagram.

TABLE 1. Experimental results of robot linear motion testing in indoor environment.

Speed (mm/s)	Movement time (s)	Actual distance (mm)	Actual speed (mm/s)	Relative error (%)
500	20	9632	498.2	0.33
	30	14616		
1000	10	8485	1008.7	0.84
	20	18572		
1200	5	3636	1229.5	2.45
	10	9782		

different motion states, which plays a significant role in the overall performance detection of the robot.

Indoor linear motion experiments are conducted. The robot performs linear motion in an indoor environment paved with ceramic tiles at speeds of 0.5m/s, 1.0m/s, and 1.2m/s under a 10kg load. The mass of the robot is relatively large, and there is a certain acceleration process. Therefore, the same process is repeated at the same time. To eliminate the impact of acceleration and deceleration processes on the experimental results, two different sets of exercise times are set. Each group of experiments is independently repeated twice. Table 1 displays the results.

According to Table 1, when the robot moves in a straight line, the error basically shows an increasing trend as the motion speed increases. However, when the speed reaches 1m/s, the average error is only 0.84%. When the speed reaches 1.2m/s, the error increases 2.45%. At this point, the robot experiences significant slipping while parking. The experiment shows that the robot can ensure good accuracy in indoor linear motion under medium to low speed working

conditions. However, the robot may experience slight tire deformation or slipping during operation, as well as acceleration and deceleration stages and control delays. Therefore, the actual distance of movement may deviate from the set value. The motion error increases with the increase of motion speed.In this linear motion experiment, control system error and external environmental interference are the main factors that cause the observed errors. Among them, the control system error mainly includes the response of the actuator and the delay of the control signal. When the motion speed increases, the inertia and dynamic characteristics of the robot change, and the response speed and signal of the control system cannot be synchronized, increasing the motion error. The external environmental interference is mainly ground friction. Due to changes in the ground friction coefficient and possible unevenness, the area and shape of the tire in contact with the ground may change, leading to directional deviation or drift of the robot during linear motion, and may lose stability, unable to maintain accurate direction and speed, resulting in errors.

The experiment of measuring the accuracy of radius free rotation is conducted. Robots perform radius free rotational motion in indoor environments at different angular velocities (0.5rad/s and 1.0rad/s). The actual angle of rotation is recorded. To eliminate the impact of acceleration and deceleration on the experimental results, two sets of time are set for each angular velocity. Table 2 illustrates the experimental results.

According to Table 2, the error of robot rotation increases with the increase of angular velocity. At the maximum angular velocity in the experiment, the maximum average error is 1.41%. The possible reasons for the error are as follows.
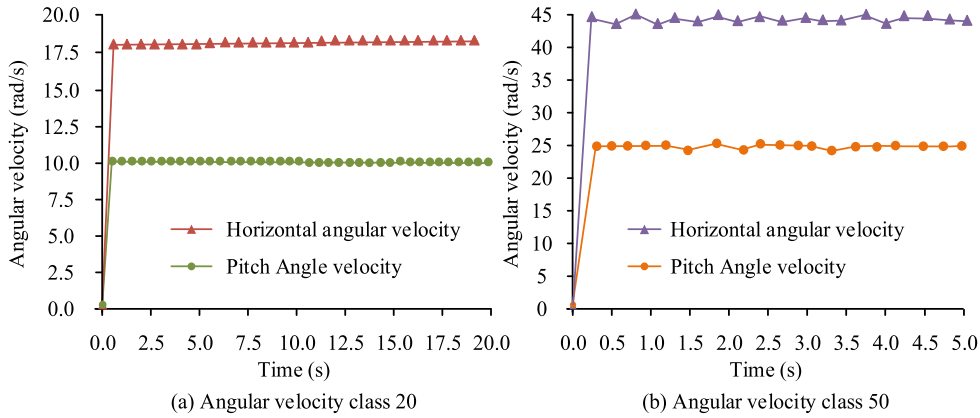
(a) Angular velocity class 20

(b) Angular velocity class 50

FIGURE 9. Curve chart of angular velocity variation of pan tilt.

(a) Head Angle positioning

(b) Visual servo response
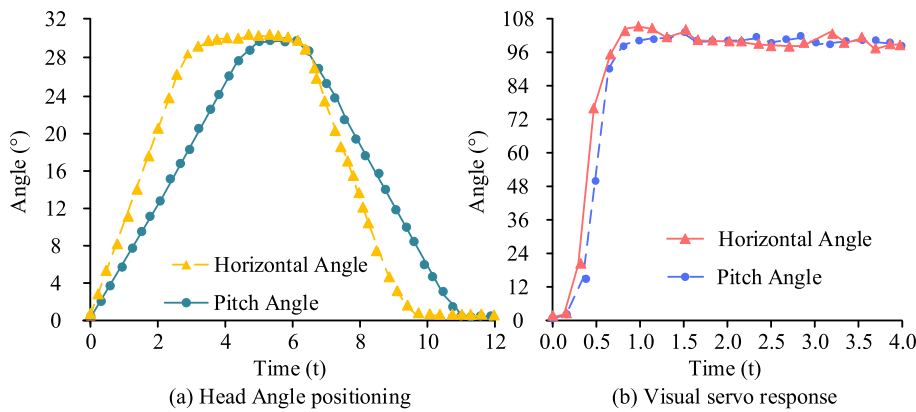
FIGURE 10. Pan tilt control response curve.

(a) Sample diagram of 17 types of equipment in the substation dataset

(b) Training loss curves for each model
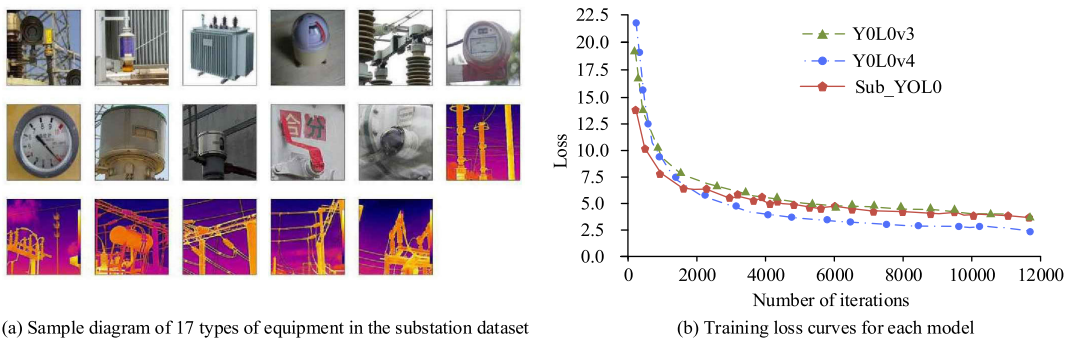
FIGURE 11. Substation equipment dataset and training loss curve.

On the one hand, there is an error between the center distance of the two wheels in the kinematic model and the actual value, resulting in a systematic error in the rotational speed. On the other hand, the position and pose errors of the two wheel installation may cause the center point of the robot to gradually shift during rotation.

In the accuracy testing of pan tilt angular velocity, the control of pan tilt angular velocity is carried out according to levels 1-100. The horizontal and pitch angular velocities of the pan tilt are tested at levels 20 and 50. The sampling time is 300 ms, and the actual rotation angle is recorded. The average velocity during the sampling time period is calculated using the angle difference as the instantaneous velocity, as shown in Figure 9.

From Figure 9, the average angular velocities of depression and elevation at level 20 are 18.1°/s and 10.05°/s respectively when the pan tilt is moving. The average angular velocity of level 50 is 45.7°/s and 25.2°/s. The error of angular velocity
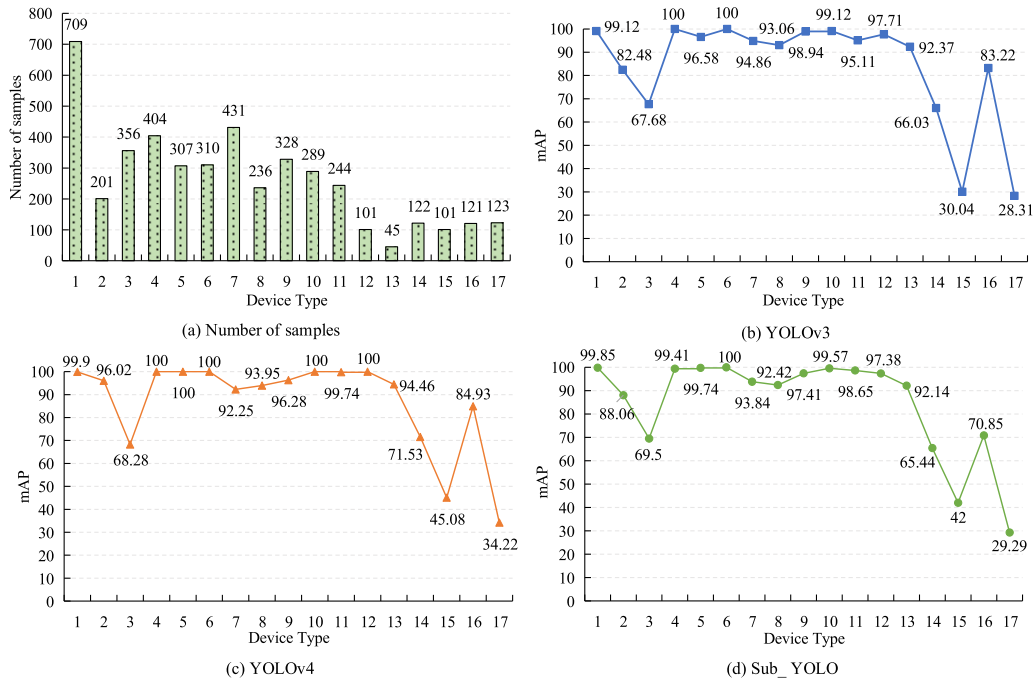
FIGURE 12. Comparison of object detection and recognition models mAP.

TABLE 2. Experimental results of robot radius free rotation test in indoor environment.

Angular velocity (rad/s)	Motion time (s)	Actual angle (°)	Actual angular velocity (rad/s)	Relative error (%)
0.5	10	285	0.497	0.82
	20	571		
1	5	274	1.015	1.41
	10	562		

increases with the increase of rotational speed. The higher the rotational speed, the more unstable the operation of the pan tilt. In the same gear, the pitch angular velocity is lower than the horizontal angular velocity, and the error is small. The speed regulation performance of the pan tilt meets the technical requirements. The medium to low speed working performance is stable.

In the accuracy test of pan tilt angle positioning, the angle is measured according to the preset position of the pan tilt. When the angular velocity level is 10, the preset position 1 is set to 0°, and the preset position 2 is set to 30°. The sampling time is set to 300ms. The sampling angle value is recorded. The PID parameters of the integrated motion control of the inspection robot and the pan tilt are debugged based on engineering experience. The horizontal axis and pitch axis are respectively input with a deviation signal of 100 pixels for data detection. The response curve of the pan tilt control is displayed in Figure 10.

In Figure 10 (a), at a 10 level angular velocity, the horizontal starting angle of the pan head is 0.27°, the starting angle

is −0.1°, the final angle is 30.27°, the starting angle is 29.98°, and the angular displacement accuracy reaches +0.1°. During operation, when a preset angle value is reached, there is a decrease in speed. After analysis, this is determined by the speed curve of the motor. By adjusting the program, the acceleration and deceleration of the corresponding speed curve can be achieved. In Figure 10 (b), the visual servo controller of the inspection robot can quickly reach a stable state. There is a certain degree of overshoot in the horizontal attitude adjustment. In the later stage of movement, there is interference. The pitch pose adjustment is stable and smooth, with good control performance. Overall, it takes 900ms for an increase and 500ms for a decrease. This method has low coupling in both horizontal and pitch positions, which can be controlled separately. Moreover, it meets the system requirements in stability, real-time performance, and other aspects.

B. MULTI SOURCE HETEROGENEOUS VISUAL DETECTION AND OBJECT RECOGNITION EXPERIMENT

The visible and infrared image data of the substation equipment is expanded. Among these two types of images, there are 3815 visible light images in 11 categories and 613 infrared images in 6 categories. The ratio of training set to test set is 9:1. The former is applied for model training, while the latter is used for model performance verification. Compared with the existing YOLO series representative algorithms YOLO 3 and YOLOv4, the sub_ The YOLO object detection and recognition algorithm is validated. On the basis of the substation equipment dataset, unified network training and model validation work is carried out. The function loss with

TABLE 3. Comparison of object detection and recognition model performance.

Serial Number	Model size	mAP	FPS
YOLOv3	236MB	83.80	16
YOLOv3-tiny	33.2MB	70.22	32
YOL Ov4	245MB	86.86	11
YOLOv4-tiny	23MB	73.47	32
Sub_ YOLO	54MB	84.44	31

FIGURE 13. Sub_ YoLo network detection results.

significant fluctuations in the early iteration training is eliminated. The loss values of 200 to 12000 iterations are taken. The loss values of each model in 200 to 12000 iterations are plotted into curves. The substation equipment dataset and training loss curve are shown in Figure 11.

From Figure 11, as the iterations increase, the loss function curves of the three algorithms gradually converge. When the convergence frequency reaches around 40000 times, the trend of the loss function curve has stabilized. Among them, YOLOv4 algorithm has the strongest convergence strength. The test set is imported into the trained model for model accuracy testing. After verification, the training loss of the study was 0.03, the training accuracy was 98%, the testing loss was 0.10, and the testing accuracy was 95%.The results are shown in Figure 12.

In Figure 12, YOLOv4 has the best detection accuracy, followed by sub_ YOLO, finally YOLOv3. The average model accuracy of visible light images reaches 94.26%, while the average accuracy of infrared images reaches 68.05%. There are significant differences in resolution, clarity, and contrast between infrared and visible light images. At the same time, obtaining infrared images has become more difficult. The samples are also smaller than that of visible light images. The model has poor fitting ability. The difficulty of image acquisition for power equipment is high, and the cost of manual annotation is high. It relies on continuous learning of hardware and time such as GPU. Therefore, by expanding the sample set and optimizing the training, the accuracy can be improved.

To verify the universality and real-time performance, a performance comparison is conducted with representative YOLO series algorithms, as displayed in Table 3.

From Table 3, the current camera sampling frame rate is 30 FPS. Sub_ YoLo is an algorithm that can achieve a detection speed of 30 FPS with an accuracy greater than 80% mAP. This method can effectively reduce the cost of edge computing and storage under the premise of ensuring accuracy and real-time, achieving the research goal. This detection framework can detect and identify information from different sources, types, and types of devices. It can learn new types of debugging data, so as to expand the types of debugging data. It overcomes the defects of the previous single mode, which has better universality and simpler advantages. Sub_ YoLo network detection results are shown in Figure 13.

Sub_ The contribution of YOLO algorithm to the entire target recognition experiment lies in providing a reliable,
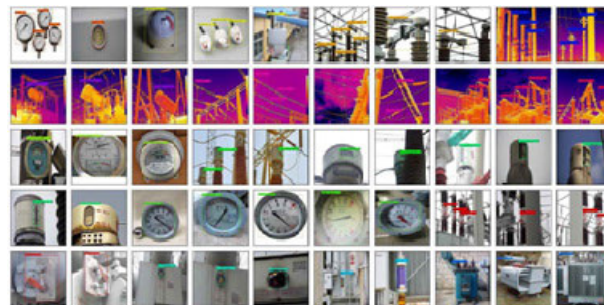
efficient, and easily adaptable machine vision detection method to diverse detection environments. On the one hand, Sub_ The YOLO algorithm can better handle input data from different sources with different characteristics, while maintaining the fast calculation speed and high-precision detection characteristics of YOLOv3 itself, it enhances the recognition and processing ability of substation equipment diversity. On the other hand, Sub_ The combination of YOLO algorithm and software systems developed based on ROS improves the efficiency of detection tasks and reduces human errors and omissions. Its advantage lies in strengthening the processing ability of multi-source heterogeneous data, improving network structure and loss function to achieve more accurate detection, and possessing the ability to quickly adapt to new environments through transfer learning.,

The results of this study indicate that as the operating speed of the gimbal increases, the error generally shows an upward trend. However, the speed adjustment performance of the current robot gimbal has met the technical requirements, and its working performance is stable at medium and low speeds. For future development, technical challenges and potential improvement suggestions mainly include four points. Firstly, in response to the industry barriers and difficulties in the data collection process of iron towers, future work should strengthen the convenience of data acquisition and reduce labor costs; Secondly, considering the insufficient performance of the algorithm in small object detection and occlusion situations, it is necessary to improve the detection algorithm in complex scenes to reduce missed and false detections; The development of potential applications in multi-spectral image fusion is still a challenge, as it can enhance the system's ability to process diverse visual data; Finally, increasing the optimization of the performance of deep learning algorithms in limited edge computing resources is still a challenge, and further research can be conducted to achieve more efficient real-time processing and recognition.

In Table 4, in terms of complexity, the model designed for research has the lowest complexity, only 130 million FLOPs, which means that the model can complete necessary detection tasks with less computational complexity, superior to other models, and especially suitable for resource constrained situations. In terms of mAP, the model designed in the study leads with 84.4% accuracy, indicating a very high

TABLE 4. Model performance comparison.

Model name	FLOPs	MAP（%）	Inference time (ms)
EfficientDet-D0	254 million	82	45
EfficientDet-D1	6.10 million	84	63
CenterNet	1.40 million	80	50
Research Design Model	1.30 million	84.4	30

TABLE 5. Comparison of benchmark models.

Model Name	MAP（%）	FPS
RetinaNet	81.6	40
Mask R-CNN	83.2	25
RefineDet	82.4	33
Research Design Model	88.4	52

accuracy in object detection, which is a crucial advantage. Although the FLOPs of CenterNet are relatively low, the mAP is 80.0%, which is the lowest among the four models, which may be slightly insufficient in some scenarios that require extremely high accuracy. In terms of inference time, the research designed model only takes 30 milliseconds, far exceeding other models, demonstrating its significant advantage in detection speed, which is extremely important for applications that require real-time processing.

In Table 5, the model designed in the study has an mAP of 88.4%, indicating excellent performance in object detection accuracy. In terms of frames per second (FPS), the research designed model performs best at a speed of 52 FPS, which means that the model can handle high frame rate video streams and is very suitable for real-time application scenarios.

Through comprehensive data comparison, it can be seen that the model designed in the study performs excellently in multiple key performance indicators. Whether it is computational complexity, detection accuracy, or inference speed, it has demonstrated its practicality and leading position in object detection tasks, especially in real-time analysis and low-power devices, with huge advantages. In addition, this performance advantage also means that the research model can have stronger adaptability, robustness, and flexibility in different environments.

V. CONCLUSION

An integrated motion control technology for dynamic acquisition of inspection data is proposed. Based on the YOLOv3 network framework, an improved multi-source heterogeneous visual detection and recognition model is designed. A path aggregation network is introduced to accelerate

information flow and integrate features at different levels. According to the results, as the motion speed increases, the error basically shows an increasing trend. In the angular velocity accuracy test and angular positioning accuracy test of the pan tilt, the speed regulation performance of the pan tilt meets the technical requirements. The working performance at medium and low speeds is stable. In the target recognition experiment, by expanding the visible and infrared images of substation equipment, the sub_YOLO object detection and recognition algorithm is validated. The YOLOv4 has the best detection accuracy, followed by sub_YOLO, finally YOLOv3. The average model accuracy of visible light images reaches 94.26%, while the average accuracy of infrared images reaches 68.05%. In addition, Sub_YoLo is an algorithm that can achieve 30 FPS detection speed when the accuracy is more than 80% mAP, effectively reducing the cost of edge computing and storage, etc. Through in-depth research on the motion control and target recognition strategies of power inspection robots, as well as the development of software systems suitable for substation robot inspection, important theoretical and practical support are provided for the practical application of power inspection robots. The experiment verifies the feasibility and accuracy of integrated motion control between the inspection robot and the pan tilt, as well as the effectiveness of multi-source heterogeneous visual detection and target recognition. It provides important technical support for the wheeled mobile robot system of power inspection. However, there are certain difficulties and industry barriers in collecting substation data, which require a lot of time and manpower. In addition, the detection performance of existing algorithms in small targets, occlusion, and other situations is not ideal, which is prone to missed and false detections. Therefore, in future work, emphasis should be placed on how to produce high-quality datasets and improve the performance of detection algorithms for multiple scenarios in substations.

In summary, this study compared the performance of multiple variants based on the YOLO algorithm in object detection tasks and found that YOLOv4 performed the best among various indicators, while Sub_ The YOLO algorithm ensures good real-time performance while maintaining high detection accuracy. Sub_ The YOLO algorithm can achieve an average accuracy of over 80% while maintaining an image processing rate of 30 frames per second, which surpasses most current models designed for such applications. The research has successfully demonstrated that high-precision and high-speed object detection can still be achieved through algorithm optimization in environments with limited computing resources, which is necessary for research in this field. At the same time, the research results provide important support for future intelligent power inspection systems, especially in terms of algorithm stability and accuracy, enhancing the system's application ability in extreme environments, and providing new possibilities for industrial automation and intelligence.

Although the model designed for research has superior performance, current research mainly focuses on standard

detection environments. However, in practical applications, the detection environment is likely to be variable and affected by special environments. Under diverse and unknown external influences, the detection performance of the model may be affected to a certain extent. Therefore, in the future, research will focus on adaptive optimization for special and ever-changing work environments, enhancing the robustness of algorithms under external influencing factors. Furthermore, a more intelligent power system detection scheme will be developed to improve power supply safety and efficiency.

REFERENCES

[1] Z. Wang, B. He, Y. Zhou, K. Liu, and C. Zhang, "Design and implementation of a cable inspection robot for cable-stayed bridges," *Robotica*, vol. 39, no. 8, pp. 1417–1433, Jan. 2021, doi: 10.1017/s0263574720001253.

[2] H. Jang, T. Y. Kim, Y. C. Lee, Y. H. Song, and H. R. Choi, "Autonomous navigation of in-pipe inspection robot using contact sensor modules," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 6, pp. 4665–4674, Dec. 2022, doi: 10.1109/TMECH.2022.3162192.

[3] J. Li, S. Jin, C. Wang, J. Xue, and X. Wang, "Weld line recognition and path planning with spherical tank inspection robots," *J. Field Robot.*, vol. 39, no. 2, pp. 131–152, Mar. 2022, doi: 10.1002/rob.22042.

[4] J. Zan, "Research on robot path perception and optimization technology based on whale optimization algorithm," *J. Comput. Cognit. Eng.*, vol. 1, no. 4, pp. 201–208, Mar. 2022, doi: 10.47852/bonviewjcce597820205514.

[5] R. K. Jain, A. Das, A. Mukherjee, D. N. Ray, and P. Karmakar, "Experimental performance of robotic inspection system for underground pipelines," *J. Inst. Eng., India, C*, vol. 102, no. 3, pp. 683–703, Apr. 2021, doi: 10.1007/s40032-021-00691-x.

[6] R. A. Ufa, Y. Y. Malkova, V. E. Rudnik, M. V. Andreev, and V. A. Borisov, "A review on distributed generation impacts on electric power system," *Int. J. Hydrogen Energy*, vol. 47, no. 47, pp. 20347–20361, Jun. 2022, doi: 10.1016/j.ijhydene.2022.04.142.

[7] Z. Ren, F. Fang, N. Yan, and Y. Wu, "State of the art in defect detection based on machine vision," *Int. J. Precis. Eng. Manuf.-Green Technol.*, vol. 9, no. 2, pp. 661–691, Mar. 2022, doi: 10.1007/s40684-021-00343-6.

[8] L. Zhu, P. Spachos, E. Pensini, and K. N. Plataniotis, "Deep learning and machine vision for food processing: A survey," *Current Res. Food Sci.*, vol. 4, pp. 233–249, Apr. 2021, doi: 10.1016/j.crfs.2021.03.009.

[9] S. A. Singh and K. A. Desai, "Automated surface defect detection framework using machine vision and convolutional neural networks," *J. Intell. Manuf.*, vol. 34, no. 4, pp. 1995–2011, Apr. 2023, doi: 10.1007/s10845-021-01878-w.

[10] Z. Zhonglin, F. Bin, L. Liquan, and Y. Encheng, "Design and function realization of nuclear power inspection robot system," *Robotica*, vol. 39, no. 1, pp. 165–180, Jan. 2021, doi: 10.1017/s0263574720000740.

[11] J. Liu, T. Zhao, and S. Dian, "General type-2 fuzzy sliding mode control for motion balance adjusting of power-line inspection robot," *Soft Comput.*, vol. 25, no. 2, pp. 1033–1047, Jan. 2021, doi: 10.1007/s00500-020-05202-1.

[12] X. Li, D. Shang, F. Li, and R. Chen, "The climbing performance analysis of a robot for power line inspection with retractable double serial manipulators," *Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci.*, vol. 236, no. 9, pp. 4946–4961, May 2022, doi: 10.1177/09544062211054984.

[13] L. Huang, G. Wu, W. Tang, and Y. Wu, "Obstacle distance measurement under varying illumination conditions based on monocular vision using a cable inspection robot," *IEEE Access*, vol. 9, pp. 55955–55973, 2021, doi: 10.1109/ACCESS.2021.3070877.

[14] Z. Liu, Y. Cai, H. Wang, L. Chen, H. Gao, Y. Jia, and Y. Li, "Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6640–6653, Jul. 2022, doi: 10.1109/TITS.2021.3059674.

[15] B. Shi, Q. Zhang, D. Wang, and Y. Li, "Synthetic aperture radar SAR image target recognition algorithm based on attention mechanism," *IEEE Access*, vol. 9, pp. 140512–140524, 2021, doi: 10.1109/ACCESS.2021.3118034.

[16] C. Du, Y. Cong, L. Zhang, D. Guo, and S. Wei, "A practical deceptive jamming method based on vulnerable location awareness adversarial attack for radar HRRP target recognition," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 2410–2424, May 2022, doi: 10.1109/TIFS.2022.3170275.

[17] Q. Wang, H. Xu, L. Yuan, and X. Wen, "Dense capsule network for SAR automatic target recognition with limited data," *Remote Sens. Lett.*, vol. 13, no. 6, pp. 533–543, Mar. 2022, doi: 10.1080/2150704x.2022.2044089.

[18] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," *Multimedia Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.

[19] W. Liu, G. Ren, R. Yu, S. Guo, J. Zhu, and L. Zhang, "Image-adaptive YOLO for object detection in adverse weather conditions," in *Proc. 36th AAAI Conf. Artif. Intell.*, vol. 36, Mar. 2022, pp. 1792–1800.

[20] S. Li, Y. Li, Y. Li, M. Li, and X. Xu, "YOLO-FIRI: Improved YOLOv5 for infrared image object detection," *IEEE Access*, vol. 9, pp. 141861–141875, 2021, doi: 10.1109/ACCESS.2021.3120870.

MINJIE ZHU was born in Wenzhou, Zhejiang, in July 1972. He received the bachelor's degree in electrical engineering and automation from Shanghai Electric Power University, in 2004, and the master's degree in electrical engineering from Zhejiang University, in 2012, specializing in power planning and construction. From 2013 to 2020, he was the Director of the Construction Department and the General Manager Assistant of Zhejiang Wenzhou Power Supply Company. From 2020 to 2023, he was the Deputy General Manager and a Party Committee Member of State Grid Zhejiang Electric Power Supply Company Taizhou Branch. He obtained one invention patent and two utility model patents. He mainly participated in four invention patents and five utility model patents. In terms of articles, he has successively published one Chinese core journal as the first author and three journals included in EI. In terms of monographs, as a member of the editorial committee, he wrote four monographs.

BINGYE ZHANG was born in Linhai, Zhejiang, in June 1981. She received the bachelor's degree in computer science and engineering from Shanghai Electric Power University, in 2003. From 2003 to 2019, she was a Technical Specialist with State Grid Taizhou Power Supply Company. From 2020 to 2023, she was the Director and the Deputy Manager of the Information and Communication Technology Department, State Grid Zhejiang Electric Power Supply Company Taizhou Branch. She is responsible for winning the first prize of Zhejiang Electric Power Company's Science and Technology Innovation Award for the project. She obtained two invention patents and one utility model patent. She has published articles, such as "Research on Intelligent Power Outage Management in Distribution Network under Major Maintenance Mode" and "Research on a Video Quality Intelligent Detection System."

CHEN ZHOU was born in Wenling, Zhejiang, in February 1980. He received the bachelor's degree in automation from Ningbo University, in 2003, and the master's degree in electrical engineering from North China Electric Power University, in 2012. His research interests include power systems and automation. From 2003 to 2011, he was the Monitor of the second (1st) class of Taizhou Power Supply Company. From 2011 to 2020, he was the Director of the Human Resources Department and Office of Taizhou Power Supply Company. From 2020 to 2022, he was the Chairperson of Hongchuang Group of Taizhou Power Supply Company. Since 2022, he has been the Deputy General Manager of Huayun Clean Energy Company, State Grid Zhejiang Electric Power Supply Company Taizhou Branch. He served as the Project Leader for the design of low-frequency AC wind turbines and the development of key technologies for power integration. He obtained three national invention patents and he published two Chinese core and EI papers. He is a member of the editorial board, wrote one monograph on distribution network construction technology.

XUEYAN WANG was born in Taizhou, Zhejiang, China, in December 1994. She received the master's degree in electrical engineering from Hunan University, in June 2020. Since November 2022, she has been the Deputy Director of the Power Grid Technology Department, State Grid Zhejiang Electric Power Supply Company Taizhou Branch and Hongchuang Group Technology Branch. She served as a technology project leader five times. She applied for over 50 invention patents, authorized over 20 invention patents, published four high-level journal articles, and completed two enterprise level gold medal product releases.

● ● ●

HONGLIANG ZOU was born in Gao'an, Jiangxi, China, in June 1982. He received the bachelor's degree in electronic information engineering majoring in new power systems and artificial intelligence, the master's degree in power electronics and power transmission, and the Ph.D. degree in power system and automation from Wuhan University, in 2004, 2007, and 2017, respectively. From 2012 to 2017, he was a Specialist in the maintenance and repair of switches, transformers, and substations with the Operation and Maintenance Department. From 2017 to 2020, he was the Deputy Director of the Operations and Maintenance Department (in charge of power transformation and planning). Since 2021, he has been the Manager of State Grid Zhejiang Electric Power Supply Company Taizhou Branch and Hongchuang Group Technology Branch. Four invention patents have been granted in the fields of artificial intelligence and power energy. He published a total of six papers in the field of power technology, including three IE conference papers, one electrical technology paper, and two high-voltage technology papers.