scientific reports

OPEN

Innovative segmentation technique for aerial power lines via amplitude stretching transform

Pengfei Xu[1,2✉], Nor Anis Asma Sulaiman[2], Yafei Ding[1] & Jiangwei Zhao[1]

Accurate segmentation of power line targets helps quickly locate faults, evaluate line conditions, and provides key image data support and analysis for the safe and stable operation of the power system.The aerial power line in segmentation due to the target is small, and the imaging reflected energy is weak, so the Unmanned Aerial Vehicle (UAV) aerial power line image is very susceptible to the interference of the environment line elements and noise, resulting in the detection of the power line target in the image of the defective, intermittent, straight line interferences and other low accuracy and real-time efficiency is not high. For this reason, this paper designs a pure amplitude stretching kernel function to form a Fourier amplitude vector field and uses this amplitude vector field to implement the stretching transformation of the amplitude field of the aerial power line image, so that the angular field after the Fourier inverse transformation can better react to the spatial domain line targets, and finally, after the Relative Total Variation (RTV) processing, the power line can be well detected. The proposed algorithm is compared with the main power line segmentation algorithms, such as Region Convolutional Neural Networks(R-CNN) and Phase Stretch Transform(PST). The average values of evaluation indicators P_{PA}, M_{MPA} and M_{MIoU} of the image segmentation results of the proposed algorithm reach 0.96, 0.96 and 0.95 respectively, and the average time lag of detection is less than 0.2s, indicating that the accuracy and real-time performance of the segmentation results of the proposed algorithm are significantly better than those of the above algorithms.

In the process of UAV inspection of power lines, power line detection is crucial to the safe flight of UAVs. Due to the complex background of aerial images and the different morphologies of power lines, which often appear in different positions with different postures and at various angles, it is difficult to distinguish power lines from similar straight lines or edges. Therefore, vision-based automatic power line detection faces the following severe challenges.

- The UAV carries visible light equipment to realize power line aerial image acquisition. During image acquisition, there are inevitably camera shaking and light changes, resulting in image blurring and large changes in color intensity, thus affecting the accurate extraction of power line features.
- Power lines in aerial images are elongated and sometimes appear as discontinuous lines such as intermittent and fuzzy, which further increases the difficulty of effective feature extraction for power lines.
- UAV inspection images are characterized by the presence of complex background clutter accompanied by many false targets, such as vegetation, fields, rivers, and buildings with straight lines in the background that are similar in appearance to the power lines, making it difficult to distinguish between the power lines and the false targets in the background.
- Power line images captured by UAV inspection lack significant linear features and sufficient pixel ratios. Compared to targets such as buildings and tall vegetation, power lines occupy a small pixel ratio, and the vast majority of pixel points in power line images are non-power lines (background), where other linear objects in the background introduce noise, further complicating power line detection.
- Aerial images used for power line inspection are made more difficult in extreme weather, such as dense fog, rain, snow, ice, and dust.

[1]Henan International Joint Laboratory of Machine Vision and Intelligent Systems, Department of Information Engineering, Pingdingshan University, Pingdingshan 467000, Henan, China. [2]Center of Sustainable in Software Engineering, Faculty of Engineering, Built Environment & Information Technology, SEGI University, Kota Damansara, Malaysia. ✉email: 3221@pdsu.edu.cn

1

In recent years, many scholars have proposed many algorithms for power line segmentation and extraction, and the main methods are summarized as the Hough transform, the Radon transform, the Line Segment Detector (LSD) method, the Freeman chain code, the phase grouping, the Deep Convolutional Neural Network (DCNN) classification, and algorithms that improve on the above algorithms[1]. Li et al.[2] developed a Pulse Coupled Neural Network (PCNN) fusion algorithm combining filter and Hough transform to detect power lines, which helps to reduce the false detection rate and algorithm complexity. However, it is easy to misdetect or miss detection in complex scenes. Zhang et al.[3] realized automatic detection and tracking of power lines. The algorithm is computationally large, and the image processing time is long. Yuan et al.[4] realized linking and fitting of broken and underlinked straight lines, and the extraction of power lines in complex geographic environments requires the adjustment of the threshold range of the discrete polar angle in the aerial image. Cao et al.[5] utilized the Hough transform and least squares method to detect the location of power lines and fit the power lines, the width of the power lines needs to be reasonable. Nasseri et al.[6] seted the Sobel operator and particle swarm optimization algorithm to improve the Hough transform. Zhao et al.[7] used a multi-scale linear target reinforcement algorithm and angular constraints Radon transform for the identification of high-voltage power transmission lines, which has very good noise immunity, interference immunity, and real-time performance. Approximate horizontal power lines and vertical interfering objects are required as a priori conditions. Tian et al.[8] utilized a priori knowledge constraints and the Radon transform to extract straight lines, which reduced the rate of incorrect extraction of power lines and improved the efficiency of power line recognition. The complexity of the algorithm is high, and the range of power line extraction has certain constraints, but there are still some straight-line interferences. Fu et al.[9] used Canny edge detection and LSD algorithm to remove noise well in complex backgrounds, and accurately extract nearly parallel and locally symmetric power lines. Wu et al.[10] accurately recognized power lines in complex backgrounds by using the LSD algorithm and found the endpoints of straight line segments, and the detection accuracy rate is above 92%. However, it is difficult to recognize power lines accurately and effectively in the case of weak visual features of power lines. Luo et al.[11] improved the LSD algorithm by using near-infrared images, which has a higher detection rate, lower false alarm rate, and better robustness than the LSD algorithm, and the algorithm's computation is larger, which requires more time. Zhang et al.[12] realized that auxiliary power line detection by power pylons with a higher detection rate and lower false alarm rate. detection rate and low false alarm rate, the power line spatial mapping may be invalid. Teplyakov L et al.[13] used the LSD algorithm and line segment fitting algorithm, which can quickly extract the complete power line under the noise interference, but it needs to be based on the selection of the threshold parameter to decide whether the detection performance is good or bad. Zhang et al.[14] solved the broken strand and splitting problem occurring in the detection of power lines. Searching the 4 or 8 neighborhood pixel points of each pixel point in the aerial image is not only computationally intensive but also requires more storage space. Li et al.[15] effectively solved the problem of misdetection, breakage, and omission of straight lines in the power line image under the complex background environment by utilizing the pixel gradient and the distance grouping. Zhao et al.[16] implemented a power line extraction algorithm with localized contextual information, and the algorithm can be implemented in The algorithm can realize the accurate extraction of power lines under various aerial photography scenes, and has high engineering application value. The algorithm is susceptible to noise, resulting in pixel phase mutation and incomplete power line extraction.

Gubbi et al.[17] implemented four convolutional layers and two fully connected layers, and the power line detection F1-score was improved by 3.6% compared with that of the GoogLeNet pre-trained model. The input image has to be chunked and needs to be converted into gradient histogram HoG vectors. Yetgin et al.[18] implemented end-to-end detection of power lines in aerial images. The exact location of the power lines could not be determined. Liu et al.[19] improved the Faster R-CNN and the category detection mAP reached more than 92%, which is 8% higher than that of the original Faster R-CNN model. Extracting power line features also extracts a large amount of background noise, which affects the accurate localization of power lines. Jiang et al.[20] implemented the Spatial Attention (SA)-Unet model based on the stage attention mechanism, and the pixel accuracy and the average intersection and merger ratio of the SA-Unet model are better than those of the FCN, SegNet, and U-Net networks. Wang et al.[21] introduced Deeplab v3 and Decoder to extract transmission lines, and the mIoU of the DeepLab v3 + Decoder network was improved by 2.3% compared to the DeepLab v3 + network. Madaan et al.[22] implemented front-end and context modules and concluded that the model used outperforms the Fully Convolutional Networks (FCN)−8 model. Nguyen V N et al.[23] solved the problem of detected power line discontinuities using the LS-Net network model. Chen M et al.[24] implemented self-supervised learning combined with the SaSnet model for power line extraction, which outperforms the existing optimal methods in terms of accuracy and speed. However, these algorithms require large training samples, have high requirements on hardware resources, different labeling criteria for each dataset, and no uniform measure for the network model. Choi H et al.[25] proposed a power line detection algorithm using a weakly supervised learning method to reduce the labeling cost for dataset generation. Yan J et al.[26] proposed a novel method for detecting broken strands on power lines in images captured by UAVs. Hu C et al.[27] applied the YOLOv5s algorithm to improve the detection precision and classification accuracy of the insulator self-explosions. Zhao W et al.[28] constructed a power line dataset using UAV images and classified the data according to the image clutter (IC). A method combining line detection and semantic segmentation is used. Son H S et al.[29] dealt with the power line recognition problem for the safety of agricultural spraying drones in agricultural environments. Wang D et al.[30] proposed a power line detection algorithm for aerial images based on adjustable direction filtering. In this paper, the detection and segmentation of power lines in UAV multi-spectral images using a convolutional neural network is proposed[31].

However, due to the power line target is small, the imaging reflected energy is weak, and easily affected by noise, coupled with the environment of the line elements of the mixed interference, resulting in power line imaging noise is serious, the imaging is weak and fuzzy, which is not easy to distinguish. This makes the above detection algorithms exist in the training time is long. Additionally, because of the existence of straight-line

interfering objects, the power line inspection results are not continuous and are less than the optimal extraction results. To improve the stability and accuracy of detection, detection algorithms must be able to reliably detect power lines in complex backgrounds and multiple application scenarios. To enhance the power line target expression capability under limited hardware resources. To improve real-time detection, the efficiency of power line detection must be improved under the conditions of limited computing power and limited memory occupation.

The Fourier transform tool links the spatial domain with the frequency domain, and has certain advantages for the frequency domain algorithms for problems such as power line detection. PST[32] was developed based on the time-stretching transform of analog signals, but it can also be applied to the processing of digital signals or images; its essential characteristic is to use phase data after divergence stretching, or the information highlights and enhances the high-frequency features of an image and converts high-frequency feature expressions based on image intensity values into representations through stretching, which is a new perspective or theoretical method for extracting and expressing image features. It is a method of extracting and expressing image features from a new perspective or theory by converting high-frequency features based on image intensity values into a phase representation after stretching. At present, the PST has been used in digital image edge feature detection[33], feature enhancement of visually impaired images[34], digital image compression[35], etc. However, the main theoretical phase distortion stretching transform has not been analyzed in depth in the paper, and the phase-shifted kernel function initially designed by M.H. Asghari and B. Jalali has been applied to concentric circles of equal radii in the plane of the frequency domain (u, v). On the concentric circles, the phase shifts of each spectrum are equal, which causes the detected edges to contain considerable noise, especially those isolated fine high-frequency noise that is mistaken for high-frequency components in the image and retained, which makes the subsequent thresholding process difficult.

To improve the stability, and accuracy and adapt to the real-time demand of aerial power line segmentation. In this paper, a pure amplitude stretch transform kernel function model is proposed based on the PST, the theoretical derivation process of the pure amplitude stretch transform kernel function is given, and the kernel function is applied to the aerial power line image segmentation. After the experimental results are verified and the data analysis of pixel accuracy, average pixel accuracy and average intersection and merger ratio of evaluating the indicators of the image segmentation, the very good results of aerial power line segmentation are obtained.

Theoretical foundation derivation and result analysis
Introduction to PST
In 2015, M.H. Asghari and B. Jalali proposed a digital image transform method inspired by a physical phenomenon called the PST, which simulates the propagation process of electromagnetic waves in a diffractive medium with a distorted dispersion dielectric function. This PST simulates the diffraction process with an all-pass phase filter $H(\omega) = \exp\{j\beta(\omega)\}$ with a specific frequency-dependent dispersion, where the group delay of the phase $\beta(\omega)\tau(\omega) = \partial\beta(\omega)/\partial\omega$ is a linear or sublinear function of the "S" type, which allows to reshape the shape of the complex field of the signal prior to signal sampling and digitization on the analog side. The shape of the signal complex field can be reshaped before sampling and digitizing the signal at the analog end, thus compressing the bandwidth of the analog signal without increasing the duration of the signal in the time domain, i.e., reducing the time-bandwidth product (TBP) of the signal, which "solves the two problems inherent in the traditional Nyquist uniform sampling itself"[32]: First, for a given sampling rate, conventional Nyquist uniform sampling can only capture twice the maximum frequency component of the signal; second, when there is redundancy in the analog signal, the conventional Nyquist uniform sampling process results in a much larger number of samples than is necessary (because the portion of the signal that is lower than the Nyquist frequency is oversampled).

PST, which is applied before analog-to-digital conversion, can also be applied in the field of digital signal processing, and its process for digital image edge detection is shown in Fig. 1, in which the original image first undergoes a local low-pass filter kernel function to achieve smoothing, and and then implements a nonlinear phase operation of the frequency function in the frequency domain, which is known as PST, and finally achieves the edge detection through the post-processing of thresholding and morphological filtering.

The mathematical model of the frequency domain phase stretch transform PST is[32]:

$$A(m, n) = \angle IFFT2\left\{\widetilde{K}\left(u, v\right)\cdot\ \widetilde{L}\left(u, v\right)\cdot\ FFT2[B(m, n)]\right\} \tag{1}$$
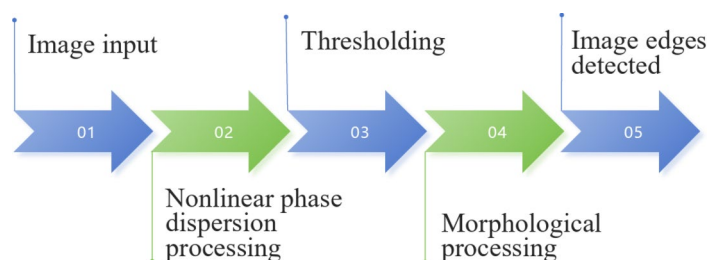

Fig. 1. Schematic diagram of the process of detecting image edges via the PST.

Where,$A(m, n)$ denotes the angular image, "∠" denotes the angle-taking operation,$B(m, n)$ denotes the original input image, FFT2 and IFFT2 denote the two-dimensional Fast Fourier Transform and Inverse Transform, respectively, and (u, v) denotes the frequency variable. $\tilde{L}(u, v)$ is the frequency response of the locally smoothed low-pass filter, and $\tilde{K}(u, v) = e^{-j\varphi(u,v)}$ is the frequency-dependent nonlinear phase distortion kernel function. ф (u, v) is a nonlinear function of the frequency variable.

Introduction to PST kernel functions

Although arbitrary phase kernel functions can be considered in PST, according to the results of[32], the derivative of the kernel phase function $\varphi(u, v)$, i.e., the group delay, is a linear or sublinear function of the frequency variable, and a simple example of such a phase kernel function is the "S" type of arctangent function. For simplicity, if this phase twisting operation is isotropic in the frequency domain plane, the degree of twisting is related only to the polar radius r in the polar coordinate system of the o-UV frequency plane but not to the polar angle θ; i.e., it is assumed that the prototypical phase-stretching kernel function of the PST is circularly symmetric with respect to the frequency variable:

$$\varphi'(u, v) = \varphi'_{polar}(r, \theta) = \varphi'_{polar}(r) \tag{2}$$

where r is the polar diameter in the frequency plane o-uv polar coordinate system and θ is the polar angle, which is related to the uv frequency variable by: $r = \sqrt{u^2 + v^2}$. $u = r * \cos\theta$, $v = r * \sin\theta$, $\theta = tan^{-1}(\frac{v}{u})$. If the derivative of $\varphi'_{polar}(r)$ with respect to r is required to be an inverse tangent function of type S, we have

$$\frac{d\phi_{polar}}{dr} = tan^{-1}(r) \tag{3}$$

Note that the UV-frequency plane of the image after Fourier transform is a finite region, so it can be solved according to Eq. (3)$\varphi_{polar}(r)$:

$$\varphi_{polar}(r) = \int_0^r tan^{-1}(x)dx = x\tan^{-1}(x)\,|_0^r - \int_0^x \frac{x}{1 + x^2}dx = r\tan^{-1}r - \frac{1}{2}\ln(1 + r^2) \tag{4}$$

Normalizing the phase function in Eq. (4) givesφ_N:

$$\varphi_N(r) = \frac{r\tan^{-1}r - 1/2 \cdot \ln(1 + r^2)}{r_{max}\tan^{-1}r_{max} - 1/2 \cdot \ln(1 + r^2_{max})} \tag{5}$$

With respect to the phase function in Eq. (5), the phase stretching strength S (strength) and warping W (warp) in the nonlinear warped stretching transform are added to obtain the final PST transform with the strength parameter S and the warping parameter W with the phase translational kernel function $\varphi_N(r, W, S)$:

$$\varphi_N(r, W, S) = S \cdot \frac{Wr\tan^{-1}(W \cdot r) - 1/2 \cdot \ln(1 + (W \cdot r)^2)}{W \cdot r_{max}\tan^{-1}(W \cdot r_{max}) - 1/2 \cdot \ln(1 + (W \cdot r_{max})^2)} \tag{6}$$

where$\tan^{-1}(\cdot)$ denotes the inverse tangent function,$\ln(\cdot)$ is the natural logarithm, andr_{max} represents the maximum frequency polar path of the UV frequency plane. Equation (6) is the phase stretching kernel function used in the literature[32].

Derivation of the phase image A (m, n)

From the Eq. (6) ,we obtain:

$$\varphi_N(C, t) = C \cdot \left\{ t \cdot \tan^{-1}(t) - \frac{1}{2}\ln[1 + (t)^2] \right\} \tag{7}$$

where $t = W \cdot r, C = S/W \cdot r_{max} \cdot tan^{-1}(W.r_{max}) - 0.5ln^{[1+(W \cdot r_{max}) \cdot 2]}$.
   so $A(m, n) = \angle IFFT2\{e^{j\varphi_N(C,t)} \cdot FFT2[B(m, n)]\}$,
$A(m, n) = \angle IFFT2\{e^{j\varphi_N(S,W,r)} \cdot FFT2[B(m, n)]\}$,
   Let $D(m, n) = e^{j\varphi_N(C,t)}$ (8)
   So $A(m, n) = \angle IFFT2\{D(m, n) \cdot FFT2[B(m, n)]\}$ (9)

$$A(m, n) = \angle IFFT2\{FFT2[E(m, n)] \cdot FFT2[B(m, n)]\} \tag{10}$$

According to the convolution theorem, it is clear that.
   $E(x, y) * B(x, y) \Leftrightarrow FFT2[E(m, n)] \cdot FFT2[B(m, n)]$
   so we obtain

$$E(x, y) * B(x, y) = IFFT2\{FFT2[E(m, n)] \cdot FFT2[B(m, n)]\} \tag{11}$$

$$A(m, n) = \angle E(x, y) * B(x, y) \tag{12}$$

where B(x, y) is the input image and E(x, y) is the spatial matrix corresponding to the computed phase stretching kernel.

The above results show that taking an angle after the phase stretching transformation is equivalent to "taking" an angle by convolving the spatial domain "complex matrix" corresponding to the "complex stretching kernel" in the frequency domain with the image. The gradient of the input image is "reflected" into the angle image formed after "taking" the angle through convolution, so the theoretical basis for the "angle image" after phase stretching to realize the function of edge extraction may lie in this. Therefore, the theoretical basis of the edge extraction function can be realized by the phase stretching of the "angle image". The different selections of parameters W and S may cause the gradient of the image to be reflected in the angle image in different proportions, so the edge detection effect is also different.

Theoretical derivation and analysis of pure amplitude stretching kernel functions

From Eq. (1) and the principle that the angular information of the phase image A(m, n) of the PST of 2.3 can be used for edge extraction, we can see:

$$A(m,n) = \angle IFFT2\{\tilde{K}(u,v) \cdot FFT2[B(m,n)]\} \tag{13}$$

can be simplified as:

$$A(x) = \angle IFFT\{\tilde{K}(u) \cdot FFT[B(x)]\} \tag{14}$$

$$\tilde{B}(u) = \int B(x)e^{-jux} \cdot dx \tag{15}$$

$$B(x) = \int \tilde{B}(u)e^{jux} \cdot du \tag{16}$$

$$A(x) = \angle IFFT\{\tilde{K}(u) \cdot \tilde{B}(u)\} \tag{17}$$

Set $\tilde{K}(u) = (1 + u)$ (18)
   So $A(x) = \angle IFFT\{(1 + u) \cdot FFT[B(x)]\}$ (19)

$$A(x) = \angle IFFT\{(1 + u) \cdot \tilde{B}(u)\} \tag{20}$$

$$A(x) = \angle IFFT\{\tilde{B}(u) + u \cdot \tilde{B}(u)\} \tag{21}$$

$$A(x) = \angle\{IFFT[\tilde{B}(u)] + IFFT[u \cdot \tilde{B}(u)]\} \tag{22}$$

$$A(x) = \angle\{B(x) - jB'(x)\} \tag{23}$$

The pure amplitude-transformed angular image $A(x) = -\arctan\frac{B'(x)}{B(x)} \approx -\frac{B'(x)}{B(x)}$ is obtained. For the pure amplitude stretching transform, in 2D image processing, there is:$\tilde{K}(u,v) = (1 + \sqrt{u^2 + v^2})$, if it is expressed in polar coordinates:

$$\tilde{K}(u,v) = \tilde{K}(r\cos\theta, r\sin\theta) = (1 + r) \triangleq \tilde{K}(r) \tag{24}$$

Normalization:
   $\tilde{K}_N(r) = \frac{(1+r)}{1+r\max}, \tilde{K}_N(r,S,W) = S \cdot \frac{(1+W \cdot r)}{1+W \cdot r\max}$ (25)

Therefore, the pure amplitude stretching kernel function with strength parameter S and twist parameter W is obtained:

$$\tilde{K}_N(r,S,W) = S \cdot \frac{(1 + W \cdot r)}{1 + W \cdot r\max} \tag{26}$$

Corresponding angular images:

$$A(m,n) = \angle IFFT2\{\tilde{K}(u,v) \cdot FFT2[B(m,n)]\} \tag{27}$$

$$A(m,n) = \angle IFFT2\{\tilde{K}_N(r,W,S) \cdot \tilde{B}(u,v)\} \tag{28}$$

If $\tilde{K}(u) = (1 + u^2)$ (29)

The improved phase filter stretching kernel function with stretching strength parameter S and twisting parameter W is obtained:

$$\tilde{K}_N(r,S,W) = S \cdot \frac{(1 + W \cdot r^2)}{(1 + W \cdot r_{\max}^2)} \tag{30}$$

Experimental analysis of the kernel function for pure amplitude stretching transformations

The u of Eq. (16) can be transformed to the kth power of u, which can be expressed as follows: $\tilde{K}(u) = (1 + u^k)$. For different values of K and the tensile strength parameter S, the experimental results obtained are different, and the specific effects are shown in Fig. 2.

From the result of the 7th power stretching, it can also be seen that the higher the power of the pure amplitude stretching is, the better the detection of the "harder bending" edges. As shown in the figure, the curved edges curved edges are extracted very finely and perfectly. This is because the higher the power is, the more the higher-order derivatives in the image can be characterized. The higher-order derivatives of a function are the curved information, singular information, and higher-order nonsmooth information in the function, but they also produce more fine noise. For this reason, the stock amplitude stretching transform has to choose the appropriate power parameter K and stretching strength parameter S in the specific target segmentation algorithm.

1. **Algorithmic flow and experimental analysis of pure amplitude stretching transform for segmentation of power lines in aerial photography**.

The flowchart of the algorithm

The main steps of aerial power line segmentation include grayscale processing, filtering processing, pure amplitude stretch transformation edge extraction, and KTV processing. The specific algorithm flow chart is shown in Fig. 3:
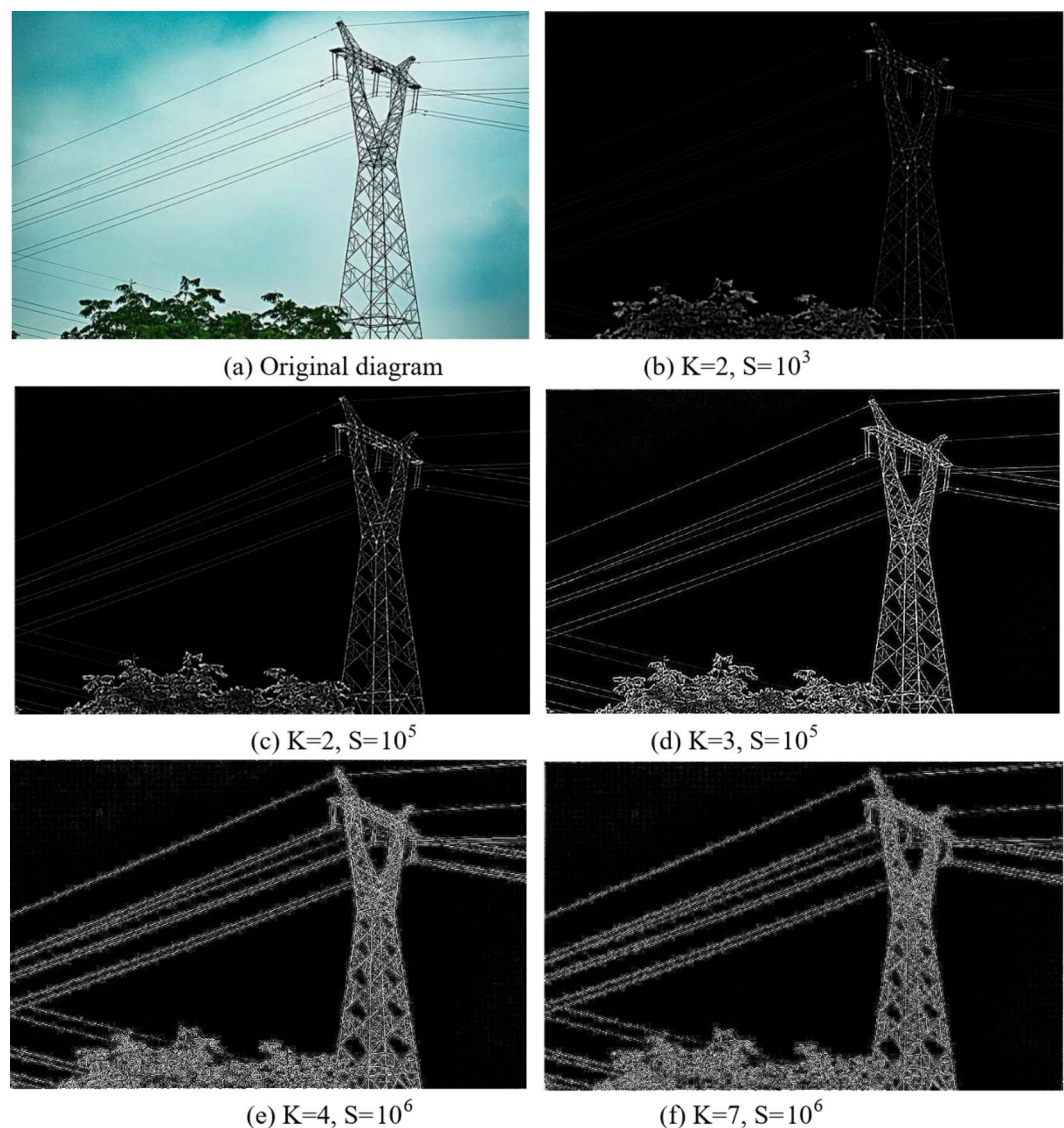

(a) Original diagram

(b) K=2, S=10^3

(c) K=2, S=10^5

(d) K=3, S=10^5

(e) K=4, S=10^6

(f) K=7, S=10^6

Fig. 2. Experimental results corresponding to different values of K and tensile strength S.
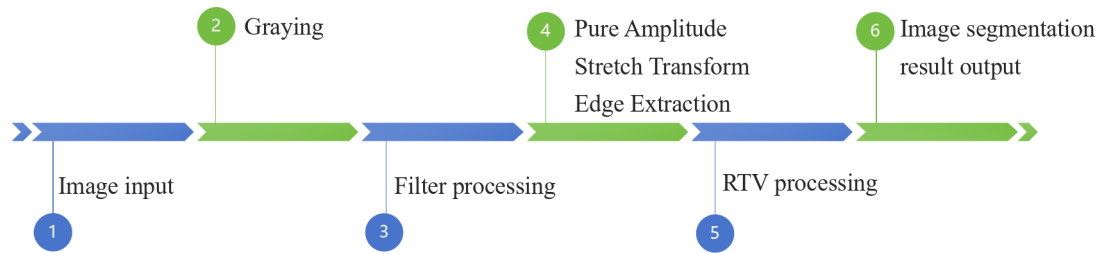
Fig. 3. Flowchart of the power line segmentation algorithm for aerial photography.

Relative total variation (RTV)

The purpose of image sharpening is to highlight meaningful large-scale structural features, such as contour edges and structural details, and to suppress unimportant fine-grained textures, such as irregular, cluttered, and recurring fine-grained patterns. To achieve the sharpened image enhancement effect, the extracted and enhanced edge and important structure images need to be superimposed with the original image so that the resulting image highlights the edges and important structural features, which is conducive to subsequent specific applications. However, the superimposed image will exhibit noise and edge burrs, which are caused by insufficient fine texture suppression. For this reason, this paper utilizes relative total variation RTV to implement postprocessing on the superimposed image, and good results are obtained.

The relative total variance RTV can measure and clearly distinguish important structures from fine textures well because of its metric criterion on the relative total variance, which is determined by two sets of metrics, the pixel-by-pixel windowed total variance measures $D_x(p)$, $D_y(p)$ and the windowed intrinsic variance measures $L_x(p)$, $L_y(p)$[36]:

$$D_x(p) = \sum_{q \in R(p)} g_{p,q} \cdot | (\partial_x S)_q |, \quad D_y(p) = \sum_{q \in R(p)} g_{p,q} \cdot | (\partial_y S)_q | \tag{31}$$

where S can be tentatively considered the input image, R(p) is a rectangular local neighborhood window centered at pixel p, q is any pixel in R(p), and $g_{p,q}$ is a weighting factor defined based on the spatial similarity relationship.

Drawing on the ability of relative total variation (RTV) to optimize the separation of major structures from weak and fine textures, the RTV step is introduced into the pure amplitude stretch transform image feature extraction algorithm in this paper, which allows the sharpened image to better highlight the major edge contours and structural textures and suppress the weak and fine textures.

Comparison of experimental effects

The image segmentation effect graphs obtained from this paper's algorithm and the algorithms based on the PST, YOLOv5, and R-CNN algorithms are analyzed against each other.

The main technical indicators involved in power line segmentation based on YOLOv5 are as follows: The dataset primarily consists of the Aerial Power Line Dataset, the TTPLA dataset, and data collected by the research team, covering power line images from different regions, lighting conditions, weather patterns, and shooting angles. The dataset records the target category and bounding box information. A diverse and comprehensive dataset helps reduce overfitting and improves the model's generalization ability in practical applications. Data augmentation plays a crucial role in aerial power line detection, effectively enhancing the model's robustness. By adjusting brightness, contrast, and saturation, different lighting conditions are simulated to achieve color enhancement. Gaussian noise is added to improve the model's performance on low-quality images. The choice of hyperparameters directly affects the training results of YOLOv5. The initial learning rate can be set to 0.001, with the Cosine Annealing learning rate scheduler gradually reducing it. The training cycle is set to 100 epochs, with early termination if necessary to save computing resources. The optimizer can be selected as SGD, with a momentum of 0.937 and weight decay of 0.0005. Proper adjustment of the IoU threshold (0.5) and enabling gradient accumulation can improve accuracy and prevent running out of video memory.

The main technical indicators involved in power line segmentation based on R-CNN are as follows: The dataset used is essentially the same as the one used for power line segmentation based on YOLOv5. For the team's aerial images, the rotation operation can be set between −15° and 15° to simulate changes in the angle of the power line under different flight directions. The zoom range is set to 0.8 to 1.2 to accommodate different shooting distances. Translation and cropping randomly alter the image area or target position to present various perspectives. In terms of color enhancement, brightness, contrast, and saturation can be flexibly adjusted. Through these augmentation methods, the diversity of the training data is greatly enriched, enabling accurate identification and segmentation of both conventional and special aerial power line images. The hyperparameter settings for R-CNN are as follows: The initial learning rate is usually set to 0.001, with a learning rate decay method applied to optimize the convergence process. The batch size is set to 16 or 32, and the training cycle is typically set to 50 to 100 epochs, adjusted according to the dataset and results. The optimizer is usually SGD or Adam, with a momentum of 0.9, and the balance coefficient between bounding box regression loss and classification loss is typically 0.5. Reasonable hyperparameter settings can help improve the convergence speed and detection accuracy of the model.

The resulting comparative results are shown in Figs. 4, 5, 6, 7 and 8.

z`Power line segmentation effect analysis

To effectively evaluate the segmentation performance of the algorithm on the image, three commonly used evaluation metrics in image segmentation, namely, image pixel accuracy precision, mean pixel accuracy, and mean intersection and merger ratio[37], are selected for analysis.

The image pixel accuracy precision reflects the ratio of the number of pixels accurately predicted to the total number of pixels, which is calculated as follows:

$$P_{PA} = \frac{\sum_{i=0}^{N} n_{ii}}{\sum_{i=0}^{N} \sum_{j=0}^{N} n_{ij}} \tag{32}$$

where n_{ii} is the number of true pixels of class i, and n_{ij} is the number of pixels of class i misidentified as class j. The number of pixels of class i misidentified as class j is given in the table below.

The mean pixel accuracy reflects the proportion of pixels correctly predicted by the category, for which the average value is taken, which is calculated as follows:

$$M_{MPA} = \frac{1}{N+1} \sum_{i=0}^{N} \frac{n_{ii}}{\sum_{j=0}^{N} n_{ij}} \tag{33}$$

The average intersection ratio reflects the ratio of the correctly predicted pixel area to all the predicted pixel areas, and is calculated by the following formula:

$$M_{MIoU} = \frac{1}{N+1} \sum_{i=0}^{N} \frac{n_{ii}}{\sum_{j=0}^{N} n_{ij} + \sum_{j=0}^{N} (n_{ij} - n_{jj})} \tag{34}$$

The values of P_{PA}, M_{MPA} and M_{MIoU} for the four image segmentation methods of PST, YOLOv5, R-CNN, and the proposed algorithm for Figs. 4, 5, 6 and 7 are shown in Tables 1, 2 and 3, respectively.

From the results of Tables 1, 2 and 3; Figs. 8 and 9, and Fig. 10, it can be seen that the values of the evaluation indicators P_{PA}, M_{MPA} and M_{MIoU} of the image segmentation results of this paper's algorithm are higher than the values of the image segmentation algorithms based on the PST, R-CNN, etc., which fully demonstrates that this paper's algorithm is more effective in suppressing noise. When segmenting the aerial power line image containing noise and unclear edges, the chances of over-segmentation, under-segmentation, and incomplete

(a) Original Figure 1

(b) Original Figure 2

(c) Original Figure 3

(d) Original Figure 4

Fig. 4. Original diagram.

(a) PST processing effect: S=1.48,W=10.24   (b) YOLOv5 processing effect

(c) R-CNN processing effect graph   (d) Processing effect graph of the proposed: $K=3$, $S=10^5$
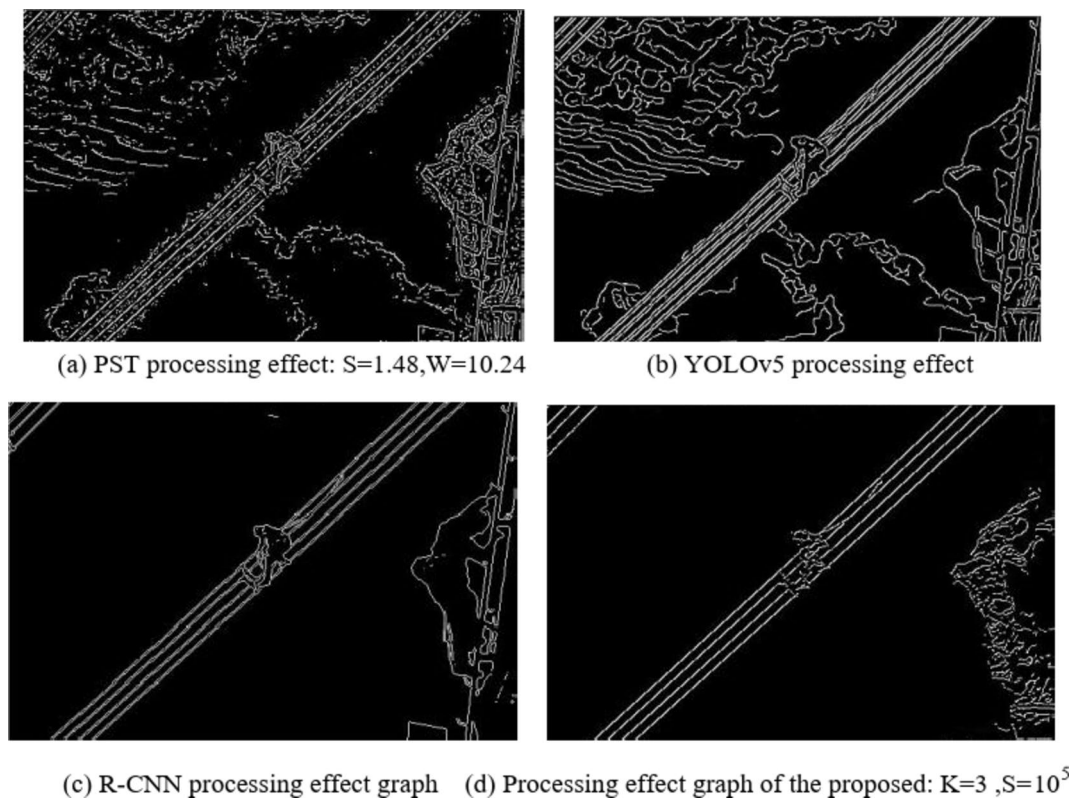
Fig. 5. Comparison of the effects of various algorithms in Fig. 4(a).

segmentation are reduced, and due to the enhanced image edges, the segmentation results are closer to the ideal contour, and better segmentation results are obtained. PST-based image segmentation is better for the edges of the image, but the effect of suppressing the edge noise of the power line is not good. YOLOv5-based image segmentation under the influence of noise will appear more over-segmentation and other inaccurate segmentation results; R-CNN-based image segmentation appears to be insensitive to the edges of the image, and there are too many under-segmentation and other erroneous results.

To further illustrate the advantages of the algorithm in this paper, 100 power line detection maps were taken aerially with a UAV, and four image segmentation algorithms, such as PST and R-CNN, were used to segment the 100 images, and the values of the evaluation metrics P_{PA}, M_{MPA} and M_{MIoU} were obtained from as shown in Figs. 9 and 10, and Fig. 11.

The results of comparing the average P_{PA}, M_{MPA} and M_{MIoU} values of this paper's algorithm with those of the PST, YOLOv5, and R-CNN algorithms are shown in Tables 4 and 5, and Table 6 respectively.

The results of comparing the average time lag of the proposed algorithm with those of the PST, YOLOv5, and R-CNN algorithms are shown in Table 7.

According to the average values of the objective evaluation indicators P_{PA}, M_{MPA} and M_{MIoU}, the power line segmentation results of the proposed algorithm are significantly better than those of the PST, YOLOv5, and R-CNN algorithms. Compared with those of the image separation algorithms based on the PST and other image segmentation algorithms, the average values of the evaluation indicators of P_{PA}, M_{MPA} and M_{MIoU} reach 0.96, 0.96 and 0.95, respectively. According to Table 7, it can be concluded that the time lag of this paper's algorithm detection is on average less than 0.2 s, and the results can be applied to aerial power line inspection and detection. The effectiveness and stability of this paper's algorithm are further verified.

Summary

This paper first reviews the application of mainstream image segmentation techniques in UAV aerial photography power line segmentation and summarizes the characteristics of different methods and relevant experimental results. Second, the theoretical derivation process based on the pure amplitude stretching transform kernel function and its application in aerial photography power line segmentation are described. Finally, the results are compared and analyzed with the experimental results. The results show that the algorithm in this paper can extract power lines more accurately and in real time under a variety of backgrounds, which has a certain value for engineering applications and has a certain reference significance for future research on power line extraction. However, further research is needed in the following areas:

In UAV inspection of transmission lines, in addition to visible light images, there are thermal images, ultraviolet images and LiDAR images, etc. Image data from different data sources suit power line characteristics in various scenarios. Most power line inspection methods are only for visible light images. Effective use of multi-
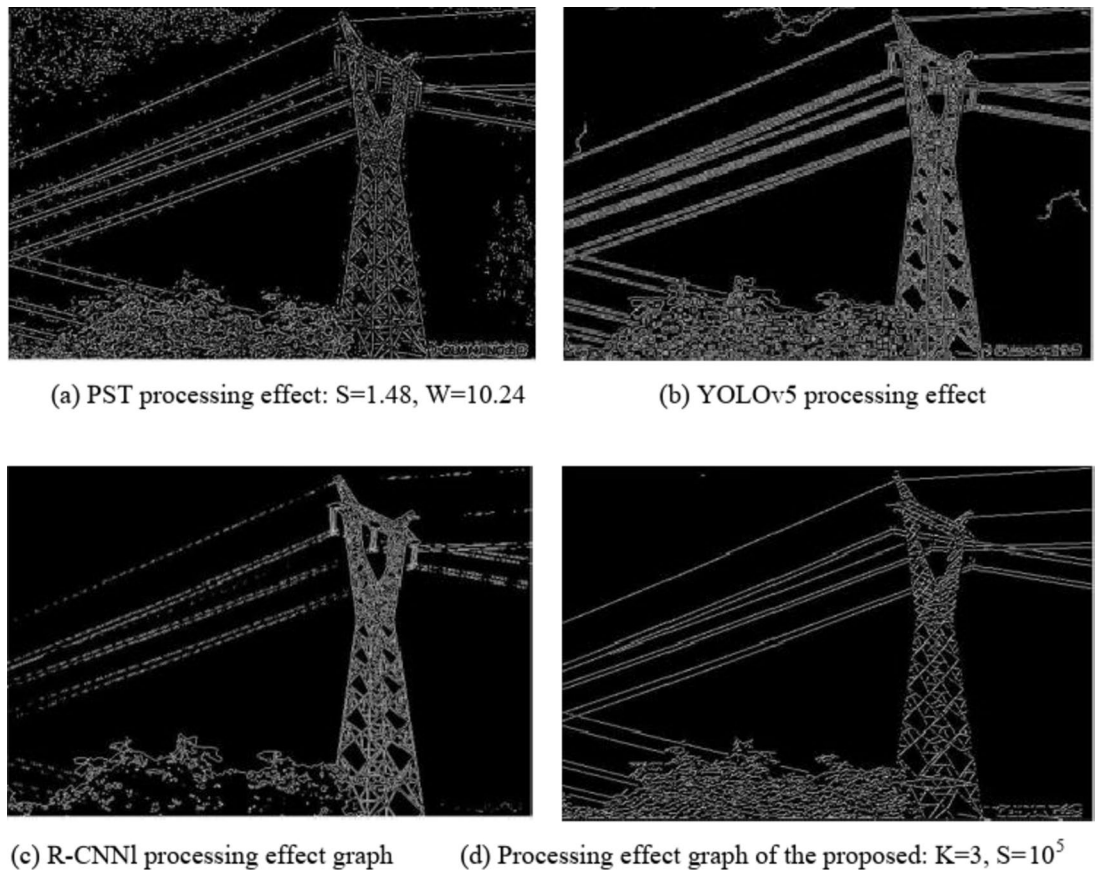
(a) PST processing effect: S=1.48, W=10.24   (b) YOLOv5 processing effect

(c) R-CNNl processing effect graph   (d) Processing effect graph of the proposed: K=3, S=10^5

Fig. 6. Comparison of the effects of various algorithms in Fig. 4(b).

source image data fusion can realize power line state recognition under multi-application scenarios. How to effectively utilize multi-source image data fusion to realize power line status recognition in multiple application scenarios.

Combination of traditional image processing methods and deep learning methods. Accurate extraction of power line features is one of the key tasks for UAV inspection of transmission lines, and power lines in aerial images lack obvious linear features, and it is usually difficult for power lines to be separated from the background in complex backgrounds. How can traditional image processing methods be combined with deep learning methods to effectively extract power lines in complex backgrounds? One approach is to first use deep learning methods to detect the power line region and then apply traditional image processing techniques to extract the power line features.

For power line image detection in diverse environments such as urban, mountain, water, and forest settings, different deep learning models are typically applied to account for the specific challenges posed by each environment. Since a single model often fails to perform well across all these scenarios, integrating multiple models can significantly enhance the ability to detect power lines in various conditions.
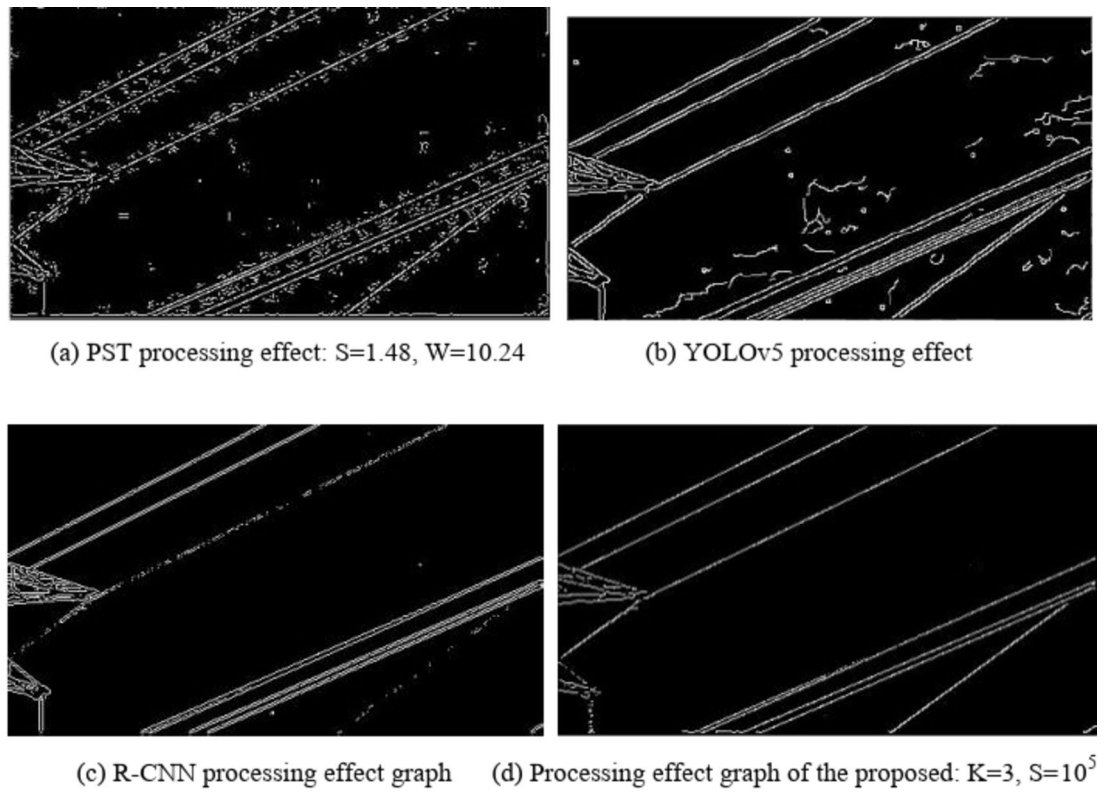
(a) PST processing effect: S=1.48, W=10.24   (b) YOLOv5 processing effect

(c) R-CNN processing effect graph   (d) Processing effect graph of the proposed: K=3, S=10^5

Fig. 7. Comparison of the effects of the various algorithms in Fig. 4(c).

(a) PST processing effect: S=1.48, W=10.24   (b) YOLOv5 processing effect

(c) R-CNN processing effect graph   (d) Processing effect graph of the proposed: K=3, S=10^5
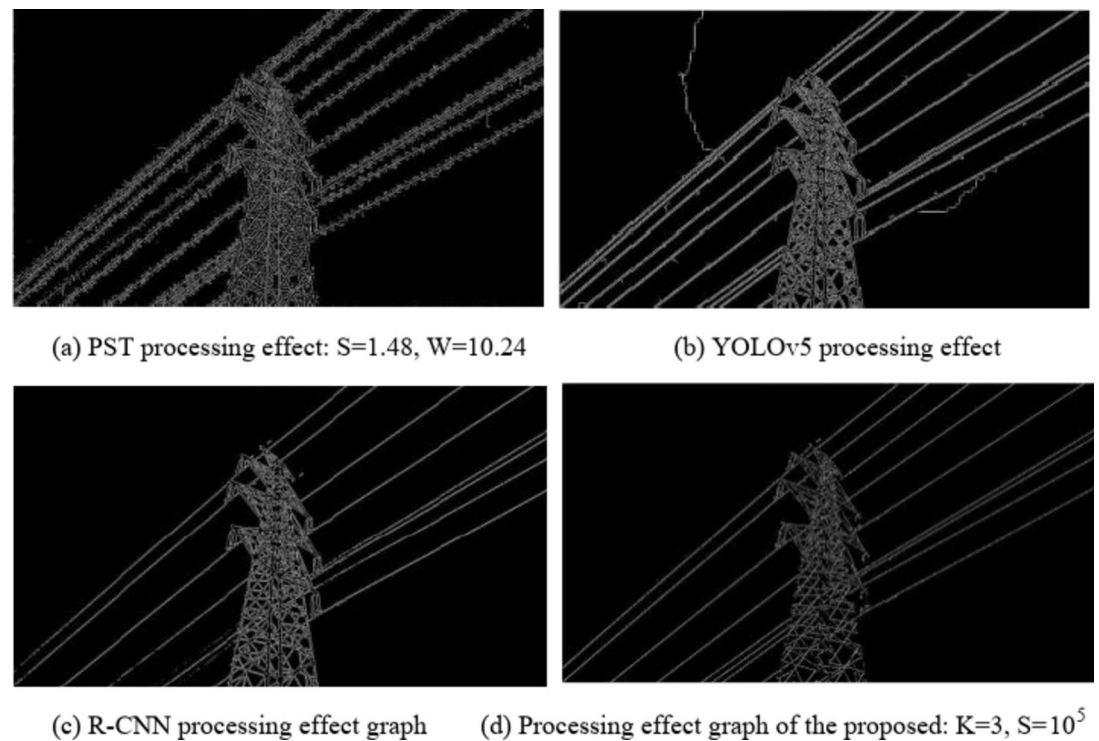
Fig. 8. Comparison of the effects of the various algorithms in Fig. 4(d).

	PST	YOLOv5	R-CNN	The proposed method
Figure 4	0.86	0.91	0.92	0.97
Figure 5	0.88	0.89	0.93	0.95
Figure 6	0.85	0.92	0.91	0.96
Figure 7	0.89	0.90	0.90	0.96

Table 1. P_{PA} Valuess for four different segmentation algorithms.

	PST	YOLOv5	R-CNN	The proposed method
Figure 4	0.87	0.89	0.92	0.96
Figure 5	0.89	0.91	0.90	0.97
Figure 6	0.84	0.88	0.88	0.95
Figure 7	0.87	0.90	0.90	0.96

Table 2. M_{MPA} Valuess for four different segmentation algorithms.

	PST	YOLOv5	R-CNN	The proposed method
Figure 4	0.86	0.89	0.88	0.95
Figure 5	0.89	0.91	0.89	0.96
Figure 6	0.85	0.90	0.90	0.98
Figure 7	0.88	0.92	0.92	0.94

Table 3. M_{MIoU} Valuess for four different segmentation algorithms.
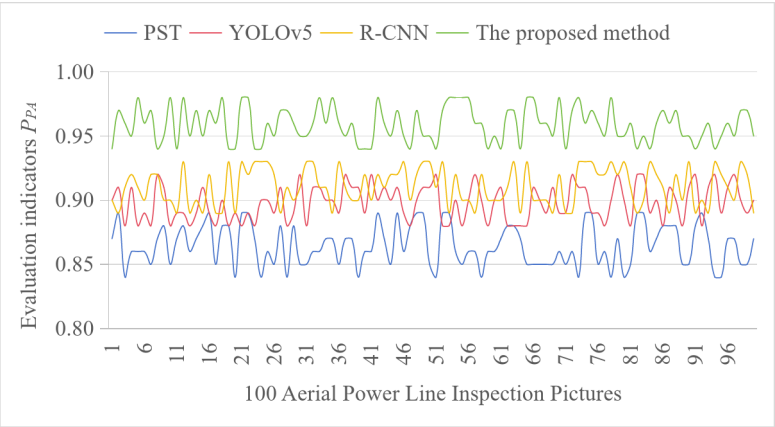
Fig. 9. Evaluation indicators P_{PA} Values.
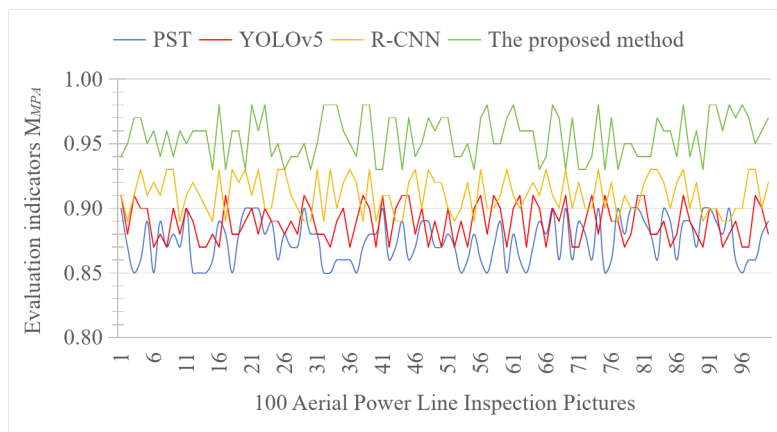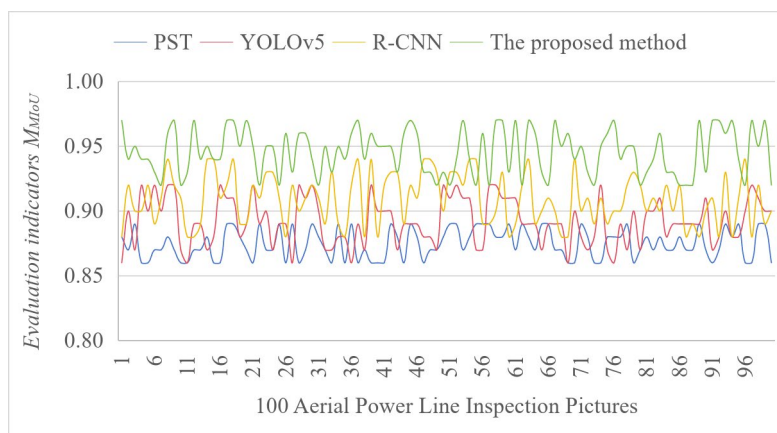
Fig. 10. Evaluation indicators M_{MPA} Values.

Fig. 11. Evaluation indicators M_{MIoU} Values.

arithmetic	Average P_{PA} value
PST	0.86
YOLOv5	0.9
R-CNN	0.91
The proposed method	0.96

Table 4. Average P_{PA} values for different algorithms.

arithmetic	Average M_{MPA} value
PST	0.88
YOLOv5	0.89
R-CNN	0.91
The proposed method	0.96

Table 5. Average M_{MPA} values for different algorithms.

arithmetic	Average M_{MIoU} value
PST	0.87
YOLOv5	0.89
R-CNN	0.91
The proposed method	0.95

Table 6. Average M_{MIoU} values for different algorithms.

arithmetic	Average time lag (seconds)
PST	0.6
YOLOv5	0.9
R-CNN	1.1
The proposed method	0.15

Table 7. Average time lag of different algorithms.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.https://github.com/lanbaoshi110/papers/tree/master.

References
1. Chuanyang, L. I. U., Yiquan, W. U. & Jingjing, L. I. U. Research progress of power line detection methods in UAV aerial images[J]. *Chin. J. Image Graphics.* **28** (10), 3025–3048 (2023).
2. Li, Z. et al. Towards automatic power line detection for a UAV surveillance system using pulse coupled neural filter and an improved Hough transform[J]. *Mach. Vis. Appl.* **21**, 677–686 (2010).
3. Zhang, J. et al. High speed automatic power line detection and tracking for a UAV-based inspection[C]//2012 International Conference on Industrial Control and Electronics Engineering. IEEE, : 266–269. (2012).
4. Chenxin, Y. U. A. N. et al. Power line extraction based on improved Hough transform[J]. *Beijing Surveying Mapp.* **32** (6), 730–733 (2018).
5. Cao Haopeng, Z. et al. Power line detection based on Hough transform and overall least squares[J]. *Collection*, 10. (2018).
6. Nasseri, M. H. et al. Power line detection and tracking using hough transform and particle filter[C]//2018 6th RSI International Conference on Robotics and Mechatronics (IcRoM). ieee, : 130–134. (2018).
7. Lipo, Z. H. A. O. et al. Algorithm for real-time detection and recognition of high-voltage lines for patrol drones[J]. *Small Microcomputer Syst.* **33** (4), 882–886 (2012).
8. Tian, F., Wang, Y. & Zhu, L. Power line recognition and tracking method for UAVs inspection[C]//2015 IEEE international conference on information and automation. ieee, : 2136–2141. (2015).
9. Fu, L., Lu, S. & Engineering, A. Obstacle detection algorithms for aviation[C]//2011 IEEE International Conference on Computer Science and IEEE,. 4: 710–714. (2011).
10. Wu, D. L. et al. A vision-based system for power transmission facilities detection[J]. *Appl. Mech. Mater.* **423**, 2547–2554 (2013).
11. Luo, X. et al. Object-aware power line detection using color and near-infrared images[J]. *IEEE Trans. Aerosp. Electron. Syst.* **50** (2), 1374–1389 (2014).
12. Zhang, J. et al. Pylon line spatial correlation assisted transmission line detection[J]. *IEEE Trans. Aerosp. Electron. Syst.* **50** (4), 2890–2905 (2014).
13. Teplyakov, L., Erlygin, L., Shvets, E. & Lsdnet Trainable modification of lsd algorithm for real-time line segment detection[J]. *IEEE Access.* **10**, 45256–45265 (2022).
14. Zhang, Y. et al. Research on automatic extraction of railway catenary power lines under complex background based on RBCT algorithm[J]. *High. Volt Eng.* **48**, 2234–2243 (2022).
15. Pengyue, L. I. et al. Research on the algorithm of linear class power line detection in UAV imagery[J]. *Electron. Meas. Technol.* **42** (19), 148–153 (2019).
16. Le, Z. H. A. O. et al. Power line extraction algorithm based on local context information[J]. *High. Voltage Technol.* **47** (7), 2553–2563 (2021).
17. Gubbi, J., Varghese, A. & Balamuralidhar, P. A new deep learning architecture for detection of long linear infrastructure[C]//2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). ieee, : 207–210. (2017).
18. Yetgin, Ö. E., Benligiray, B. & Gerek, Ö. N. Power line recognition from aerial images with deep learning[J]. *IEEE Trans. Aerosp. Electron. Syst.* **55** (5), 2241–2252 (2018).
19. Zhao, L. I. U. et al. Improved faster R-CNN-based high voltage cable target detection method[J]. *J. Intell. Syst.* **14** (4), 627–634 (2019).
20. Zhenbang, J. I. A. N. G. & Kuansheng, Z. O. U. Power line extraction algorithm based on stage attention mechanism[J]. *J. Data Acquisition Processing/Shu Ju Cai Ji Yu Chu Li*, **36**(4). (2021).
21. Wang Xuewen. *Research on image-based Semantic Segmentation Technology for Transmission Lines [D]* (Zhejiang University, 2019).
22. Madaan, R., Maturana, D. & Scherer, S. Wire detection using synthetic data and dilated convolutional networks for unmanned aerial vehicles[C]//2017 IEEE/ RSJ International Conference on Intelligent Robots and Systems (IROS). ieee, : 3487–3494. (2017).
23. Nguyen, V. N., Jenssen, R. & Roverso, D. LS-Net: fast single-shot line-segment detector[J]. *Mach. Vis. Appl.* **32** (1), 12 (2021).
24. Chen, M. et al. Small and strong: power line segmentation network in real time based on self-supervised learning[J]. Proc. CSEE, 42: 1365–1375. (2022).
25. Choi, H. et al. Weakly supervised power line detection algorithm using a recursive noisy label update with refined broken line segments[J]. *Expert Syst. Appl.* **165**, 113895 (2021).

26. Yan, J. et al. A real-time strand breakage detection method for Power Line inspection with UAVs[J]. *Drones* **7** (9), 574 (2023).

27. Hu, C. et al. Research on an improved detection algorithm based on yolov5s for power line self-exploding insulators[J]. *Electronics* **12** (17), 3675 (2023).

28. Zhao, W., Dong, Q. & Zuo, Z. A method combining line detection and semantic segmentation for power line extraction from unmanned aerial vehicle images[J]. *Remote Sens.* **14** (6), 1367 (2022).

29. Son, H. S. et al. Real-time power line detection for safe flight of agricultural spraying drones using embedded systems and deep learning[J]. *Ieee Access.* **10**, 54947–54956 (2022).

30. Wang, D. et al. Research on image recognition algorithm technology for power line detection[C]//Journal of Physics: Conference Series. IOP Publishing, 1732(1): 012082. (2021).

31. Hota, M. & Kumar, U. Power Lines Detection and Segmentation In Multi-Spectral Uav Images Using Convolutional Neural Network[C]//2020 IEEE India Geoscience and Remote Sensing Symposium (InGARSS). IEEE, : 154–157. (2020).

32. Mohammad, H., Asghari & Jalali, B. Edge Detection in Digital Images Using Dispersive Phase Stretch Transform, International Journal of Biomedical Imaging, Volume Article ID 687819, pp.1–6. (2015).

33. Suthar Madhuri, Asghari, Hossein, Jalali, Bahram, Feature Enhancement in Visually Impaired Images[J], IEEE Access, February 14,2018.

34. Mohammad, H., Asghari & Jalali, B. Discrete Anamorphic transform for Image Compression[J]. *IEEE. Signal. Process. Lett.*, **21**, 7, JULY. pp.829–833,2014.

35. Tali et al. Phase stretch transform for super-resolution localization microscopy[J]. *BIOMEDICAL Opt. EXPRESS.* **7** (10), 4198–4209 (2016).

36. Li, X. et al. [J]. Structure extraction from texture via relative total variation. *ACM Trans. Graph* 31 6, Article 139 (November 2012), 10 pages. DOI = 10.1145/2366145.2366158.

37. Yanchao, D. U. & HUANG Shan. Lightweight remote sensing image segmentation based on branch merging strategy[J]. *Comput. Eng. Des.* **43** (11), 3210–3216 (2022).

Author contributions

Pengfei Xu wrote the main content of the manuscript, completed the collection of the main experimental image material, and performed the complete testing and analysis of the experimental procedure.Nor Anis Asma Sulaiman reviewed the framework of the manuscript and provided constructive comments on the abstract and conclusions.Yafei Ding and Jiangwei Zhao collected some of the experimental images, organized the experimental data, and plotted some of the experimental results.All authors reviewed the manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.