Potato Plant Phenotypic Parameter Extraction Based on Multi-Source Data: Postprint

Authors: Hu Songtao, Ruifang Zhai, Wang Yinghua, Liu Zhi, Zhu Jianzhong, Ren He, Yang Wanneng, Song Peng

Date: 2023-05-15T00:00:00+00:00

Abstract

Crops exhibit characteristics such as diverse structures and complex growing environments. RGB image data can faithfully reflect the texture and color features of plants, while 3D point cloud data contains volumetric information of crops. Integrating RGB images and 3D point cloud data to extract 2D and 3D phenotypic parameters of crops holds significant importance for methodological research in phenomics. This study takes potato as the research object, using an RGB camera and a laser scanner to collect RGB images and 3D laser point cloud data of 50 potato plants, respectively. The segmentation accuracy of four deep learning semantic segmentation methods—OCRNet, UpNet, PaNet, and DeepLab v3+—was compared, and the OCRNet network with higher accuracy was selected to implement semantic segmentation of potato top-view images. The Mean shift clustering algorithm workflow was optimized to accomplish individual plant segmentation of potato laser point clouds, and combined with Euclidean clustering and K-Means clustering algorithms to accurately segment stems and leaves of individual potato plant point clouds. Simultaneously, a strategy utilizing numbering to establish a one-to-one correspondence between individual potato RGB images and laser point clouds was proposed, and based on this, 8 2D phenotypic parameters and 10 3D phenotypic parameters were extracted from the same potato plant from both RGB images and laser point clouds, including maximum width, perimeter, area, plant height, volume, leaf length, and leaf width. Finally, three representative and easily measurable phenotypic parameters—leaf count, plant height, and maximum width—were selected for accuracy evaluation, with Mean Absolute Percentage Error (MAPE) values of 8.6%, 8.3%, and 6.0%, respectively, Root Mean Square Error (RMSE) values of 1.371 leaves, 3.2 cm, and 1.86 cm, respectively, and coefficient of determination R2 values of 0.93, 0.95, and 0.91, respectively. The accuracy evaluation results demonstrate that the extracted phenotypic parameters can accurately and efficiently reflect the growth status of potatoes, and that com-

bining potato RGB image data with 3D laser point cloud data can fully leverage the advantages of rich texture and color features from RGB images and volumetric information from 3D point clouds, achieving high-precision, non-destructive extraction of 2D and 3D phenotypic parameters of potato plants. The findings of this study can not only provide important technical support for potato cultivation and breeding, but also offer strong support for research based on phenotypic data.

Full Text

Extraction of Potato Plant Phenotypic Parameters Based on Multi-Source Data

HU Songtao[1], ZHAI Ruifang[1]*, WANG Yinghua[2], LIU Zhi[1], ZHU Jianzhong[1], REN He[1], YANG Wanneng[2],[3], SONG Peng[2],[3]

[1]College of Informatics, Huazhong Agricultural University, Wuhan, Hubei, China
[2]College of Plant Science & Technology, Huazhong Agricultural University, Wuhan, Hubei, China
[3]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei, China

---

Abstract

Crops exhibit diverse structural characteristics and complex growth environments. RGB image data can accurately reflect the texture and color features of plants, while three-dimensional point cloud data contains volumetric information about crops. Combining RGB images and 3D point cloud data enables the extraction of both two-dimensional and three-dimensional phenotypic parameters, which is of great significance for phenomics research. This study focused on potato plants, using cameras and laser scanners to collect RGB images and 3D laser point cloud data, respectively. Four deep learning semantic segmentation methods were compared, and OCRNet was selected for semantic segmentation of potato top-view images due to its higher accuracy. The clustering algorithm workflow was optimized, with Mean Shift completing single-plant segmentation of potato laser point clouds, while Euclidean clustering and K-Means clustering algorithms accurately segmented stems and leaves from individual potato plant point clouds. Additionally, a strategy was proposed to establish a one-to-one correspondence between individual potato plant images and laser point clouds using plant numbering. Based on this correspondence, 8 two-dimensional phenotypic parameters and 10 three-dimensional phenotypic parameters were extracted from the same potato plant, including maximum width, perimeter, area, plant height, volume, leaf length, and leaf width. Finally, three representative and easily measurable phenotypic parameters—leaf number, plant height, and

maximum width—were selected for accuracy evaluation. The Mean Absolute Percentage Error (MAPE) values were 8.6%, 8.3%, and 6.0%, respectively; the Root Mean Square Error (RMSE) values were 1.371 leaves, 3.2 cm, and 1.86 cm, respectively; and the coefficients of determination (R^2) were 0.93, 0.95, and 0.91, respectively. The accuracy evaluation results demonstrate that the extracted phenotypic parameters can accurately and efficiently reflect potato growth status. Combining RGB image data with 3D laser point cloud data fully leverages the advantages of rich texture and color features from images and volumetric information from 3D point clouds, achieving high-precision, non-destructive extraction of both 2D and 3D phenotypic parameters for potato plants. The research outcomes can provide important technical support for potato cultivation and breeding, as well as strong support for phenotypic data-based research.

Keywords: LiDAR; multi-source data; clustering segmentation; 3D phenotyping; OCRNet; laser point cloud; deep learning

1. Introduction

Crop phenotypes are used to quantitatively evaluate various plant traits and are of great significance for modern agricultural production. Early crop phenotyping was performed manually—for example, using tape measures for plant height and water displacement methods for volume measurement. These methods were not only inefficient but also involved destructive measurements, making continuous monitoring of crop development difficult. Therefore, exploring an efficient, non-destructive method that can utilize different data sources to obtain plant phenotypic parameters has become a key research direction in crop phenotyping. Among various approaches, methods based on RGB images and laser point clouds have been widely applied.

In high-throughput crop phenotyping, RGB images can reflect plant morphology while containing rich color and texture features, making them the preferred data source for phenotypic parameter extraction. Germain et al. distinguished gray value differences in images to extract wheat ears from 2D images with good results. Wang et al. used local texture energy filters to obtain corn leaf edge contours and successfully segmented corn leaves from images. Loresco et al. employed K-nearest neighbor algorithms to segment lettuce images in four different color spaces (RGB, CIELab, YCbCr, and HSV), finding that the CIELab color space provided the best segmentation performance for crop stage identification in smart farm applications.

In recent years, deep learning has become a popular method for image processing, overcoming the limitations of traditional image processing methods that require manual feature design and extraction while providing higher accuracy and universality. Grinblat et al. used Convolutional Neural Networks (CNNs) to identify and classify leaves of three legume species (white beans, soybeans, and red beans), achieving an overall accuracy of 96.9%. Lee et al. similarly used

CNNs to successfully train a model for effectively distinguishing 44 plant species, with accuracy reaching up to 99.5%. Numerous excellent network models for pixel-level segmentation and object detection have emerged and been widely applied in plant phenotyping. Fan et al. deployed the YOLOv4 model on a field phenotyping robot to measure corn stalk diameter. Wang et al. used transfer learning and Mask R-CNN models to detect and segment watermelon plants.

Point cloud data has also attracted attention from phenotyping researchers due to its rich geometric features and minimal environmental influence. There are two acquisition methods: active 3D measurement (e.g., laser scanning) and passive 3D measurement (reconstruction from image sequences). Laser point clouds offer advantages including high efficiency, high precision, and more uniform point distribution compared to reconstructed point clouds. Processing algorithms for point clouds are crucial for extracting crop phenotypic parameters, with segmentation being particularly important. Wang et al. used a Kinect v2.0 device to extract plant point clouds and applied a concavity-based clustering segmentation algorithm to complete leaf segmentation for trait measurement. Yang et al. used a handheld laser scanner to acquire point cloud data of cotton plants at 12 time points, achieving dynamic measurement of phenotypic parameters including plant height, leaf length, leaf width, and volume. These methods require small data volumes and have simple algorithm structures, making them widely applicable.

Potato is a versatile crop with high water use efficiency, a long industrial chain, and large planting area, playing an important role in alleviating food security pressure. In recent years, potato phenotypic parameter extraction has gradually attracted attention from domestic researchers. For RGB images, Zhang et al. improved the YOLOv4 model to achieve recognition and counting of potato tubers in complex environments. Zhao et al. used the Faster R-CNN network for potato leaf disease detection. For laser point clouds, Tan et al. used depth cameras to acquire potato tuber point clouds and obtained 3D positioning of potato bud eyes.

To combine the advantages of RGB images and laser point clouds, this study used RGB cameras and laser scanners to collect potato RGB images and point cloud data, extracting 18 phenotypic parameters from individual plants. The accuracy evaluation results demonstrate that combining RGB images and laser point clouds can simultaneously obtain texture and morphological features of the same potato plant while overcoming the limitations of single data modalities.

2. Materials and Methods

2.1 Technical Route This study used RGB cameras and laser scanners to collect potato RGB images and point cloud data, establishing a correspondence between them through plant numbering to extract phenotypic parameters from individual potato plants. The overall technical workflow is shown in Figure 1 [Figure 1: see original paper].

2.2 RGB Image and Point Cloud Data Acquisition Potato plants were potted in a greenhouse at the Huazhong Agricultural University phenotyping platform in April 2022, with 800 pots (one plant per pot). Data collection began after growth stabilized. This study collected RGB images and point cloud data from 50 potato plants for phenotypic parameter extraction. Additionally, 4,000 RGB images were collected from 750 potato plants on May 15 and May 19 for deep learning semantic segmentation network training.

2.2.1 RGB Image Acquisition Potato growth and development are closely related to the canopy, which is the primary site for photosynthesis and respiration and directly reflects plant status. Top-view RGB images captured from above maximize canopy feature acquisition. This study used an automated conveying system at the Huazhong Agricultural University crop phenotyping platform to periodically collect images in a darkroom environment [FIGURE:2(a)]. The RGB camera (Basler aca5472, 20 MP) was mounted at the top of the darkroom [FIGURE:2(b)], 1.5 m above the conveyor line to ensure complete coverage. The system consisted of three roller conveyors for moving plants into the darkroom, adjusting position for imaging, and moving them out.

2.2.2 Laser Point Cloud Data Acquisition Laser point cloud data were collected in groups of 10 plants across 9 stations arranged from low to high [FIGURE:2(c)]. To reduce occlusion of stem and lower leaf point clouds, five stations had laser emitters parallel to the plants (blue squares in [FIGURE:2(d)]), while four stations were slightly elevated (green squares in [FIGURE:2(d)]). Scanner parameters significantly affect results, requiring experimental optimization of resolution and quality settings. Using six potato plants, six control groups with different resolution/quality settings were scanned on the same day. Based on visual quality, registration accuracy, and time efficiency, a resolution of 1/4$×$976,000 points/s and quality of 4X provided good results with high efficiency.

2.3 Image Semantic Segmentation OCRNet is a semantic segmentation network based on fully convolutional networks [Figure 3: see original paper]. It uses object-contextual representation to transform complex pixel classification into higher-level classification problems, improving accuracy through three steps: (1) using a backbone network for semantic segmentation to obtain pixel and category features; (2) expanding category and pixel features into 2D vectors for fusion; and (3) calculating correlation vectors through self-attention mechanisms and concatenating them with pixel features to obtain object-category contextual information. OCRNet concentrates contextual information on objects, reducing interference and improving pixel classification accuracy, making it particularly suitable for complex potato RGB images with rich semantic information.

2.4 Point Cloud Segmentation

2.4.1 Single-Plant Segmentation The Mean Shift algorithm is sensitive to data density and was optimized for this study. Direct application can misclassify sparse middle regions as separate layers. The optimized workflow projects point clouds onto the XoY plane parallel to the ground, creating a 2D discrete point set for clustering before restoring to 3D, avoiding misclassification. This optimization reduced drift iterations by 45,432 and saved 195,180 ms compared to the original method.

Mean Shift requires setting an iteration threshold (typically 0.00001) and search radius. An excessively large radius merges different clusters, while too small a radius splits clusters incorrectly. Parameter validation experiments determined that a search radius of 0.08 yielded optimal segmentation results.

2.4.2 Organ Segmentation Accurate leaf and stem segmentation enables precise extraction of leaf phenotypic parameters without plant destruction. Potato plants exhibit distinct characteristics: lower growth point leaves are evenly distributed with large stem-leaf spacing and similar density, while upper leaves are dense with small spacing and significant noise. Therefore, organ segmentation was performed in two steps [Figure 5: see original paper]:

Step 1: Euclidean Cluster Extraction Algorithm, sensitive to distance, segmented lower leaves. Parameters include neighbor search radius, minimum cluster points, and maximum cluster points. Minimum/maximum points exclude noise (maximum $>$ half of total points, minimum $\leq 10\%$). The neighbor search radius relates to point cloud density and uniformity. After validation, optimal parameters were: minimum 1,000 points, maximum 200,000 points, and radius of 5 mm.

Step 2: K-Means clustering segmented remaining upper leaves from stems. When leaf density features were obvious, one iteration with K = number of upper leaves + 1 sufficed. For dense canopies, multiple K-Means iterations were needed, with K set to remaining leaf count + 1 each time. This study used up to two K-Means iterations for upper leaf-stem separation.

2.5 Establishing Correspondence Between RGB Images and Laser Point Clouds RGB images and laser point clouds were collected asynchronously to improve efficiency and flexibility. A numbering strategy established one-to-one correspondence: RGB images were collected sequentially by number using the automated system, while point cloud numbers were extracted through:

1. Color threshold filtering (RGB average in range 140-180) [FIGURE:6(b)]
2. Clustering algorithm to locate number tags [FIGURE:6(c)]
3. Calculating average normal vector of tag point cloud [FIGURE:6(d)]
4. Projecting tag point cloud onto plane perpendicular to average normal vector [FIGURE:6(e)]

5. Image processing of 2D projection to extract digital numbers [FIGURE:6(f-g)]

Traditional binarization uses fixed thresholds and cannot handle uneven illumination [FIGURE:7(b)]. The Sauvola algorithm was adopted for local adaptive binarization based on neighborhood gray values and standard deviation, effectively suppressing shadow effects [FIGURE:7(c)]. The dynamic threshold is calculated as:

$$m(x,y) = \frac{1}{r^2} \sum_{i=x-r/2}^{x+r/2} \sum_{j=y-r/2}^{y+r/2} \text{gray}(i,j)$$

$$s(x,y) = \sqrt{\frac{1}{r^2} \sum_{i=x-r/2}^{x+r/2} \sum_{j=y-r/2}^{y+r/2} [\text{gray}(i,j) - m(x,y)]^2}$$

$$T(x,y) = m(x,y) \times \left[1 + 0.1 \times \left(\frac{s(x,y)}{R} - 1\right)\right]$$

where r is the neighborhood range, R is the dynamic range (128 for 8-bit grayscale), $m(x,y)$ is the mean gray value, and $s(x,y)$ is the standard deviation.

Deformed number contours during projection were recognized using the classic MNIST dataset and AlexNet (5 convolutional layers, 3 pooling layers, 3 fully connected layers with ReLU activation, Dropout, and local response normalization).

2.6 Extraction of Potato Plant Phenotypic Parameters Eighteen phenotypic parameters related to canopy, main stem, and leaves were extracted from corresponding RGB images and laser point clouds:

2D Morphological Features: Canopy Width (CW), Canopy Perimeter (CP), Canopy Area (CA), Convex Hull Perimeter (CHC), Convex Hull Area (CHA)

2D Texture Features: Mean Intensity (MI), Smoothness (SE), Standard Deviation (S)

3D Morphological Features: Plant Height (PH), Plant Width (PW), Plant Depth (PD), Convex Hull Volume (PV), Cylinder Volume (CV), Leaf Number (LN), Leaf Length (LL), Leaf Width (LW), Leaf Area (LA), Projection Area (PA)

2D Morphological Features: Extracted from contours [FIGURE:8(a) green line], convex hulls [FIGURE:8(a) light blue], and circumscribed circles [FIGURE:8(a) dark blue] for perimeter, area, convex hull perimeter/area, and maximum width.

2D Texture Features: Calculated using gray-level histograms to reflect roughness, directionality, and regularity:

$$M = \sum_{i=0}^{L-1} z_i P(z_i)$$

$$\sigma^2(z) = \sum_{i=0}^{L-1} (z_i - M)^2 P(z_i)$$

$$R(z) = 1 - \frac{1}{1 + \sigma^2(z)}$$

where L is the gray-level range, z_i is the gray value, and $P(z_i)$ is the histogram frequency.

3D Morphological Features: Extracted from minimum bounding boxes [FIGURE:8(b)], circumscribed cylinders [FIGURE:8(c)], convex hulls [FIGURE:8(d)], and XY-plane projections [FIGURE:8(e)]. Leaf length and width were measured using an approximation method [FIGURE:8(f)]: endpoints A and B were selected manually with distance L, midpoint C was selected, and distances D1 (A-C) and D2 (B-C) were calculated. The process iterated until the deviation between L and (D1+D2) was below threshold d.

Accuracy Evaluation: MAPE, RMSE, and R^2 were calculated against manual measurements of leaf number, plant height, and maximum width.

3. Results

3.1 Laser Scanner Parameter Settings Six control groups with different resolution/quality settings were tested. While 1/2$×976, 000points/sprovidedgoodresults, itsufferedfromdat points/s with 4X quality provided optimal density, morphological features, and efficiency [Figure 9: see original paper].

3.2 RGB Image Semantic Segmentation Four deep learning networks (OCRNet, DeepLab v3+, UPerNet, PANet) were compared using 1,600 top-view RGB images (800$×$800 pixels, 1,200 for training/validation, 400 for testing). OCRNet achieved the best performance in mean pixel accuracy and target pixel accuracy, though slightly lower in MIoU than DeepLab v3+ . By focusing contextual information on objects, OCRNet reduced noise and improved classification accuracy for complex potato plant structures. It was selected for subsequent analysis, with segmentation results shown in [Figure 10: see original paper].

3.3 Plant Point Cloud Segmentation

3.3.1 Single-Plant Segmentation The optimized Mean Shift algorithm effectively avoided mis-segmentation [FIGURE:11(a-b)]. Search radius significantly affected results: too large caused cluster merging [FIGURE:11(c)], too small caused over-segmentation [FIGURE:11(d)]. A radius of 0.08 produced optimal results [FIGURE:11(e-f)].

3.3.2 Organ Segmentation Two-step segmentation was performed [Figure 12: see original paper]. Step 1 using Euclidean clustering with optimal parameters (min 1,000 points, max 200,000 points, 5 mm radius) successfully segmented lower leaves, leaving only upper leaves connected to stems [FIGURE:12(a)]. Step 2 using K-Means clustering completed the segmentation. For dense canopies, two K-Means iterations were required [FIGURE:12(e-f)].

3.4 Correspondence Between RGB Images and Laser Point Clouds Using plant numbering, RGB images were automatically matched to numbers via the conveying system [FIGURE:13(a)]. Number extraction from point clouds was implemented on a Windows workstation (Intel Core i7-10750H, 16 GB RAM, NVIDIA RTX 2060) using Python 3.7, Pytorch 1.7.1, and CUDA 11.1. AlexNet successfully recognized numbers from projected tag images [FIGURE:13(b-c)].

3.5 Accuracy Evaluation Results Algorithm-extracted values for leaf number, plant height, and maximum width were compared with manual measurements . Accuracy evaluation showed MAPE of 8.6%, 8.3%, and 6.0%; RMSE of 1.371 leaves, 3.2 cm, and 1.86 cm; and R^2 of 0.93, 0.95, and 0.91, respectively [Figure 14: see original paper]. These results demonstrate high accuracy and validate the feasibility of the proposed method.

4. Conclusion and Outlook

This study explored the feasibility of combining 2D optical imaging and 3D laser scanning technologies to efficiently extract phenotypic parameters. RGB cameras and laser scanners were used to obtain top-view RGB images and point cloud data, enabling extraction of 18 phenotypic parameters from the same potato plants.

First, an automated conveying system enabled efficient RGB image acquisition, while experimental optimization determined optimal laser scanner parameters and station configurations for complete, information-rich point clouds.

Second, a numbering-based correspondence method was developed, allowing extraction of 18 phenotypic parameters from both RGB images and laser point clouds. This integration leveraged the advantages of rich texture/color features from images and volumetric information from point clouds.

Finally, accuracy evaluation of three representative parameters (leaf number, plant height, maximum width) demonstrated high precision with R^2 values of 0.93, 0.95, and 0.91, validating the method' s feasibility and accuracy.

However, limitations exist: the method focused only on canopy, stem, and leaf parameters during tuber formation stage, neglecting other organs, tubers, and growth stages. Future work should collect multi-temporal data of plants, flowers, and tubers for dynamic phenotyping throughout the entire growth cycle. This multi-source approach can be extended to other crops (corn, wheat, rice) and integrated with additional sensors (thermal infrared, chlorophyll fluorescence, multispectral/hyperspectral imaging) to provide new avenues for crop phenotypic parameter acquisition.

Conflict of Interest Statement: The authors declare no conflicts of interest.

References

[1] WENG Y, ZENG R, WU C M, et al. A survey on deep-learning-based plant phenotype research in agriculture[J]. Scientia sinica (vitae), 2019, 49(6): 698-716.

[2] GERMAIN C, ROUSSEAUD R, GRENIER G. Non destructive counting of wheatear with picture analysis[C]// Fifth International Conference on Image Processing and its Applications. London, UK: IET, 2002: 435-439.

[3] WANG C Y, GUO X Y, DU J J, et al. Corn leaf image segmentation method and apparatus in field environment: CN106296662B[P]. 2019-07-02.

[4] LORESCO P J M, VALENZUELA I C, DADIOS E P. Color space analysis using KNN for lettuce crop stages identification in smart farm setup[C]// TENCON 2018—2018 IEEE Region 10 Conference. Piscataway, NJ, USA: IEEE, 2019: 2040-2044.

[5] GRINBLAT G L, UZAL L C, LARESE M G, et al. Deep learning for plant identification using vein morphological patterns[J]. Computers and electronics in agriculture, 2016, 127: 418-424.

[6] LEE S H, CHAN C S, WILKIN P, et al. Deep-plant: Plant identification with convolutional neural networks[C]// 2015 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2015: 452-456.

[7] FAN Z Q, SUN N, QIU Q, et al. A high-throughput phenotyping robot for measuring stalk diameters of maize crops[C]// 2021 IEEE 11th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER). Piscataway, NJ, USA: IEEE, 2021: 128-133.

[8] WANG X T, ZHAO Z, WANG Y, et al. Intelligent detection algorithm of plant phenotype based on improved Mask R-CNN[J]. Journal of Chinese agricultural mechanization, 2022, 43(8): 151-157.

[9] ZHANG J Y, ZHAO X L, CHEN Z, et al. A review of deep learning-based semantic segmentation for point cloud[J]. IEEE access, 2019, 7: 179118-179133.

[10] ZHOU J J, GUO X Y, WU S, et al. Research progress on plant three-dimensional reconstruction based on multi-view images[J]. Journal of agricul-

tural science and technology, 2019, 21(2): 9-18.

[11] WANG Y J, WEN W L, WU S, et al. Maize plant phenotyping: Comparing 3D laser scanning, multi-view stereo reconstruction, and 3D digitizing estimates[J]. Remote sensing, 2018, 11(1): ID 63.

[12] WANG Y W, CHEN Y F. Non-destructive measurement of three-dimensional plants based on point cloud[J]. Plants (basel, Switzerland), 2020, 9(5): 571.

[13] YANG X, HU S T, WANG Y H, et al. Cotton phenotypic trait extraction using multi-temporal laser point clouds[J]. Smart agriculture, 2021, 3(1): 51-62.

[14] QU D Y, XIE K Y, JIN L P, et al. Development of potato industry and food security in China[J]. Scientia agricultura Sinica, 2005, 38(2): 358-362.

[15] ZHANG Z G, ZHANG Z D, LI J N, et al. Potato detection in complex environment based on improved YOLOv4 model[J]. Transactions of the Chinese society of agricultural engineering, 2021, 37(22): 170-178.

[16] ZHAO Y, ZHAO H, JIANG Y C, et al. Detection method of potato leaf diseases based on deep learning[J]. Journal of Chinese agricultural mechanization, 2022, 43(10): 183-189.

[17] TAN Y, TIAN F, LIU X X, et al. Potato image acquisition device based on RGB-D cameras and method for identifying and locating bud eye: CN108830272A[P]. 2018-11-16.

[18] FENG H, YANG W N, YE J L, et al. Automatic two-channel conveying device for automatic hyperspectral imaging of rice plants and control method: CN107720213B[P]. 2020-10-27.

[19] YUAN Y H, CHEN X K, CHEN X L, et al. Segmentation transformer: Object-contextual representations for semantic segmentation[J/OL]. arXiv: 1909.11065, 2019.

[20] SONG L P. Segmentation and classification of 3D point cloud data in outdoor scenes[D]. Dalian: Dalian University of Technology, 2016.

[21] RUSU R B. Semantic 3D object maps for everyday manipulation in human living environments[J]. KI-Künstliche Intelligenz, 2010, 24: 345-348.

[22] HARTIGAN J A, WONG M A. A K-means clustering algorithm[J]. Journal of the royal statistical society: Series C (applied statistics), 1979, 28(1): 100-108.

[23] FANG M, XU J Y, WANG J P, et al. A new binarization algorithm for document image[J]. Journal of Hefei university of technology (Natural Science), 2001, 24(2): 166-169.

[24] LECUN Y, CORTES C, BURGES C J C. THE MNIST DATABASE of handwritten digits[EB/OL]. (1996-11-08) [2023-01-10]. http://yann.lecun.com/exdb/mnist

[25] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Commun ACM, 2017, 60(6): 84-90.

[26] XIONG X. Study on rice ear segmentation and lossless yield prediction of field rice based on deep learning[D]. Wuhan: Huazhong University of Science and Technology, 2018.

[27] YANG W L, DUAN L F, YANG W N. Deep learning-based extraction of rice phenotypic characteristics and prediction of rice panicle weight[J]. Journal of Huazhong agricultural university, 2021, 40(1): 227-235.

[28] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// European conference on computer vision. Berlin, Germany: Springer, 2018: 833-851.

[29] XIAO T T, LIU Y C, ZHOU B L, et al. Unified perceptual parsing for scene understanding[C]// European Conference on Computer Vision. Berlin, Germany: Springer, 2018: 432-448.

[30] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 8759-8768.

Note: Figure translations are in progress. See original paper for figures.