

RESEARCH

Open Access



A robot vision navigation method using deep learning in edge computing environment

Jing Li*, Jialin Yin and Lin Deng

* Correspondence: lj@scetc.edu.cn
Department of Electrical
& Information Engineering, Sichuan
Engineering Technical College,
Deyang 618000, Sichuan, China

Abstract

In the development of modern agriculture, the intelligent use of mechanical equipment is one of the main signs for agricultural modernization. Navigation technology is the key technology for agricultural machinery to control autonomously in the operating environment, and it is a hotspot in the field of intelligent research on agricultural machinery. Facing the accuracy requirements of autonomous navigation for intelligent agricultural robots, this paper proposes a visual navigation algorithm for agricultural robots based on deep learning image understanding. The method first uses a cascaded deep convolutional network and hybrid dilated convolution fusion method to process images collected by a vision system. Then, it extracts the route of processed images based on the improved Hough transform algorithm. At the same time, the posture of agricultural robots is adjusted to realize autonomous navigation. Finally, our proposed method is verified by using non-interference experimental scenes and noisy experimental scenes. Experimental results show that the method can perform autonomous navigation in complex and noisy environments and has good practicability and applicability.

Keywords: Edge computing, Agricultural robot, Cascaded deep convolutional network, Hybrid dilated convolutional network, Improved Hough transform, Farmland image segmentation, Path extraction, Autonomous navigation

1 Introduction

The development direction of world agricultural production in the twenty-first century is shifting from traditional agriculture to modern agriculture [1, 2]. Agriculture is an important basic industry to protect the national economy. The maximum utilization of agricultural resources, maximum production, and maximum development are the keys to measuring the level of modern agriculture [3, 4]. For China, one of the checks and balances in the level of modern agricultural production is the independence and intelligence for production machinery and equipment. The development of high-level intelligent agricultural machinery is an important direction for current agricultural development [5, 6].



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

With the rapid development of electronic technology and intelligent algorithms, intelligent robots have been widely used in many fields. Its autonomy and intelligence are becoming more and more perfect. Facing the demand for efficient production in modern agriculture, intelligent robots have also attracted many attentions of agricultural researchers. As a new concept of agricultural machinery [7, 8], agricultural robots have huge economic benefits in the field of agricultural production and have broad market prospects. The timely development and development of a new generation for agricultural machinery represented by agricultural robots are of great significance for my country's transition to modern agriculture [9, 10].

At present, the existing image semantic segmentation algorithm, the network model is very complex, the parameter calculation is large, and the requirement of hardware equipment is also high. How to optimize the algorithm structure and reduce the dependence on hardware equipment is the current research focus, so as to better apply the technology in real life.

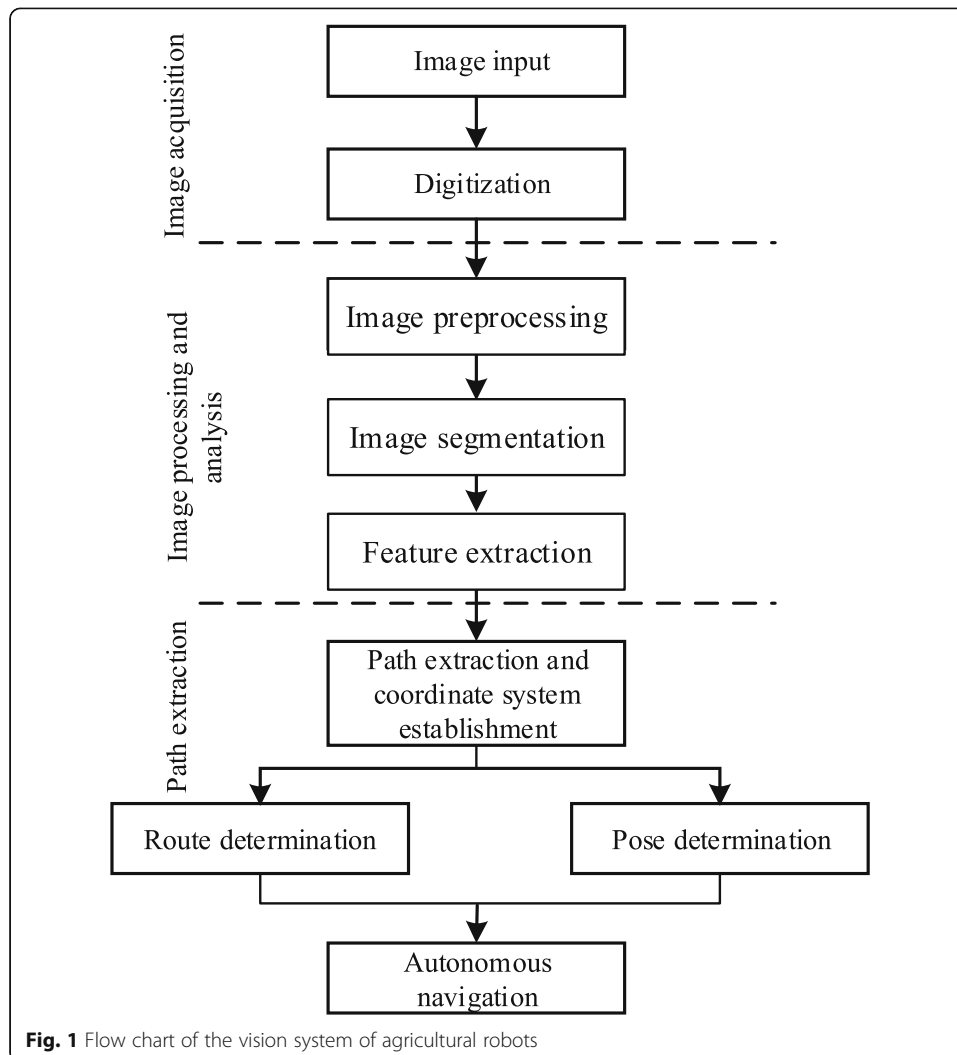
2 Related work

A visual navigation system is the core device of agricultural robots. An excellent visual navigation system can help agricultural robots to process and analyze collected images with the help of advanced intelligent algorithms or artificial intelligence algorithms. This helps robots to observe and understand the outside world and realize the intelligence and autonomy of mechanical equipment.

The robot vision system first captures a two-dimensional image of a three-dimensional external environment by an image acquisition device such as a camera. The obtained two-dimensional images are processed by intelligent algorithms to realize image segmentation, feature extraction, and other image understanding processes [11]. Finally, the symbolic description of the image itself is obtained to support agricultural robots to make decisions on the next action. The workflow is shown in Fig. 1.

Many scholars have conducted research on the visual navigation of agricultural robots. They confirmed the importance of vision systems for agricultural robots and the feasibility of practical applications [12–14]. Researchers at Kyoto University in Japan confirmed the feasibility of machine vision in agricultural mobile robot applications and extracted HIS space of images. In HIS, images were scanned with horizontal lines, and the least square method was used to identify crop spacing [15]. Han et al. used the K-means clustering algorithm to obtain crop row spacing information. And they through image comparison and evaluation judged the accuracy of image processing, in order to achieve agricultural tractor navigation [16]. Akane team discussed an image processing method that classifies collected images based on grayscale histograms. In addition, different methods were used to distinguish between traversable and non-traversable areas in the farmland to realize the navigation of agricultural vehicles [17]. Researchers such as David relied on the global positioning system and inertial navigation system. They combined a robot vision system to solve the problem of autonomous navigation for agricultural robots and realized the sustainable intensification of large-scale agriculture [18].

Due to its high reliability and detection accuracy, Hough has been used in intelligent agricultural equipment [19–21]. Chen et al. solved the problem of machine vision on the effect of multiple environmental variables on crop row recognition during the



entire growth period of lettuce and green cabbage and at the same time improves the effectiveness of the machine vision crop row recognition algorithm. This paper proposed a multi-crop row extraction algorithm based on automatic Hough transform accumulation threshold [22]; Li and other studies analyzed the principle of Hough transform proposed in image processing. They proposed to use this transformation for the processing of gravity and magnetic data. Based on the linear features contained in this identification data, it corresponded to information such as the geological body boundary and plane distribution characteristics of fault structure. The calculation and analysis of the theoretical model and actual data show that this method can extract the boundary information of gravity and magnetic data more accurately, and it had good noise robustness [23]. Olsen and Sogaard proposed a method that uses machine vision to obtain RGB three-channel images and used 2G-R-B operators to convert color images into single-channel grayscale images. Calculate the position of the center for crop gravity in the horizontal direction by analyzing images, and use the least square method to fit the spacing of crop position [24]. Qun et al. designed a greenhouse robot based on machine vision, using a watershed algorithm to segment images and convert them into a binary image. The establishment of a navigation path by Hough transform can

significantly reduce the effect of natural light and greenhouse plastic film on image segmentation in a greenhouse environment. The correct rate of road information extraction was 95.7% [25].

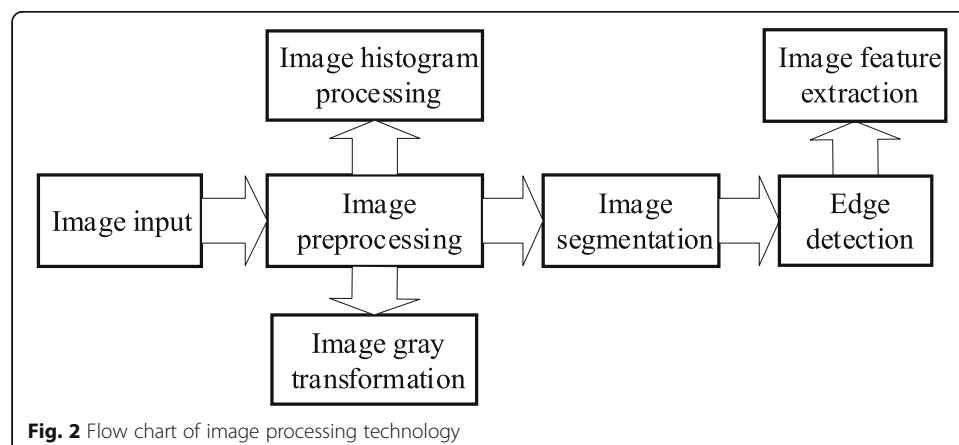
Agricultural robot research is an interdisciplinary subject, which is a comprehensive product of many fields and disciplines. The vision navigation system is just like human eyes, which is the premise of the normal and stable operation of an intelligent robot. The actual agricultural production environment is complex and diverse, so compared with other industrial robots, the accuracy of agricultural robot navigation is much higher. Therefore, it is of great significance for precision vision navigation of agricultural robots. Drawing on the existing research on autonomous navigation of crops, this paper proposes a visual navigation algorithm for agricultural robots based on deep learning image understanding. The main contributions are as follows:

- 1) Improve the Hough transform method based on subdivision algorithm, improve the calculation efficiency of the traditional Hough algorithm, and realize the effective extraction of robot path.
- 2) And the correspondence relationship between the image coordinate system and the actual scene coordinate system and the state equation are established to achieve robots' autonomous navigation posture adjustment.

The rest of this paper is organized as follows. The third section introduces vision system image processing technology, including image segmentation and edge detection technology. Section 4 introduces the technology of path extraction and pose adjustment for agricultural robots. Section 5 uses actual scenarios to verify our proposed method. Section 6 is the conclusion of this paper.

3 Image processing of farmland scenery

Efficient and good image processing is the prerequisite for agricultural robots to autonomously navigate. The main flow of image processing technology is shown in Fig. 2, which mainly includes steps such as image preprocessing, image segmentation, and feature extraction.



3.1 Image acquisition

Image acquisition is the first step in image processing. Generally, the vision system cannot directly process simulated images because collected images are simulated images. This paper uses a CCD image sensor in the robot vision system to convert analog images collected by an image acquisition device into a digital image and transmit it to the vision system computing center to ensure the goodness of acquired image attributes. That is, the position and gray scale are helpful for further research on subsequent image processing.

3.2 Image preprocessing

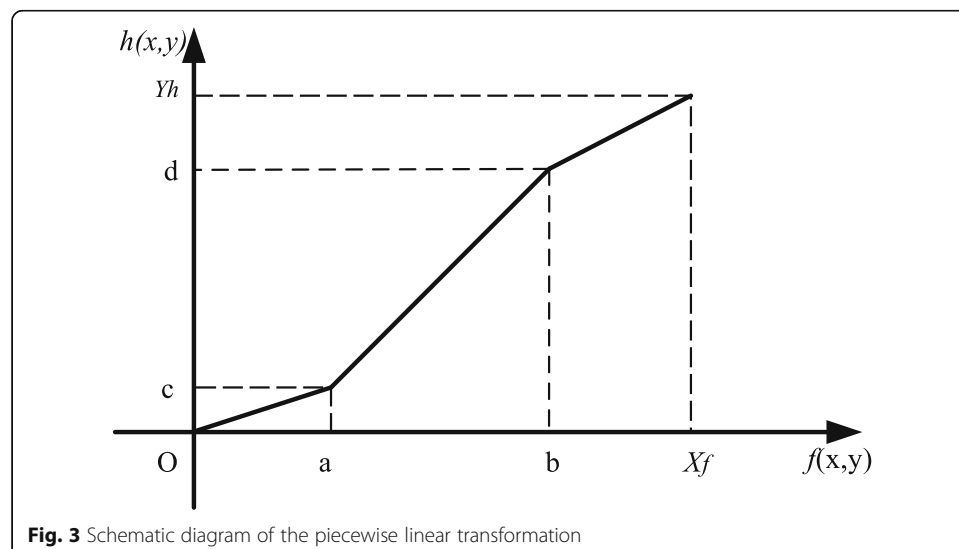
In order to provide better quality images to the vision system computing center, images are preprocessed to solve the problems of distortion and deformation caused by hardware equipment and digital-analog conversion during image acquisition and transmission.

3.2.1 Grayscale image

Image graying is an important method for image enhancement. Make targeted corrections to the pixels in images to enhance the obvious features of images. At the same time, expand the image dynamic adjustment range and contrast to make the image effect more clear and uniform.

The piecewise linear grayscale transformation is used to realize the grayscale processing of images, enhance the target grayscale interval, and suppress the non-target grayscale interval. And set the image grayscale range to $[0, X]$; the linear relationship is shown in Fig. 3.

By changing the coordinates of each inflection point and the slope of the line segment by a piecewise linear transformation, the grayscale interval can be expanded or compressed. The mathematical expression is



$$h(x, y) = \begin{cases} \frac{c}{a}f(x, y) & 0 \leq f(x, y) < a \\ \frac{(d-c)}{(b-a)}[f(x, y)-a] + c & a \leq f(x, y) < b \\ \left[\frac{(Y_h-d)}{(X_f-b)} \right] [f(x, y)-b] & b \leq f(x, y) < X \end{cases} \quad (1)$$

3.2.2 Grayscale histogram

Grayscale histogram is the simplest and most effective tool for describing grayscale values of images. It reflects the frequency of occurrence of uniform gray values, and it is the basis of image processing.

If the gray value of the gray image $h(x, y)$ is within the range of $[0, X - 1]$, the gray histogram equalization expression of image $h(x, y)$ is:

$$\eta(g_i) = \frac{n_i}{n}, \quad 0 \leq g_i \leq 1; \quad i = 0, 1, \dots, X-1 \quad (2)$$

where $\eta(g_i)$ is the probability of gray level i , g_i is the gray level of level i , n is the total number of pixels, and n_i is the number of pixels of gray level g_i .

3.3 Image segmentation

In actual farming scenes, the environment is complex and crops are diverse. It is difficult to obtain the ideal image segmentation results only by underlying feature information. It has been confirmed that deep learning technology can collect global feature information in images to obtain better segmentation results.

Based on the hybrid dilated convolution, the cascaded deep residual network is improved to complete image segmentation processing in the agricultural robot vision system. A one-to-one mapping relationship between image pixels and semantic categories is established.

The more the number of network layers in the deep convolutional neural network, the richer the level of information extraction for global feature items of images [26–28]. However, it should also be noted that with the deepening of the network layer, the gradient disappearance and network degradation caused by chain derivative in the back propagation of the network will cause the image segmentation speed and accuracy to decrease. In order to solve this problem, we add a residual structure to a deep convolutional network to increase the shortcut constant connection, which avoids the harm of segmentation processing caused by the disappearance of gradients and network degradation in deep networks. Figure 4 shows the residual structure added to the deep network.

Set the input parameter of the shallow network of the deep convolutional network to x , and the expected output value is $E(x)$. If the deep network is not improved, the input parameter x is passed to output as the initial result. The mapping function required for network learning is $F(x) = E(x) - x$, and the feature mapping is also $E(x) = F(x) + x$. After adding the residual unit and maintaining the dimension of the input and output parameter elements unchanged, the residual unit adds the parameter input elements and output elements of multiple parameter layers cascaded. Ensure that input parameters and output parameters are within a reliable range. And we, through the ReLu activation

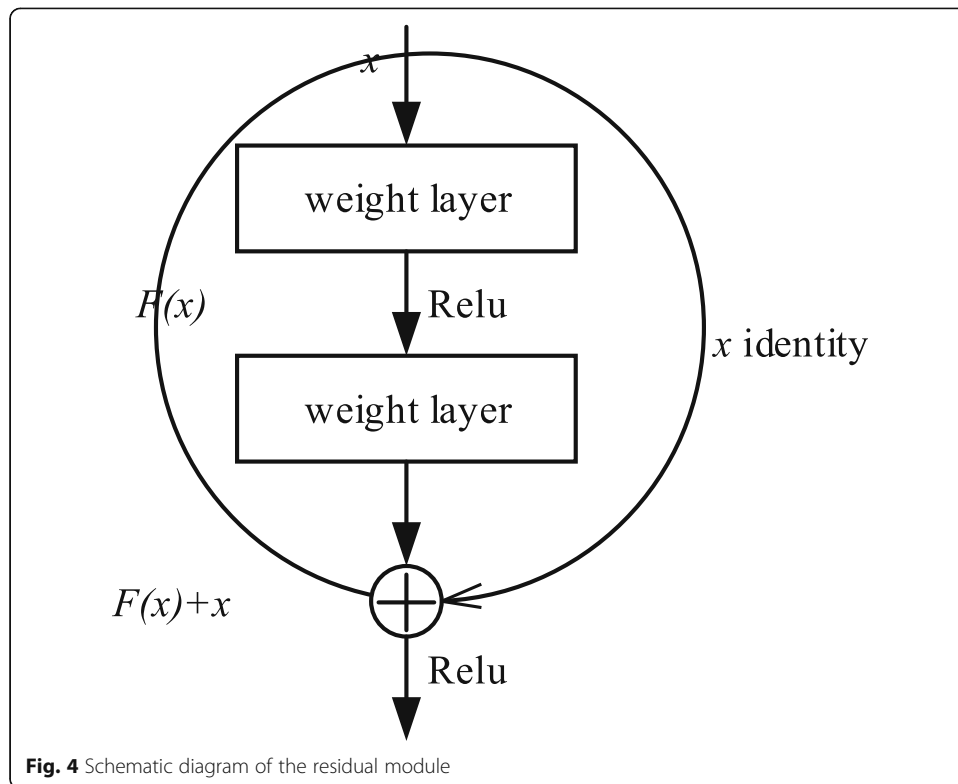


Fig. 4 Schematic diagram of the residual module

function to get the final output, reduce the impact of network gradient disappearance and mesh degradation.

We use ResNet101 as the reference network for deep networks, because of its deeper network layer core and more elaborate network structure design. The deep residual network ResNet is divided into 5 network layers. Each network layer is configured with 5 convolution modules, an average pooling layer, and a classification layer, as shown in Fig. 5. The convolution modules are conv1, conv2_x, conv3_x, conv4_x, and conv5_x. For the parameters in each convolution module, 7×7 is the size of the convolution

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7,64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10 ⁹	3.6×10 ⁹	3.8×10 ⁹	7.6×10 ⁹	11.3×10 ⁹

Fig. 5 Network configuration parameters of RESNET

kernel, and 64 is the number of channels in the convolution kernel. The brackets are a residual unit and X3 indicates that there are 3 residual units in the convolution module.

Gradient disappearance and network degradation are very serious for image segmentation results. To this end, we cascade a new convolution module conv6_x behind ResNet101 network to form a cascaded deep residual network. The network structure and network parameter settings of its convolution module are the same as conv5_x. To further extract the image features globally, consider adding the conv7_x module. However, it was found by experiments that the semantic segmentation accuracy has not been improved compared to the cascaded conv6_x module. Therefore, as shown in Fig. 6, the cascaded deep residual network is finally composed of 6 convolution modules, conv1, conv2_x, conv3_x, conv4_x, conv5_x, and conv6_x.

At the same time, using hollow convolution can increase the receptive field of the agricultural robot vision system, so as to better control image resolution [29] and fusion convolution of conv5_x and conv6_x convolution modules in ResNet network. To avoid the influence of the “grid” phenomenon in the convolutional network on segmentation results, set different void rates in the convolution module so that the receptive field can completely cover the input feature map. Taking conv5_x as an example, the module contains 3 consecutive residual units. The conv5_1 residual unit void rate is set to 1, the conv5_2 void rate is set to 2, and the conv5_3 void rate is set to 3. The conv5_x and conv5_x network structure and parameter settings are consistent. Thus, the void parameter of the residual unit in the conv6_x convolution module is set the same as conv5_x. Figure 7 is a schematic diagram of a convolution structure of a mixed cavity. The proposed model improves the cascaded deep convolution network based on the hybrid hole convolution method to solve the problem of network degradation caused by too many layers of deep network, and uses B-spline wavelet transform to detect the image edge to realize the image processing steps in the vision system, so as to provide the optimal image data support for the follow-up aircraft autonomous navigation.

3.4 Multi-resolution edge detection

The B-spline wavelet transform is used to detect the outline of the large-scale area after the above processing, and the image signal can be multi-resolution analyzed.

After processing the two-dimensional image signal, a low-pass smoothing function $\omega(x, y)$ is used to perform wavelet transformation along the x and y directions, that is, the two-dimensional image wavelet transform can be expressed as

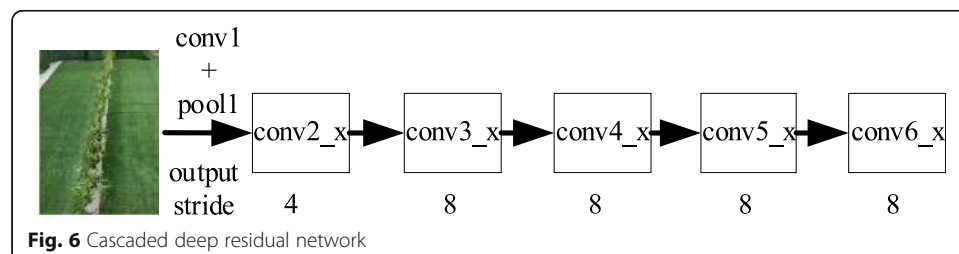


Fig. 6 Cascaded deep residual network

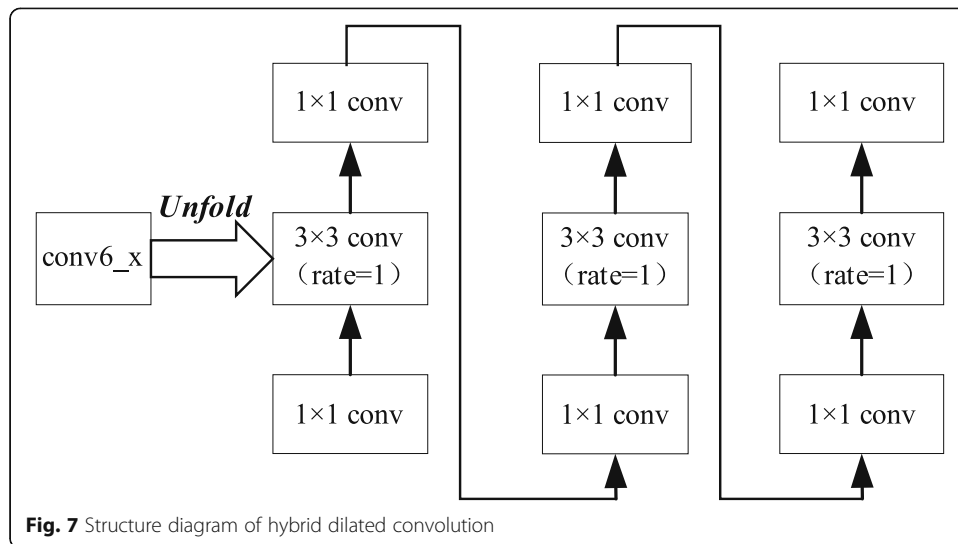


Fig. 7 Structure diagram of hybrid dilated convolution

$$\begin{cases} \phi(x, y) = \frac{\partial \omega(x, y)}{\partial x} \\ \phi(x, y) = \frac{\partial \omega(x, y)}{\partial x} \end{cases} \quad (3)$$

$$Rg(x, y) = \begin{bmatrix} R^1 g \\ R^2 g \end{bmatrix} = \begin{bmatrix} g^* \phi^1(x, y) \\ g^* \phi^1(x, y) \end{bmatrix} = \begin{bmatrix} \int g(\tau) \phi^1(x-\tau, y) d\tau \\ \int g(\tau) \phi^2(x, y-\tau) d\tau \end{bmatrix} \quad (4)$$

where $R^1 g$ and $R^2 g$ are the two variables after the image changes, which are the gradients of the two-dimensional image along x and y directions.

The time-domain two-scale equation of scale function and wavelet function is

$$\begin{cases} \frac{1}{2} \phi\left(\frac{x}{2}\right) = \sum_n p_0(n) \phi(x-n) \\ \frac{1}{2} \phi\left(\frac{x}{2}\right) = \sum_n p_0(n) \phi(x-n) \end{cases} \quad (5)$$

The two-scale equation in the frequency domain is

$$\begin{cases} \Psi(2\omega) = P_0(\omega) \Psi(\omega) \\ \Phi(2\omega) = P_1(\omega) \Phi(\omega) \end{cases} \quad (6)$$

where the wavelet function is the scale function Fourier transform

$$\Phi(\omega) = e^{-jk\frac{\omega}{2}} \left[\frac{\sin\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}} \right]^n = e^{-jk\frac{\omega}{2}} \text{sinc}^n\left(\frac{\omega}{2}\right) \quad (7)$$

P_0 and P_1 are filters corresponding to the scale function and wavelet function, respectively, according to the conservation of energy of space division.

$$P_0(z)P_0(z)^{-1} + P_1(z)P_1(z)^{-1} = 1 \quad (8)$$

where

$$\begin{cases} P_0(z) = \sum_n p_0(n)z^{-n} \\ P_1(z) = \sum_n p_1(n)z^{-n} \end{cases} \quad (9)$$

In this paper, the impulse response coefficients of the third-order B-spline wavelets ($n = 4$), $P_0(z)$, and $P_1(z)$ are shown in Table 1.

Due to the spatial separability of a two-dimensional image signal, the rows and columns can be separately subjected to wavelet transform according to the above algorithm to achieve multi-resolution edge detection.

4 Path extraction for visual navigation

The farming environment is a multi-variable time-varying and nonlinear complex system, which brings great difficulty to the intelligent robot autonomous navigation. Based on the image processing results in Section 3, improved Hough transform is used to extract the navigation path of the crop row, so that robots' posture can be adjusted in time.

4.1 Improved Hough transform

Hough transform is based on the global characteristics of images, forming a local peak at a point in the parameter space where straight line points in images are concentrated. Find and link line segments in the images.

Hough transform has the advantages of strong robustness and strong anti-noise ability. But at the same time, there is also a problem of a large amount of calculation, which will affect the real-time nature of autonomous navigation. Therefore, this paper uses the following steps to improve Hough transform:

Determine the parameter value range after changing polar coordinates. The image after image processing is $U \times V$, and the polar coordinate parameter space is (ρ, θ) , where $\sqrt{|U^2 - V^2|} \leq \rho \leq \sqrt{U^2 + V^2}$ and $0 \leq \theta \leq 180^\circ$. It is worth noting that we use every 2° to calculate, and the amount of calculation is 1/2 of traditional transformation. This is because when digitally quantizing polar coordinate parameters, if the quantization precision is too small, the effect of parameter space cohesion is not obvious. The accuracy is too large, the calculation process is cumbersome, and the calculation amount is large.

Table 1 Coefficients of wavelet filter

n	$p_0(n)$	$p_1(n)$
-3	0	-0.00006
-2	0.6341	-0.01878
-1	0.2562	-0.10210
0	0.3851	-0.51212
1	0.2600	0.51212
2	0.0618	0.19812
3	0	0.01231
4	0	0.00023

Store the sine and cosine values as an array. Store the sine and cosine values from 0 to 180° as values. When the query is needed during the calculation process, directly call the calculation, which is simple and quick.

Use refinement algorithm to improve the Hough algorithm. The refinement of the algorithm can effectively reduce the amount of data after image segmentation, thereby reducing the calculation process and shortening the calculation time.

Effectively determine the corresponding peak of parameter space and the straight line in images. First, the median filter is used to remove noises in parameter space. And a few larger peak points are detected according to the phase angle and deviation characteristics of navigation. Finally, the peak point of the navigation path is determined by statistical analysis.

4.2 Obtaining navigation parameters

Suppose the detected straight line is l_{AB} in the input image coordinate system, image vertex A coordinate is $A(x_A, y_A)$, and vertex B coordinate is $V(x_B, y_B)$. After Hough transform, the coordinates of vertices for the coordinate system are $A'(X'_A, Y'_A)$ and $B'(X'_B, Y'_B)$, and the mathematical expression of conversion relationship is:

$$\begin{cases} X'_A = \frac{L_1}{V} \left(x_A - \frac{V}{2} \right) \\ X'_B = \frac{L_2}{V} \left(x_B - \frac{V}{2} \right) \end{cases} \quad (10)$$

$$\begin{cases} Y'_A = U_1 - U_2 \\ Y'_B = 0 \end{cases} \quad (11)$$

where L is the length of the top and bottom edges for view field in the actual scene, U is the distance from agricultural intelligent device camera to the top and bottom edges of view field, and V is the width of the processed image. After vertices A' and B' of the coordinate system are obtained, the two equations of a straight line $A'B'$ can be obtained, and then, the distance and yaw angle from camera point to a straight line $A'B'$ can be obtained.

4.3 Path extraction

The steps of extracting the navigation path of the agricultural robot by the improved Hough transform algorithm are as follows:

- 1) A thinning algorithm is used to refine segmented images in the third section;
- 2) Discretize parameter space ρ and allocate memory for each;
- 3) Calculate θ step by step every 2°, and calculate ρ corresponding to (x, y) in the image to achieve one-to-one correspondence;
- 4) Use median filtering method to remove the noise points of detected images in the parameter space;
- 5) According to the phase angle and deviation characteristics of navigation, a few larger peak points are detected. Finally, the peak point of the navigation path is determined by statistical analysis.

4.4 Pose determination of robots

When agricultural robots perform normal command operation, its own posture determination is the prerequisite for navigation and agricultural operations. The values of offset angle α and offset distance γ can determine the posture of agricultural robots relative to the center line of the crop row.

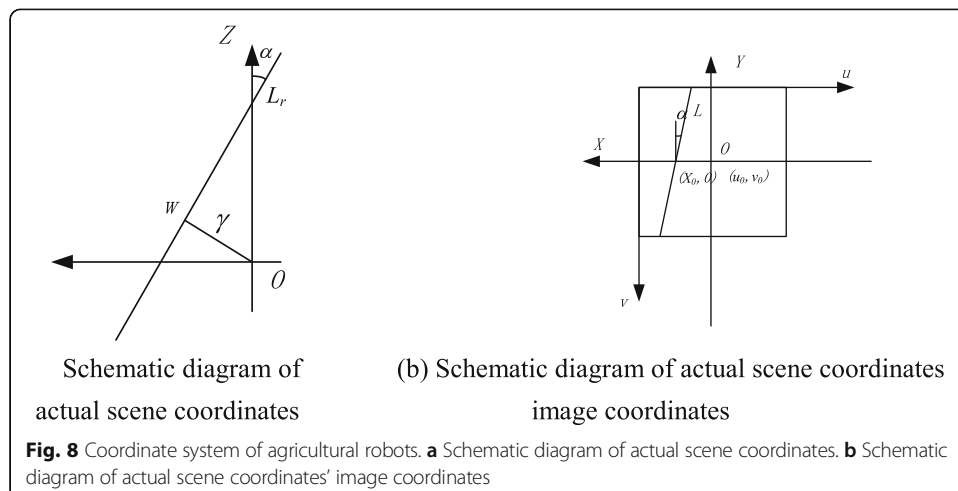
Existing studies have shown that the pose adjustment of intelligent robots can be determined according to the correlation between actual coordinates and image coordinates [30]. Figure 8a is a schematic diagram of the coordinates of the actual scene for robots. X_r axis refers to the left side of the car body in the actual scene, and Z_r is the upper side of the car body center line (car body navigation line). L_r is the center line between rows and crops; γ is the robot offset distance, it is the vertical distance from camera coordinate point to L_r ; α is the angle between robot center line and navigation line. Figure 8b is a schematic diagram of image coordinates, the u - v coordinate system is the image coordinate system in pixels, and the x - y coordinate system is the image coordinate system in millimeters. Based on homogeneous coordinates and matrix form, the mathematical expressions corresponding to pixels and sizes are:

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} -d_0 & 0 & u_0 d_x \\ 0 & -d_y & v_0 d_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{12}$$

Based further on the Hough transform, the straight line in Fig. 8b can be expressed as

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X_0 d_x \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{d_x}{d_y} \tan \alpha \\ 1 \end{bmatrix} \tag{13}$$

According to the camera perspective principle, the actual scene coordinates in image coordinates correspond to:



$$\begin{cases} X = f \frac{X_c}{Z_c} = \frac{\frac{X_0 d_x h}{f \sin \beta} - k \tan \alpha}{\frac{h}{\sin \beta} + k \cos \beta} \\ X = f \frac{Y_c}{Z_c} = \frac{k f \sin \beta}{\frac{h}{\sin \beta} + k \cos \beta} \end{cases} \quad (14)$$

where k is any real number, and the angle formed by the horizontal line of camera β and the smallest observation point to the ground.

Then, offset angle α and offset distance γ of the agricultural robot are respectively:

$$\alpha = \arctan\left(\frac{\sigma_y}{\sigma_x} \tan \alpha \sin \beta - \frac{X_0}{\sigma_x} \cos \beta\right) \quad (15)$$

$$\begin{aligned} \gamma &= |X \cos \alpha + Z \sin \alpha| \\ &= \left| \frac{X d_x h}{f \sin \beta} \cos \alpha + \frac{h}{\tan \beta} \sin \alpha \right| \end{aligned} \quad (16)$$

where h is the distance between the camera and ground. σ_x and σ_y are the scale factors $\sigma_x = \frac{f}{d_x}$ $\sigma_y = \frac{f}{d_y}$ in the image coordinate system.

5 Results

The experimental equipment of this paper is a Tesla K80 GPU host, and the experimental environment is Ubuntu16.04. The code is written based on Tensorflow, a deep learning framework. The camera equipment is Bumblebee2, a stereo vision product produced by Point Grey Research (PGR). The software environment is operating system Chinese Windows 10, English version software Microsoft Visual Studio 2012. The main programming language is C#.

This section tests the visual navigation system and analyzes the data in the test. This test is divided into posture measurement error test, non-interference navigation test, and weed background navigation test.

5.1 Measurement and analysis for pose errors

The actual working environment is more complicated, and it is difficult to measure the pose of robots. Therefore, the accuracy of robot pose calculation directly affects subsequent control actions. Choose to test the simulated rice seedlings in the laboratory. The experimental design is as follows:

- (1) The deviation of fixed phase angle is 0, that is, keep the robot's median line parallel to the actual direction of advance, and choose to move the robot perpendicular to the direction of the seedling row. Keep the displacement deviation range as $[-40, 40]$, the recorded data displacement interval is 10mm, and the recorded data is shown in Table 2. The calculated standard deviation is 0.312mm.
- (2) The fixed displacement deviation is 0. While keeping the center of robots in a straight line with the actual direction of advancement, rotate robots so that the median line and the centerline of the seedling row form a certain angle. Keep the displacement deviation range as $[-10, 10]$, the phase angle interval of recorded data

Table 2 Deviation between measured and calculated distance

Experiment serial number	Experimental result (d/mm)	Calculated result (d/mm)
1	-40	-41.67
2	-30	-31.09
3	-20	-20.71
4	-10	-10.01
5	0	0.12
6	10	9.91
7	20	19.91
8	30	31.01
9	40	40.12

is 5° , and the recorded data is shown in Table 3. The calculated standard deviation is 0.121° .

From the above test data, it can be seen that the results obtained by pose calculation are consistent with the measured results. And the standard deviation is small, which satisfies the measurement requirements.

5.2 Non-interference navigation test

The experiment uses plastic to simulate seedling rows and simulates the farming environment under ideal conditions indoors for testing. Its purpose is to verify whether the navigation first extraction is correct in image processing, so as to confirm whether the visual navigation system is effective.

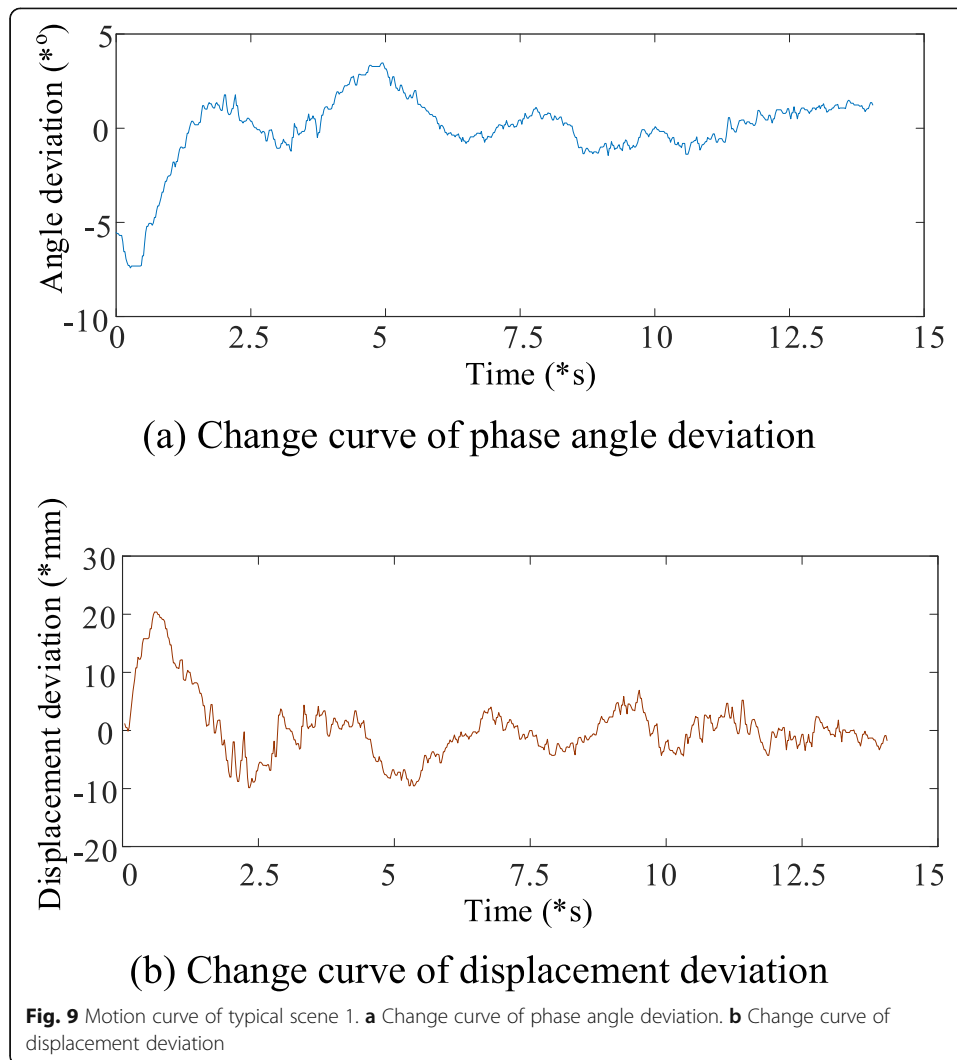
Take the initial angle deviation and position deviation as $(-5^\circ, 0\text{mm})$, $(5^\circ, -5\text{mm})$ two initial states for analysis, draw the curve of its movement process, and analyze the test results.

Scenario 1: The robot motion curve of initial position angle deviation -5° and displacement deviation 0mm is shown in Fig. 9. The phase angle returns to 0° in about 1.53s, and the displacement deviation is 2.11 to reach 0mm deviation. And in the later movement, due to the vibration of robots, the jitter alternates between positive and negative, so it does not affect the overall effect.

Scenario 2: The motion curve of initial phase angle deviation 5° and displacement deviation -5mm is shown in Fig. 10. The angle deviation is large at the beginning of the movement, and the angle deviation quickly attenuates after the movement starts. The displacement deviation reaches the peak value at about 2.61s, and the angle deviation

Table 3 Deviation between measured and calculated distance

Experiment serial number	Experimental result (d/mm)	Calculated result (d/mm)
1	-10	-10.18
2	-5	-4.09
3	0	0.11
4	5	5.13
5	10	10.21

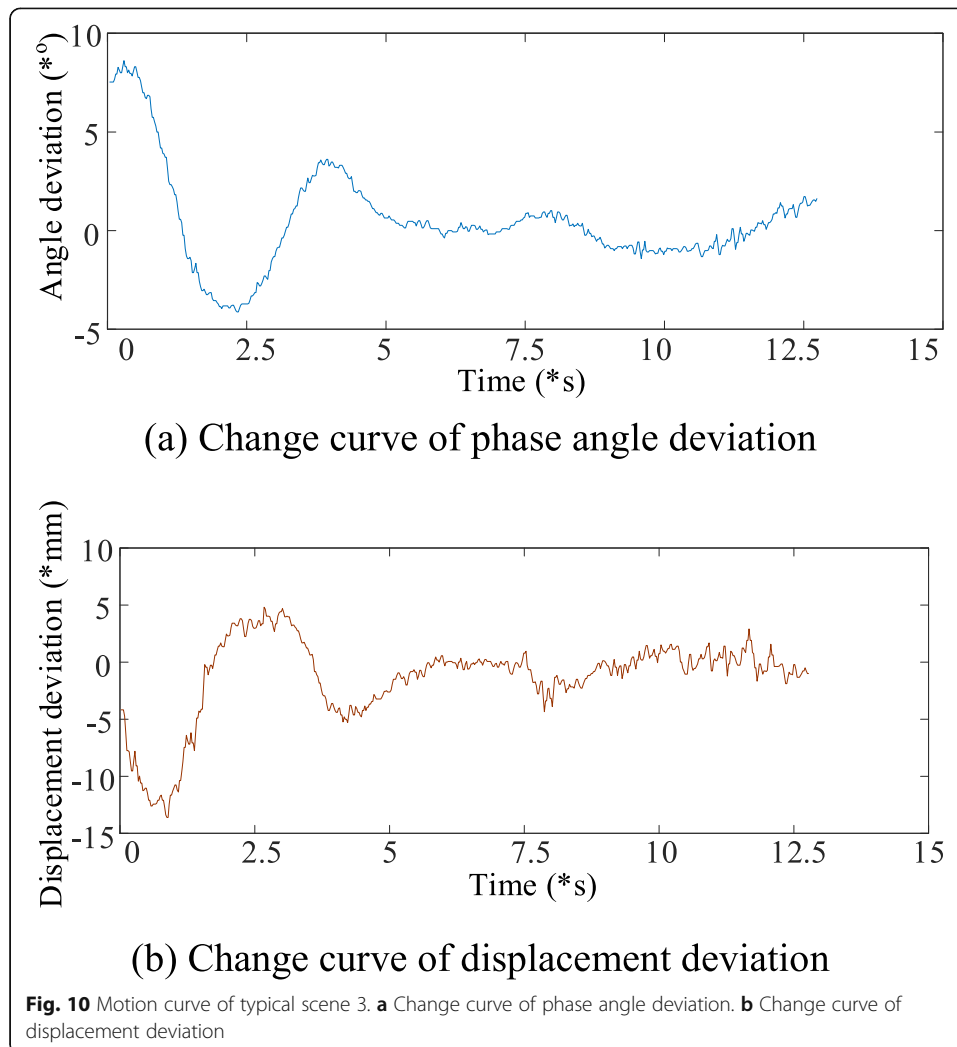


reaches the minimum value at about 2.34s. After several fluctuations, the angle deviation and position deviation both decay to 0.

The non-interference navigation test result proves that the method proposed in this paper can effectively set the navigation line of the seedling row and keep the robot posture and timely and effective adjustment, which can meet the accuracy of autonomous navigation of agricultural robots.

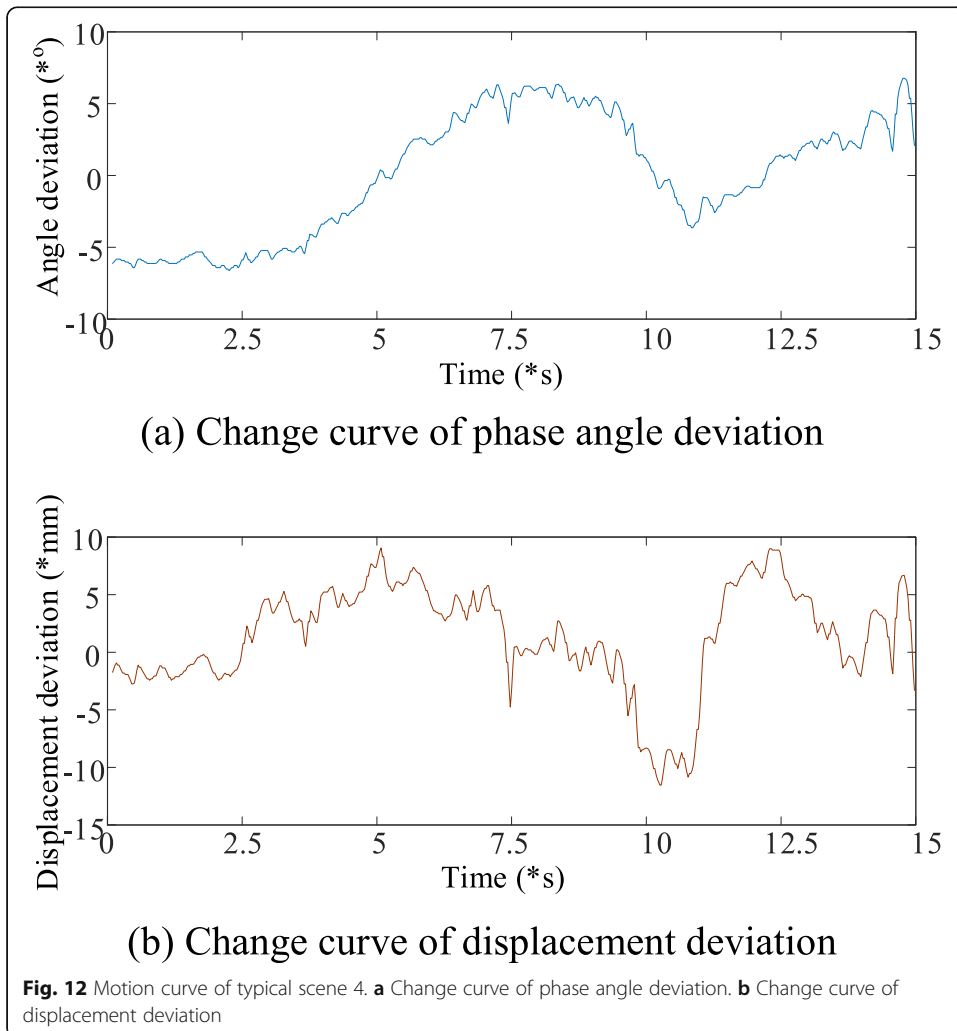
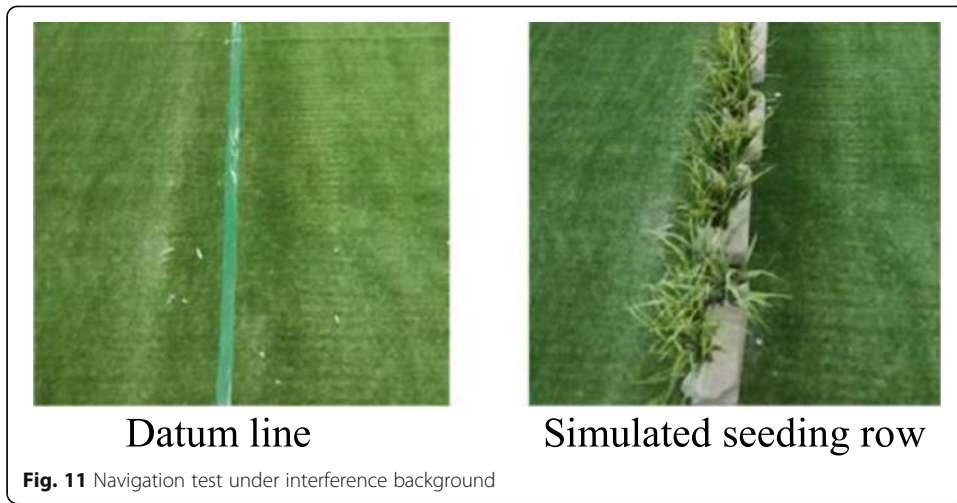
5.3 Weed background navigation test

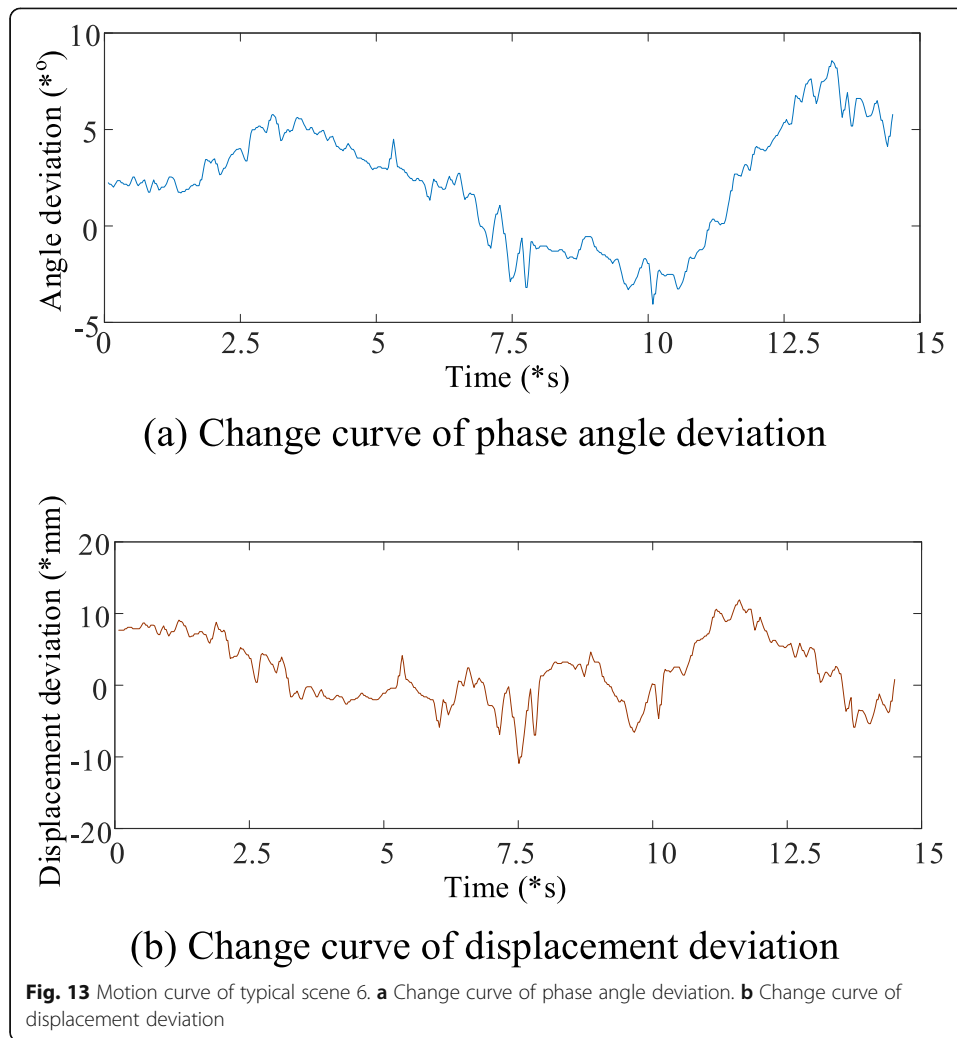
In order to verify the feasibility of this method in this paper in actual farming scenarios, the actual environment is simulated in the laboratory, and the layout scenario is shown in Fig. 11. Artificial turf is used to simulate the most complex paddy field environment in the farming environment including duckweed and waterweed. Due to the inability to accurately measure the displacement deviation and phase angle deviation in the manual layout scenario, this paper selected two random combinations for experimental analysis.



Scenario 3: The motion curve of initial phase angle deviation -6° and displacement deviation -2.3mm is shown in Fig. 12. The phase angle deviation converges to 0 and continues to increase in the opposite direction 5 s after the movement starts. The displacement deviation converges to 0° in 2.6s, the standard deviation of phase angle deviation is 4.21° , and the displacement deviation is 5.31mm . Because the background color is similar to the seedling color, there is still noise after image processing. This causes feature point extraction and clustering errors, resulting in unstable navigation line parameters. But in general, it can still travel along the seedling column and does not step on the seedling.

Scenario 4: The motion curve of initial phase angle deviation 2.3° and displacement deviation 8.12mm is shown in Fig. 13. As shown in Fig. 13, the displacement deviation converges to 0 at 3.1s due to the large displacement deviation relative to phase angle deviation and reaches the extreme value when phase angle deviation is 3.2s. At 11s, the displacement deviation increases in the positive direction. In order to correct the displacement deviation, the phase angle deviation is corrected. The phase angle deviation also increases to correct displacement deviation, and eventually, the displacement deviation converges to 0.





The navigation test is carried out in the presence of background noise, and the proposed method can still accurately extract the navigation line when the background noise is large. By setting the coefficients in time and walking along the set route, the feasibility and practicability of the proposed method for autonomous navigation in complex farming environments are confirmed. The proposed model improves the cascaded deep convolution network based on the hybrid hole convolution method to solve the problem of network degradation caused by too many layers of deep network.

6 Results and discussion

Facing the accuracy requirements of autonomous navigation of intelligent agricultural robots, this paper proposes an agricultural robot visual navigation algorithm based on deep learning image understanding. The algorithm mainly includes two aspects of image processing and visual navigation path extraction. In the processing of collected images, collected images are processed based on a cascaded deep convolutional network and hybrid dilated convolution method, which provides optimal image data support for the subsequent autonomous navigation of robots. Moreover, the Hough transform method is improved based on the subdivision algorithm in visual navigation

path extraction. And the correspondence relationship between the image coordinate system and the actual scene coordinate system and state equation are established to achieve robots' autonomous navigation posture adjustment. The experimental results show that our proposed method embodies rapid response characteristics at the same time in the non-interference scene and complex noise scene to ensure normal and stable operation of agricultural robots. The focus of future research will be to explore the adaptability of the proposed algorithm and agricultural robots in the market to improve algorithm scalability.

However, limited by the author's level, the proposed algorithm still cannot get very accurate segmentation results for object boundary and small object segmentation. To solve this problem, we can consider using deeper network structure in the future, such as Resnet152, Densenet169, Densenet201, etc.; we can also consider fusing other deep learning technologies to complete image semantic segmentation tasks, such as a new variant of recurrent neural network RNN, counter network GAN, etc.

Acknowledgements

We wish to express our appreciation to the reviewers for their helpful suggestions which greatly improved the presentation of this paper.

Authors' contributions

The main idea of this paper is proposed by JL. The algorithm design and experimental environment construction are jointly completed by JL and JY. The experimental verification was completed by all the three authors. The writing of the article is jointly completed by JL and JY. And the writing guidance and English polish are completed by LD. The authors read and approved the final manuscript.

Funding

This work was not supported by any funding projects.

Availability of data and materials

The data included in this paper are available without any restriction.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 21 February 2021 Accepted: 6 May 2021

Published online: 17 May 2021

References

1. G.Y. Wang, Route choice of rural economic development in offshore areas from the perspective of modern agriculture. *J. Coastal Res.* **98**, 247–250 (2019)
2. G. Ozkan, I.B. Gurbuz, E. Nasirov, A greener future: the addictive role of technology in enhancing ecoliteracy in rural community. *Fresenius Environ. Bull.* **29**(6), 4372–4378 (2020)
3. C.W. Bac, E.J. van Henten, J. Hemming, et al., Harvesting robots for high-value crops: state-of-the-art review and challenges ahead. *J. Field Robot.* **31**(6), 888–911 (2014)
4. A.I. Hong-juan, J.I.A.N.G. He-ping, Studies on evaluating modern agricultural development level in Xinjiang based on factor analysis method. *J. Agric. Sci. Technol.* **14**(4), 157–164 (2015)
5. S. Dong, Z. Yuan, C. Gu, et al., Research on intelligent agricultural machinery control platform based on multi-discipline technology integration. *Transact. Chinese Soc. Agric. Eng.* **33**(8), 1–11 (2017)
6. Z. Zhi-Hui, L.I. Quan-Xin, The evaluation index system of agricultural modernization based on the perspective of human-oriented development. *J. Anhui Agric. Sci.* **44**(6), 254–257, 287 (2016)
7. J. Wang, Y.T. Zhu, Z.B. Chen, et al., Auto-steering based precise coordination method for in-field multi-operation of farm machinery. *Int. J. Agric. Biol. Eng.* **11**(5), 174–181 (2018)
8. J. Mao, Z. Cao, H. Wang, B. Zhang, Z. Guo and W. Niu, "Agricultural Robot Navigation Path Recognition Based on K-means Algorithm for Large-Scale Image Segmentation," 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 1233–1237 (2019). <https://doi.org/10.1109/ICIEA.2019.8834296>.
9. K. Kurashiki, M. Aguilar, S. Soontornvanichkit, Visual navigation of a wheeled mobile robot using front image in semi-structured environment. *J. Robot. Mechatronics* **27**(4), 392–400 (2015)
10. Y.F. Zhou, N. Chen, The LAP under facility disruptions during early post-earthquake rescue using PSO-GA hybrid algorithm. *Fresenius Environ. Bull.* **28**(12A), 9906–9914 (2019)
11. X.Y. Wang, Q.Y. Wang, H.Y. Yang, et al., Color image segmentation using automatic pixel classification with support vector machine. *Neurocomput.* **74**(18), 3898–3911 (2011)

12. U. Alganci, The use of broadband vegetation indices in cultivated land detection with Landsat 8 OLI multi-temporal images. *Fresenius Environ. Bull.* **28**(2), 739–744 (2019)
13. S.S. Mehta, T.F. Burks, W.E. Dixon, Vision-based localization of a wheeled mobile robot for greenhouse applications: a daisy-chaining approach. *Comput. Electron. Agric.* **63**(1), 28–37 (2008)
14. K.A.Y.A. Latif Gürkan, Z. Kaynakci-Elinc, C. Yücedağ, et al., Environmental outdoor plant preferences: a practical approach for choosing outdoor plants in urban or suburban residential areas in Antalya, Turkey. *Fresenius Environ. Bull.* **27**(12), 7945–7952 (2018)
15. T. Torii, T. Takamizawa, T. Okamoto, et al., Crop row tracking by autonomous vehicle using machine vision (part I). *J. JSAE* **62**(2), 41–48 (2000)
16. S. Han, Q. Zhang, B. Ni, et al., A guidance directrix approach to vision-based vehicle guidance systems. *Comput. Electron. Agric.* **43**(3), 179–195 (2004)
17. Akane T, Ryohei M, Michihisa I, et al. Image processing for ridge/furrow discrimination for autonomous agricultural vehicles navigation [C]// IFAC Conference on Modelling and Control in Agriculture, 2013:27-30, (2003).
18. B. David, U. Ben, W. Gordon, et al., Vision-based obstacle detection and navigation for an agricultural robot. *J. Field Rob.* **33**(8), 1107–1130 (2016)
19. G. Lin, X. Zou, L. Luo, et al., Detection of winding orchard path through improving random sample consensus algorithm. *Nongye Gongcheng Xuebao/Transact. Chinese Soc. Agric. Eng.* **31**(4), 168–174(7) (2015)
20. X. Changyi, Z. Lihua, L. Minzan, et al., Apple detection from apple tree image based on BP neural network and Hough transform. *Int. J. Agric. Biol. Eng.* **8**(6), 46–53 (2015)
21. W. Wera, F.F. Veronika, D. Christian, et al., Crop row detection on tiny plants with the pattern Hough transform. *IEEE Robot. Automation Lett.* **3**(4), 3394–3401 (2018)
22. Z.W. Chen, W. Li, W.Q. Zhang, et al., Vegetable crop row extraction method based on accumulation threshold of Hough transformation. *Transact. Chinese Soc. Agric. Eng.* **35**(22), 314–322 (2019)
23. F. Li, L.F. Wang, H. Hui, et al., Linear features extraction of gravity and magnetic data based on Hough transform. *Chinese J. Eng. Geophys* **015**(001), 22–31 (2018)
24. H.T. Sogaard, H.J. Olsen, Determination of crop rows by image analysis without segmentation. *Comput. Electron. Agric.* **38**(2), 141–158 (2003)
25. Z. Qun, S. Jian, C. Gaohua, et al., Path recognition algorithm and experiment of greenhouse robot by visual navigation. *Res. Exploration Lab* **37**(5), 14–17 (2018)
26. Y.F. Zhou, H.X. Yu, Z. Li, J.F. Su, C.S. Liu, Robust optimization of a distribution network location-routing problem under carbon trading policies. *IEEE Access* **8**(1), 46288–46306 (2020)
27. H. Zhou, A. Han, H. Yang, et al., Edge gradient feature and long distance dependency for image semantic segmentation. *Comput. Vis. IET* **13**(1), 53–60 (2019)
28. L. Yunwu, X. Junjie, L. Dexiong, et al., Field road scene recognition in hilly regions based on improved dilated convolutional networks. *Transact. Chinese Soc. Agric. Eng.* **35**(7), 150–159 (2019)
29. Z. Zhang, X. Wang, C. Jung, DCSR: dilated convolutions for single image super-resolution. *IEEE Transact. Image Process.* **28**(4), 1625–1635 (2019)
30. S. Ulrich, J.Z. Sasiadek, I. Barkana, Nonlinear adaptive output feedback control of flexible-joint space manipulators with joint stiffness uncertainties. *J. Guidance Control Dyn.* **37**(6), 1961–1975 (2014)

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
