Plantorgan hunter: a deep learning-based framework for quantitative profiling plant subcellular morphology

Xuping Feng
Zhejiang University
Zeyu Yu
Zhejiang University
Hui Fang
Zhejiang University
Hangjin Jiang
Zhejiang University
Guofeng Yang
Zhejiang University
Liting Chen
Nanjing Agricultural University
Xinran Zhou
Nanjing Agricultural University
Bing Hu
Nanjing Agricultural University
Chun Qin
Nanjing Agricultural University
Gang Hu
Nanjing Agricultural University
Guipei Xing
Nanjing Agricultural University
Boxi Zhao
Nanjing Agricultural University
Yongqiang Shi
Zhejiang University
Yong He
Zhejiang University
Feng Liu (liufeng@njau.edu.cn)
Nanjing Agricultural University
Feng Liu

University of Melbourne

Bo Han

Hong Kong Baptist University https://orcid.org/0000-0002-6338-0958

Article

Keywords:

Posted Date: July 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1811819/v1

Plantorgan hunter: a deep learning-based framework for quantitative profiling plant subcellular morphology

Xuping Feng[1,3#], Zeyu Yu[1,3#], Huifang[2,3], Hangjin Jiang[4], Guofeng Yang [1,3], Liting Chen[5], Xinran Zhou[5], Bing Hu[6], Chun Qin[6], Gang Hu[6], Guipei Xing[6], Boxi Zhao[5], Yongqiang Shi[3], Yong He[3]*, Feng Liu[5]*, Feng Liu[7], and Bo Han[8]

[1] The Rural Development Academy &Agricultural Experiment Station, Zhejiang University, Hangzhou 310058, China

[2] Huzhou Institute of Zhejiang University, Huzhou 313000, Zhejiang, China;

[3] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China;

[4] Center for Data Science, Zhejiang University, Hangzhou 310058, China;

[5] College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, P. R. China.

[6] College of Life Sciences, Biological experiment teaching center, Nanjing Agricultural University, Nanjing 210095, P. R. China.

[7] School of Mathematics and Statistics, University of Melbourne

[8] Department of Computer Science, Hong Kong Baptist University

[#] Xuping Feng and Zeyu Yu contributed equally to this work.

[*]Correspondence and requests for materials should be addressed to Yong He (yhe@zju.edu.cn) or Feng Liu (liufeng@njau.edu.cn)

Abstract

Accurate delineation of plant cell organelles from electron microscope images is essential to understand subcellular behaviors and functions. Here, we develop a deep learning pipeline, organelle segmentation network (OrgSegNet) for pixel-wise segmentation to identify chloroplasts, mitochondria, nuclei, and vacuoles. OrgSegNet was evaluated on a large manually-annotated dataset of 6371 organelles collected from 13 plant species, and achieved a state-of-the-art segmentation performance of these organelles. To generalize the morphological characteristics of plant organelles, we defined three morphological metrics (shape-complexity, electron-density, and area), and released an open-source web tool "Plantorgan Hunter" allowing quantitative profiling of subcellular morphology. The functionalities of Plantorgan Hunter can be easily operated, and we believe that it will increase the efficiency and productivity of plant subcellular morphological characteristics for the plant science community.

Main

Cells are the basic functional units of plants and contain organelles that perform the various specific functions needed. The morphology and number of most organelles in each cell are dynamic. Some changes occur in the process of organelle biogenesis, segregation or senescence during the normal cell cycle, while other changes occur as an adaptive response to stress conditions to optimize organelle function. For example, chloroplasts, the sites of photosynthesis, can sense changes in environment conditions[1] and are quickly and seriously influenced. Death of the cell vacuole is often accompanied by a gradually decrease of cytoplasmic content and a concomitant increase in lytic vacuoles that engulf various cytoplasmic organelles[2]. Changes in the morphology of mitochondria reflect alterations in the cellular energy status because of changes in nutrient level, redox state, or darkness stress[3]. Alterations in the shape and size of nuclei were observed in response to various disease, as well as ploidy change and senescence[4]. Therefore, in order to better understand the behavior and function of cell, it is essential to accurately delineate the organelles and identify their morphological changes.

Advancement in the acquisition and analysis of transmission electron microscope

images (TEMs) has allowed biologists to study plant cell ultrastructure in unprecedented detail. However, considerable skill and knowledge is needed to obtain high-quality TEM data and to label biological structures inside a cell for analysis. The irregularity of plant cell organelles makes rigorous statistics and quantitative analysis of individual organelle tedious and troublesome. Thus, it is in great need of user-friendly tools for processing TEMs that are adaptable and suitable to use with multiple plant species and parameters.

The fundamental problem in comprehensively quantifying the phenotype of organelles is to accurately segment objects in the TEMs. The threshold-based techniques or analyses of the image histogram were proposed to segment glial cells[5] which require time-consuming user interventions to guarantee accurate segmentation. In seeking an automated response to these challenges, deep learning (DL) methods provided excellent segmentation of various biological images in an unbiased and automated manner needed for biomedical application[6-8]. Complementary to the endeavor to decipher the cellular information in biomedical imaging, similar type analysis of plant cell phenotypes would accelerate the identification of critical mechanisms in plants that could help address food security challenges. Recent quantification of plant stomata by DL indicates that this approach could be extremely useful in plant cell studies[9,10], but its application for plant organelle interpretation at the subcellular levels has not been explored. Analysis of stomata is relatively simple, because leaf epidermal micrographs contain only two cell types (stomata and pavement cells) and the leaf epidermal cells can be easily separated from the background on the basis of intensity. Reliable automated segmentation of plant organelles in TEM stacks is much difficult not only because a typical TEM dataset contains a heterogeneous admixture of organelle morphologies across the sample, but also because of image noisy caused by uneven illumination. Despite the importance of quantifying changes in subcellular morphology, a dedicated analysis tool for understanding these TEM stacks at the subcellular scale does not yet exist to the best of our knowledge. It is well known that robust and accurate DL models are always built on a large number of reliable datasets. Indeed, there is a scarcity of suitable training data sets for plant organelle

analysis.

The main objective of this work was therefore to develop a robust DL pipeline to segment complicated objects in TEMs, including chloroplasts, mitochondria, nuclei, and vacuoles, providing efficient and automated quantification. The proposed hierarchical strategy to comprehensively characterize the location and phenotype of organelles in plant cell was used to evaluate genetic mutants and responses to heavy metal stress. Additional contributions are: (1) the first publicly available dataset of TEMs with a large number of plant organelles manually-annotated; (2) a demonstration of biological relevance extracting and analyzing the morphological characteristics (i.e. organelle shape-complexity, electron-density, and area), (3) a user-friendly website https://www.cropopen.com/#/Cell, which allows researchers to submit TEM data of plant cells and quickly identify changes in cell morphology.

Results

An overview of AI system (Plantorgan hunter) for segmentation of subcellular structure and quantified analysis of morphology is illustrated in Fig. 1a. Our AI system is a useful tool for plant science community to accurately identify and quantitatively measure morphological characteries of plant organelles, which is scalable to TEM dataset of many plant species. Precise organelle segmentation is the first procedure. We designed a continuous depth-wise separable convolution (DSC) architecture focusing on improving performance for segmenting small sized organelles like mitochondria (Fig. 1b). Three morphological metrics (shape-complexity, electron-density, and area) were defined to capture the diversity of an organelle. The organelle shape characteristics are represented by the completeness of the visibility graphs that is defined by placing nodes equidistantly along the contour (Fig. 1c).

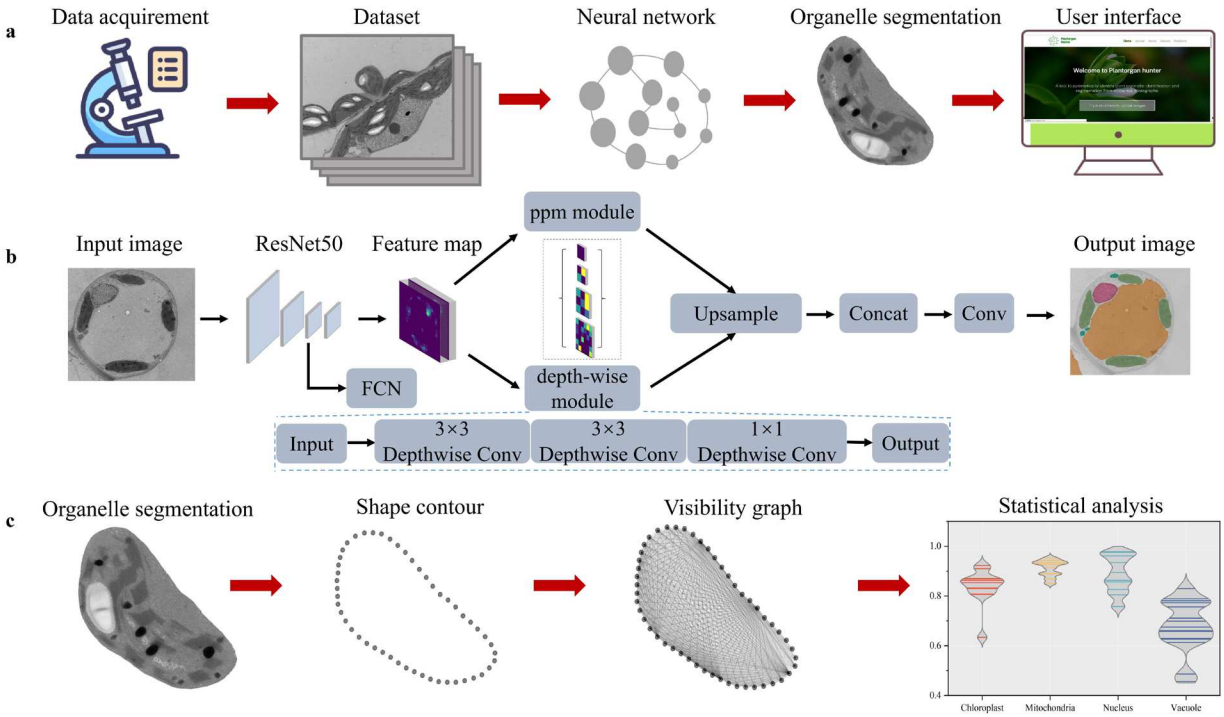

Fig. 1 Overview of Plantorgan hunter for segmenting and characterizing organelle morphology. **a**, Flowchart of image processing and data analyses for measuring organelle morphological features. **b**, Deep learning network structure for segmenting organelles. **c**, Quantitative descriptions of organelle morphological features after segmentation.

OrgSegNet for plant cell organelle segmentation

A disadvantage of DL methods is that they typically require massive amounts of labeled training data. We firstly generated large amounts of pixel-level labelled TEMs to develop a robust and universal model to identify microscopic organelles. Our TEM dataset was created from 13 plant species with 3149 chloroplasts, 1507 mitochondria, 123 nuclei, and 1592 vacuoles annotations (Fig. 2a,b). This TEM dataset had various magnifications from 1200× to 20000×. Chloroplasts are typically spindle-shaped and 3-10 μm in size. Small mitochondria are generally ovoid with a diameter of 0.5-1.0 μm, and so have low pixel size in our TEM dataset. Vacuoles of various sizes were observed with low electron-density and surrounded by a membrane (tonoplast). The nucleus of most plant cells is either round or oval with a diameter 1-10 μm. Disintegration of some internal structures can lead to organelles, especially mitochondria, being no longer intact and this increases the difficulty of recognition. Because there are big differences

in the appearance of these organelles TEM data, it is important to collect a large-scale, well-curated dataset in order to establish an AI system with good robust generalization.

We trained a DL model, OrgSegNet (Fig. 1b), established on these 673 high-quality TEMs to address the problem of accurate segmentation. We evaluated the accuracy of segmentation on a testing dataset of 135 images. For each organelle, segmentation accuracy was evaluated by two metrics, Precision and Recall, which are calculated based on true positives, false positive and false negative (Fig. 2c,d). OrgSegNet fully captured the characteristics of chloroplasts, with Precision and Recall of 0.9828 and 0.9780, respectively. In total, 97.80% region of interest (ROI) were successfully predicted, 1.72% ROI resulted from under-segmentation and 2.20% ROI resulted from over-segmentation. Small mitochondria had very similar values of Precision and Recall (0.8262 and 0.8535, respectively). Compared to chloroplasts and mitochondria, OrgSegNet tended to over segment vacuoles (Precision and Recall were 0.7700 and 0.9203, respectively) and miss partial nuclei (Precision and Recall were 0.9396 and 0.8664, respectively). As TEM magnification can be adjusted in a broad range, we also evaluated the segmentation produced by our system on images of different magnification for each organelle (Supplementary Fig. 1). It was observed that the performance of mitochondria segmentation was very sensitive to TEM magnification. Next, we used the OrgSegNet to segment each organelle by visualizing the performance over a single image with direct magnification of 3000× (Fig. 2e). The quality of these prediction suggested that OrgSegNet could capture the real size of chloroplasts, mitochondria, nuclei, and vacuoles in the plant cell.

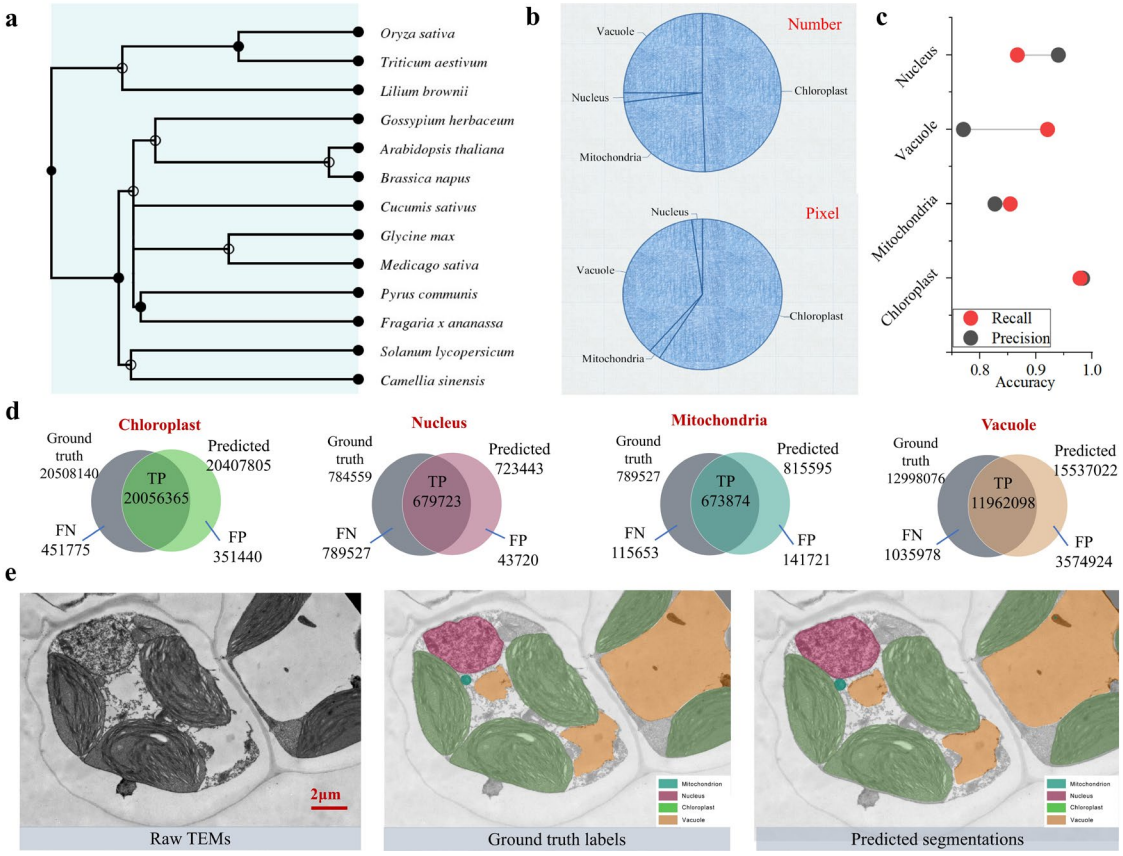

Fig. 2 Dataset information and OrgSegNet segmentation results. **a**, An underlying training set of high-resolution electron microscope images (TEMs) annotated for chloroplasts, mitochondria, nuclei, and vacuoles was collected from 13 plant species. **b**, Annotated number and pixel area of each organelle according to dataset. **c**, Recall and Precision value of OrgSegNet for each organelle segmentation. **d**, Venn diagrams showing the performance of OrgSegNet for the organelles detecting as true positives (TP), false positive (FP) and false negative (FN) in the testing data. **e**, Representative example of input, ground-truth label and predicted image generated by OrgSegNet. The ground-truth label and predicted image is overlain onto the input image with different color for different organelles.

Comparison of OrgSegNet with classical DL image segmentation algorithms

To comprehensively evaluate the performance of OrgSegNet, we compared it with some classical DL image segmentation algorithms: PSPNet[11], DeepLabv3+[12], DANet[13], DNLNet[14], PSPU-Net[15], Swin Transformer (Swin)[16], and for both accuracy and speed (Fig. 3a). All these segmentation algorithms are shown to be useful for classifying image pixels, but have not so far been used for segmenting subcellular structures. The

following evaluation metrics were used (see “Methods”): mean intersection over union (mIoU), mean class accuracy (mACC), overall accuracy (OA), parameters (Params), number of float-point operations (FLOPs), and computing time for inference (Time). OrgSegNet had the best performance with highest accuracy (mIoU = 81.30%) and top 2 fastest computing speed (costing 67 seconds for inference). The models, DANet, PSPU-Net and DNLNet, did not match the performance of OrgSegNet when trained on this underlying training set, while Swin, PSPNet and DeepLabv3+ provided reliable segmentation. In addition to segmentation accuracy, the methods that we benchmarked varied substantially in their computational time. PSPU-Net and Swin suffered from being time-consuming when inferenced on the same data. OrgSegNet was only 4.47% slower than PSPNet, despite a significant model capacity increase, and was 2 times faster than Swin. This comparison was not affected by the metrics we selected, as we detected similar trends for mAcc, OA, Params, and FLOPs (Supplementary Table 1). Totally, OrgSegNet yielded 81.30% mIoU using only 52.93M parameters and 356G FLOPs outperforming its peers.

Next, we used the algorithms to segment organelle by visualizing the performance comparisons between OrgSegNet and other benchmark segmentation algorithms (Swin, PSPNet and DeepLabv3+) over two single images (Fig. 3b). For each image, segmentation accuracy was evaluated by the F1 score calculated from overlap between prediction and ground-truth area. In line with their lowest F1 score, Swin and DeepLabv3+ had organelles that were partially detected or misidentificatified. OrgSegNet segmented organelles most effectively, with F1 score of 0.9725 and 0.9755, respectively. In comparison, PSPNet gave a slightly lower F1 score, and failed to identify small-sized mitochondria.

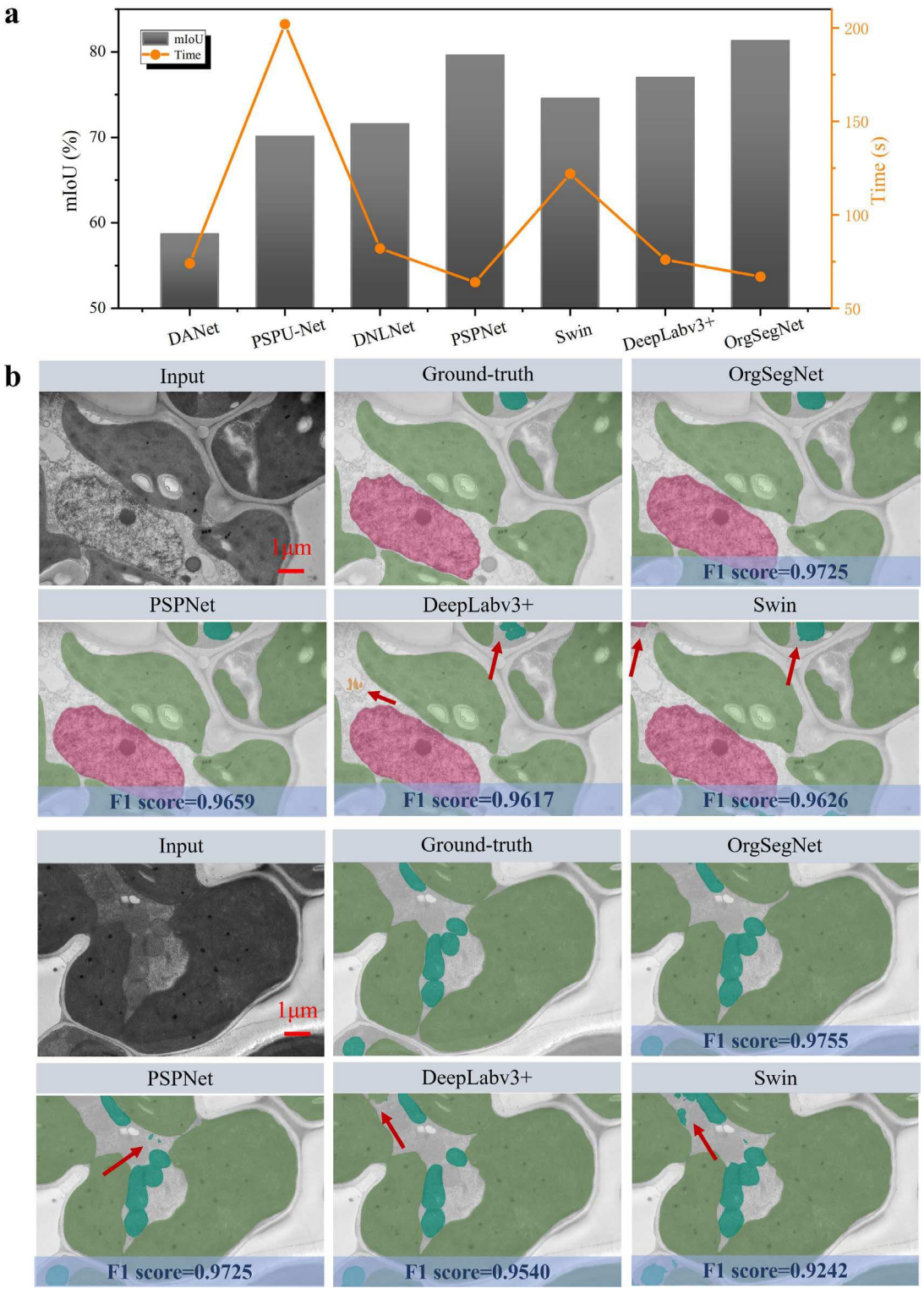

Fig. 3 Detail computational comparison of OrgSegNet and previous algorithms. **a**, Performance of seven deep learning models based on accuracy and computing time in inference process. Accuracy is presented by mean intersection over union (mIoU) and computing time was calculated by inference time on test dataset (Time). **b**, Two

representative examples of segmentation results from OrgSegNet, Swin, PSPNet and DeepLabv3+ with direct magnifications of 5000×. The F1 score is shown for each image. The red arrows indicate the recognition errors.

To capture small sized organelles like mitochondria, we added a continuous DSC structure parallel with the pyramid pooling module (see “Methods”). We therefore compared the segmentation performance of OrgSegNet and PSPNet across a range of TEM data under different direct magnifications from 2500×to 10000× (Supplementary Table 2 and Supplementary Fig. 2). Both models gave better segmentation of mitochondria at higher magnifications, but OrgSegNet substantially outperformed PSPNet providing a higher-accuracy of segmentation and capturing the true size of mitochondria. Under high magnification, both approaches gave impressive segmentation. While at low magnification, OrgSegNet had the capability to precisely follow the contours of mitochondria, and was superior than PSPNet. Some zoom-in views of segmentation results are shown in Fig. 4. In some cases, it was evident that OrgSegNet was also challenged to segment and merged some mitochondria when a number of targets were close to each other (Fig. 4b,c).

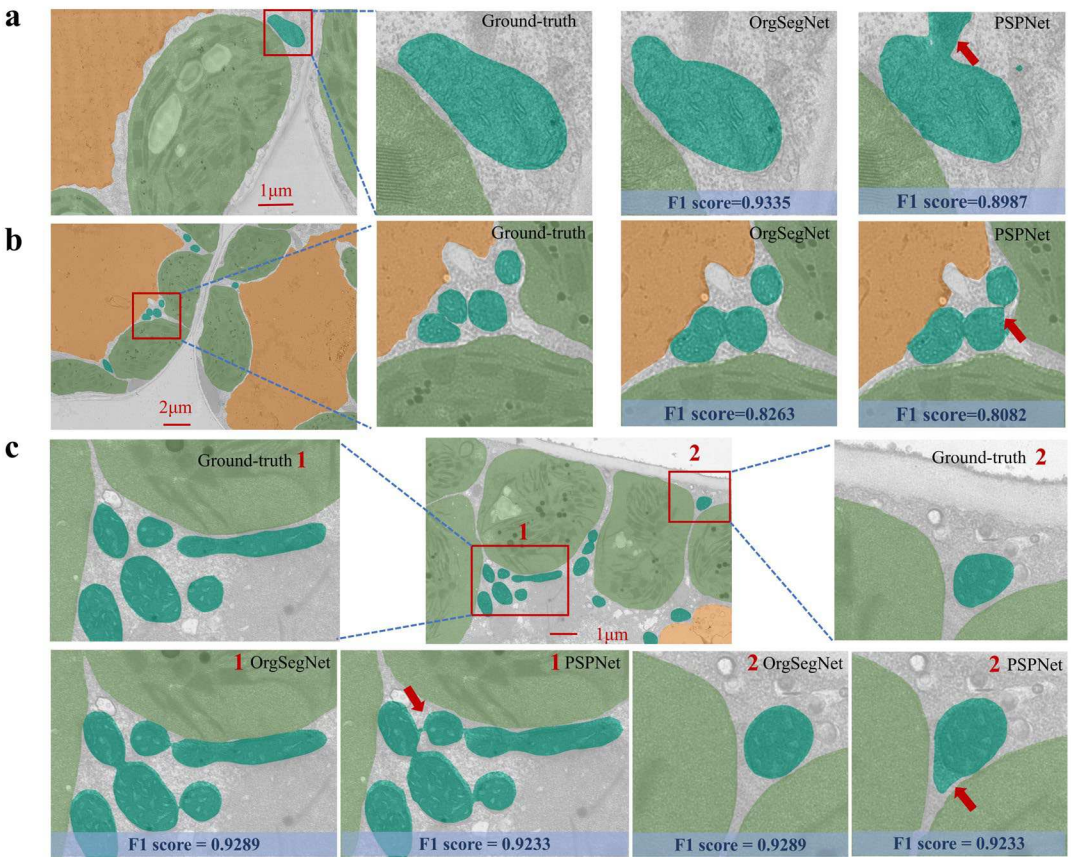

Fig. 4 Zoom-in representative TEM images of diverse morphology of small mitochondria segmented by OrgSegNet and PSPNet at magnifications of 6000× (**a**), 2500× (**b**), 5000× (**c**). The F1 score is shown for each image. The red arrows indicate the recognition errors.

OrgSegNet captured the morphological characteries of organelles similar to human decision-marking

To better understand the inference process of DL method and ensure its reliability for segmentation, a class activation map (CAM) was generated[17] to visualize the input image region and see what the neurons have learned (Fig. 5 and Supplementary Fig. 3-6). By simply upsampling the CAM to the input image size under different magnifications, the class-specific regions most relevant to the particular organelle were identified. For instance, chloroplasts contain thylakoid stacking structures and black dot-like structures called osmiophiles. The highlighted region in the CAM overlaps with the most apparent thylakoids in chloroplasts, indicating that thylakoid regions are the most important ones for machine learning recognition (Supplementary Fig. 3). The internal folds of mitochondria were specifically highlight and recognized as the features

for segmentation, while the ability reduced at low magnification, consistent with the results that mitochondria segmentation performance decreased with decreasing magnification (Supplementary Fig. 4). Over-segmentation tended to happen for vacuoles as indicated by the false positive and false negative regions. Typical structural features of vacuoles for expert recognition are the lumen with low electron-density enclosed by a double membrane. Similarly, the DL model paid more attention to the double membrane structure, as illustrated in the Supplementary Fig. 5 with a hot region on membrane structure but low attention to its interior. OrgSegNet can misidentify the vacuole when only a small part of membrane structure appears in the TEM data. It is well known that the interior of the vacuole with its low electron-density looks very similar to the area of intercellular space, which made the vacuole segmentation task difficult. The CAM highlighted the dense material textures as the membrane-bound nucleus class-specific discriminative regions for segmentation (Supplementary Fig. 6). In this context, OrgSegNet would lose some nucleus information, because machine learning did not much focus on the characteristics of double membrane structure. In summary, the results showed that our neural network was able to capture the characteristics of organelles, such as texture of electron-density, membrane structure and thylakoids, and thus specifically recognize each organelle in a manner similar to the decision-marking of a plant scientist.

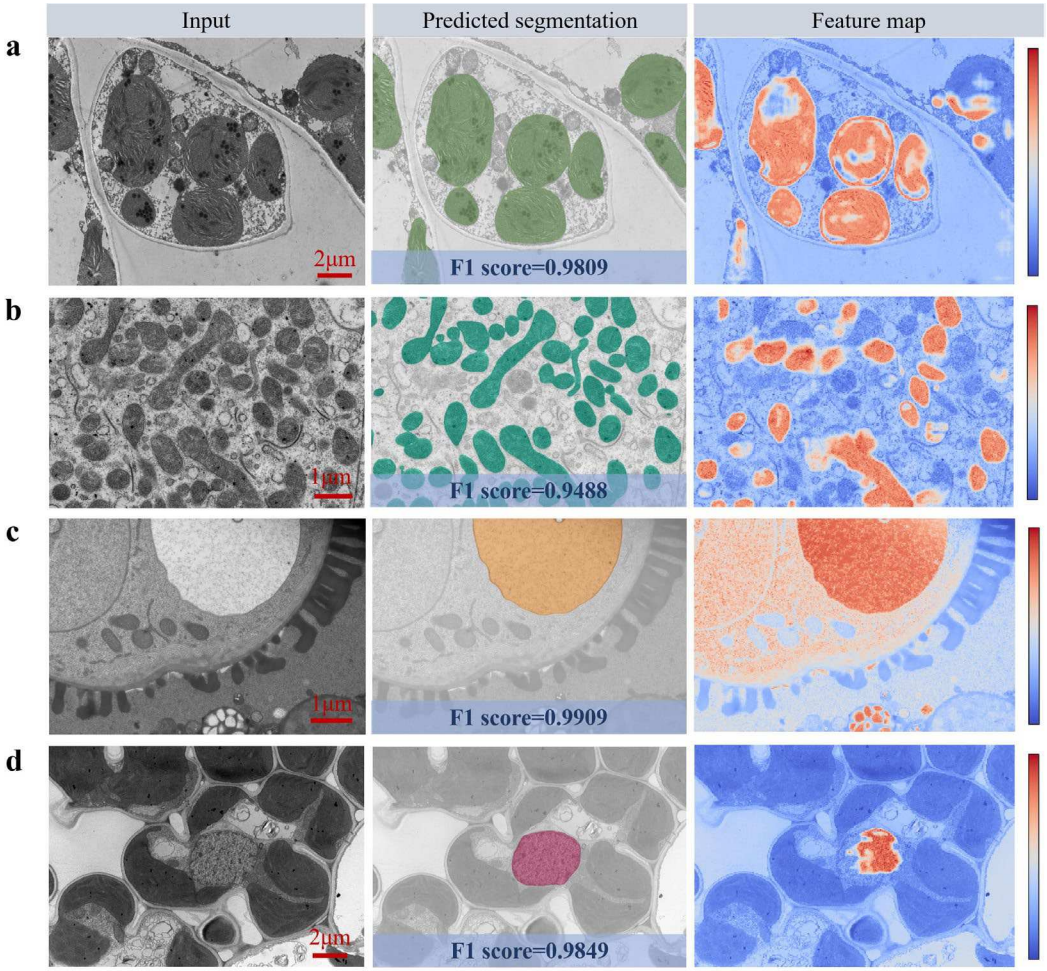

Fig. 5 Example of chloroplast (**a**), mitochondria (**b**), vacuole (**c**) and nucleus (**d**) segmentation results and class activation map visualization learned by OrgSegNet. F1 score is shown for each image. First column, input image; second column, predicted segmentation; last column, class activation map.

Morphological metrics quantitative analysis and biological application

To comprehensively characterize organelle phenotype, we defined three morphological metrics (shape-complexity, electron-density, and area) to capture the morphological complexities of each organelle following segmentation. More than 10 independent samples of each organelle were selected for analysis (Supplementary Fig. 7). We used a novel visibility graphs method[18] to quantitatively describe organelle shape-complexity such as the spindle shape of chloroplast, ovoid mitochondria and oval nucleus. The electron-density relationship of the organelle contents was calculated based on the grayscale discrepancy related to bright background areas (see "Methods"). The area was obtained from the relationship between the number of pixel points of the

organelle and the scale bar of the TEM image. Because of large differences in the cross-section samples of plant leaves under the electron microscope, the shapes of each organelle varied considerably (Supplementary Fig. 12), but using our method it appears that each type of organelle has its organelle-type phenotype (Supplementary Fig. 7). The morphological metrics of individual organelles provides a way to quantitatively compare the same type of organelle at different times. The size and shape-complexity of vacuoles are extremely dynamic and pleiomorphic with irregular features. Overall, our network-based framework can accurately quantify characteristics of organelles and offer additional functionality to that of contending segmentation method.

Finally, we used our AI system to provide a few examples of the application of chloroplast segmentation, which are proved to be reliable, to examine the phenotypes of plants at the subcellular level (Fig. 6). The first examples were two mutants of rice. Sample 1 was a thermo-sensitive rice albino leaf mutant developed under low temperature (20–24 °C)[19]. In TEM images no typical chloroplasts could be seen in the albinotic leaves, and our AI could not identify any chloroplast structures (Supplementary Fig. 8). Sample 2 was young leaf chlorosis 1 (ylc1) acquired from a 60Co-irradiated population. Young leaves of the ylc1 mutant showed decreased levels of chlorophyll and lutein compared to corresponding wild type[20], and its TEM analysis revealed that the thylakoid lamellar structures were obviously loosely arranged (Supplementary Fig. 9). Compared to the qualitative phenotypic descriptions, our AI system quantified the chloroplast phenotypes of mutants, with electron-density increase of 13.28%, area decrease of 45.97%, and shape-complexity change of 22.51% compared to corresponding wild type, respectively (Fig. 6c).

Analysis of plant mechanisms explored to heavy metal condition can help us evaluate their phytoremediation capability for remediating heavy metal polluted soil. The last examples were to analyze differences in the copper resistance of *Medicago sativa* and *Elsholtzia haichowensis*. The TEM data showed that the chloroplast structure in *M. sativa* leaves was destroyed, while the chloroplast structure of well-known copper tolerant plant *E. haichowensis* was not changed significantly under the same copper treatment (Supplementary Fig. 10 &11). This phenomenon was also supported by the

quantitatively measurement of these morphological metrics in treated versus untreated control chloroplasts (Fig. 6d,e). Following copper treatment, electron-density and area were substantially decreased in *M. sativa*, but there were no statistically significant differences in any of the morphological metrics in *E. haichowensis*. These findings illustrate the importance of quantifying the complexity of chloroplast morphology.

We used theses external data (rice mutants, *M. sativa* and *E. haichowensis* that was not in the internal testing set) to further estimate whether OrgSegNet can generalize well to the completely unseen data (Fig. 6f,g). The segmentation results indicated that we had almost the same F1 score as on the internal testing set (supplementary Fig. 8-11 and Fig. 2), suggesting that OrgSegNet is capable of high robust genitalization even when chloroplasts are not abnormal, which is particularly important for its application. Normal chloroplast structures had an average F1 score of 0.9729, while abnormal chloroplasts with small area and loosely arranged thylakoid lamellar structures had a comparable score of 0.9574.

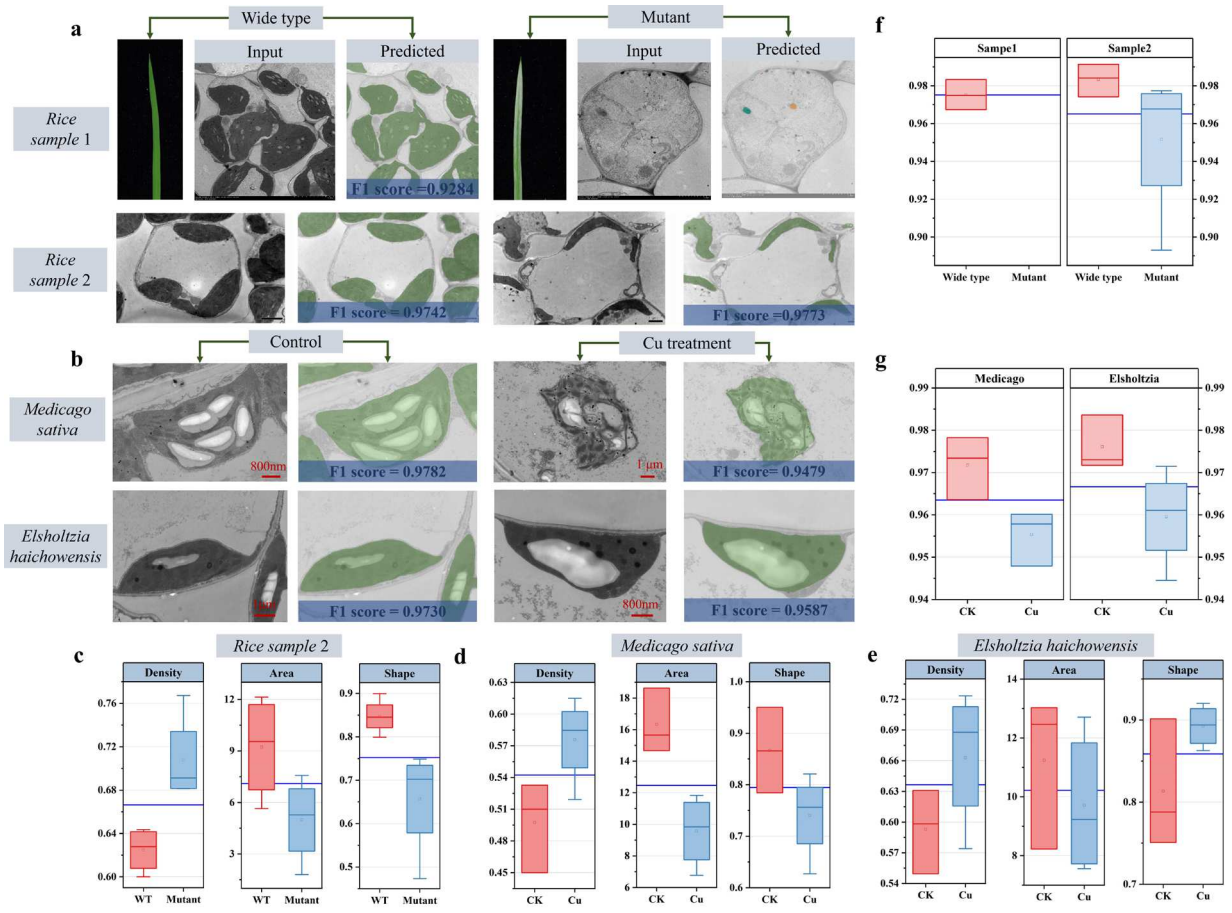

Fig. 6 Application of our AI system for quantitatively evaluating large-scale TEMs of plant chloroplast changes. **a,b,** Representative images of OrgSegNet segmentation predictions using TEMs from two rice mutants (**a**) and copper treated *Medicago sativa and Elsholtzia haichowensis* (**b**). **c,d,e,** Quantitatively evaluating of the morphological characteries of chloroplast in rice mutant (**c**), copper treated *M. sativa* (**d**) and copper treated *E. haichowensis* (n≥3) (**e**). **f,g** Segmentation performance of OrgSegNet on mutants (**f**) and copper-treated plants (**g**) based on F1 score (n≥3). Wt, wide type; Ck, untreated control; Cu, treated with 1600 mmol of copper ions.

Discussion

Despite the advantages of TEMs for characterizing organelles at the nanoscale, there is a lack of computer-aided tools to extract and understand the image data. To address this issue, we developed an AI system that makes several contributions to facilitate the development and application of digital image-based subcellar analysis for the plant science community.

Effective and accuracy segmentation of target organelle is essential to any valuable downstream biological image inferences can be drawn. Segmentation of organelles in high-resolution TEMs is challenging because of the moderately low contrast between some organelles and the cytoplasm or intercellular substance. Moreover, different type of organelles are spatially staggered in TEM dataset, and the appearance of the same type of organelle changes as plants develop or respond to external challenges. Our study shows that Plantorgan hunter is able to solve these challenges at nanometer levels and can provide quantitative metrics to describe subcellular morphology. The OrgSegNet pipeline proved to be better than other state-of-the-art segmentation models both in terms of calculating speed and with high segmentation accuracy (Fig. 3). The OrgSegNet captured the real size of chloroplasts, mitochondria, nuclei, and vacuoles in the plant cells and also performed well on an unseen image set with abnormal chloroplasts and from different plant species (Fig. 6), thus demonstrating the scalability of our method. With the added continuous DSC structure parallel to the pyramid pooling module (ppm), OrgSegNet can capture more detailed structure information than PSPNet (Fig. 4). This fine-tuning greatly improves the recognition of small structures at low magnification, especially for the identification of mitochondria with smaller morphologies. Of course, this module can also improve the sampling accuracy of the model to a certain extent, and ensure that the model can better fit the boundaries of the organelles. By looking into the black-box decision made by OrgSegNet (Fig. 5), we demonstrated that our AI system recognized the characters of chloroplasts, mitochondria, nuclei, and vacuoles, resembling human decision-marking and proving the reliability of the model.

In addition, we generalized quantitative information about three morphological metrics of single cellular organelles, directly after automatic segmentation. Here, organelle shape-complexity was described as the concept of a visibility graph, and the electron-density of the organelle matrix was defined by the grayscale histograms difference between organelles and bright background areas. Apparently, chloroplasts, mitochondria, vacuoles and nuclei each have their own shape-complexity, electron-density, and area, and can be distinguished based on our method (supplementary Fig.

7). The extraced information was successfully applied to reveal the subcellular phenotype of plants under multiple situations (Fig. 6). We were thus able to demonstrate and quantify the different chloroplast phenotypes related to genetic mutation and heavy metal stress. The ability to quantify such metrics at the level of individual organelles from high-through TEM data can provide new insights into plant cell biology.

In principle, creating large expert-curated and annotated dataset with widely varying organelle morphology will guarantee the accuracy of artificial intelligence-based methods. The TEM dataset collected from 13 plant species cost us more than 300,000 RMB, and required >700 person-hours with 6371 organelles annotated manually, which is extremely time-consuming and expensive. Although we adopted some common data augmentation methods, such as random clipping, random flipping, random photometric distortion methods, to lessen this restriction for training. Obviously, the characteristic of the data itself had the biggest effect on segmentation performance. While images at low magnification level often contained full views of plant cell organelle patterns, images at high magnification level mapped organelles in clearer detail but led to the structure of some organelles become incomplete. Normally, there is only one nucleus per cell and this limited the source of nuclei in the training data. This issue is likely to avoidable by enhancing size and quality of training TEM dataset preparation. Alternatively, genitive adversarial network like architectures might have the potential to reduce the requirement for large numbers of manual annotations[21], which need more computer scientist attend to the biological issue.

Finally, we deployed the OrgSegNet model on open-source platform https://cropopen.com/#/Cell, to improve the intuitive and friendly use for non-specialist and those who are not as familiar with machine learning techniques. We strongly believe that the open science and data sharing will accelerate such interdisciplinary research that needs knowledge of biology, computer science, microscopy and statistics. We have therefore made all the labeled training datasets, code and trained models available to the scientific community under a permissive license. Obviously, there is more work to do to explore our data and improve the segmentation performance, and also to include other organelles that are currently not analyzed.

Methods

Preparation of microscopy images

To create a dataset of popular plant organelle segmentations, leaf tissues from various plant species were selected for ultrastructure observation including rice (*Oryza sativa*), wheat (*Triticum aestivum*), lily (*Lilium brownii*), cotton (*Gossypium herbaceum*), arabidopsis (*Arabidopsis thaliana*), rape (*Brassica napus*), cucumber (*Cucumis sativus*), soybean (*Glycine max*), alfalfa (*Medicago sativa*), pear (*Pyrus communis*), strawberry (*Fragaria × ananassa*), tomato (*Solanum lycopersicum*) and sasanqua (*Camellia oleifera*) (Fig. 2a). Plant leaf tissue was first fixed in 2.5% glutaraldehyde buffered with 0.2M phosphate buffer (pH 7.2) at 4°C overnight followed by post fixation with 2% osmium tetroxide at 4°C for 3 h. The samples were thoroughly washed with the phosphate buffer, dehydrated in an acetone series and then embedded in Epon812 resin. After polymerization, resin blocks were sectioned using an ultra-microtome (Ultra-microtome UC 7, Leica, Wetzlar Germany), then stained with uranyl acetate and lead citrate solution. Microscopic observation was performed using a transmission electron microscope (H-7650, Hitachi, Tokyo, Japan) at an acceleration voltage of 80 kV.

Datasets description

The full dataset totaled 673 high-resolution transmission electron microscopy images with size of 4800×3730 pixels. Here, four plant cell organelles (chloroplast, mitochondria, nucleus, and vacuole) were studied, and their individual instance masks in each TEM were labeled by life science experts using labelme software. These organelles come in various sizes and have different characteristics at different nanometer levels (Supplemental Fig. 12).

OrgSegNet workflow

The workflow of OrgSegNet includes the ResNet50 module[22] for feature extraction, the ppm module and the depth-wise module for feature processing, an upsampling module and a segmentation module for the output of segmentation results. The detailed architecture of OrgSegNet is presented in Supplementary Fig. 13, and some of the hyper-parameters were set with reference to previous work[11,23].

For faster training of the model, we used the parameters of the model trained on the cityscapes dataset as initialization parameters instead of using a random initialization method[24]. Considering the balance between segmentation accuracy and GPU load, before model training, we reduced the collected TEM images to the size of 800×600 pixels by bilinear interpolation. In order to increase the generalization and stability of the model, we adopted random clipping, random flipping, random photometric distortion methods for data augmentation. 80% of the dataset was used as training set, and the remaining 20% was used for model validation and testing.

In OrgSegNet, the ResNet50 is used as a feature extraction network, and the data is output through the network into a feature map with a size of 2048×100×75 (Supplemental Fig. 13a). Subsequently, the feature map passes through the ppm module and the depth-wise module (Supplemental Fig. 13b,c).

In the ppm module, the feature map is pooled to 1×1, 2×2, 3×3 and 6×6 sizes respectively, which is used to extract features of different dimensions and generate context priors. In detail, AdaptiveAvgPool layer is used to implement the pooling function, which can adaptively adjust the kernel size, padding and stride size of the pooling layer while generating the output of the specified size. After a convolution module with the kernel size of 1×1, the number of channels is adjusted to 1/4 of the the original number of channels, which is used to adjust the number of channels, preserve the original features and reduce the amount of computation. Subsequently, all the feature maps are merged and upsampled to the original feature map size. Up to this step, the prior information of the context is extracted and fused by the ppm module.

In the depth-wise module, the feature map goes through successive DSC layers with convolution kernels of 3×3, 3×3 and 1×1. In this step, the feature map is further refined, the first DSC layer is used to adjust the number of channels of the feature map, and the second and third convolution layer take into account the processing of detailed features and global features. A DSC layer can be divided into a depth-wise convolution layer and a pointwise convolution layer[25]. Unlike ordinary convolution operation, one convolution kernel of DSC is responsible for one channel. That is to say, a channel is operated by only one convolution kernel. The pointwise convolution is actually a

convolution with a kernel size of 1×1, which adjusts the number of channels of the feature map and fuses the feature map. The feature map generated from the depth-wise module contains high-level semantic information in the original image, which is helpful for the recovery of fine structures so that tiny organelles, such as small mitochondria, can be predicted.

Subsequently, the feature maps output from the ResNet50, ppm module and depth-wise module are combined, and fed into a convolution layer to get the classification results.

During model training, due to the imbalance of categories, especially the number of nucleus in dataset was much smaller than the other three types of organelles, the weighted cross-entropy loss function was used to supervise the probability map of the output of the last layer. The weights of background, chloroplast, vacuole, mitochondria and nucleus were 1, 1, 2, 5 and 8, respectively, which were based on multiple trials and quantitative relationships between organelles.

$$Loss = -w_c \sum_{i=1}^{N} \sum_{c=1}^{C} y_{ic} \, log \, \bar{y}_{ic} \quad (1)$$

Where, N is the total number of samples, C is the total number of classes, y_{ic} is the indicator that the i^{th} sample belongs to the c^{th} class, $\bar{y}_{ic}$ is the predicted probability that the model associates the i^{th} input with class, and w_c is the weight of class c.

In order to ensure the stability of model training, an FCN network was introduced as an auxiliary training network, and optimized by the weighted cross-entropy loss function. Auxiliary loss branch helped optimize the learning process and allowed the model to achieve the best performance. The compute method of the final loss of backpropagation:

$$Loss = main\ Loss + aux\ Loss * \alpha \quad (2)$$

Where the α was 0.4 according to previous work[11], *main Loss* and *aux Loss* were the loss of the OrgSegNet and the loss of the FCN, respectively.

The SGD optimizer was used to update the network parameters to speed up the model fitting, where the initial learning rate was set to 0.01, and the learning rate decay strategy was that the learning rate decays by 0.0002 each epoch. The model stoped after 800 epochs of training.

OrgSegNet postprocessing

The output from the model was the feature map segmentation results with labels, with categories ranging from 1 to 5, representing background, chloroplast, mitochondria, vacuole and nucleus, respectively. We created a set of colors for each label, mapped the output labels to the original image, and transformed them to the standard RGB image format. In order to accurately segment each organelle, the organelle images corresponding to the five types of labels were divided into five layers, and a series of image processing algorithms were applied on each layer:

1. Remove objects with an area smaller than a certain threshold (set to 1/100 of the image area).

2. Connect regions and remove holes.

3. Calculate the contour of each organelle in each layer.

4. Take out the contour of a single organelle in the layer and use it to calculate the mask.

5. Overlay the organelle mask on the original image to get the segmentation result of a single organelle.

Model performance evaluation

The output of OrgSegNet was quantitatively evaluated and compared with the most popular DL methods: PSPNet, DeepLabv3+, Danet, DNLNet, PSPU-Net, Swin. The target segmentation performance was evaluated by metrices mIoU, mACC and OA, while computing speed was evaluated by metrices Params, FLOPs, and Time. They are defined as following:

$$mIoU = \frac{1}{C}\sum_i^C \frac{TP_i}{TP_i+FP_i+FN_i} \quad (3)$$

$$mAcc = \frac{1}{C}\sum_i^C \frac{TP_i}{TP_i+TN_i+FP_i+FN_i} \quad (4)$$

$$OA = \frac{\sum_i^C TP_i}{\sum_i^C TP_i+FN_i} \quad (5)$$

Where TP_i, TN_i, FP_i, FN_i refer to TP, TN, FP and FN of the i^{th} class sample respectively, and C refers to the total number of classes.

FLOPs reflects the running speed of the model to a certain extent, which is a relatively simple indicator and a reference quantity for the model inference time. For

most of the models involving convolutional structures, the proportion of FLOPs involved in the activation layer and the pooling layer is very small (less than 0.1% FLOPs). Therefore, the following only focus on the compution of FLOPs in the convolutional layer. Theoretically, for convolutional layer, it consists of an addition operation and a multiplication operation, respectively. When bias is not used by default, we have:

$$FLOPs\ of\ Conv = (2C_{in}K^2 - 1)HWC_{out} \tag{6}$$

$$FLOPs\ of\ DSC\ = (2C_{in}K^2 - 1)HW + C_{in}C_{out}HW \tag{7}$$

where, H, W and C_{in} are height, weight and number of channels of the input feature map, K is the kernel size, and C_{out} is the number of output channels.

The amounts of parameters of the neural network are mainly contributed by the convolutional layer:

$$params\ of\ Conv = C_{out}K^2C_{in} \tag{8}$$

$$params\ of\ DSC\ = K^2C_{in} + C_{in}C_{out} \tag{9}$$

The evaluation metric Time is the cost of the model inference process on the test dataset on the computer consist of the Windows10 64-bit operating system, Inter (R) Core (TM) i5-10400f CPU, Graphic Processing Unit (3090 GPU) and 32GB RAM. All the models were created based on python 3.7, pytorch 1.8,1.

Quantification of morphological features

The morphological metrics (shape-complexity, electron-density and area) of each organelle are given by the post-processing part of OrgSegNet. The area was obtained from the relationship between the number of pixel points of the organelle and the scale bar of TEM image.

$$Area = N \times S^2 \tag{10}$$

Where, N is the total number of pixels of the organelle and S is the ratio of the scale bar to its pixel length.

Since it is difficult to accurately obtain the electron-density value of a certain part from a TEM image, we counted the grayscale histograms of the segmented organelles and the grayscale histograms of the bright background areas. It was found that the distribution of grayscale histograms from each organelle were totally different

(Supplementary Fig. 14). Therefore, gaussian distribution curves were fitted to these histograms, and the peak of the curve was selected as the representative of density value. Then, the ratio of the peaks of the two regions was used as the density of the organelle. Density reflects the electron-density relationship of the organelle contents, which is a value from 0 to 1. The closer to 1, the greater the electron-density of the organelle.

$$y = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{11}$$

Where μ is the mean and σ is the standard deviation.

$$Density = \frac{\mu_o}{\mu_b} \tag{12}$$

Where, μ_o presents the gray value corresponding to the peak value of the Gaussian distribution fitting curve of the gray histogram of the organelle. μ_b presents the gray value corresponding to the peak value of the Gaussian distribution fitting curve of the gray histogram of the cell background.

In our research, organelle shape-complexity was defined by the visibility graph[18]. The process of constructing the visibility graph is as follows:

1. Segment the organelles and extract the contours of the organelles.

2. Place nodes equidistantly on the contours.

3. If the nodes are visible, create connections.

4. Calculates the euclidean distance between two nodes, where the coordinate of the node is (x, y).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 + y_2)^2} \tag{13}$$

5. Scale the euclidean distance between nodes that are visible to each other between 0 and 1 and assign it to the edge as a weight.

$$w_i = \frac{d_i}{d_{max}} \tag{14}$$

Where, d_i is euclidean distance between each node, d_{max} is maximum distance between nodes in a graph, w_i *is* weight of the edge between two nodes.

We quantified cell shape-complexity δ by computing the ratio between the number of nodes and the number of edges in the graph.

$$\delta = \frac{2m}{n(n-1)} \tag{15}$$

Where, *m* and *n* are the numbers of all edges and nodes in the graph, respectively.

Web service graphical user interface

In this step, we have deployed the OrgSegNet model on our web portal https://cropopen.com/#/Cell (Supplemental Fig. 15), which hosts the complete OrgSegNet pipeline. The back-end of the web portal is developed in a mixture of C# and Python, and the front end of the web page is developed using the Vue progressive framework and JavaScript language. All data are operated based on the MySQL database management system. The web back-end server provides abundant computing resources, and it can dynamically allocate computing resources according to the amount of data submitted by users, and return the computing results in a short time. Users can simply process the images to meet the input criteria of the system according to the guidelines of the website's documentation, then upload the images and obtain three morphological metrics (shape-complexity, electron-density and area) of every organelle through e-mail for free.

Generating predictions using local deployments

Although the OrgSegNet pipeline deployed in the cloud can well implement the segmentation task of plant organelles and is friendly to users who lack experience in image processing and DL methods, the functions implemented on the website are only the tip of the iceberg of what we can achieve. In addition, the portal website is also not suitable for additional processing technology changes, and it is difficult to meet the additional needs of user customization. Therefore, we additionally provide a local deployment of OrgSegNet for possible future workflow integration and development of new models. We provide jupyter notebook and python files of the original example. In addition, the environment installation requirements and process are detailed in the notes to help users reproduce and iterate on features. Finally, we provide all resource guides and model migration guides on the github website to help users get started quickly.

Data and code availability

The plant organelle dataset for the current study is available in the Sicence Data Bank repository, https://www.scidb.cn/s/EBvqei. The code is available in github,

https://github.com/yzy0102/Plantorgan-hunter-OrgSegNet.

References

1. Kirchhoff, H. Chloroplast ultrastructure in plants. *New Phytologist* **223**, 565–574 (2019).
2. Shimada, T. *et al.* Plant vacuoles. *Annu. Rev. Plant Biol* **69**, 123–145 (2018).
3. Broda, M., Millar, A. H. & Van Aken, O. Mitophagy: a mechanism for plant growth and survival. *Trends in plant science* **23**, 434–450 (2018).
4. Goswami, R. *et al.* Mechanical shielding in plant nuclei. *Current Biology* **30**, 2013–2025 (2020).
5. Healy, S., McMahon, J., Owens, P., Dockery, P. & FitzGerald, U. Threshold-based segmentation of fluorescent and chromogenic images of microglia, astrocytes and oligodendrocytes in FIJI. *Journal of neuroscience methods* **295**, 87–103 (2018).
6. Greenwald, N. F. *et al.* Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature biotechnology* **40**, 555–565 (2022).
7. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nature methods* **18**, 100–106 (2021).
8. Moen, E. *et al.* Deep learning for cellular image analysis. *Nature methods* **16**, 1233–1246 (2019).
9. Fetter, K. C., Eberhardt, S., Barclay, R. S., Wing, S. & Keller, S. R. StomataCounter: a neural network for automatic stomata identification and counting. *New Phytologist* **223**, 1671–1681 (2019).
10. Li, S. *et al.* LeafNet: a tool for segmenting and quantifying stomata and pavement cells. *The Plant Cell* **34**, 1171–1188 (2022).
11. Zhao, H., Shi, J., Qi, X., Wang, X. & Jia, J. Pyramid scene parsing network. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2881–2890 (2017).
12. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. in *Proceedings of the European conference on computer vision (ECCV)* 801–818 (2018).
13. Xue, H. *et al.* Danet: Divergent activation for weakly supervised object localization. in *Proceedings of the IEEE/CVF International Conference on Computer Vision* 6589–6598 (2019).
14. Yin, M. *et al.* Disentangled non-local neural networks. in *European Conference on Computer Vision* 191–207 (Springer, 2020).
15. Pérez-Pelegrí, M. *et al.* PSPU-net for automatic short axis cine MRI segmentation of left and right ventricles. in *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)* 1048–1053 (IEEE, 2020).
16. Liu, Z. *et al.* Swin transformer: Hierarchical vision transformer using shifted windows. in *Proceedings of the IEEE/CVF International Conference on Computer Vision* 10012–10022 (2021).
17. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. & Torralba, A. Learning deep features for discriminative localization. in *Proceedings of the IEEE conference on*

computer vision and pattern recognition 2921–2929 (2016).
18. Nowak, J. *et al.* A network-based framework for shape analysis enables accurate characterization of leaf epidermal cells. *Nature communications* **12**, 1–13 (2021).
19. Liu, Y. *et al.* Phenotypic characterization and gene mapping of a thermo-sensitive albino leaf mutant tsa1 in rice. *Acta Agronomica Sinica* **42**, 1754–1763 (2016).
20. Zhou, K. *et al.* Young Leaf Chlorosis 1, a chloroplast-localized gene required for chlorophyll and lutein accumulation during early leaf development in rice. *Planta* **237**, 279–292 (2013).
21. Tschuchnig, M. E., Oostingh, G. J. & Gadermayr, M. Generative adversarial networks in digital pathology: a survey on trends and future potential. *Patterns* **1**, 100089 (2020).
22. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 770–778 (2016).
23. He, T. *et al.* Bag of tricks for image classification with convolutional neural networks. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* 558–567 (2019).
24. Cordts, M. *et al.* The cityscapes dataset for semantic urban scene understanding. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 3213–3223 (2016).
25. Chollet, F. Xception: Deep learning with depthwise separable convolutions. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 1251–1258 (2017).

Acknowkedgements

This work was supported by the Key R&D Program of Zhejiang (2022C02032), the National Natural Science Fundation of China (32071895), and the Fundamental Research Funds for the Central Universities, China (K20210123).

Author information

Authors and Affiliations

The Rural Development Academy &Agricultural Experiment Station, Zhejiang University, Hangzhou 310058, China

Xuping Feng

Huzhou Institute of Zhejiang University, Huzhou 313000, Zhejiang, China

Huifang

College of Biosystems Engineering and Food Science, Zhejiang University,

Hangzhou 310058, China

Xuping Feng, Zeyu Yu, Guofeng Yang, Yongqiang Shi, Yong He

Center for Data Science, Zhejiang University, Hangzhou 310058, China

Hangjin Jiang

College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, P. R. China.

Liting Chen, Xinran Zhou, Bing Hu, Guipei Xing, Boxi Zhao, Feng Liu

College of Life Sciences, Biological experiment teaching center, Nanjing Agricultural University, Nanjing, China.

Bing Hu, Chun Qin, Gang Hu, Guipei Xing

School of Mathematics and Statistics, University of Melbourne, Victoria, Australia.

Feng Liu

Department of Computer Science, Hong Kong Baptist University, China

Bo Han

Contributions

# Xuping Feng and Zeyu Yu contributed equally to this work.

Xuping Feng and Feng Liu designed and conceptualized the study. Zeyu Yu implemented the algorithm and analyzed the image data. Huifan, Feng Liu and, Bo Han deployed the OrgSegNet model on web portal. Hangjin Jiang, Yong He and Guofeng Yang review the manuscript. Liting Chen, Xinran Zhou, Boxi Zhao, Yongqiang Shi, Feng Liu labled the dataset. Bing Hu, Chun Qin, Gang Hu, Guipei Xing prepared of microscopy images.

Corresponding author

Yong He (yhe@zju.edu.cn) and Feng Liu (liufeng@njau.edu.cn)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalDataset.docx
- SupplementaryInformation.docx