A cable tension measurement method for transmission lines based on micro-vibration broadband phase motion magnification and deep learning

Huang Zhiming*, Wang Shuo, Wen Hongbing and Li Zixin

Jiangmen Power Supply Bureau, Guangdong Power Grid Co., Ltd., Jiangmen, Guangdong, China

Introduction: Cable components are widely used in transmission lines, and their tension values and variations are critical factors affecting the intrinsic safety of these lines. Thus, tension monitoring becomes a priority during both construction and operational maintenance. Traditional cable tension measurement methods suffer from limitations such as low accuracy, stringent environmental requirements, and difficulties in live-line monitoring, resulting in a lack of universality for application in transmission lines.

Methods: This paper utilizes visual image technology and Broadband Phase Motion Magnification to amplify the micro-vibration amplitude and enhance the vibration images of transmission line cable- type components under environmental excitation. Furthermore, this study develops a combined segmentation algorithm using the U-Net network architecture and level set loss entropy to accurately capture the centroid motion trajectory of cables, thereby precisely extracting the vibration displacement time series. Finally, spectrum analysis is applied to invert the self-vibration characteristic parameters of the components and establish a tension calculation model.

Results: Experimental verification shows that the proposed method can precisely capture the micro-vibration signals induced by environmental excitation. The tension calculation results, when compared to standard sensor data, have a deviation of no more than 8%.

Discussion: This method successfully establishes a non-contact, high-precision measurement system for cable-type components, providing a new technical pathway for intelligent monitoring during the construction and maintenance of transmission lines.

KEYWORDS

transmission lines, micro vibration, tension measurement, deep learning, image recognition, vibration frequency

1 Introduction

Transmission lines are crucial national infrastructure, and the structural safety during both their construction and operation is vital to the overall reliability of the power system (Yao et al., 2021). Tension members are widely distributed in the structural components of transmission line projects, including various conductors and ground wires, guy wires of transmission towers, as well as stay cables and anchor cables during construction, all of which are types of tension members. Under operational conditions, these tension members primarily bear large axial

tensile forces, and their tension values fluctuate significantly due to environmental loads such as wind, temperature changes, and ice accumulation, as well as material characteristics changes like stress relaxation, creep, and broken strands. These fluctuations in tension directly impact the overall safety and stability of the structure (Zhenya, 2009; We et al., 2015). Therefore, identifying the modal parameters of the tension members and measuring their tension values have always been key aspects of structural health monitoring during both construction and operation (Dengke et al., 2015). The tension members in transmission line projects can be classified into two main categories: one category bears electrical functions such as power transmission and lightning protection, mainly consisting of conductors and ground wires; the other category stabilizes structures like towers and brackets, primarily consisting of tower guy wires, stay cables, and anchor cables. For existing structures, it is difficult to measure the tension of these members by installing tension sensors, so quick and accurate tension measurement has always been a challenge in transmission line construction and operation.

Currently, cable tension monitoring and detection can be classified into contact-based and non-contact methods. In complex field environments, technicians often have to rely on manually shaking the cable (Yanbao et al., 2022) and subjectively judging its tightness based on experience. With the development of sensing technology, researchers and engineers have begun exploring cable tension detection using various sensors. References (Zhanghua et al., 2023; Xiongjun et al., 2016) installed highly sensitive vibration pickups on stay cables to acquire their natural frequencies, establishing vibration equations based on the elastic distributed mass and cable vibration direction to calculate cable tension. This method can yield relatively accurate tension results but requires contact-based vibration sensors. In transmission line engineering, as cables are typically energized and susceptible to various environmental factors like terrain and weather, contact-based measurements are often impractical, lack universality, and do not align with the expected trend toward intelligent construction and maintenance of future transmission lines (Haitao et al., 2023). Consequently, contact-based measurement methods have gradually fallen out of use for tension detection in transmission lines.

In contrast, non-contact detection methods have developed rapidly in recent years due to advantages such as ease of installation and debugging, and no interference with transmission lines (Jian et al., 2022; Qiao et al., 2022; Longqiang et al., 2023). Among these, vision-based tension detection is an emerging non-contact intelligent detection technology that has been widely adopted in this field owing to its high measurement accuracy, convenient installation, and good real-time performance. As early as 2015, Banfu et al. (2015) conducted research on cable force testing using moving target image tracking. They used industrial cameras and smartphones to capture vibration videos of steel strands, proposing morphological image processing to extract displacement time-history for calculating cable tension. However, this method relies on binarization to extract morphological features of the strand, which may fail to accurately identify features when image pixel grayscale values are high. Lan et al. (2022) applied this approach to measure tension in transmission guyed towers, employing Canny edge detection and sub-pixel localization to address issues of rough strand edges and low positioning accuracy. While this method can achieve relatively precise tension values, it requires manual excitation to impart initial

velocity or displacement to the cable to induce sufficiently large free vibrations before capturing the vibration images. Summarizing previous non-contact measurement research, two main factors affect measurement accuracy: first, insufficient robustness of the measurement method, which can lead to significant errors in practical engineering environments; second, difficulties in directly applying artificial excitation to transmission line cables, whether for primary or secondary components, due to their working environment and energized state. The only viable solution is to identify the visual images of minor vibrations of cable components under ambient excitation, enabling analysis of natural frequency and tension calculation. Therefore, leveraging minor vibrations under environmental excitation for modal analysis is key to achieving rapid, accurate, and non-contact tension measurement for transmission line cable components.

In recent years, phase amplification technology based on machine vision has gradually developed, providing a solution for identifying small movements. Reference (Zhang Yuhang et al., 2021) applied phase amplification algorithms to measure bridge stay cables, using manually placed markers for edge localization to obtain displacement curves and then calculate cable tension using the frequency method. This method can effectively measure the vibration frequency of cables by recognizing artificial markers, but it requires the prior determination of the natural frequency. Building upon this, Xuan et al. (2023) and colleagues adopted broadband phase amplification technology, which amplifies small vibrations across a wide frequency band without requiring prior knowledge of the frequency. They then used template matching to determine the displacement-time history of the object, enabling the identification of the natural frequency. These studies all amplified small movements and achieved relatively accurate experimental results, validating the feasibility of using amplification algorithms for structural vibration identification. However, unlike the objects studied in these references, the tension members in transmission lines are typically made of twisted strands, with rougher edges. The presence of video noise and issues such as image signal step transitions caused by the phase pyramid algorithm can lead to artifacts, which significantly interfere with the accurate extraction of the displacement of tension members and the subsequent data processing, ultimately affecting the calculation accuracy.

With the development of artificial intelligence, deep learning algorithms have expanded the possibilities for identifying and removing artifacts (He et al., 2023; Chuankai et al., 2021; Lee et al., 2023; Yang et al., 2021). Kim et al. (2019) proposed combining the level-set image processing method with deep learning semantic segmentation networks, and using backpropagation of the loss functions from both methods to enhance the network's segmentation performance. Yang and Dengke (2022) applied this method to seismic arrival picking, solving the problem of poor recognition accuracy in conventional machine vision.

To identify the natural frequencies and calculate the tension of transmission line tension members under environmental excitation, this paper proposes using broadband phase video amplification technology. The small vibrations of the tension members are amplified within a specific frequency band, responding to environmental excitation. To effectively remove artifacts, a deep learning semantic segmentation network is introduced to segment and cluster the movement of tension members. Finally, edge fitting
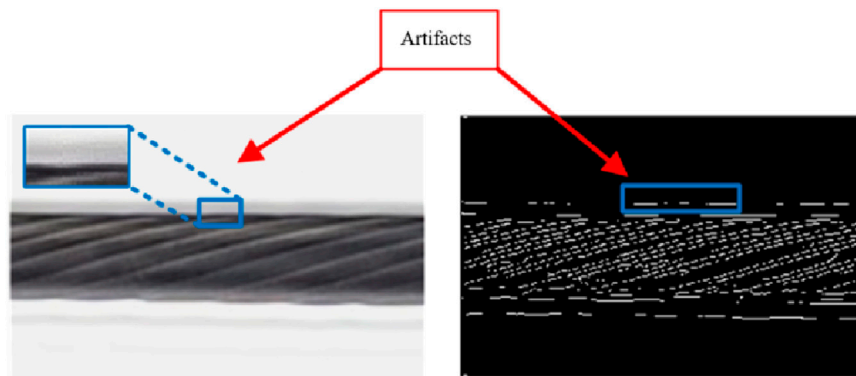
FIGURE 1
Video Frame Amplified by BPMM Algorithm.

is applied to improve accuracy and extract the displacement-time history of the centroid of the tension member. The target vibration frequency is obtained through spectral transformation, and the tension is calculated using the frequency method, thereby achieving non-contact tension measurement of transmission line tension members. This approach, by combining phase amplification with deep learning for artifact removal and segmentation, can significantly enhance the accuracy and reliability of tension measurements in real-world transmission line environments, addressing the challenges posed by noise and rough edges in tension member images.

2 Tension measurement method based on micro-vibration monitoring

The research method proposed in this paper is as follows: The sub-band signals of vibration video images in different directions and scales are extracted using a complex-valued controllable pyramid. The broadband phase amplification algorithm is applied to process the target motion video and enhance its motion state. A deep learning network-based U-Net and level-set entropy loss joint segmentation method is employed, with images magnified by different amplification factors and captured at various angles used as the dataset for training. This allows the network to learn the pixel information of tension members, extracting the relevant pixel information of transmission line tension members from various image pixels. To improve the accuracy of edge pixel information, a least-squares fitting algorithm is used to smooth the edge pixels, extracting the pixel coordinates of the centroid of the tension member and performing statistical analysis. The modal frequency is obtained using the Fast Fourier Transform (FFT) algorithm, and finally, the tension of the member is calculated using the frequency method.

2.1 Broadband phase-based motion amplification algorithm

The vibration amplitude of tension members in transmission line projects is very small under environmental excitation, and the

resulting displacement is not significant. Traditional visual image processing techniques based on edge recognition algorithms struggle to detect such small vibrations and cannot eliminate noise interference. The broadband phase-based motion magnification (BPMM) algorithm has shown good performance in identifying and amplifying small movements. The complex-valued controllable pyramid is the main processing step in the broadband phase-based motion amplification algorithm. It decomposes the video of vibrating objects into individual frames and then breaks down each frame into sub-band signals in different directions and scales, thereby extracting phase information related to object displacement. The 2D Gabor filter within the pyramid can effectively handle noise in image videos and enhance the signal-to-noise ratio of the vibration video image. The motion of the cable component, after being decomposed by the complex-valued controllable pyramid, can be defined by image intensity. Assuming the image intensity of the cable component is f (x,y), when the image pixels of the cable component experience small vibrations along the x-axis and y-axis at time t, denoted as δ (x,t) and δ (y,t), the image intensity can be represented as shown in Equation 1:

$$f(x + \delta(x,t), y + \delta(y,t)) \tag{1}$$

The vibration signal of the tension member can be decomposed into a sum of sine waves at all frequencies using Fourier series.

When the tension member in the video has not started experiencing small vibrations, i.e., at t = 0, the image intensity is:

$$f(x,y) = \sum_{\omega=-\infty}^{\infty} A\omega(x,y,0)e^{i\omega(x,y)} \tag{2}$$

When the video starts capturing and the tension member begins to experience small vibrations, i.e., at t > 0, the image intensity is:

$$f(x + \delta(x,t), y + \delta(y,t)) = \sum_{\omega=-\infty}^{\infty} A_\omega(x,y,t)e^{i\omega(x+\delta(x,t),y+\delta(y,t))} \tag{3}$$

Here, $A_\omega(x,y,t)$ represents the vibration amplitude of the tension member; ω represents the harmonic frequency; $e^{i\omega(x,y)}$ and
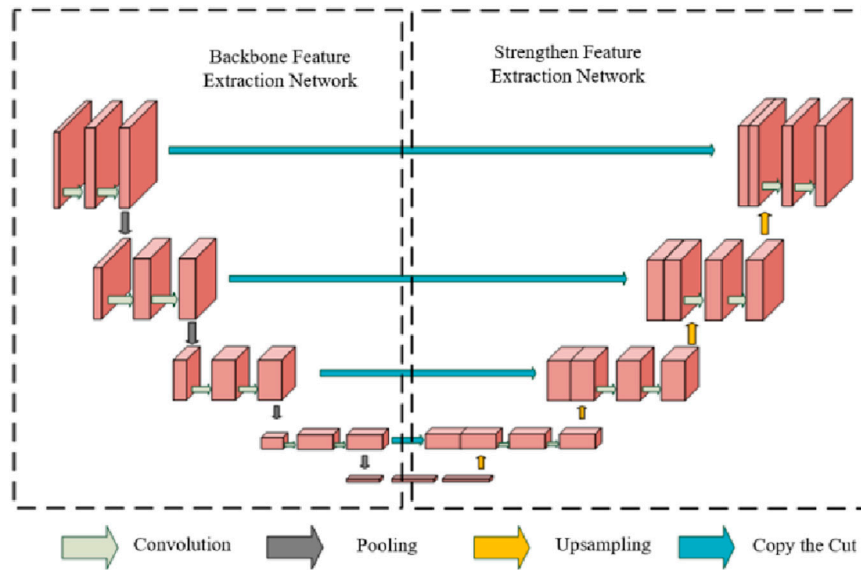
FIGURE 2
U-Net network structure.

$e^{i\omega(x+\delta(x,t),y+\delta(y,t))}$ represents the phase information of the tension member's image; By subtracting Equation 2 from Equation 3, the image motion phase—P (x,y,t) difference of the tension member at the ω harmonic frequency can be obtained, as shown in Equation 4:

$$P(x, y, t) = \omega(\delta(x,t), \delta(y,t)) \tag{4}$$

According to the time-shifting property of the Fourier transform, changes in the phase information can be used to achieve variations in the image motion information. By multiplying the expression in Equation 4 by an amplification factor α, the local phase amplification is achieved, as shown in Equation 5:

$$\bar{P}(x, y, t) = \alpha\omega(\delta(x,t), \delta(y,t)) \tag{5}$$

The final amplified image intensity is given by Equation 6:

$$f\left(x + (1 + \partial)\delta(x,t), y + (1 + \alpha)\delta(y,t)\right)$$
$$= \sum_{\omega = -\infty}^{\infty} A_\omega(x, y, t) e^{i\omega\left(x+(1+\partial)\delta(x,t),y+(1+\alpha)\delta(y,t)\right)} \tag{6}$$

The algorithm amplifies the small motion in the image domain by increasing the phase difference. Here, ω represents controlled through filtering, and when using the amplification algorithm, the filtering range must be greater than the frequency range of the natural frequency of the tension member. In this paper, the measured fundamental frequency of the transmission line tension member is used as the result for solving via the frequency method. Since most of the first-order natural frequencies of tension members lie between 0 and 20 Hz, to ensure both the amplification effect and experimental accuracy, the filtering range in this study is set to 0–20 Hz.

## 2.2 An image segmentation network

For video images of cable components undergoing small vibrations, the broad-band phase magnification algorithm is applied. During the phase magnification process, artifacts around the target are generated due to phase constraints from the complex-valued controllable pyramid, and environmental noise interference arises from camera-based image capture.

To extract the displacement of cable vibrations amidst environmental noise and artifacts, traditional image recognition techniques such as edge detection struggle to differentiate the edge motion information of the cable component, which is surrounded by artifacts, making it difficult for subsequent tension calculations, as shown in Figure 1.

Therefore, this paper proposes the use of semantic segmentation techniques from deep learning to extract cable images, remove environmental noise, and eliminate the artifacts caused by the phase magnification technique.

2.2.1 U-net network

With the development of computer vision technology, computers have already achieved the ability to automatically classify various objects in images through learning. Semantic segmentation technology in deep learning is a technique that achieves object segmentation by classifying each pixel of an object. The U-Net network is a precise and efficient semantic segmentation network (Ronneberger et al., 2015). It improves upon the FCN network by using a concatenation method to combine deep and shallow image features instead of the summation method used in FCN (Long et al., 2015). The advantage of U-Net lies in its ability to achieve high segmentation accuracy with a smaller dataset and shorter training time, which addresses the issue of limited dataset availability in power transmission line engineering.

FIGURE 3
Centroid selection after fitting.
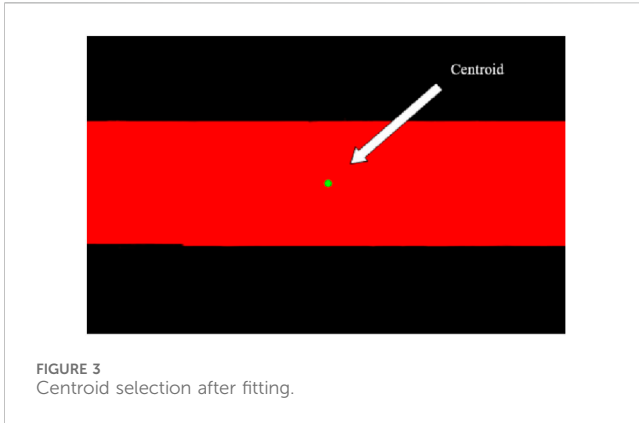
The U-Net network structure is a U-shaped symmetric architecture. The left side consists of an encoder made up of four convolution blocks. In each convolution block, convolutional layers and pooling layers downsample the image information. As the sampling channels of the image pixels are doubled, the feature map size is reduced to half of the original. The right side structure is the decoder, which has a similar design to the encoder. The key difference is that the pooling layers in the convolution blocks are replaced by deconvolution layers, which restore the image dimensions. Through upsampling, the decoder classifies and segments each pixel of the image. Skip connections concatenate the processed images with the same pixel size from both sides of the network, achieving the fusion of edge information across different scales. Finally, a single convolution layer outputs the semantic segmentation results. The structure is shown in Figure 2.

2.2.2 Semantic segmentation nerwork integrated with level sets

Level set methods, as traditional image segmentation techniques, have been widely used in engineering practices, typically for segmenting foreground and background. Therefore, they are well-suited for segmenting a single target, such as transmission line conductor components. The level set method constructs an energy function based on image pixel information, such as grayscale values, and iteratively moves the segmentation line toward the target object until the segmentation is completed. To improve the accuracy of edge segmentation in magnified images, a global level set segmentation method is combined with the loss entropy from semantic segmentation. The level set energy function EC is shown in Equation 7:

$$E_C = \lambda_1 \iint |I(x,y) - c_{l,1}|^2 H(\phi) dx dy + \\ \lambda_2 \iint |I(x,y) - c_{l,2}|^2 (1 - H(\phi)) dx dy \quad (7)$$

The parameters $c_{l,1}$, $c_{l,2}$ and $H_\varepsilon(\phi)$ in Equation 7 are given by Equations 8–10, respectively.

$$c_{l,1} = \frac{\int_{K_l} K_l(x,y) H_\varepsilon(\phi(x,y)) dx dy}{\int_{K_l} H_\varepsilon(\phi(x,y)) dx dy} \quad (8)$$

$$c_{l,2} = \frac{\int_{K_l} K_l(x,y)(1 - H_\varepsilon(\phi(x,y))) dx dy}{\int_{K_l} (1 - H_\varepsilon(\phi(x,y))) dx dy} \quad (9)$$

$$H_\varepsilon(\phi) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\frac{z}{\varepsilon}\right) \quad (10)$$

Here, K (x,y) represents the Gaussian kernel function, $\lambda1$, $\lambda1$, and ω represents the weights, c_{l1} and c_{l2} represents the mean pixel grayscale values inside and outside the segmentation boundary, ϕ represents the level set distance function, H(ϕ) represents the approximate expression of the step function, and ε represents a constant, ε tends to be close to 0.

Deep learning-based semantic segmentation methods can output intuitive segmentation results, but they are prone to issues with data imbalance. Therefore, it is necessary to introduce a loss function to adjust the deep learning segmentation model. Using only the loss value from the U-Net semantic segmentation network for error backpropagation can result in unclear boundary segmentation of the enlarged object. To address this, the level set energy functional is introduced as a loss term for backpropagation, which enhances robustness against noise and the contours of the target object.

The loss function of the level set is as shown in Equation 11:

$$L_s = \frac{1}{2n}\sum_{i=1}^{n} \|y_l - e_l\|^2 \quad (11)$$

Here, y_l represents the pixel true value, and e_l represents the pixel predicted value. The final goal of this study is to segment the cable images for result processing. Therefore, artifacts and background are considered noise, and the segmentation task is converted into a binary classification with a single-label task. For the case of evenly distributed binary classification samples, we use the Cross-Entropy loss function L_u, as shown in Equation 12:

$$L_u = -\sum_{i=1}^{C} y(x_i)\log(q(x_i)) \quad (12)$$

Here, C represents the number of samples, y (x_i) represents the true pixel, and q (x_i) represents the predicted pixel probability. In this paper, a binary classification method is used, so the cross-entropy is Equation 13:

$$L_U = -(y\log\hat{y} + (1-y)\log(1-\hat{y})) \quad (13)$$

Here, y represents the true distribution probability of the pixel for the transmission line component, $\hat{y}$ represents the predicted probability of the pixel for the transmission line component by the U-Net network, and 1-y and 1-$\hat{y}$ represent the true distribution probability of the pixels for non-transmission line components (background, artifacts) and the predicted probability of the non-transmission line components' pixels, respectively.

The final loss function is shown in Equation 14:

$$L_R = E_C + \beta L_U \quad (14)$$

Here, β represents the weight value.

2.3 Vibration and centroid extraction of the cable component

2.3.1 Edge fitting and centroid extraction of the cable component

When the cable component is subjected to environmental excitation, it generates random vibrations in the normal

FIGURE 4
Overall process.

FIGURE 5
Test site.

FIGURE 6
Part of the test data set.
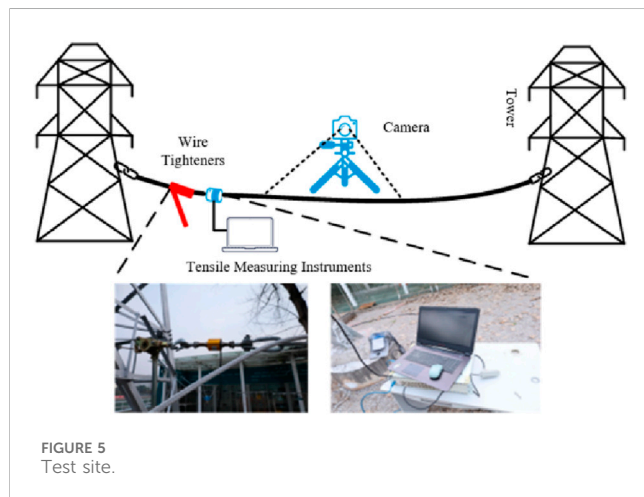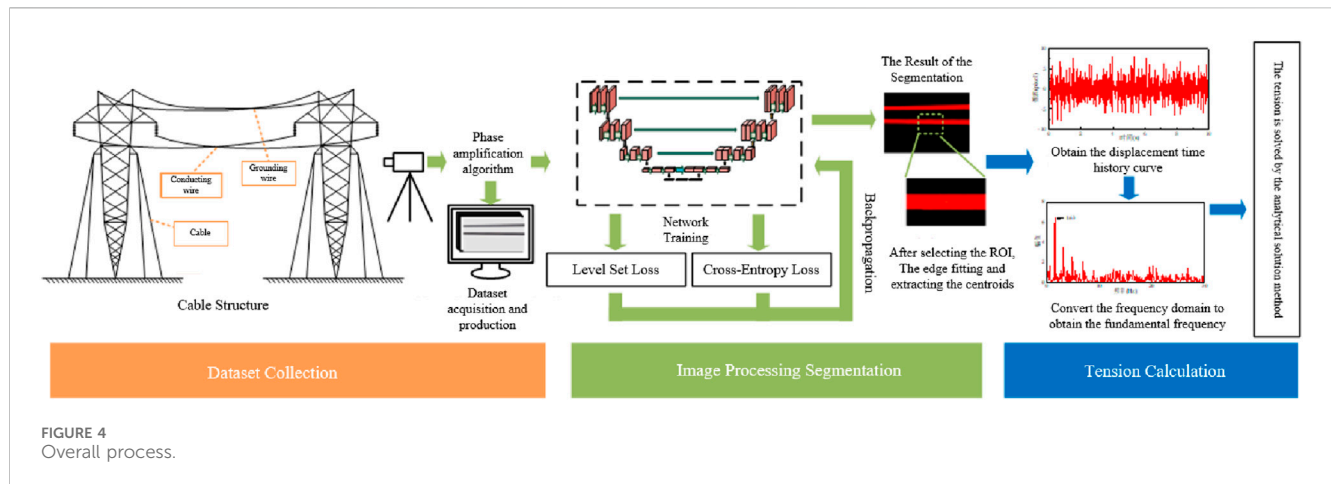
direction. During image capture, this can be viewed as motion on a plane. Based on this, a displacement coordinate system for the cable can be established, and the cable component can be segmented using a semantic segmentation network. To extract its displacement, this paper uses the centroid of the binary image after segmentation as a statistical value.

Since the edge of the cable component is not a smooth straight line, using the centroid as a reference point to extract its displacement coordinates through the semantic segmentation network may introduce some errors. Therefore, this paper employs the least squares method to fit the edge points of the cable component. The fitting formula is as follows in Equation 15:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \qquad (15)$$

Here, $\alpha_0, \alpha_1, \alpha_2 \ldots, \alpha_n$ represents the coefficients of the fitting equation, x and f(x) represent the coordinate information of the cable component's edge.

The fitted image recognition can reach sub-pixel accuracy. After selecting the Region of Interest (ROI), the center of centroid of the fitted cable image is found. After binarizing the image, the result is shown in Figure 3. The red pixels (representing the identified cable structure pixels) are located at the average y-coordinate of the vertical axis, while the green point represents the extracted center of centroid. This green
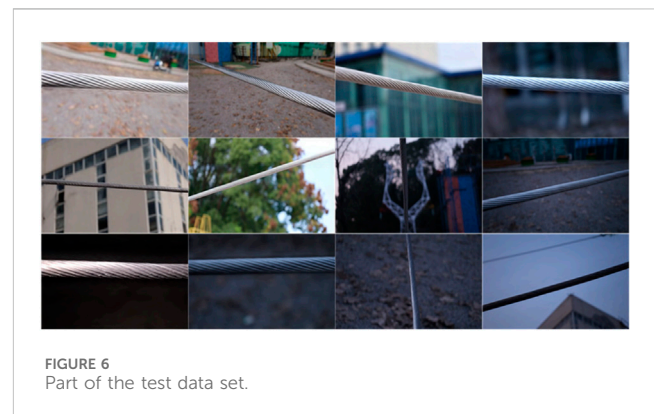
point is used as the center of centroid, and the segmented images, frame by frame, are processed to track its displacement over time. The displacement time series can then be transformed into frequency domain data using the Fast Fourier Transform (FFT) method.

2.3.2 Frequency method measurement

Under environmental excitation, transmission line cables undergo small random transverse vibrations (Yunlong et al., 2023). Treating the environmental excitation as white noise, the modal parameter identification peak-picking method can be used to recognize the natural frequency of transmission line cable components.

Using cable vibration theory, the relationship between the vibration frequency and tension of transmission line cables can be established. Let the direction of the cable installation be the x-axis, and the direction perpendicular to the cable be the y-axis. The vibration equilibrium equation of the cable can be derived as Equation 16:

$$T\frac{\partial^2 y}{\partial x^2} - m\frac{\partial^2 y}{\partial t^2} - EJ\frac{\partial^4 y}{\partial x^4} = 0 \qquad (16)$$

Here, T represents the tension of the cable component; t is time; m is the centroid per unit length of the cable; EJ represents the bending stiffness of the cable component. When the cable is relatively long, its bending stiffness can be neglected. The above equation can be simplified to the classical wave Equation 17:

TABLE 1 Different magnification tension identification.

Zoom scale	Accelerometer tension (kN)	Vision-based tension (kN)	Error (%)
5	15.485	14.620	5.59
10	15.485	14.930	3.58
15	15.485	14.838	4.18
20	15.485	16.409	−5.97

FIGURE 7
Loss curve.

FIGURE 8
Confusion matrix.
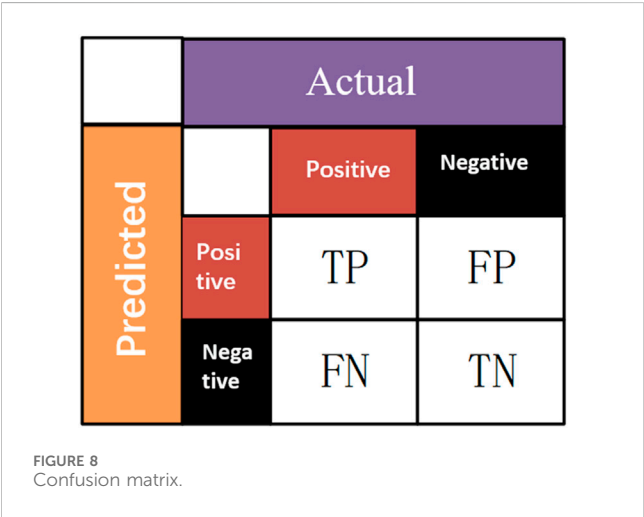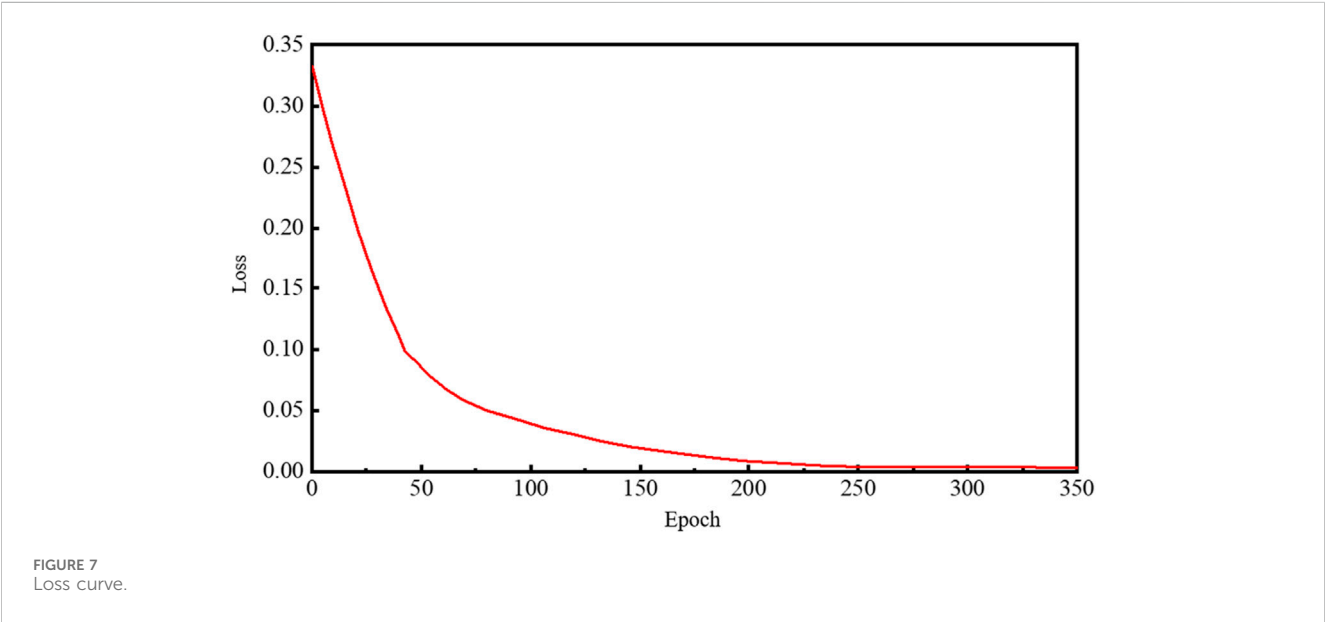
$$T\frac{\partial^2 y}{\partial x^2} - m\frac{\partial^2 y}{\partial t^2} = 0 \tag{17}$$

By using the method of separation of variables, the spatial equation of the above equation can be solved. Since the two ends of the cable are fixed (displacement is 0 at $x = 0$ and $x = I$), the relationship between the angular frequency ω and the tension T can be obtained as Equation 18:

$$\omega = \frac{n\pi}{l}\sqrt{\frac{T}{m}} \tag{18}$$

From the natural frequency $f_n = \frac{\omega}{2\pi}$, $EJ << TL^2$, the bending stiffness of such members can be considered negligible. The following can be obtained as Equation 19:

$$T = \frac{4ml^2 f n^2}{n^2} \tag{19}$$

Here, n is the order of the natural frequency of the cable vibration; f_n is the nth-order natural frequency of the cable, and l is the length of the cable.

2.4 Tension calculation process

The method for calculating the tension of transmission line cables in this paper is divided into three parts, as shown in Figure 4:

Step 1: Dataset Collection: Due to the lack of semantic segmentation datasets for transmission line cables, this paper creates a new dataset. Video images of cable vibrations under environmental excitation are collected and magnified using the broadband phase magnification algorithm. The magnified video images are then decomposed frame by frame, and part of the images are

TABLE 2 Evaluation index comparison.

Network	Loss	Evaluation metrics	
		MIoU	Accuracy
FCN	0.092	0.832	0.867
U-net	0.013	0.924	0.974

manually annotated with the edges of the cable components to create the dataset.

Step 2: Network Processing and Segmentation: The manually annotated dataset is input into a network combining the level-set and U-Net loss functions for training, achieving network fitting.

Step 3: Tension Calculation: The magnified video images are decomposed frame by frame and processed through the trained network. The semantic segmentation network converts the magnified cable image into a binary image.

After selecting the ROI area, the center of centroid of the image is used to extract the cable displacement time series, which is then transformed into the frequency domain. The tension is calculated using the frequency method formula.
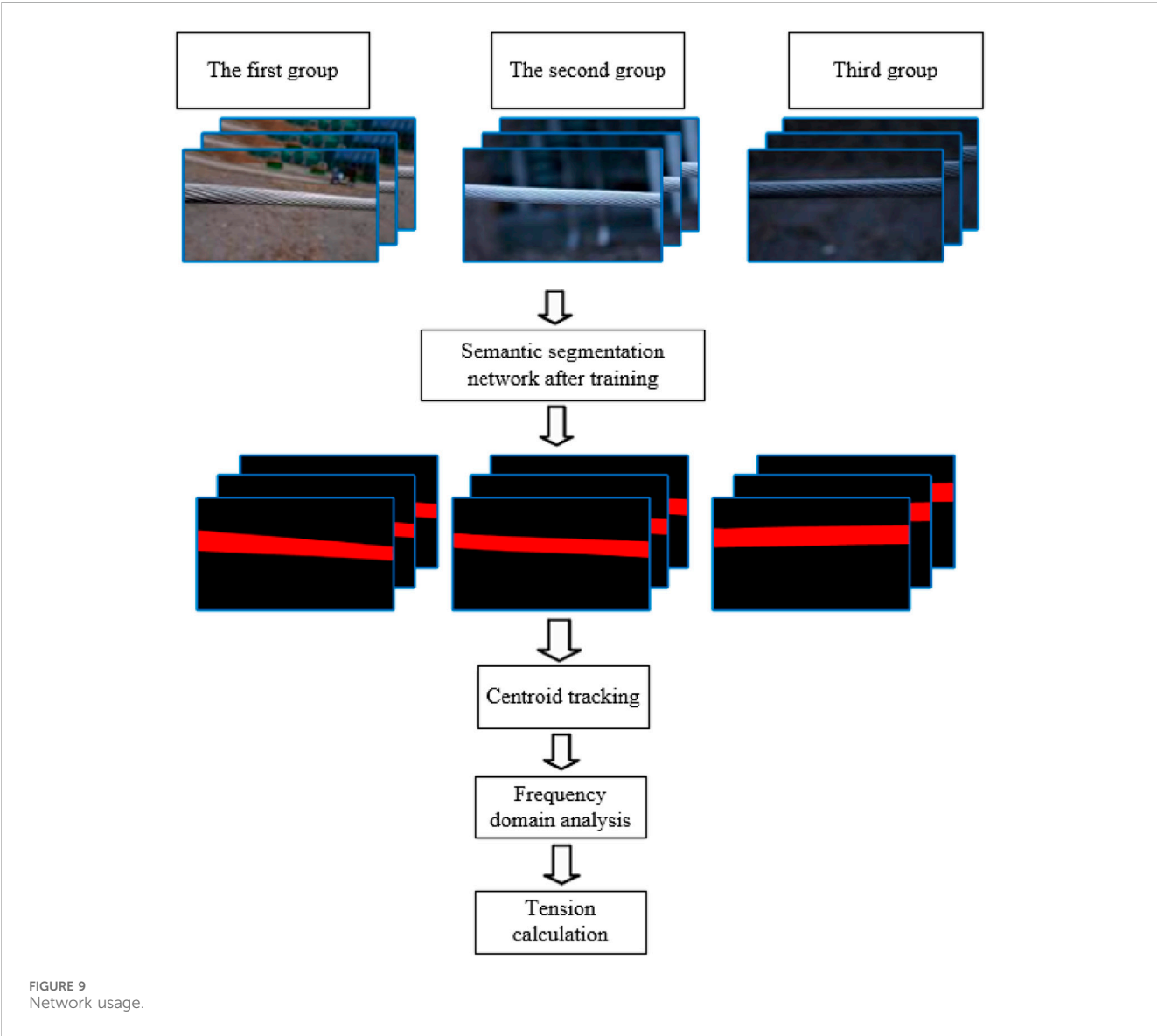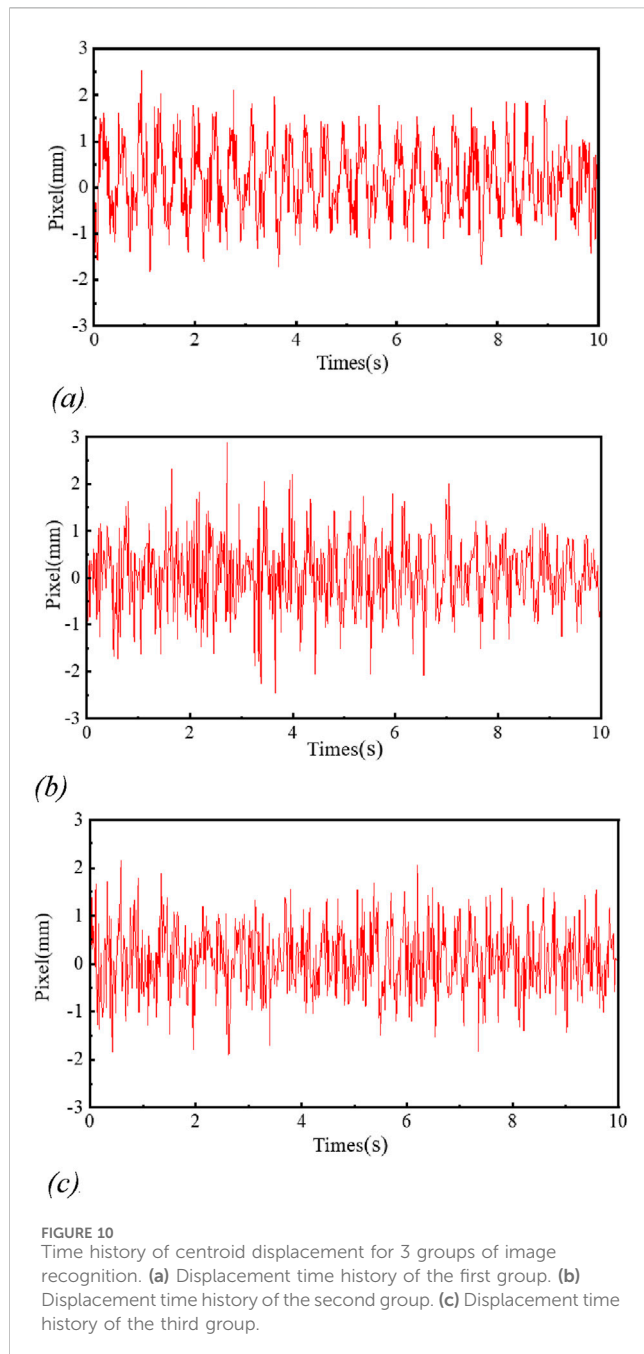
3 Expeprimental validation

3.1 Dataset creation

The dataset was collected at an outdoor experimental site, as shown in Figure 5, where steel-core aluminum-stranded wires were installed between two iron towers. The steel-core aluminum-stranded wire consists of 24 aluminum strands and 7 steel strands, with an outer diameter of 16.67 mm and a unit length centroid of 0.549 kg/m. These wires were used to simulate the operating conditions of overhead conductors, and the tension of the wires was adjusted by controlling the tensioning device.

FIGURE 9
Network usage.

FIGURE 10
Time history of centroid displacement for 3 groups of image recognition. **(a)** Displacement time history of the first group. **(b)** Displacement time history of the second group. **(c)** Displacement time history of the third group.

To verify the accuracy of the tension measurement method proposed in this paper, a tension sensor was installed between the line clamp and the tensioner to obtain the force values, which were then compared with the tension results obtained from the proposed method.

A camera (Canon EOS RP) with a frame rate of 60 frames per second and a resolution of 1920 × 1,080 was used for the video capture. To reduce image distortion caused by the lens, the camera was positioned approximately 1 m away from the target wire during each capture. The shooting angle in this article is kept as parallel as possible to the suspension components. When it is not possible to shoot in parallel, drones will be used to assist in the filming. The dataset was divided into three groups, with tension controlled by the tensioner. The groups were filmed under three tension conditions:

2673N, 7703N, and 9261N. For each group, 7 sets of 10-s vibration video images of the wire under environmental excitation were captured from different angles and lighting conditions. During the experimental testing period, the environmental excitation was predominantly aeolian vibration, with frequencies ranging from 10 Hz to 40 Hz. The lighting conditions were controlled to be well-lit in the morning and insufficiently lit in the evening, with a filming ratio of 5:2 for daytime and evening shots.

The 21 sets of 10-s video images collected were magnified using the BPMM algorithm with a magnification factor of 10. To balance the workload and network fitting accuracy, one frame was selected every 30 frames from the magnified video images to create the dataset, resulting in a total of 420 original images. A part of the dataset is shown in Figure 6. The edges of the wire in the images were manually annotated using Labelme, and the original JSON files were converted into JPG images with mask masks, which were then fed into the network for training.

To validate the robustness of the proposed method across different magnification scales, we selected zoom levels of 5, 10, 15, and 20 as comparative conditions. Videos at these four different magnification scales were input into the network, and the centroid pixel variation was statistically calculated to determine the tension. The measured results were compared with those from the acceleration sensor, as shown in Table 1.

Based on the results presented in Table 1, the proposed method effectively extracts the pixel-level morphological features of cable components even when affected by artifacts and accurately calculates their tension values, with all errors falling within acceptable ranges for engineering applications. The data indicate that under ambient excitation, cables undergo minor vibrations. When the zoom scale is relatively small ($\alpha < 10$), the measurement error is more pronounced. At zoom scales of 10 or 15, the cable tension can be measured with higher accuracy. Therefore, for cable vibration analysis, it is concluded that maintaining a zoom scale around 10 enables accurate identification while avoiding image distortion caused by excessive magnification.

To prevent overfitting, the initial dataset was augmented by random rotations, random cropping, and random resizing, generating a total of 3220 augmented images. The original dataset and augmented dataset were combined, totaling 3640 images, and they were randomly split into training and validation sets at an 8:2 ratio.

3.2 Network training

The experiment in this paper was conducted on a Windows system, based on the GPU version of the PyTorch deep learning framework, using an RTX3060Ti 8G graphics card as the experimental environment. The labeled training set images were input into the network for training, with the learning parameters configured as a learning rate of 0.001, a maximum of 500 epochs, and a batch size of 2. To avoid overfitting due to excessive training, the early stopping method was used to control the network. Training was stopped after 8 consecutive epochs without improvement in the validation loss. The total training time for the network was 6 h and 17 min, and the network stabilized and converged.

TABLE 3 Results of 4 groups of experiments.

Serial number	Tension sensor displays the tension value/N	Image recognition frequency value/Hz	Image recognition tension value/N	Error/%
1	2673	1.267	2608	2.43
2	7703	2.148	7496	2.69
3	9261	2.345	8934	3.53
4	9261	2.292	8534	7.84

The U-Net experiment was conducted in a Windows environment utilizing a GPU-accelerated Py Torch deep learning framework, with an RTX 3060 Ti 8G graphics card serving as the hardware platform. The annotated training set images were fed into the network for training with the following hyperparameters: a learning rate of 0.001, maximum of 500 epochs, and a batch size of 4. An early stopping mechanism was implemented, halting the training process if the validation loss showed no improvement over eight consecutive epochs. The total training duration was 6 h and 27 min, by which point the network had stabilized and converged. The corresponding loss trajectory during training is illustrated in Figure 7.

The four values in the confusion matrix are critical parameters for evaluating the performance of the network model, as illustrated in Figure 8.

Specifically: TP denotes the number of samples correctly predicted as positive; FP represents the number of samples incorrectly predicted as positive; FN indicates the number of samples incorrectly predicted as negative; TN corresponds to the number of samples correctly predicted as negative.

The metric MIoU (Mean Intersection over Union) quantifies model accuracy by measuring the overlap between the predicted segmentation and the ground truth labels. It is mathematically defined as Equation 20:

$$MIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{TP}{FP + FN + TP} \qquad (20)$$

Here, k denotes the total number of classes.

The metric **Accuracy** quantifies the proportion of correctly predicted samples relative to the total number of samples. It is calculated using the Equation 21:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (21)$$

The training evaluation metrics for the FCN and U-Net networks are presented in Table 2 below.

As shown in Table 2, the improved U-Net achieves superior performance over the FCN network in terms of both MIoU and Accuracy, validating that its adoption can significantly enhance segmentation performance.

3.3 Measurement results

After the network training is completed, the magnified video images under three different tension levels are decomposed frame by frame and input into the trained semantic segmentation network.

The displacement of their centroids is then statistically measured. The measurement process is shown in Figure 9.

The centroid displacement time histories for the three groups are shown in Figure 10. From the displacement time history graph, it can be observed that the maximum centroid displacement for the three groups of magnified images is about 3 pixels. The corresponding frequency spectra of the three groups after fast Fourier transform (FFT) are shown in Figures 11a–c. The first peak value in the spectrum corresponds to the fundamental frequency of the conductor. By substituting this value into Equation 17, the calculation results are presented in Table 3. The measurement errors for the three groups of tension are all below 4%, verifying the effectiveness of the proposed method for micro-vibration tension identification.

To explore the effect of lighting on the proposed method, a fourth group of experiments was added. The video for this group was captured in the evening under poor lighting conditions, and the tension sensor reading was 9261N. The magnified image videos were decomposed frame by frame and input into the network. The results for the fourth group are shown in Table 3. By calculating the error between the image recognition results and the tension sensor readings for the fourth group, it was found that the error was relatively large, reaching 7.84%. This could be due to the semantic segmentation network's poor recognition ability under low lighting conditions, where the image pixel contrast is not high.

To verify the applicability of the proposed method for detecting small vibrations, a video image captured during a well-lit period was tested, as shown in Figure 12. The video was magnified using the broadband phase algorithm and input into the semantic segmentation network to extract the centroid. The displacement time histories before and after amplification were compared, as shown in Figure 12a.

By comparing the centroid displacement maps before and after amplification, it can be observed that the vibration amplitude of the conductor without the amplification algorithm is relatively small under environmental excitation. The image displacement is within 0.6 pixels, and in many cases, no displacement change is detected, leading to little difference in pixel displacement between adjacent frames. However, after applying broadband phase amplification, the centroid displacement extracted through semantic segmentation can reach approximately 2 pixels.

The Fourier-transformed frequency spectrum analysis, as shown in Figure 12b, indicates that the frequency spectrum of the image not processed by the magnification algorithm struggles to detect the fundamental frequency peak. However, the image processed by the method presented in this paper allows for the identification of the natural frequency through peak detection, enabling tension calculation. This validates that the proposed method can effectively amplify small vibrations and accurately extract the displacement-time history, thus solving the issue of machine
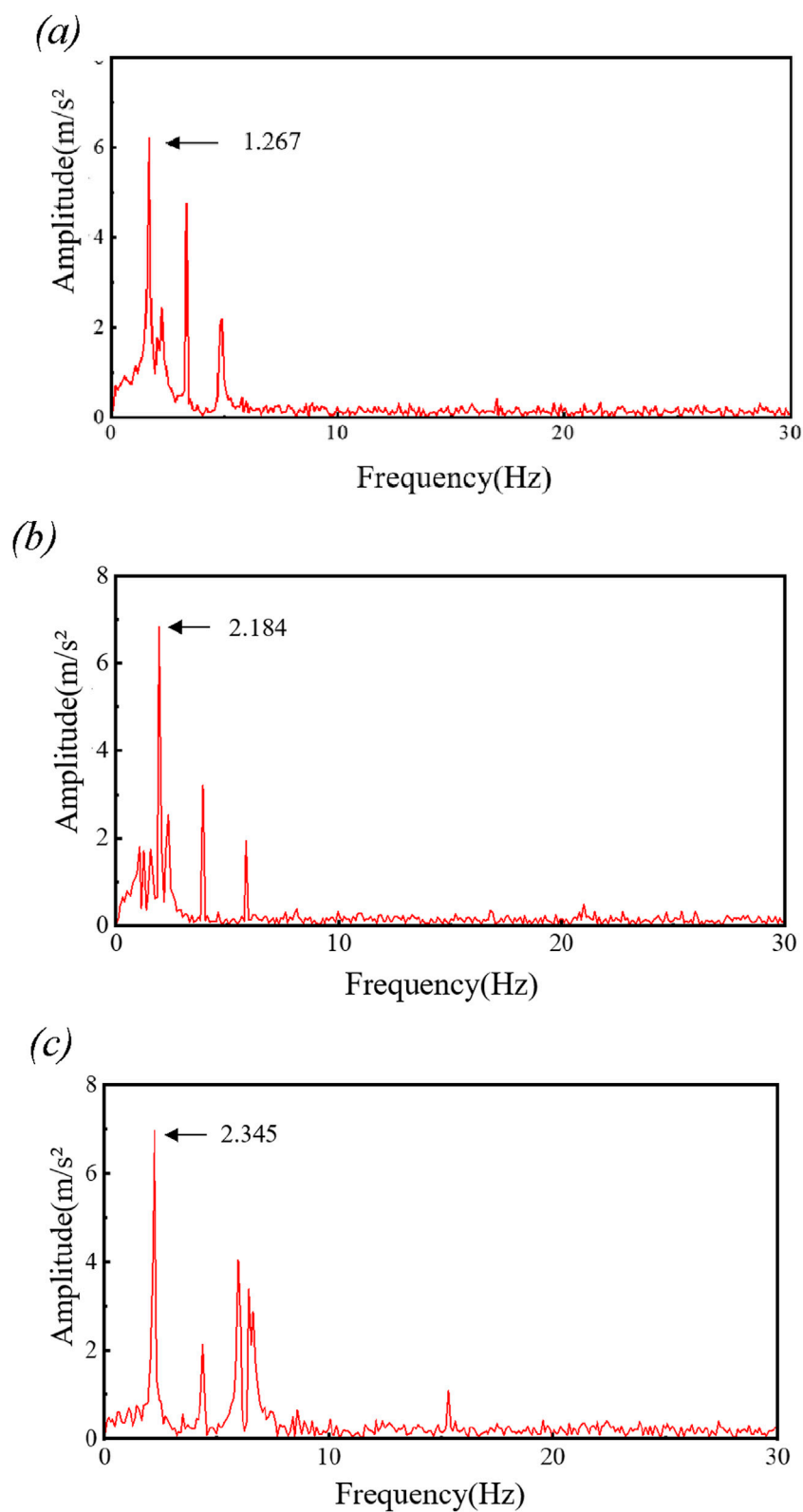
FIGURE 11
3 groups of image recognition spectrum. **(a)** Spectrum of the first group. **(b)** Spectrum of the second group. **(c)** Spectrum of the third group.
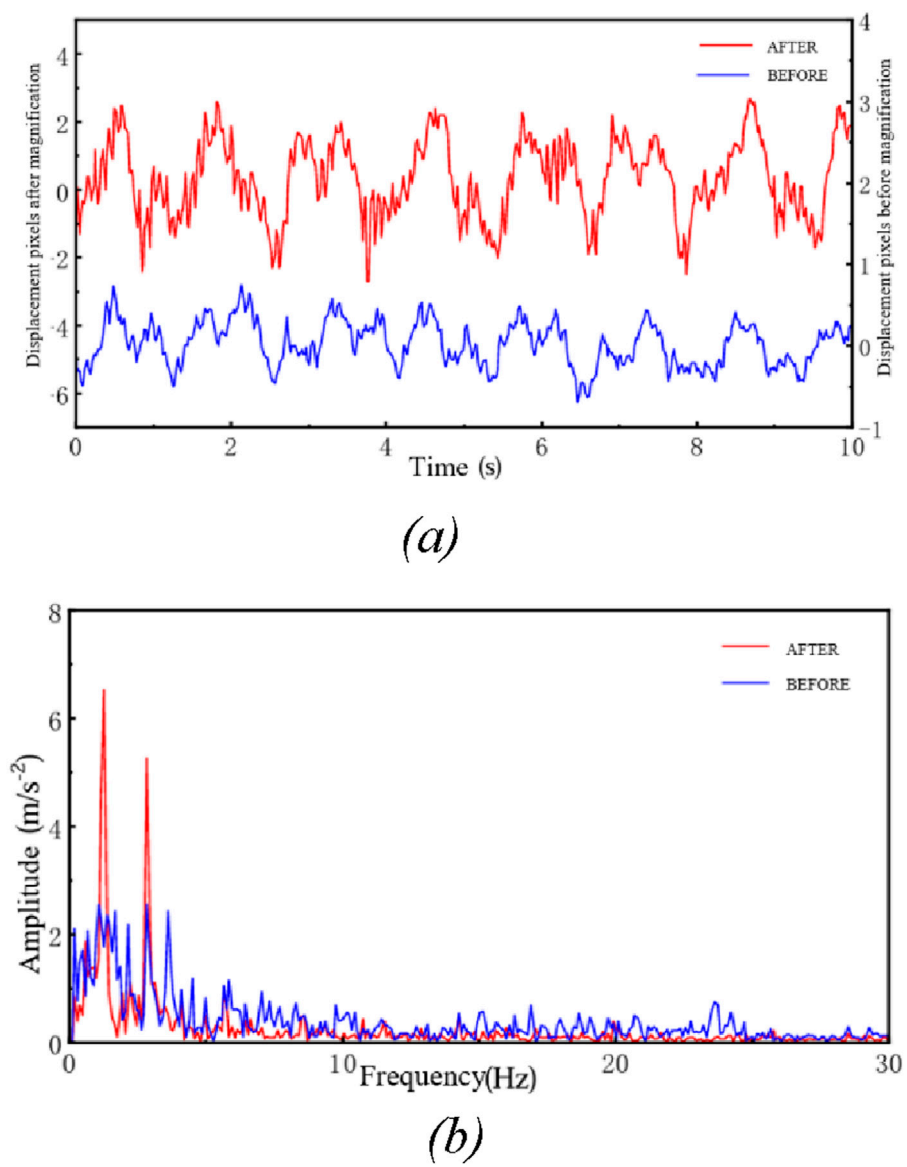
FIGURE 12
Network segmentation results before and after magnification algorithm. **(a)** Comparison of displacement time histories before and after magnification. **(b)** Comparison of frequency spectra before and after magnification.
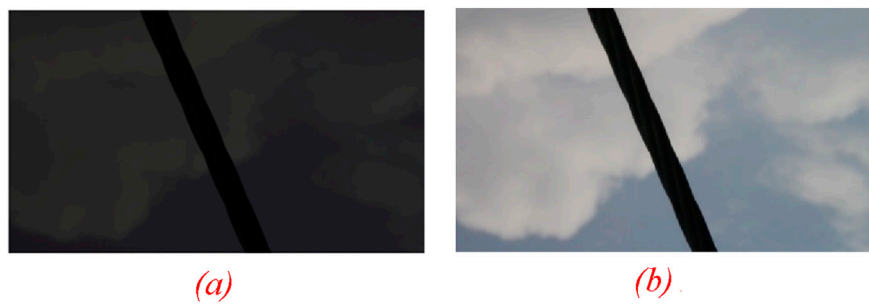

FIGURE 13
Performance Comparison of Low-light Adaptation: Pre-vs. Post-Improvement. **(a)** Original low-light image. **(b)** Enhanced low-light image.

vision recognition unable to obtain accurate natural frequencies due to the small displacement pixels in images without magnification.

The primary source of error in the proposed method stems from a significant reduction in contrast between the conductor and the background, which leads to a decreased image signal-to-noise ratio (SNR). Experimental data indicate that during evening hours, the grayscale difference between the conductor and the background drops from 80 to 120 under normal lighting conditions to 20–40, accompanied by an approximately 70% attenuation in edge gradient magnitude. This degradation in image quality directly compromises the accuracy of subsequent edge detection and centroid localization. Furthermore, despite the use of fixed mounting brackets, minor environmental vibrations can still induce image jitter, thereby introducing additional measurement errors.

3.4 Improved adaptation to low-light conditions

To address the performance degradation of image segmentation under low-light conditions, this paper introduces an adaptive contrast enhancement mechanism and an illumination-invariant feature extraction method during the image preprocessing stage, as shown in Figure 13. These improvements enhance feature extraction in low-light scenarios at the level of physical imaging mechanisms.

3.4.1 Adaptive contrast enhancement mechanism

The Contrast-Limited Adaptive Histogram Equalization (CLAHE) method is employed, which operates by dividing the image into sub-regions and performing histogram equalization independently on each. This approach effectively enhances local contrast without amplifying noise. The process can be mathematically represented as follows:

$$I_{\text{enhanced}}(x, y) = \text{CLAHE}\left(I_{\text{original}}(x, y), \text{clipLimit}, \text{tileGridSize}\right)$$

where the clip limit is set to 2.0 and the tile grid size is 8×8. This configuration maximizes contrast enhancement while suppressing noise amplification. By constraining the height of local histograms, it limits the degree of contrast enhancement to prevent over-enhancement in homogeneous regions.

3.4.2 Illumination-invariant feature extraction

Based on the Retinex theory, the image is decomposed into an illumination component and a reflection component as Equation 22:

$$I(x, y) = L(x, y) \cdot R(x, y) \qquad (22)$$

The two components are separated via logarithmic transformation and Gaussian filtering, with emphasis on enhancing the intrinsic object features contained within the reflection component as Equation 23:

$$R(x, y) = \exp\left(\log I(x, y) - \log\left[G_\sigma(x, y)^* I(x, y)\right]\right) \qquad (23)$$

Where G represents the Gaussian kernel function as Equation 24:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \qquad (24)$$

Multi-scale processing (using σ_1, σ_2, σ_3) is applied to integrate detail information from different frequency bands, resulting in the final reflection component as Equation 25:

$$R(x, y) = \sum_{k=1}^{3} w_k R_k(x, y) \qquad (25)$$

Following the aforementioned enhancement process, experimental results demonstrate that the segmentation error under low-light conditions is reduced from 7.84% to 3.75%. This improvement significantly enhances the all-weather applicability of the method and provides reliable technical support for nighttime maintenance of transmission lines.

4 Conclusion

1. This paper achieves high-precision, non-contact tension measurement for transmission line cable components by accurately extracting the centroid of the cable member using a segmentation method that combines the deep learning U-Net network with a level set loss entropy function. This enables precise acquisition of the vibration displacement time history. Subsequently, the natural frequency is obtained through frequency domain analysis, from which the cable tension is derived. The measured cable tension values exhibit errors within 8% when compared to sensor-based measurements.

2. The non-contact cable tension measurement method proposed in this paper demonstrates errors within 8% compared to sensor measurements. Furthermore, by applying image enhancement techniques to improve the contrast between the conductor and the background in images captured under low-light conditions, the measurement error under such conditions was reduced from 7.84% to 3.75%.

3. Limited by the current number of learning samples, the generalization capability of the proposed algorithm under extreme operating conditions requires further improvement. In subsequent research, we will focus on constructing a large-scale and diverse visual database for transmission lines. This will involve systematically collecting conductor images under various meteorological conditions, geographical environments, and line structure configurations, thereby further enhancing the generalization performance of the proposed method in complex scenarios.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

HZ: Conceptualization, Writing – original draft, Writing – review and editing. WS: Data curation,

Conflict of interest

Authors HZ, WS, WH, and LZ were employed by Jiangmen Power Supply Bureau, Guangdong Power Grid Co., Ltd.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Banfu, Y., Zechu, C., and Zigang, Z. (2015). Cable force identification based on noncontact photogrammetry system. *J. Hunan Univ. Nat. Sci.* 42 (11), 105–110. doi:10.16339/j.cnki.hdxbzkb.2015.11.035

Chuankai, Y., Kong Zhizha, K., and Xie Qiannan, X. (2021). Image recognition method for transmission line based on the DeepLab v3+ deep convolutional network. *Electr. Power Eng. Technol.* 40 (04), 189–194. doi:10.12158/j.2096-3203.2021.04.027

Dengke, Y., Zhengliang, L., and Jinghua, S. (2015). Wind tunnel tests for effect of guy pretension on mechanical properties of UHV cross-rope suspension tower-line. *J. Vib. Shock* 34 (13), 163–168. doi:10.13465/j.cnki.jvs.2015.13.028

Haitao, Z., Haojun, L., and Yuanyong, K. (2023). Research on laser point cloud data slicing technology for power grid line inspection based on enhanced octree. *High. Volt. Eng.* 49 (S1), 97–102. doi:10.13336/j.1003-6520.hve.20231164

He, L., Daiyong, Y., and Chunming, Z. (2023). Extraction of galloping characteristics of overhead line in distribution network by using deep learning method. *Proceedings of the CSU-EPSA* 35 (5), 89–94. doi:10.19635/j.cnki.csu-epsa.001225

Jian, W., Xiang, W., and Zhimin, Z. (2022). Nano millimeter wave radar for bridge cable tension measurement. *J. Natl. Univ. Def. Technol.* 44 (02), 118–122. doi:10.11887/j.cn.202202015

Kim, Y., Kim, S., Kim, T., and Kim, C. (2019). CNN-based semantic segmentation using level set loss. *ArXiv* 1752–1760. doi:10.1109/wacv.2019.00191

Lan, J., Ruoheng, C., and Bo, T. (2022). Tension test method for guy cables of transmission line guyed tower based on image edge recognition. *High. Volt. Eng.* 48 (11), 4469–4477. doi:10.13336/j.1003-6520.hve.20220417

Lee, H., Yoon, H., and Kim, S. (2023). Vibration detection of stay-cable from low-quality CCTV images using deep-learning-based dehazing and semantic segmentation algorithms. *Eng. Struct.* 29. doi:10.1016/j.engstruct.2023.116567

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 07-12 June 2015 (IEEE), 431–440. doi:10.1109/TPAMI.2016.2572683

Longqiang, G., Yunze, H., and Xu, D. (2023). Volumetric measurement of idar based on visual correction. *Chin. J. Sci. Instrum.* 44 (10), 48–59. doi:10.19650/j.cnki.cjsi.J2311797

Qiao, H., Yichao, W., and Yuan, R. (2022). Cable force measurement of cable-stayed bridge based on microwave interferometric radar. *J. Hohai Univ. Nat. Sci.* 50 (06), 144–151. doi:10.3876/j.issn.10001980.2022.06.019

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci.* 9351, 234–241. doi:10.1007/978-3-319-24574-4_28

Wengang, Y., Zhangqi, W., and Bowen, Z. (2015). Time history analysis on wind-induced response of uhv guyed single-mast transmission tower-line system. *Proc. CSEE* 35 (12), 3182–3191. doi:10.13334/j.0258-8013.pcsee.2015.12.032

Xiongjun, H., Yongchao, Y., and Xiang, X. (2016). A vibration frequency method based on finite element equivalent cable length. *Bridge Constr.* 46 (06), 40–44.

Xuan, K., Kui, L., and Lu, D. (2023). Structural frequency identification based on broad-band phase-basedmotion magnification and computer vision. *CCEJ* 56 (10), 105–117. doi:10.15951/j.tmgcxb.22060550

Yanbao, H., Wei, L., and Ruijian, P. (2022). Review on intelligent live-line maintenance technology applied on power transmission lines. *Electr. Power Autom. Equip.* 42 (02), 163–175. doi:10.16081/j.epae.202110015

Yang, L., and Dengke, H. (2022). First-break picking technology based on level set method and U-Net. *J. Min. Sci. Technol.* 7 (04), 437–445. doi:10.19606/j.cnki.jmst.2022.04.005

Yang, H., Xu, H., and Jiao, C. (2021). Semantic image segmentation based cable vibration frequency visual monitoring using modified convolutional neural network with pixel-wise weighting strategy. *Remote Sens.* 13 (8), 1466. doi:10.3390/rs13081466

Yao, Z., Ao-Han, W., and Hong, Z. (2021). Overview of smart grid development in China. *Power Syst. Prot. Control* 49 (05), 180–187. doi:10.19783/j.cnki.pspc.200573

Yunlong, Z., Jiayuan, Z., and Xuesong, Q. (2023). Spectrum-driven methods for modal parameter identification of bridge under environmental excitation. *J. Jilin Univ. (Eng. Tech. Ed.)* 53 (06), 1580–1591. doi:10.13229/j.cnki.jdxbgxb.20230077

Zhang Yuhang, Z., Su Cheng, S., and Yichuan, D. (2021). A eulerian video magnification based cable tension identification method for bridge structures. *J. Graph.* 42 (6), 941–947. doi:10.11996/JG.j.2095-302X.2021060941

Zhanghua, X., Qihuang, C., and Youqin, L. (2023). Grouted prestressed cable force calculation method based on vibration theory. *Chin. J. Appl. Mech.*, 1–14.

Zhenya, L. (2009). "Electromagnetic environment of UHVDC transmission project," in *UHVDC transmission technology series*. Beijing, China: China Electric Power Press.