An automated method for stem diameter measurement based on laser module and deep learning

Sheng Wang[1,2], Rao Li[1,2], Huan Li[1,2], Xiaowen Ma[1,2], Qiang Ji[3], Fu Xu[1,2] and Hongping Fu[1,2*]

Abstract

Background Measuring stem diameter (SD) is a crucial foundation for forest resource management, but current methods require expert personnel and are time-consuming and costly. In this study, we proposed a novel device and method for automatic SD measurement using an image sensor and a laser module. Firstly, the laser module generated a spot on the tree stem that could be used as reference information for measuring SD. Secondly, an end-to-end model was performed to identify the trunk contour in the panchromatic image from the image sensor. Finally, SD was calculated from the linear relationship between the trunk contour and the spot diameter in pixels.

Results We conducted SD measurements in three natural scenarios with different land cover types: transitional woodland/shrub, mixed forest, and green urban area. The SD values varied from 2.00 cm to 89.00 cm across these scenarios. Compared with the field tape measurements, the SD data measured by our method showed high consistency in different natural scenarios. The absolute mean error was 0.36 cm and the root mean square error was 0.45 cm. Our integrated device is low cost, portable, and without the assistance of a tripod. Compared to most studies, our method demonstrated better versatility and exhibited higher performance.

Conclusion Our method achieved the automatic, efficient and accurate measurement of SD in natural scenarios. In the future, the device will be further explored to be integrated into autonomous mobile robots for more scenarios.

Keywords Measurement, Forest inventory, Laser module, Image sensor, Deep learning

Background

Stem Diameter (SD) is a key parameter for estimating standing timber volume [1], assessing economic value [2], and planning silvicultural interventions [3]. Larger trees are normally measured using the diameter at breast height (DBH) [4, 5], whereas for trees below breast height including multi-stemmed trees [6], shrubs [7] and saplings [8], measurements are generally taken below the most common location where the stem section forms multiple leaders [5, 9]. In such vegetation, the traditional methods of measuring SD require trained personnel to determine the location and angle of the measurement. Automatic and cost-effective methods of measuring have become much-needed tools for forest inventories.

The main challenge in automatically measuring DBH of individual tree largely lies in the limitations of conventional measurement methods. In earlier studies, foresters manually measured trees using altimeter, tape measure, and diameter tape [10–12]. These methods are relatively reliable and low-cost, but they are also time-consuming, labor-intensive, and error-prone [13–16].

*Correspondence:
Hongping Fu
fhongping@bjfu.edu.cn
[1] School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China
[2] Engineering Research Center for Forestry-oriented Intelligent Information Processing, National Forestry and Grassland Administration, Beijing 100083, China
[3] Space Star Technology Co., Ltd., Beijing 100048, China

Other studies have compared and evaluated different manual measurement methods, such as the angle gauge [17], Bitterlich sector fork [18, 19], electronic tree measuring fork [20] and Bilt-more stick [21]. The measurement method based on projection geometry improves the efficiency. However, the accuracy of the device depends on the forester experience, and additional auxiliary tools are required to indicate altitude [10, 22]. Similarly, some methods were designed based on optical measurement principles such as optical calipers [23, 24], optical forks [25]. The measurement uncertainty increases with DBH. The inaccessibility of the target area and the limitations of the device accuracy constrain the application of above methods.

Some semi-automatic methods based on light detection and ranging (LiDAR) have gained popularity recently. Terrestrial laser scanning (TLS) is proved to be a promising solution for deriving DBH from TLS data through either direct geometric fitting or tree stem modeling and separation [26–29]. The advantage of TLS data is that it can capture forest data in detail and enable time series analysis. However, the device is high-cost, complex data processing and demands high expertise [30]. Backpack and vehicle-based LiDAR systems are spatially flexible [31, 32], but vehicle-based LiDAR is constrained by difficult terrain and available roads, while the stability of backpack LiDAR system is affected by irregular movement. Despite the acceptable accuracy of LiDAR-based methods, automation of these methods still faces the challenges of cost of hardware, complexity of data processing, and portability.

To achieve high automation in acquiring tree structure information, close-range photography technology and a segmentation and fitting algorithm based on point cloud data are widely applied. However, some challenges still exist. The automated methods to generate dense point clouds for estimating DBH using close-range photography are attempted, but are susceptible to light conditions [33–35]. Gao et al. [29] modeled the forest based on structure from motion photogrammetry to automatically estimate DBH by circular fitting. The method is economical but not applicable to scenarios with deviations in tree and circularity. Machine vision methods can obtain the pixel size of objects from rich image information. Wu et al. [36] used machine vision and close-range photogrammetry to measure the DBH of multiple trees from an image taken by a smartphone. The method is convenient, efficient and has great potential for development to bring to a wide range of users. However, the coordinate system conversion in photogrammetry is very complicated and cannot guarantee the accuracy of transformation from 2D images to 3D coordinates [37]. The tree tilt angle, ground slope and photographic distance limit its use in daily practice.

This paper aims to achieve automated stem diameter measurements by integrating a laser module and an image sensor. It also considers relevant factors to reduce the professional and economic costs of the method. The accuracy of the method is evaluated and analyzed in different scenarios to verify its generality and feasibility.

Materials and methods
Device description
The device proposed in this paper is used to estimate the stem diameter of trees in the field, including image sensor, laser module, development board, stepper motor, GPS receiver, touch screen, and acrylic board. The image sensor was used to capture panchromatic images of the target tree while the laser module formed a fixed size spot on the tree trunk. Then, analysis module calculated the panchromatic image to obtain the SD. The device is highly integrated, eliminating the need for tripods and other external auxiliary devices (Fig. 1). Table 1 lists the core parameters of the laser module and image sensor used in this device.

The development board has an embedded system module with a GPU and CPU built-in. When the device is running, the embedded system module reads live video from the image sensor interface and displays it on the touch screen. The image sensor and the laser module are coaxial. The horizontal angle between the image sensor and the optical axis of the laser module is fixed. The laser module is situated right above the image sensor and is controlled by a stepper motor that rotates at a modest speed. When the spot emerges on the tree trunk, the operator clicks a touch screen button to capture an image of the target tree. In order to avoid repeated measurements, the GPS receiver is in charge of recording the location of the target tree.

Workflow
The logical structure of our automated approach consists of three parts (Fig. 2), i.e., the spot detection algorithm (SDA), our improved U^2-Net, and the analysis module. The SDA provides reference information for the analysis module and key point information for our improved U^2-Net. Our improved U^2-Net obtains visual saliency map and segments trunk contour. Finally, the analysis module combines the reference information and the trunk contour information to calculate the SD.

Algorithms
SDA
The SDA is proposed to retrieve the location of the spot centroid and to calculate the number of pixels in the
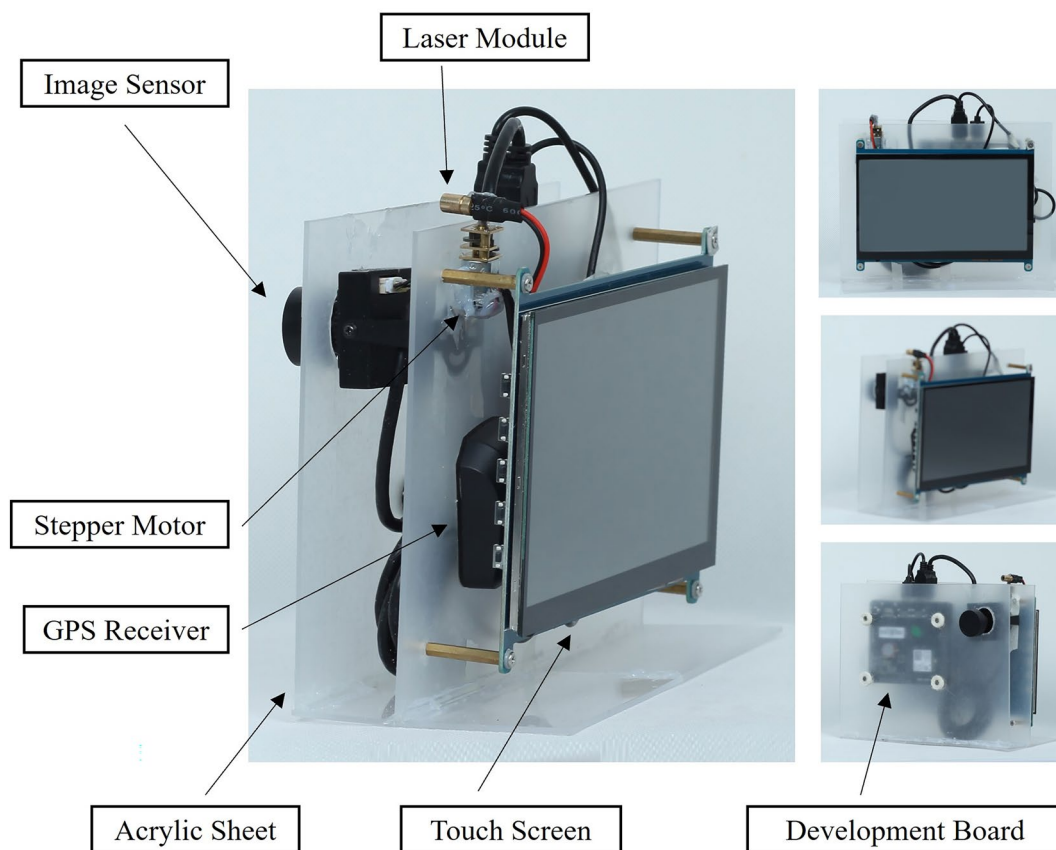
Fig. 1 The structure of the developed device, including image sensor, laser module, development board, stepper motor, GPS receiver, touch screen, and acrylic sheet. The battery is embedded internally. The housing of the device is made of acrylic sheet

spot diameter. Reducing the impact of lighting noise on the algorithm of panchromatic images is a key issue [38, 39]. The angle α between the image sensor and the laser module also causes the spot to form a slight distortion in original image. Therefore, SDA keeps the image features invariant during processing by constructing a pixel coordinate system, and detects the spot by multi-scale images and circular fitting methods.

SDA establishes a pixel coordinate system for the original image, which is composed of black and white regions (Fig. 3a). The gray areas are the uneven lighting conditions, the yellow area is the spot, and the white areas are the backgrounds. We perform an opening operation on the original image. The opening operation is divided into two steps: first, erosion is used to eliminate small blobs. Then, the dilation is used to regenerate the size of the original object. The difference between the original image and the image after the opening operation is represented by the circular structural elements (CSE). The CSE include the spot and the disturbing factors. To remove the disturbing factors close to the image edge, each image is cropped to 15/16

of the original scale, and the output is shown in Fig. 3c. In order to balance the enhancement of the spot with the suppression of disturbing factors, we only iterate on the above steps twice.

The spot is usually an isolated spot after morphological processing and is closest to the center of the image compared to most disturbing factors. Retrieving the spot from the CSE is the key to SDA. The algorithm considered in most of this study is based on the argument of the minimum problem, as shown in Eq. 1.

$$\underset{x \in S}{\arg\min} \ \|F(x)\|, \tag{1}$$

where S is the centroids of CSE. When the distance between factor x and the image centroid is minimum, the minimum value can be obtained in linear space F. The cropping operation of the algorithm changes the position of the spot. In order to keep the invariance of the image features, the variable λ is introduced in the calculation process for computing the optimal solution $\Omega^*(x)$, as shown in Eq. 2. $C(x)$ is the centroid coordinate of the image.

Wang *et al. Plant Methods*   (2023) 19:68

Page 4 of 16

Table 1 Core parameters of the device components

Module	Core parameters	Values
Image sensor	Image size	1920 × 1080 pixels
	Focal length	6.00 mm
	Image sensor type	CMOS
	Vertical field of View	60°
	Infrared filter	980 nm Narrow Band
	Price	$ 32.83
Laser module	Wavelength	980 nm
	Spot size	3.00–30.00 mm
	Output power	300 mw
	Angle of divergence	1.05°
	Working distance	0.05-30.00 m
	Price	$ 27.07
Development board	CPU	H2 Quad-core Cortex-A7 H.265/HEVC 1080P
	GPU	Mali400MP2 GPU @600MHz
	Memory(SDRAM)	512MB DDR3
	Power supply	5V 3A
	Price	$ 22.21
Stepper motor	Actual rated input voltage	3 V–5 V
	Speed	15 RPM
	Motor size	16.0 mm * 12.0 mm * 9.9 mm
	Price	$ 0.50
GPS receiver	Voltage	5 V
	Size	48.0 mm * 35.0 mm * 13.5 mm
	Receiving mode	GPS,BeiDou
	Price	$ 3.50
Touch screen	Resolution	320*480 DOTS
	Interface	HDMI
	Price	$ 8.26
Acrylic sheet	Thickness	2 mm
	Width	300 mm
	Length	200 mm
	Price	$ 0.56
Battery	Capacity	3000 mA
	Voltage	5 V
	Price	$ 9.07

$$\Omega^*(x) := F(x) + \lambda C(x). \tag{2}$$

The mathematical distribution of grayscale values in the original images was explored in order to separate spot and disturbing factors more precisely, as shown in Eq. 3.

$$p \sim N(\mu, \sigma), \tag{3}$$

where N is the number of pixels with pixel value p, μ is the mean, and σ is the standard deviation. Pixels of different nature are labled by thresholds during processing.

The pixel values p greater than the threshold are assigned as η. Others are set to 0 (Eq. 4).

$$p = \begin{cases} \eta, & \text{if } p > \frac{1}{2}\eta + \mu, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The gray value of the spot in the image decreases in a gradient from the centroid to the edge [40, 41]. The SDA selects the pixel points in the region with large variation
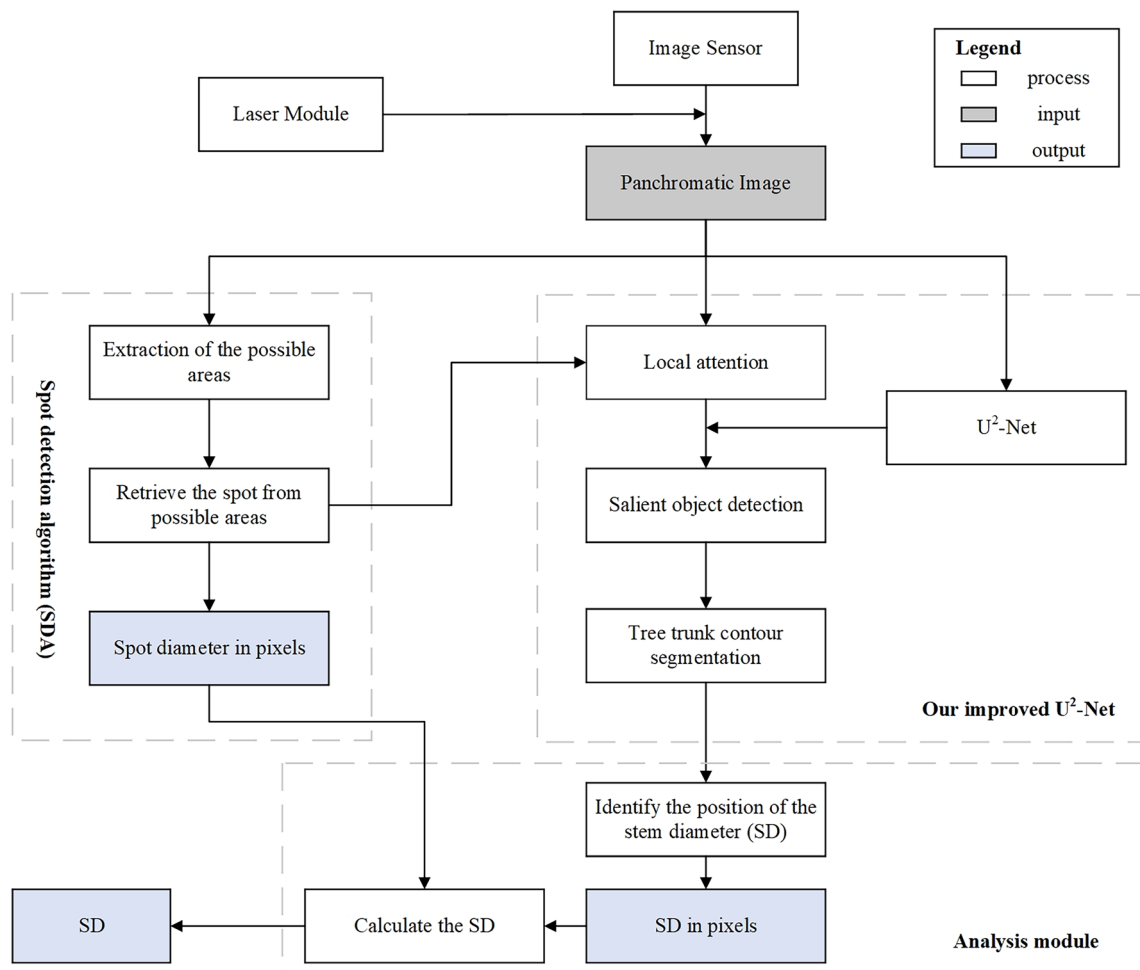
Wang *et al. Plant Methods*   (2023) 19:68

Page 5 of 16

Fig. 2 Workflow diagram of the device, including three sub-processes: the spot detection algorithm (SDA), our improved U²-Net, and the analysis module

of light intensity as the composition of linear space $\Gamma(y)$. The least squares method performs circular fitting on the linear space $\Gamma(y)$ of the optimal solution $\Omega^*(x)$ and estimates the pixel size of spot diameter S', as shown in Eq. 5.

$$S' = \frac{2}{k}\|\Gamma(y) - \Omega^*(x)\|_2. \tag{5}$$

In Eq. 5, k is the number of vectors in linear space $\Gamma(y)$. The larger the n, the more accurately the circle is fitted. Therefore, we experimentally obtained the convergence values n to make the value of the linear space $\Gamma(y)$ finite (Fig. 3d).

Our improved U²-Net

Compared to most state-of-the-art networks such as DPNet [42] and RCSB [43], the U²-Net has low memory and computational requirements [44]. The U²-Net is a two-level nested U-shaped structure that can keep the image features unchanged. The ReSidual U-block (RSU) module at the bottom level is designed to integrate receptive fields at different scales to capture more contextual information at different scales, while the top level ensures depth and reduces computation [44].

The input feature map x is converted into an intermediate map $F(x)$ in the weight layer. In Fig. 4a, U-block is a U-net-like symmetric encoder-decoder structure that learns to extract and encode the multiscale contextual information $U(F(x))$ from the intermediate feature map $F(x)$, where U represents a U-net-like structure. The RSU module has a residual connection of multi-scale features and local feature fusion, which can be represented as $U(F(x)) + F(x)$. Due to the complex natural environment and high-resolution images, the larger the receptive field of RSU module is, the richer the local and global features will be. Otherwise, the computational redundancy will be increased. The output features of SDA provide the RSU module with information about the spot. Therefore, we
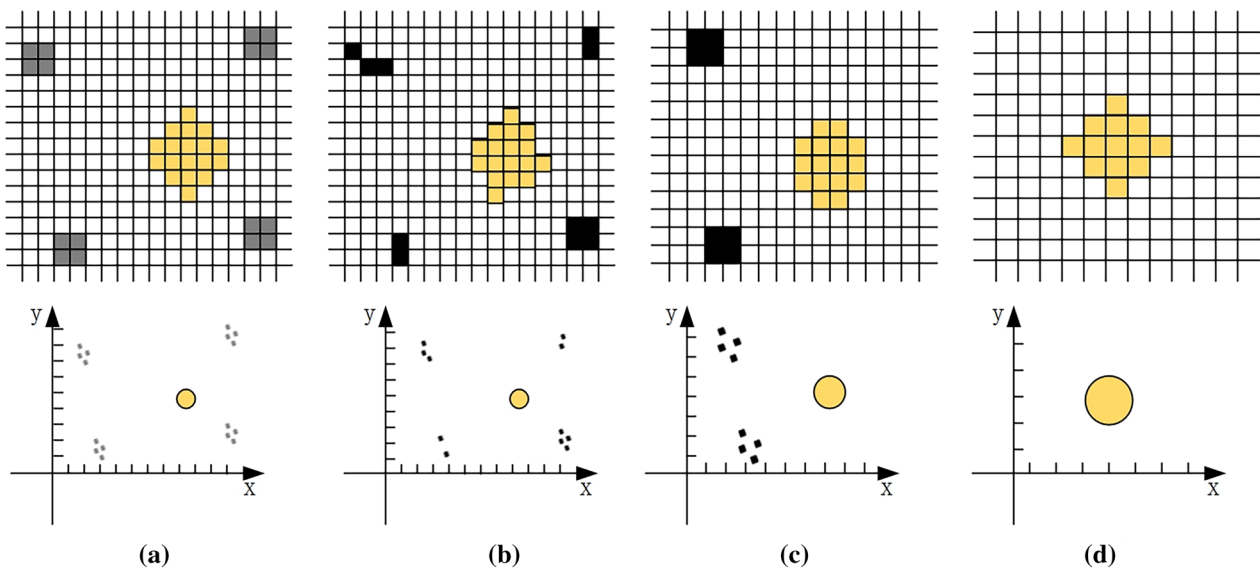
Wang *et al. Plant Methods*   (2023) 19:68

Page 6 of 16

Fig. 3 Diagram of circular structural elements in the spot detection algorithm (SDA) and their variation in the coordinate system. The gray is uneven light, the yellow represents the spot and the black is the disturbing factor. **a** The original image and its coordinate system. **b** The original image and coordinate system after opening operation. **c** Cropped image and coordinate system. **d** Spot position image and coordinate system
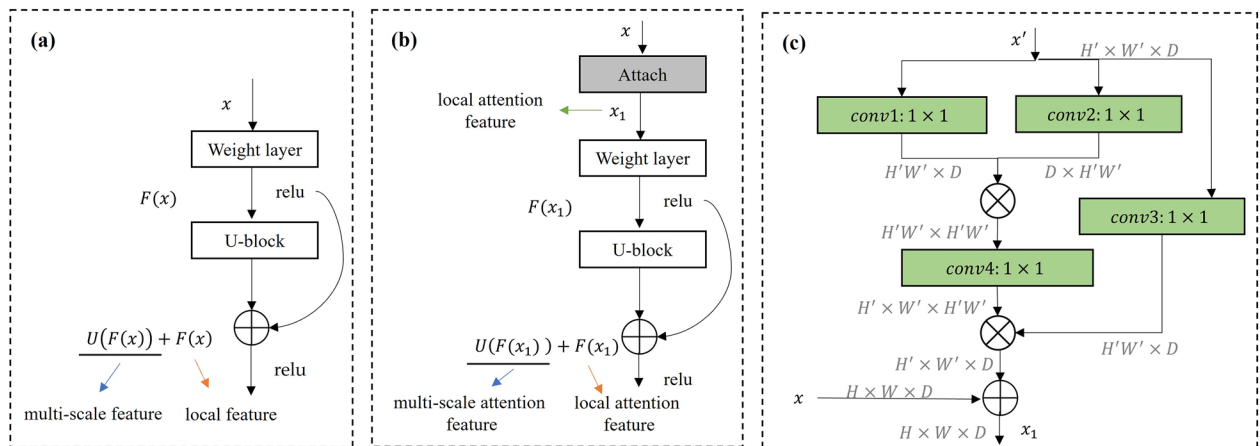

Fig. 4 Comparison of the ReSidual U-block (RSU) and our Attach-ReSidual U-block (A-RSU). **a** RSU. **b** A-RSU. **c** Attach

add the *Attach* module, which extracts the local attention feature x_1 of input feature map x (Fig. 4b). Our Attach-ReSidual U-block (A-RSU) has a novel residual connection which fuses local attention features and the multi-scale attention features by the summation: $U(F(x_1)) + F(x_1)$.

The six stages encoder of U^2-Net consists of n RSU ($n = 6$), with different RSU processing feature maps at different spatial resolutions [44]. We extract the area around the spot centroid as local feature map x' to be input into the different A-RSU, respectively. The relationship between the local feature map x' and the input feature map x is shown in Eq. 6. Where H, W are the height and width of input feature x, respectively, and H'_m, W'_m are the height and width of the m-th local feature map x'.

$$\frac{H'_m W'_m}{HW} = \frac{m}{n} \tag{6}$$

The structure of the *Attach* module consists of four convolutional layers and the matrix multiplications (Fig. 4c). First, the local feature map x' are linearly mapped to obtain three features, and the three features are reshaped. Then, the matrix multiplication is applied to the features

Wang *et al. Plant Methods*   (2023) 19:68

Page 7 of 16

output from *conv*1 and *conv*2. The features output from the matrix multiplication are further applied to the 1×1 convolution to recover the dimensionality of the features. Finally, the result of matrix multiplication of the output from *conv*4 and *conv*3 is added to the input feature map x, and then output x_1.

We trained our improved U^2-Net on the trunk dataset of 1600 images created by ourselves, which consisted of 800 panchromatic images and 800 ground truth. The ground truth has the same spatial resolution as the panchromatic images, with either the pixel value of 0 or 255. The 255 indicates the foreground salient object pixels, while the 0 indicates the background pixels (Fig. 5). The Adam optimiser [45] and hyper parameters [44] are set to default values (initial learning rate $= 1e - 3$, betas $= (0.9, 0.999)$, eps $= 1e - 8$, weight decay $= 0$). After 1000 iterations (batch size $= 64$), the loss function has converged. In the training process, our training loss function is defined as:

$$L = \sum_{m=1}^{M} \omega_{side}^m l_{side}^m \phi^m + \omega_{fuse}^m l_{fuse}^m \tag{7}$$

where l_{side}^m is the loss function of the side saliency map and l_{fuse}^m is the loss of the final fusion output saliency map of our A-RSU. ω_{side}^m and ω_{fuse}^m are the weights of each loss term. In addition, an weight ϕ^m is added in this study to improve the adaptability of the loss term. The smaller the resolution of the local feature map x', the higher the probability that the *Attach* module captures valid information. The weight ϕ^m is determined based on the resolution proportion of the local feature map x' and the input feature map x.

$$\phi^m = \begin{cases} 1 - \frac{R(x'_m)}{R(x)}, & n < m, \\ 1, & n = m \end{cases} \tag{8}$$

For the case when the local feature map x' and the input feature map x are the same, we keep the original weights as shown in Eq. 8. $R(x'_m)$ is the resolution of the m-th local feature map and $R(x)$ is the resolution of the input feature map x. Finally, the final fused features are used as the output of the model.

Analysis module

The analysis module of our method is shown in Fig. 6. The SD is calculated from the tree trunk contour information (object properties, pixel width) extracted by our improved U^2-Net and spot diameter. S is the SD and R is the spot diameter. S' and R' are the SD in pixels and spot diameter in pixels mapped to the image, respectively. d represents the distance between the device and the target tree. h is SD height. The α is the angle formed by the image sensor and the spot. The slight optical distortion caused by α ($91°$) hardly affects the SD calculation.

In the photogrammetry principle and projective geometry theory, the relation of S, S', R and R' can be expressed as Eq. 9.

$$\frac{S'}{S} = \frac{R'}{R} \tag{9}$$

For the measurement of tree DBH, the height in pixels h' can be calculated as in Eq. 10.
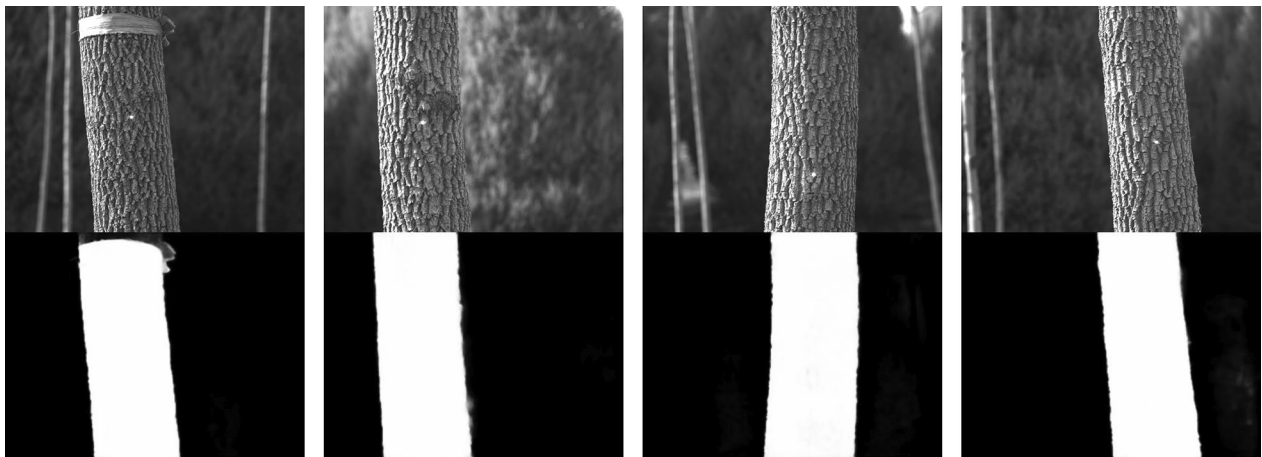
$$\frac{h'}{R'} = \frac{h}{R} \tag{10}$$

Fig. 5 Training and testing of our improved U^2-Net

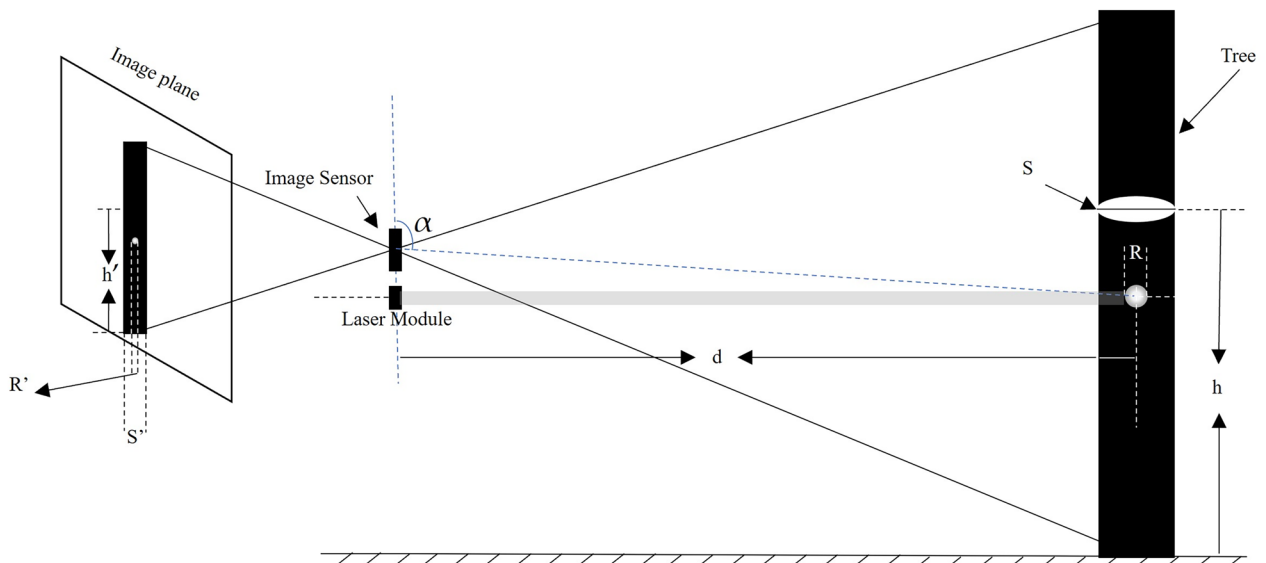Wang *et al. Plant Methods*   (2023) 19:68

Page 8 of 16

Fig. 6 Diagram of the device for SD measurement. The laser module is parallel to the ground, and the image sensor is coaxial with the laser collimator
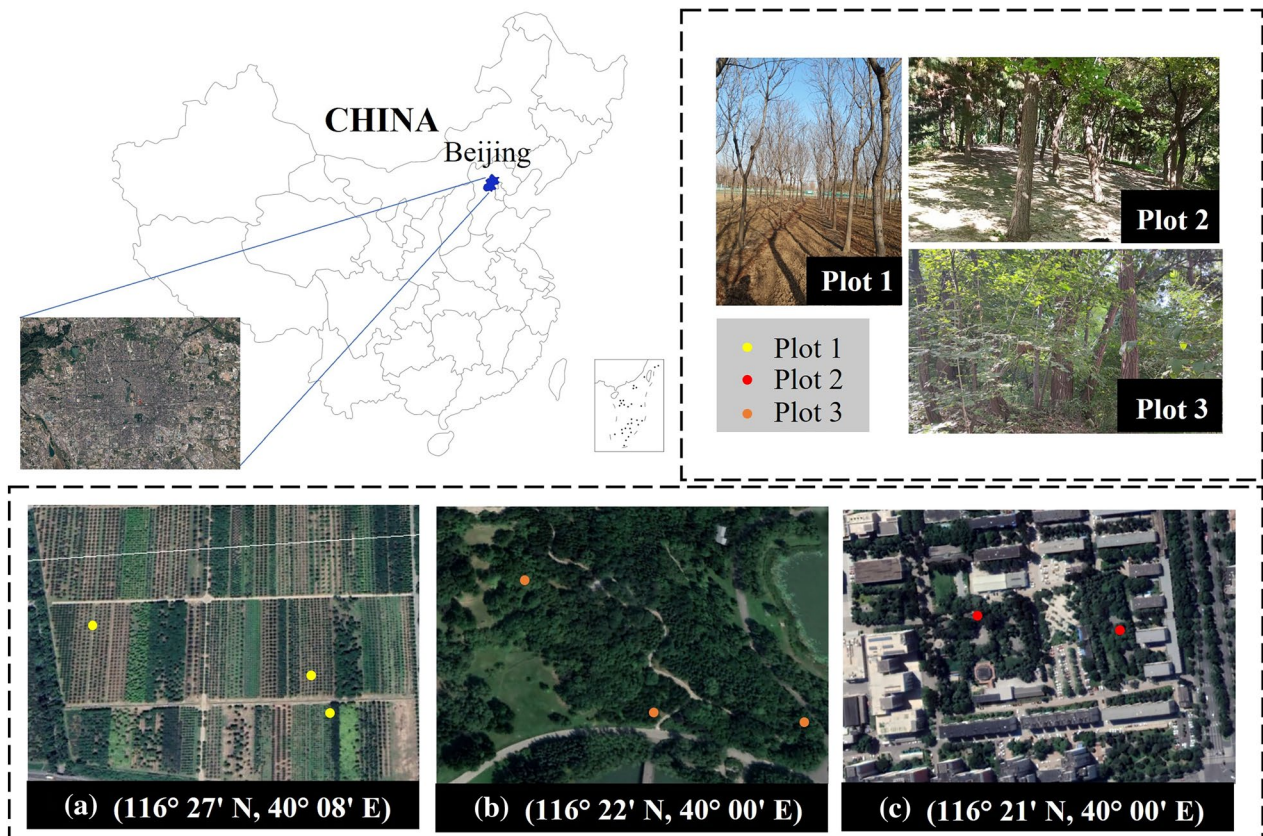

Fig. 7 The study area includes three regions in Beijing, China, where blue indicates the provinces. **a**–**c** are the geographical distribution in Google Earth imageries, and Plot 1, Plot 2, and Plot 3 are the field images respectively

Wang *et al. Plant Methods* (2023) 19:68

Page 9 of 16

Experimental data

The experimental data were primarily obtained from three areas in Beijing, China, i.e., transitional woodland/shrub, mixed forest, and green urban area (Fig. 7). Transitional woodland/shrub is in Changping District (116°27′N, 40°08′E), including *Styphnolobium japonicum*, *Koelreuteria paniculata*, and *Buxus sinica*, dominated by saplings and shrubs, with flat terrain. Mixed forest is in Chaoyang District (116°22′N, 40°00′E), including *Ginkgo biloba*, *Pinus tabuliformis*, and *Acer truncatum*, with rugged terrain and difficult to approach in some areas. The green urban area is in Haidian District (116°21′N, 40°00′E), including *Robinia pseudoacacia* (*L.*) and *Lonicera maackii*, dominated by

trees with shrubs, with overall flat terrain and undulating in some areas.

In the field measurements, we take tape measure data as reference values to evaluate the accuracy of our method. The distribution of measurement data is shown in Fig. 7. In the transition woodland/shrub, 129 trees were measured, with the largest proportion of trees with SD ranging from 7.50 cm to 15.50 cm. 355 trees were measured in the mixed forest, with a height of 1.30 m and the largest proportion of trees with SD ranging from 11.50 cm to 61.50 cm (Fig. 8). 42 trees were measured in the green urban area, with different types of trees measured at different heights. The largest proportion of SD was between 11.50 cm and 31.50 cm (Fig. 8). The operator keeps the
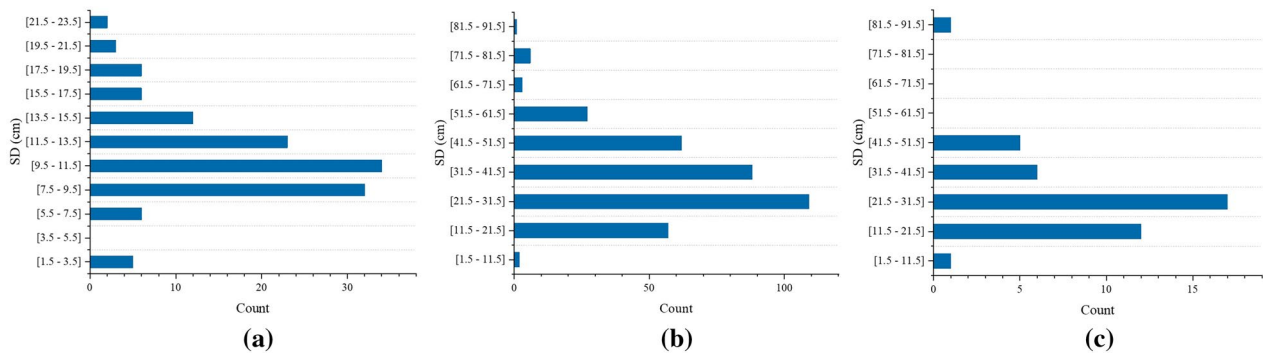

Fig. 8 The frequency distribution of stem diameters in field measurements, where (**a**–**c**) correspond to three areas of transitional woodland/shrub, mixed forest, and green urban area, respectively
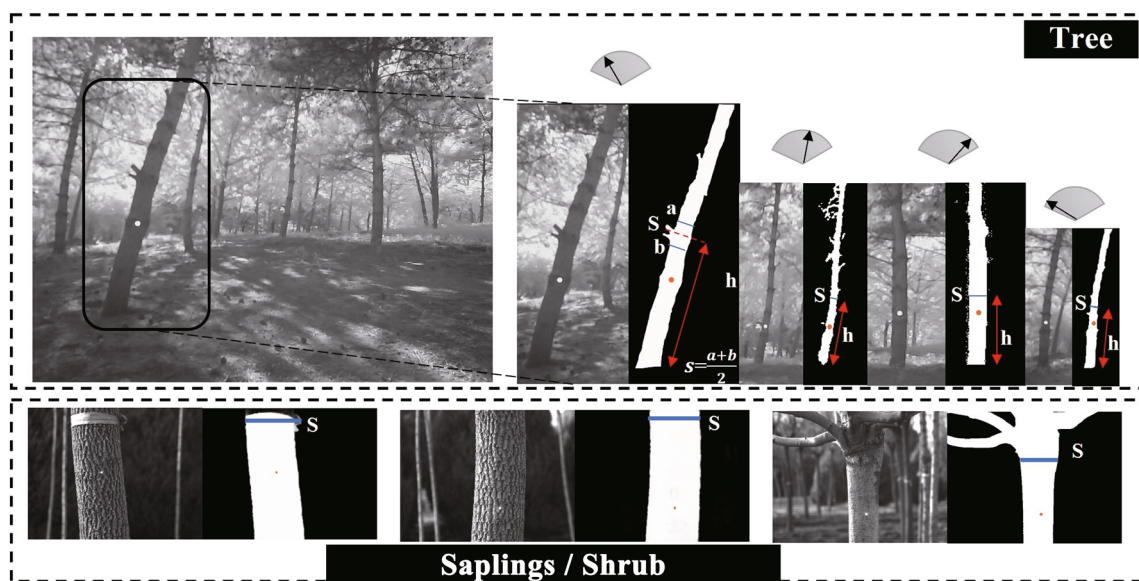

Fig. 9 Illustration of measurements of our method for different stem diameters, where the shade on the trunk affects the calculation results. The rotation of the laser module creates visible spots on different trees

acrylic sheet at the bottom of the device parallel to the ground. The laser module and image sensor are controlled by stepper motors rotating horizontally at a modest speed, allowing multiple trees to be measured at one site (Fig. 9). In transition woodland/shrub, we mark the measurement locations manually on the touch screen. The analysis module in our method calculates the SD of saplings or multi-stemmed trees combining the measurement locations. In mixed forest and green urban area, we do not need to mark measurement locations.

Evaluation metrics

In this study, the primary interest is to evaluate the performance of our method in different scenarios. For this purpose, different experiments were designed to verify the effectiveness of our method process and results. We introduce structural measurement (S_m) to evaluate structural similarity [46], enhanced-alignment measurement (E_m) to evaluate global statistics and local pixel matching information [47], and F-measure (F_m) to evaluate image-level accuracy. F-measure calculation is shown in Eq. 11. The precision P and recall R are two metrics widely used in computing to evaluate the quality of results. P is the precision of a model, while R reflects the completeness of a model. F_m is a weighted summation average of P and R. β represents parameter.

$$F_m = (1 + \beta^2) \times \frac{P \times R}{\beta^2 \times P + R} \tag{11}$$

In addition, Eq. 12 is introduced to improve the adequacy of evaluation measures for trunk contour images [48], where ω is the weight.

$$F_m^\omega = (1 + \beta^2) \times \frac{P^\omega \times R^\omega}{(\beta^2 \times P^\omega) + R^\omega} \tag{12}$$

To compare the differences between reference and measured values, relative error (RE), root mean square error ($RMSE$), mean squared error (MSE), mean absolute error (MAE), and the coefficient of determination (R^2) are calculated by the Eq. 13. The RE is the ratio of the absolute error of the measurement to the actual SD S_i multiplied by 100%. The number of samples is *num*. S_i denotes the i-th calculated SD. $\widehat{S}$ denotes reference values.

$$\begin{cases} RE = \frac{|S_i - \widehat{S}|}{S_i} \times 100 \\ MSE = \frac{1}{num} \sum_{i=1}^{num} (S_i - \widehat{S})^2 \\ RMSE = \sqrt{\frac{\sum_{i=1}^{num}(S_i - \widehat{S})^2}{n}} \\ MAE = \frac{1}{num} \sum_{i=1}^{num} |(S_i - \widehat{S})| \\ R^2 = 1 - \frac{\sum_i (\widehat{S} - S_i)^2}{\sum_i (\bar{S} - S_i)^2} \end{cases} \tag{13}$$

Results

Trunk identification evaluation

In the accuracy evaluation results of model, U^2-Net results in MAE is 0.99×10^{-2}, S_m is 0.96, F_m^ω is 0.96, and mean E_m is 0.97. Our improved model has MAE of 0.76×10^{-2}, S_m of 0.98, F_m^ω of 0.99, and mean E_m of 0.99 (Fig. 10). Among the 526 samples obtained in the field measurements, only four samples in which occlusion affected the accuracy of trunk contour identification, further affecting the results of SD measurements (Fig. 10). Overall, most of the samples can be accurately identified, with only four images producing significant deviations of +23 pixels, -25 pixels, +27 pixels, and +24 pixels, respectively.

Field measurements

We compared the measurement deviations in different plots (Fig. 11, Table 2). The SD in the transition forest area/shrub ranged from 2.00 cm to 24.00 cm, with a MAE of 0.32 cm. Eight trees out of 129 had an absolute deviation greater than 0.80 cm and the minimum deviation was 0.11 cm. The SD in the mixed forest ranged from 10.00 cm to 82.00 cm, with MAE of 0.38 cm. The maximum absolute deviation was 1.34 cm out of 355 trees, with the minimum absolute deviation being 0.10 cm. The SD of one tree in the green urban area was 89.00 cm. The other trees ranged from 8.00 cm to 48.00 cm, with a MAE of 0.39 cm, with no significant difference.

To evaluate the correlation between the measured and reference values, linear regression and correlation analyses were done in different plots (Fig. 12). The results showed that the measured and reference values were significantly correlated. The MAE, MSE, $RMSE$ and
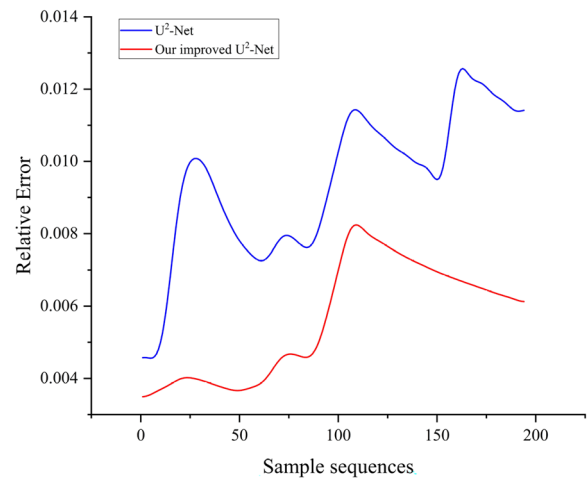

Fig. 10 Comparison of the relative error of U^2-Net and our improved U^2-Net

Wang *et al. Plant Methods* (2023) 19:68
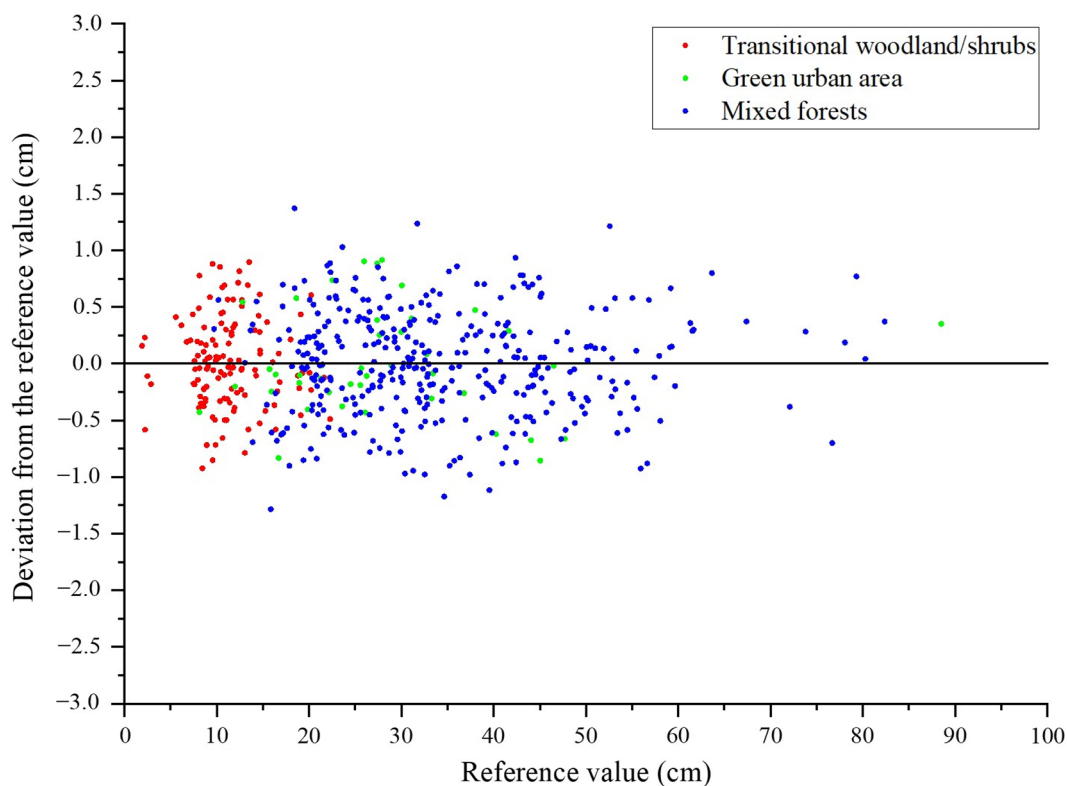
Page 11 of 16

Fig. 11 Deviation between the measured SD and the reference value measured by the tape measure

R^2 indicated that the measured values were close to the reference values (Table 2). We also excluded the four samples with significant differences and performed independent sample t-test on the overall sample. p was less than 0.01 at significance level equal to 0.05. There was no significant difference between the measured and reference values.

Optimal spot size
To verify the effect of spot size on measurement accuracy, a standard cylinder of 12 cm diameter was tested using different spot sizes. We found that spot sizes of 3 mm and 30 mm resulted in the lowest relative error (Fig. 13). With this in mind, we use a 3 mm spot size laser module in transitional woodland/shrub. For mixed forest and green urban area, we use laser modules with the 30 mm spot size, as most trees are larger than 10 cm.

Discussion
Characteristics of the device
We developed a device based on image sensor and laser module to estimate SD of individual tree and evaluate the accuracy of measurements by comparing their corresponding field measurements. Different from the work of Fan et al. [49] and Song et al. [37], the collimated beam

emitted by the laser module keeps the spot shape constant in natural scenes and can be used as a reference and anchor point. Image sensor can retain texture and shape features in images. This study evaluated the accuracy of the measurement method in different natural scenarios. The results show that our device can automatically locate the measurement position and measure SD with high accuracy.

Our device is designed not to preserve the color characteristics and spatial relationship characteristics of the image. Many image-based devices measure spatial relationships among trees and tree height in addition to DBH [36, 50, 51]. Some studies use the color features of the images for tree species recognition [52, 53]. In contrast to these studies, our device focuses on SD measurements.

Influence of the natural scenarios on the method
The two key steps in measuring SD in the field are localizing and estimating the diameter. Fan et al. [49] determined the measurement position by manually obtaining the base point of the target tree, while Wu et al. [36] determined it by using the spatial relationship features of the image and the complex algorithm of coordinate system transformation. In transitional woodland/shrub, the spot with 3.00 mm diameter was adopted as a reference
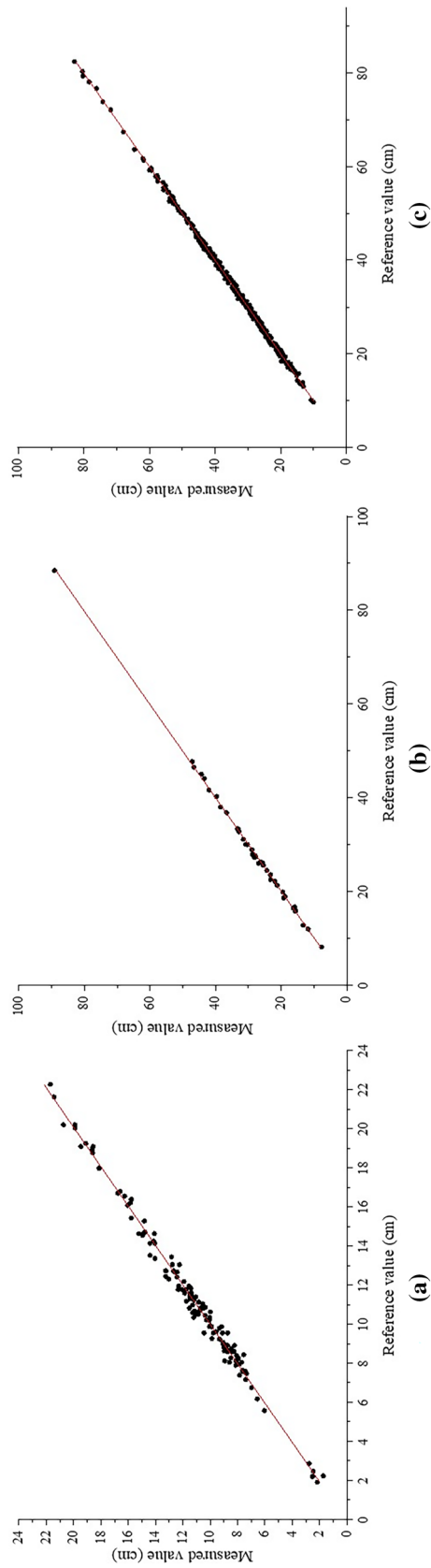
Fig. 12 Comparison of the accuracy of SD measurements under different plots. **a** Transitional woodland/shrub. **b** Mixed forest. **c** Green urban area

Wang *et al. Plant Methods*   (2023) 19:68

Page 13 of 16

Table 2 Measurement accuracy at different measurement locations

	MAE (cm)	MSE (cm)	RMSE (cm)	R^2
Transitional woodland/shrub	0.32	0.16	0.40	0.9947
Mixed forest	0.38	0.21	0.46	0.9994
Green urban area	0.39	0.22	0.47	0.9993
All stems	0.36	0.20	0.45	0.9995
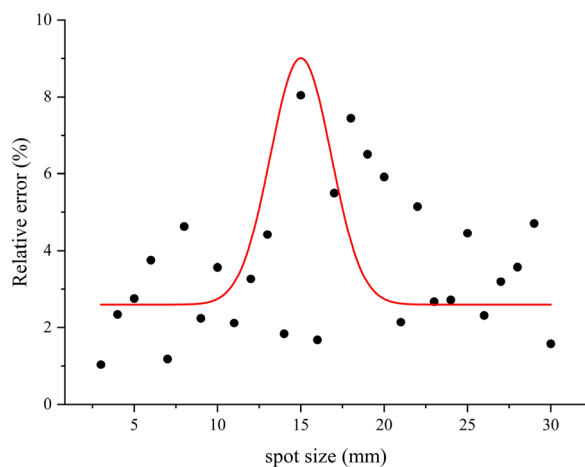
Fig. 13 Effect of different spot sizes on the measurement accuracy of a standard cylinder with a diameter of 12 cm

to calculate the SD. The measurement position was determined using the significant change in trunk contour width. For mixed forests and green urban areas, the measurement location was determined based on a linear relationship between spot diameter (30.00 mm) and measurement height. In addition, the spots with wavelengths of 680 nm and 980 nm were compared during the measurement. The spot with a wavelength of 980 nm could fit the perception of a digital camera. The spot with a wavelength of 680 nm was difficult to avoid the interference of direct sunlight.

Uneven lighting conditions on trees can lead to obvious errors in tree contour extraction. Only texture and shape features of the image are retained in the panchromatic images, which can well avoid this problem. In addition, bark surface grooves, unevenness, color differences, and ambient lighting conditions are all factors that affect the ability of the algorithm to extract salient image features. Our improved U^2-Net extracted the contour of the tree, the u-shaped structure could keep the output scale and input scale consistent. A-RSU module could overcome the influence of similar texture on model extraction (Fig. 10).

The analysis of major error sources

As can be visually observed in Table 2, Figs. 11, 12, there is a strong relationship between the reference and measured values, where *MAE* is 3.60 mm and *RMSE* is 4.50 mm. In many previous studies, the results of SD calculations varied widely. Bayati et al. [54] performed 3D reconstruction of trees using near-earth photography combined with remote sensing and computer vision technology. The *RMSE* of measured DBH was 52.00 mm and relative bias was 0.28. Mulverhill et al. [55] obtained highly overlapping spherical images from different locations by two 12-megapixel cameras to generate high-quality point clouds for each target location. The *RMSE* of measured DBH was estimated to be 48.00 mm and relative bias to be 0.20. Ucar et al. [56] estimated the DBH of trees based on the iPhone X (Apple Inc.) and measurement app, and the average deviation is 3.60 mm at a distance of 1.50 m from the tree. Wu et al. [36] photographed multiple trees with smart phones and extracted tree contours by visual segmentation method. The measured DBH produced *RSME* of 2.17 mm and *MAE* of 15.10 mm. Song et al [37]. constructed an integrated device of digital camera and LiDAR and *RMSE* of measured DBH of 3.07 mm and bias of 0.06 mm. The accuracy of our method is consistent with most image-based measurement methods (Fig. 11).

The data in this study were obtained from natural scenarios, and 96.60% of the DBH values were between 2.00 cm and 60.00 cm. Among them, 0.01% had a deviation greater than 1.00 cm (Fig. 11). The error would not increase with increasing SD. The overall trend was consistent with many studies [36, 37, 57]. The uncertainty of the measurement error comes largely from the shading object. In addition, the diffraction of light and the angle of the beam to the trunk are also factors that affect the accuracy of the measurement.

The detection of target tree contour is also a crucial factor affecting the measurement accuracy. U^2-Net can process high resolution images with lower memory and computing costs to make the network deeper and provide excellent performance in complex environments. In our study, the U^2-Net can extract tree contours accurately. However, when it comes to processing semantic information, U^2-Net achieves sub-optimal performance. When there are branches, leaves or other shielding objects on the target tree, U^2-Net will have obvious object detection errors. In order to solve this problem, we added an *Attach* module to enhance its attention mechanism through the spot location of SDA. In addition, the number of training samples is also a factor affecting significant target detection. Although 1,600 images were used for training, the limited number of samples limited the detection capability of our improved U^2-Net. In the future, more data of

different tree species should be collected and attention mechanism should be further developed.

Application scenarios of the device

In the field measurement, factors such as weight and economic cost of the device are important factors in device selection. In this section, we compare the weight, economic cost, integration, and measurement range of the method and provide some suggestions for application scenarios.

Weight of device. The total weight of our device is 0.28 kg. The weight of LiDAR is between 0.20 kg and 15.00 kg. The weight of digital camera ranges from 0.30 kg to 0.80 kg. In previous work, the weight of the device of Ma et al. [58] is about 3.00 kg. The device of Song et al. [37] is only about 0.60 kg. The weight of our device is minimal relative to these devices, making it easy for foresters to measure and carry.

Economic cost. The cheapest measuring device used in nursery survey work is a tape measure, which usually sells for about 5.00 dollars. The price of LiDAR tends to be proportional to the measurement accuracy and range, usually above 23,000.00 dollars. The total cost of the device for Song et al. [37] was about 2,000.00 dollars. The economic cost of passive measurement methods using smartphones, on the other hand, is between about 300.00 dollars and 3,000.00 dollars [36, 58]. The hardware cost of our device is $104.00 and the core module is $59.00.

Device integration. The image sensor and laser module are both economical and integrated. The model size of U^2-Net is 4.70 MB, the inference speed is 40 FPS, and the SDA is only 0.50 MB. The method meets the conditions and limitations of running on a development board or Raspberry Pi. Further weight reduction and significant cost reduction will be achieved in the future with the mass production of the equipment.

Measuring range. The device is able to work under different lighting conditions and complex environments, and the working distance of 1.50 m to 10.00 m can ensure the accuracy of the device in measuring SD. For saplings, shrubs and multi-stemmed trees, close measurements are best. For measuring DBH of trees, measuring at a distance of 5.00 m from the target tree can ensure a balance of accuracy and efficiency.

Application scenarios. Our device reduces operational or computational errors and significantly improves measurement efficiency compared to conventional tape. Unlike LiDAR and other photogrammetric means, our device does not measure tree SD at the plot level. Based on the explicit purpose of the device development, our device is focused on SD measurements of trees in forests. It is worth mentioning that the device can be applied to similar forestry or agricultural scenarios, such as SD measurement of trees along urban roads [59].

Limitations

Our method is not adapted to large scale measurement of SD of trees. Moreover, the measurement of stem diameter of irregular trees is a challenge, such as trees that lean or have roots above the ground. In addition, the effect of the physical properties of the laser on the measurement accuracy is a topic that deserves to be explored.

Conclusion

In this paper, we presented a novel device and method to measure SD of trees in natural scenes based on deep learning and a low-cost laser module. We explored the generality and automation of the method in depth and compared it with conventional methods. Our method requires less human intervention and can perform automatic SD measurement without touching the trees. We proposed an algorithm for spot detection to provide reference information for SD calculations and improved the U^2-Net to more accurately determine the linear relationship between spot and trunk contour. The field measurement results showed that our method achieved acceptable accuracy with *MAE*, *MSE*, and *RMSE* of 0.36 cm, 0.20 cm, and 0.45 cm, respectively. However, some factors such as shading still limited the measurement accuracy of the method. In the future, we will explore more rigorous image-based measurement models and obtain more forest structure parameters.

Abbreviations

SD	Stem diameter
DBH	Diameter at breast height
3D	Three-dimension
2D	Two-dimension
LiDAR	Light detection and ranging
TLS	Terrestrial laser scanning
SDA	Spot detection algorithm

Author contributions

Conceptualization, SW; methodology, SW; software, SW; validation, RL, SW and HL; formal analysis, HL and SW; investigation, XM; resources, RL and HL; data curation, QJ; writing—original draft preparation, SW; writing—review and editing, SW and FX; visualization, SW and FX; supervision, HF; project administration, FX and HF; funding acquisition, FX and HF. All authors have read and agreed to the published version of the manuscript.

Wang *et al. Plant Methods*　　(2023) 19:68

Page 15 of 16

Availability of data and materials
Part of the datasets will be made publicly available upon acceptance of the study, but not for commercial use. The authors can provide the full data if reasonably requested.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Schlich W. Manual of forestry, vol. iii—forest management. London: Bradbury, Agnew, & Co. Ld.; 1895. p. 33–40.
2. Magarik YAS, Roman LA, Henning JG. How should we measure the DBH of multi-stemmed urban trees? Urban For Urban Greening. 2020;47: 126481.
3. Clough B, Dixon P, Dalhaus O. Allometric relationships for estimating biomass in multi-stemmed mangrove trees. Aust J Bot. 1997;45(6):1023–31.
4. Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, Henry M, Martínez-Yrízar A, Mugasha WA, Muller-Landau HC, Mencuccini M, Nelson BW, Ngomanda A, Nogueira EM, Ortiz-Malavassi E, Pélissier R, Ploton P, Ryan CM, Saldarriaga JG, Vieilledent G. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Change Biol. 2014;20(10):3177–90.
5. West PW. Stem diameter. In: West PW, editor. Tree and forest measurement. Berlin, Heidelberg: Springer; 2004. p. 13–8.
6. Tietema T. Biomass determination of fuelwood trees and bushes of Botswana, Southern Africa. For Ecol Manag. 1993;60(3):257–69.
7. Bukoski JJ, Broadhead JS, Donato DC, Murdiyarso D, Gregoire TG. The use of mixed effects models for obtaining low-cost ecosystem carbon stock estimates in Mangroves of the Asia-Pacific. PLoS ONE. 2017;12(1):0169096.
8. Weiskittel AR, Hann DW, Kershaw JA Jr, Vanclay JK. Forest Growth and Yield Modeling. New Haven: John Wiley & Sons; 2011.
9. Paul KI, Larmour JS, Roxburgh SH, England JR, Davies MJ, Luck HD. Measurements of stem diameter: implications for individual- and stand-level errors. Environ Monit Assess. 2017;189(8):1–14.
10. Clark NA, Wynne RH, Schmoldt DL. A review of past research on dendrometers. For Sci. 2000;46(4):570–6.
11. Drew DM, Downes GM. The use of precision dendrometers in research on daily stem size and wood property variation: a review. EuroDendro 2008: The long history of wood utilization. 2009;27(2):159–72.
12. Luoma V, Saarinen N, Wulder MA, White JC, Vastaranta M, Holopainen M, Hyyppä J. Assessing precision in conventional field measurements of individual tree attributes. Forests. 2017;8(2):38.
13. Elzinga C, Shearer RC, Elzinga G. Observer variation in tree diameter measurements. West J Appl For. 2005;20(2):134–7.
14. Kitahara F, Mizoue N, Yoshida S. Effects of training for inexperienced surveyors on data quality of tree diameter and height measurements. Silva Fennica. 2010;44(4):657–67.
15. Berger A, Gschwantner T, Gabler K, Schadauer K. Analysis of tree measurement errors in the Austrian National Forest Inventory. Aust J For Sci. 2012;129(3–4):153–81.
16. Butt N, Slade E, Thompson J, Malhi Y, Riutta T. Quantifying the sampling error in tree census measurements by volunteers and its effect on carbon stock estimates. Ecol Appl. 2013;23(4):936–43.
17. Bitterlich W. Die Winkelzählprobe. Forstwissenschaftliches Centralblatt. 1952;71(7):215–25.
18. Weaver SA, Ucar Z, Bettinger P, Merry K, Faw K, Cieszewski CJ. Assessing the accuracy of tree diameter measurements collected at a distance. Croatian J For Eng. 2015;36(1):73–83.
19. Liu S, Bitterlich W, Cieszewski CJ, Zasada MJ. Comparing the use of three dendrometers for measuring diameters at breast height. South J Appl For. 2011;35(3):136–41.
20. Binot J-M, Pothier D, Lebel J. Comparison of relative accuracy and time requirement between the caliper, the diameter tape and an electronic tree measuring fork. For Chron. 1995;71(2):197–200.
21. Moran LA, Williams RA. Field note-comparison of three dendrometers in measuring diameter at breast height field note. North J Appl For. 2002;19(1):28–33.
22. Sun L, Fang L, Weng Y, Zheng S. An integrated method for coding trees, measuring tree diameter, and estimating tree positions. Sensors. 2020;20(1):144.
23. Robbins WC, Young HE. A field trial of optical calipers. For Chron. 1973;49(1):41–2.
24. Eller RC, Keister TD. The Breithaupt Todis Dendrometer—an analysis. South J Appl For. 1979;3(1):29–32.
25. Parker RC. Nondestructive sampling applications of the tele-Relaskop in forest inventory. South J Appl For. 1997;21(2):75–83.
26. Heinzel J, Huber MO. Tree stem diameter estimation from volumetric TLS image data. Remote Sens. 2017;9(6):614.
27. Newnham GJ, Armston JD, Calders K, Disney MI, Lovell JL, Schaaf CB, Strahler AH, Danson FM. Terrestrial laser scanning for plot-scale forest measurement. Curr For Reports. 2015;1(4):239–51.
28. Zhang C, Yang G, Jiang Y, Xu B, Li X, Zhu Y, Lei L, Chen R, Dong Z, Yang H. Apple tree branch information extraction from terrestrial laser scanning and backpack-LiDAR. Remote Sens. 2020;12(21):3592.
29. Gao Q, Kan J. Automatic forest DBH measurement based on structure from motion photogrammetry. Remote Sens. 2022;14(9):2064.
30. Liang X, Kankare V, Yu X, Hyyppä J, Holopainen M. Automated stem curve measurement using terrestrial laser scanning. IEEE Trans Geosci Remote Sens. 2014;52(3):1739–48.
31. Yu Y, Li J, Guan H, Wang C, Yu J. Semiautomated extraction of street light poles from mobile LiDAR point-clouds. IEEE Trans Geosci Remote Sens. 2015;53(3):1374–86.
32. Xie Y, Zhang J, Chen X, Pang S, Zeng H, Shen Z. Accuracy assessment and error analysis for diameter at breast height measurement of trees obtained using a novel backpack LiDAR system. For Ecosyst. 2020;7(1):33.
33. Eliopoulos NJ, Shen Y, Nguyen ML, Arora V, Zhang Y, Shao G, Woeste K, Lu Y-H. Rapid tree diameter computation with terrestrial stereoscopic photogrammetry. J For. 2020;118(4):355–61.
34. Marzulli MI, Raumonen P, Greco R, Persia M, Tartarino P. Estimating tree stem diameters and volume from smartphone photogrammetric point clouds. Forestry Int J For Res. 2020;93(3):411–29.
35. Gollob C, Ritter T, Kraßnitzer R, Tockner A, Nothdurft A. Measurement of forest inventory parameters with apple iPad pro and integrated LiDAR technology. Remote Sens. 2021;13(16):3129.
36. Wu X, Zhou S, Xu A, Chen B. Passive measurement method of tree diameter at breast height using a smartphone. Comput Electron Agric. 2019;163: 104875.
37. Song C, Yang B, Zhang L, Wu D. A handheld device for measuring the diameter at breast height of individual trees using laser ranging and deep-learning based image recognition. Plant Methods. 2021;17(1):67.
38. Zhu J, Xu Z, Fu D, Hu C. Laser spot center detection and comparison test. Photonic Sens. 2019;9(1):49–52.
39. Ma T-B, Wu Q, Du F, Hu W-K, Ding Y-J. Spot image segmentation of lifting container vibration based on improved threshold method and mathematical morphology. Shock Vib. 2021;2021:9590547.
40. Liu X, Lu Z, Wang X, Ba D, Zhu C. Micrometer accuracy method for small-scale laser focal spot centroid measurement. Opt Laser Technol. 2015;66:58–62. https://doi.org/10.1016/j.optlastec.2014.07.016.

41. Xin L, Xu L, Cao Z. Laser spot center location by using the gradient-based and least square algorithms. In: 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2013. p. 1242–1245. https://doi.org/10.1109/I2MTC.2013.6555612.

42. Wu Z, Li S, Chen C, Qin H, Hao A. Salient object detection via dynamic scale routing. IEEE Trans Image Process. 2022;31:6649–63. https://doi.org/10.1109/TIP.2022.3214332. arxiv:2210.13821.

43. Ke YY, Tsubono T. Recursive Contour-Saliency Blending Network for Accurate Salient Object Detection. In: 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), IEEE, Waikoloa, HI, USA; 2022. p. 1360–1370. https://doi.org/10.1109/WACV51458.2022.00143.

44. Qin X, Zhang Z, Huang C, Dehghan M, Zaiane OR, Jagersand M. U-2-net: going deeper with nested U-structure for salient object detection. Pattern Recogn. 2020;106: 107404.

45. Kingma DP, Ba J. Adam: a method for stochastic optimization. arXiv 2017. arxiv:1412.6980. https://doi.org/10.48550/arXiv.1412.6980.

46. Cheng M-M, Fan D-P. Structure-measure: a new way to evaluate foreground maps. Int J Comput Vision. 2021;129(9):2622–38.

47. Fan D-P, Gong C, Cao Y, Ren B, Cheng M-M, Borji A. Enhanced-alignment measure for binary foreground map evaluation. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence. IJCAI'18, AAAI Press; Stockholm, Sweden. 2018. p. 698–704.

48. Margolin R, Zelnik-Manor L, Tal A. How to Evaluate Foreground Maps. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014. p. 248–255.

49. Fan Y, Feng Z, Mannan A, Khan T, Shen C, Saeed S. Estimating tree position, diameter at breast height, and tree height in real-time using a mobile phone with RGB-D SLAM. Remote Sens. 2018;10(11):1845.

50. Iizuka K, Yonehara T, Itoh M, Kosugi Y. Estimating tree height and diameter at breast height (DBH) from digital surface models and orthophotos obtained with an unmanned aerial system for a Japanese Cypress (Chamaecyparis obtusa) forest. Remote Sens. 2018;10(1):13.

51. Mokroš M, Liang X, Surový P, Valent P, Čerňava J, Chudý F, Tunák D, Saloň Š, Merganič J. Evaluation of close-range photogrammetry image collection methods for estimating tree diameters. ISPRS Int J Geo Inf. 2018;7(3):93.

52. Tuominen S, Nasi R, Honkavaara E, Balazs A, Hakala T, Viljanen N, Polonen I, Saari H, Ojanen H. Assessment of classifiers and remote sensing features of hyperspectral imagery and stereo-photogrammetric point clouds for recognition of tree species in a forest area of high species diversity. Remote Sens. 2018;10(5):714.

53. Elmas B. Identifying species of trees through bark images by convolutional neural networks with transfer learning method. J Faculty Eng Arch Gazi Univ. 2021;36(3):1254–69.

54. Bayati H, Najafi A, Vahidi J, Gholamali Jalali S. 3D reconstruction of uneven-aged forest in single tree scale using digital camera and SfM-MVS technique. Scand J For Res. 2021;36(2–3):210–20.

55. Mulverhill C, Coops NC, Tompalski P, Bater CW. Digital terrestrial photogrammetry to enhance field-based forest inventory across stand conditions. Can J Remote Sens. 2020;46(5):622–39.

56. Ucar Z, Degermenci AS, Zengin H, Bettinger P. Evaluating the accuracy of remote dendrometers in tree diameter measurements at breast height. Croatian J For Eng. 2022;43(1):185–97.

57. Heo HK, Lee DK, Park JH, Thorne JH. Estimating the heights and diameters at breast height of trees in an urban park and along a street using mobile LiDAR. Landscape Ecol Eng. 2019;15(3):253–63.

58. Campos MB, Tommaselli AMG, Honkavaara E, Prol FDS, Kaartinen H, El Issaoui A, Hakala T. A backpack-mounted omnidirectional camera with off-the-shelf navigation sensors for mobile terrestrial mapping: development and forest application. Sensors. 2018;18(3):827.

59. Berveglieri A, Tommaselli A, Liang X, Honkavaara E. Photogrammetric measurement of tree stems from vertical fisheye images. Scand J For Res. 2017;32(8).

Publisher's Note