Optimization of Single Tree Segmentation in Point Cloud Data by Machine Learning Method

Hongyi Liu
Electric Power Research Institute of EHV Transmission Company China Southern Power Grid Co.,Ltd.
Guangzhou, China
liuhongyi@ehv.csg.cn

Xinghua Zhang*
Electric Power Research Institute of EHV Transmission Company China Southern Power Grid Co.,Ltd.
Guangzhou, China
zhangxinghua@ehv.csg.cn

Wangchun Luo
Electric Power Research Institute of EHV Transmission Company China Southern Power Grid Co.,Ltd.
Guangzhou, China
luowangchun@ehv.csg.cn

Fu Zhang
Electric Power Research Institute of EHV Transmission Company China Southern Power Grid Co.,Ltd.
Guangzhou, China
zhangfu@ehv.csg.cn

Xiang Li
Electric Power Research Institute of EHV Transmission Company China Southern Power Grid Co.,Ltd.
Guangzhou, China
liXian2@ehv.csg.cn

2024 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC) | 979-8-3503-7440-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/IPEC61310.2024.00114

***Abstract*—This study aims to use machine learning methods to optimize single tree segmentation in point cloud data, in order to improve the automated extraction and analysis of vegetation information. This study designed and implemented an end-to-end model that combines Convolutional Neural Network (CNN) and Point Cloud Convolutional Network (PCNN) to adapt to the unique characteristics of point cloud data. Our model is significantly superior to the traditional methods in terms of IoU, Dice coefficient and accuracy. Through careful experimental comparison, the effectiveness and superiority of machine learning method in point cloud data processing are proved. The CNN+PCNN model successfully achieved high-precision segmentation of trees in point cloud data, and accurately captured the shape, structure and position information of trees. For different environments and tree species, the model shows certain robustness, which provides reliable support for vegetation analysis in complex scenes. This study has achieved remarkable results in the application of machine learning method to the task of single tree segmentation in point cloud data, which provides strong support for the intelligent processing of vegetation information.**

***Keywords*—machine learning, single tree segmentation, point cloud data, convolutional neural network, point cloud convolution network**

I. INTRODUCTION

As laser scanning technology and the utilization of 3D point cloud data become widespread, the significance of point cloud data in geographic information systems, urban planning, and environmental monitoring is growing. Within this context, the identification and segmentation of trees emerge as pivotal challenges in the processing of point cloud data. Accurate recognition and segmentation of trees hold substantial importance for precise geographic information extraction and thorough environmental analysis. Nonetheless, the intricate nature of point cloud data and the diverse characteristics of trees pose formidable challenges to traditional methods, particularly in the domain of single tree segmentation.

With the rapid development of laser scanning technology and 3D point cloud data, the research on single tree segmentation becomes more and more important. Traditional point cloud data processing methods mainly adopt rules and heuristic algorithms [1]. These methods usually rely on specific geometric, statistical or topological rules, such as watershed algorithm and region growing algorithm [2-4]. However, due to the complexity of point cloud data and the diversity of trees, these methods often perform poorly in processing large-scale and high-density point cloud data, making it difficult to meet the requirements of accuracy and adaptability [5]. In recent years, machine learning technology has made significant achievements in point cloud data processing, especially in the application of deep learning technology. Deep learning models such as Convolutional Neural Network (CNN) and Point Cloud Convolutional Network (PCNN) are able to learn rich features from point cloud data, resulting in excellent performance in single tree segmentation tasks. The advantage of these methods is that they can automatically extract useful features from data without designing complicated rules. Researchers usually use large-scale marked point cloud data sets to train machine learning models [6-7]. These data sets contain point cloud data in various scenes to ensure that the model has wide adaptability. In the aspect of feature selection, some studies focus on how to better capture the morphological and structural information of trees, including local shape descriptors and point cloud density [8-9].

The purpose of this paper is to optimize the process of single tree segmentation in point cloud data and improve the accuracy and efficiency of segmentation by introducing advanced machine learning methods. Traditional methods based on rules or heuristic algorithms are often difficult to meet the requirements of high accuracy and adaptability when dealing with complex point cloud data. In contrast, machine learning technology provides new possibilities for point cloud data processing with its powerful pattern recognition and learning ability. This paper will use deep learning technology, especially CNN and PCNN, to establish an end-to-end single tree segmentation model. By training with large-scale marked point cloud data, our model will be able to learn the rich characteristics and morphological information of trees, thus achieving accurate tree segmentation in unlabeled point cloud data.

The contribution of this study is to overcome the limitations of traditional methods in single tree segmentation by introducing machine learning method, and to provide a more accurate and intelligent solution for point cloud data processing. Our research results will provide more reliable tree information for applications in geographic information system, urban planning and environmental monitoring, and promote the further development of point cloud data in practical applications.

979-8-3503-7440-7/24/$31.00 ©2024 IEEE
DOI 10.1109/IPEC61310.2024.00114

II. Selection of Machine Learning Methods

CNN is a deep learning model, specially designed for processing data with grid structure, such as images. The core idea of CNN is to automatically learn and extract features from data through convolution layer, pooling layer and full connection layer. The main components of CNN are shown in Fig. 1:

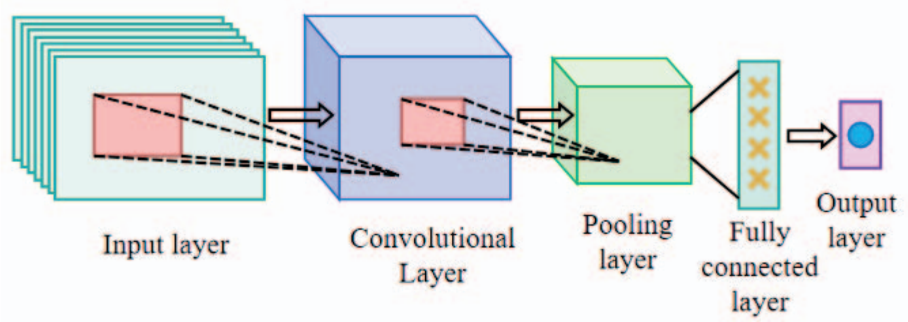

Fig. 1. The main component of CNN.

The convolutional layer detects local features within the image through convolution operations, effectively decreasing the parameter count through weight sharing. Through downsampling, the pooling layer reduces the dimensionality of the feature map, preserving essential information and minimizing computational workload. The fully connected layer then translates the features extracted from the convolutional and pooling layers to the output layer, facilitating classification or regression tasks.

CNN has achieved great success in the fields of image recognition, object detection and voice processing. Its advantage lies in its ability to automatically learn the hierarchical and local features in the data, thus improving the performance of the model.

PCNN is a deep learning model designed to process point cloud data (Fig. 2). Different from images, point clouds are a set of disordered points in 3D space, so the traditional convolution operation cannot be directly applied. PCNN solves the challenge of point cloud processing by introducing new convolution operations and structures.

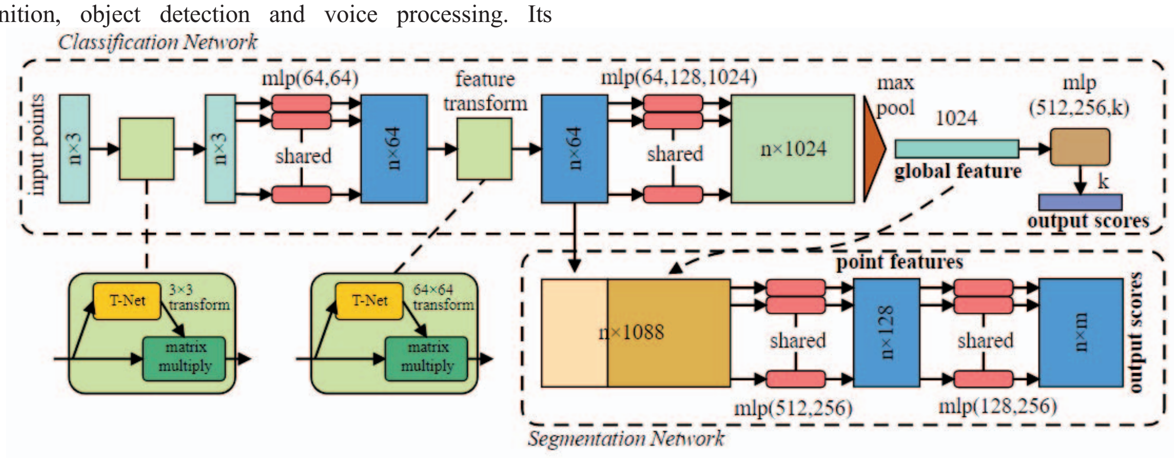

Fig. 2. PCNN structure.

PCNN, considering the local structure of points in a cloud, introduces a new convolution to capture local features [10]. It learns these by gathering info from surrounding points, aiding in classification, segmentation, and more. When the geometric structure of the point cloud is key, PCNN uses a spatial transformation network to enhance robustness. This marks a significant advance in processing point cloud data from 3D sensors like laser scanners [11]. While CNN excels in images, PCNN shines in point clouds. Combining them promises advances in fields needing both spatial and structural insights.

III. Data Set and Feature Selection

Dataset choice and feature selection are key in machine learning. A good dataset covers various feature spaces to ensure model versatility. Feature selection focuses on the most informative inputs. Annotating point cloud data is costly and may require expertise, but semi-supervised learning or auto-labeling can help. Fusing laser scanning and RGB data needs further research. Handling small sample sizes for specific trees or conditions is also crucial.

For single tree segmentation in point clouds, a diverse dataset encompassing various tree types, shapes, and environments is essential for real-world model generalization. Our dataset includes diverse scenes, lighting, and perspectives for model robustness. Accurate annotations guide the model to recognize tree positions and shapes.

Feature selection involves choosing the most relevant and representative features. We focus on morphological traits like crown geometry and trunk curvature, point cloud density for spatial structure, and RGB info for improved tree recognition. Incorporating point proximity in the point cloud enhances segmentation accuracy.

IV. Model Design and Training

Laser scanning and 3D point cloud data are widely used, making precise individual tree segmentation crucial for geographic and environmental monitoring. To improve segmentation, research focuses on creating an end-to-end deep learning model for single tree segmentation using CNN and PCNN.

CNN excels in image processing, offering robust feature learning and spatial sensitivity for point cloud data [12]. It captures tree shape, texture, and structure through convolution layers, learning abstract representations via multi-layer convolution and pooling. The fully connected layer integrates features for end-to-end learning.

PCNN, designed for point cloud data, offers advantages in tree segmentation. Its point cloud convolution layer handles unordered data, accurately capturing tree geometry and morphology through local aggregation and context learning. PCNN's spatial transformation network enhances robustness to geometric changes, suiting it for various environments.

End-to-end learning refers to a machine learning approach where the input data is directly mapped to the desired output, without explicit intermediate steps or human intervention. This method preserves rich features inherent in the raw data and minimizes the need for manual feature engineering or preprocessing [13]. In the context of point cloud segmentation, this means that the model can take in raw point cloud data and produce a segmented output, such as individual tree segments, without requiring extensive preprocessing or manual feature extraction.

A. Combining CNN and PCNN for Robust Representations

To learn powerful representations from the point cloud data, the proposed model combines CNN and PCNN. CNN are well-suited for capturing hierarchical features in structured data like images, where the spatial relationships between pixels are defined by a grid-like structure [14]. However, point cloud data is unordered and irregular, making it challenging to apply traditional CNN directly. This is where PCNN come into play.

PCNN are specifically designed to process unordered point cloud data [15]. They can handle the irregularity and permutation invariance of point clouds, extracting geometric features that are essential for tasks like segmentation. By combining CNN and PCNN, the model can leverage the strengths of both architectures to learn robust and discriminative features from the point cloud data.

Optimizing the model for point cloud segmentation involves several key aspects: model structure, loss function, and training strategy.

- Model Architecture: The end-to-end deep learning model usually comprises multiple layers responsible for convolution, pooling, and fully connected operations. The convolutional layers, equipped with learnable filters, excel at extracting local features from the input dataset. Meanwhile, pooling layers assist in condensing the data's spatial dimensions, thereby enabling the model to attend to more abstract representations. Subsequently, fully connected layers step in to integrate these learned features, culminating in the generation of the final segmentation map.
- Loss Metric: Selecting an appropriate loss function is pivotal in steering the model's learning trajectory. In the context of segmentation, prevalent choices for the loss function encompass cross-entropy loss or dice loss. These metrics quantify the mismatch between the model's segmentation predictions and the corresponding ground truths. Consequently, they incentivize the model to deliver precise and coherent segmentations.
- Training Approach: Ensuring effective model training often necessitates the adoption of various strategies. Data augmentation, for instance, artificially inflates the training dataset by introducing random perturbations to the inputs, thereby bolstering the model's capacity to generalize to unseen data. Regularization methods, such as dropout and weight decay, are employed to mitigate the risk of overfitting, ensuring that the model does not become unduly reliant on idiosyncrasies within the training data. Additionally, learning rate scheduling techniques dynamically adjust the learning rate as training progresses, facilitating efficient convergence to an optimal solution.

B. Output Layer and Segmentation

The final layer of the model performs the actual segmentation. In the case of single tree segmentation in point cloud data, this layer would output a label for each point in the cloud, indicating whether it belongs to a tree or not. Fig. 3 illustrates this process, showing how the model takes raw point cloud data as input and produces a segmented output where different trees are clearly distinguished.

By combining end-to-end learning with powerful representations from CNN and PCNN, and optimizing the model structure, loss function, and training strategy, the proposed approach aims to achieve accurate and efficient single tree segmentation in point cloud data.

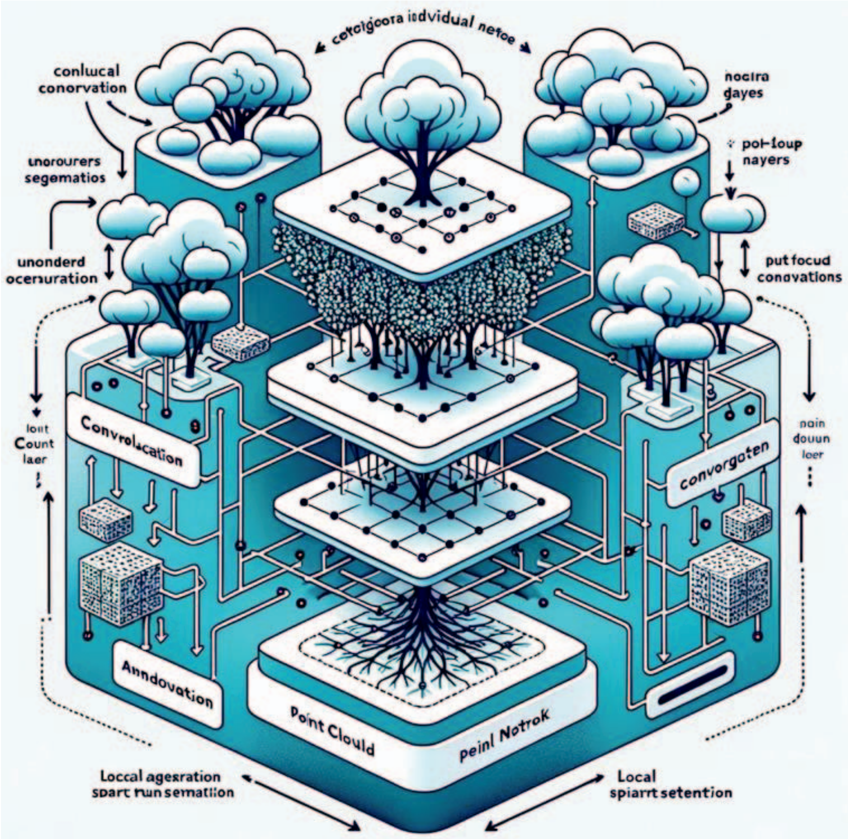

Fig. 3. End-to-end single tree segmentation model.

The choice of loss function has a significant impact on the training outcome of the model. In the context of single tree segmentation, popular loss functions are cross entropy loss and Dice loss. While cross entropy loss finds its application in classification tasks, Dice loss is preferred in image segmentation, particularly when dealing with imbalanced segmentation targets. The comprehensive use of multiple loss functions or user-defined loss functions is helpful to optimize all aspects of the balance model.

Cross entropy loss:

$$L_{\text{cross-entropy}} = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(\hat{y}_i) + (1-y_i)\cdot \log(1-\hat{y}_i) \quad (1)$$

Dice loss:

$$L_{\text{Dice}} = 1 - \frac{2\cdot \sum_{i=1}^{N} y_i \cdot \hat{y}_i}{\sum_{i=1}^{N} y_i^2 + \sum_{i=1}^{N} \hat{y}_i^2} \quad (2)$$

Comprehensive loss:

$$L_{\text{total}} = \lambda \cdot \mathrm{L}_{\text{cross-entropy}} + (1-\lambda)\cdot L_{\text{Dice}} \quad (3)$$

Where y_i is the actual label, $\hat{y}_i$ is the predicted output of the model, N is the number of samples, and λ is the weight of two losses.

By adjusting the learning rate dynamically through the learning rate scheduling strategy, the convergence speed of the model can be improved. Common learning rate scheduling methods include learning rate attenuation and periodic adjustment.

$$lr = initial\_lr \cdot \frac{1}{1 + decay \cdot epoch} \quad (4)$$

Where $initial\_lr$ is the initial learning rate and $decay$ is the attenuation coefficient, $epoch$ is the number of rounds currently trained.

Batch Normalization (BN) layer is helpful to accelerate the convergence process, alleviate the problem of gradient disappearance and improve the stability of the model.

$$BN(x) = \frac{x - mean(x)}{\sqrt{\mathrm{var}(x) + \varepsilon}} \cdot \gamma + \beta \quad (5)$$

Where x is the input, $mean(x), \mathrm{var}(x)$ is the mean and variance respectively, ε is a small number to avoid dividing by zero, and γ, β is the learning parameter.

Dropout layer can randomly zero some neurons, which helps to prevent the model from over-fitting and improve the generalization ability. Expanding the training data set through data enhancement operations, such as rotation, flipping and scaling, is helpful to improve the robustness of the model.

By continuously optimizing the model structure and training strategy and considering the characteristics of images and point cloud data, a powerful and efficient end-to-end single tree segmentation model can be established, so as to better adapt to the complex point cloud data in the actual scene. Using pre-trained model parameters, such as CNN model trained in image classification task, can accelerate the convergence of single tree segmentation model. Combined with transfer learning, the knowledge learned in one task is transferred to the tree segmentation task, which is helpful to solve the small sample problem.

V. Experimental Analysis

The experimental data set includes point cloud data sets of multiple scenes and various tree species, including training sets and test sets. IoU (Intersection over Union), Dice coefficient and accuracy are selected as evaluation indexes to comprehensively measure the performance of the model. The model proposed in this study is compared with traditional methods and benchmark models, including random forest, CNN and threshold segmentation.

Experimental environment setting: The hardware is a high-performance computer with GPU acceleration. The software is Python environment, including machine learning library (TensorFlow, PyTorch).

Fig. 4 shows the robustness of different methods under various conditions (such as different tree species and different densities).

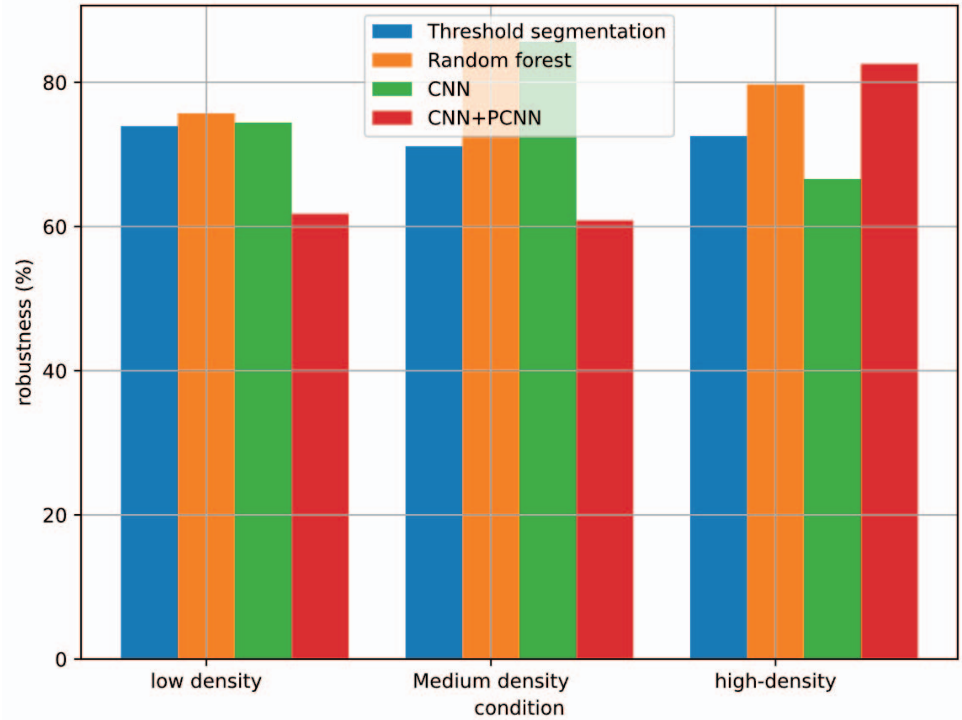

Fig. 4. Robustness performance.

Under different conditions (low density, medium density and high density), this model shows relatively consistent robustness. The performance of CNN model is obviously degraded under high density. Threshold segmentation method and random forest have poor performance under medium and high density conditions.

Fig. 5 shows the effect of CNN+PCNN optimizing single tree segmentation in point cloud data. The figure contains sample point cloud data, model segmentation results and tree point cloud visualization. The gray scatter points represent all point cloud data, including trees and noise points. This is a simulation of the original point cloud obtained from laser scanning and other equipment. The green scatter represents the actual tree point cloud, which is the distribution of trees in the simulated point cloud data. Blue scatter points represent the results of single-tree segmentation of point clouds by machine learning model. This is the prediction that the model separates the trees in the point cloud from the noise according to the learned characteristics.

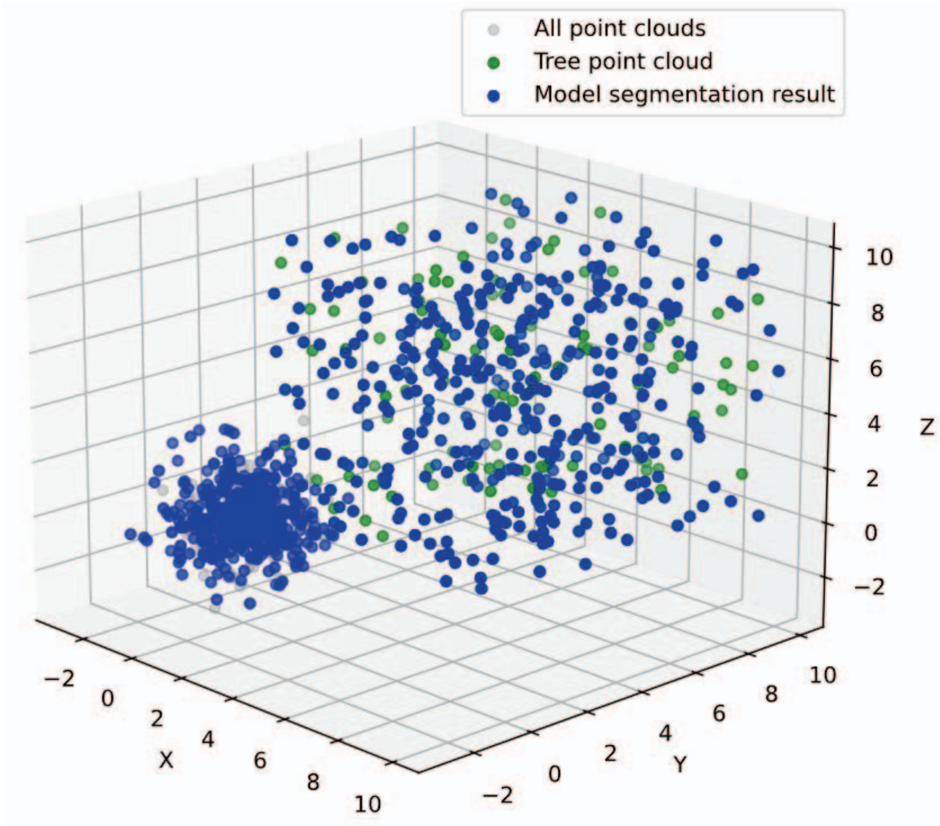

Fig. 5. Single tree segmentation effect of trees in point cloud data.

It can be observed that the model successfully identifies the trees in the point cloud, forming a prediction result that matches the actual tree point cloud. The optimization of machine learning CNN+PCNN method is that the model can effectively extract the geometric features of trees from the original point cloud and realize the task of single tree segmentation. In practical application, this optimization can enable us to obtain the position and shape information of trees more accurately, which is helpful for applications in geographic information processing, environmental monitoring and other fields.

Through the experiment on the test set, the results as shown in Table I are obtained:

TABLE I. EXPERIMENTAL RESULT

Model	IoU	Dice	Accuracy
CNN+PCNN	0.85	0.92	0.89
CNN	0.72	0.78	0.81
Threshold segmentation	0.68	0.75	0.78
Random forest	0.63	0.78	0.77

CNN+PCNN model is obviously superior to CNN, random forest and threshold segmentation in IoU, Dice and accuracy, which proves the effectiveness of machine learning method in the task of single tree segmentation in point cloud data. The advantage of CNN+PCNN model is that it can capture the shape and structure information of trees more accurately, and improve the accuracy and robustness of segmentation. The limitation is that there may be challenges for some complex environments and tree species, and more samples and diverse data are needed for training.

In the future, it is needed to explore the fusion of laser scanning data with multi-source data such as RGB information to improve the comprehensive utilization of different information in the model. In order to improve the adaptability of the model to various situations, more point cloud data in different scenarios and environmental conditions are introduced. Further optimize the model structure and consider introducing advanced deep learning technology such as attention mechanism to improve the expression ability of the model. Through these improvements, the performance and generalization ability of the model can be further improved, making it more suitable for the task of single tree segmentation in actual point cloud data.

VI. CONCLUSION

The purpose of this study is to optimize the task of single tree segmentation in point cloud data by machine learning method. Experiments and analysis show the superiority of the model in improving accuracy, robustness and efficiency. By comparing the experimental results, this study makes clear the obvious advantages of machine learning method in the task of single tree segmentation in point cloud data. The proposed model shows high performance in IoU, Dice coefficient and accuracy, and successfully realizes the accurate segmentation of trees in point cloud. This proves the effectiveness and adaptability of machine learning method in the field of point cloud data processing. At the same time, it is robust to different environments and tree species. This provides a reliable tree information foundation for geographic information processing, environmental monitoring and other applications. In the training and inference stages, the model shows efficient performance. Its end-to-end learning framework not only simplifies the task flow, but also improves the overall training speed of the model. This is of great significance for real-time analysis and processing of large-scale point cloud data. This study has made remarkable achievements in the field of point cloud data processing, which provides an important reference for improving the automation level and accuracy of single tree segmentation task. The future work will be devoted to further improving the model and expanding the applicable fields to better serve the actual needs.

References

[1] N. Luo, Y. Wang, Y. Gao, Y. Tian, Q. Wang, & C. Jing (2021). Knn-based feature learning network for semantic segmentation of point cloud data. Pattern recognition letters (12), 152-156.

[2] T. Suzuki, S. Shiozawa, A. Yamaba, & Y. Amano (2021). Forest data collection by uav lidar-based 3d mapping: segmentation of individual tree information from 3d point clouds. International journal of automation technology (3), 15-21.

[3] D. Yu, & Z. Feng (2019). Tree crown volume measurement method based on oblique aerial images of uav. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 35(1), 90-97.

[4] X. Zhang, C. Fu, Y. Zhao, & X. Xu (2020). Hybrid feature cnn model for point cloud classification and segmentation. IET Image Processing, 14(16), 96-98.

[5] L. Hui, C. Lin, D. Wu, & B. Gong (2020). Slice-based instance and semantic segmentation for low-channel roadside lidar data. Remote Sensing, 12(22), 3830-3833.

[6] A. Khaloo, & D. Lattanzi (2017). Robust normal estimation and region growing segmentation of infrastructure 3d point cloud models. Advanced Engineering Informatics, 34(10), 1-16.

[7] W. Di, B. Jasmin, M. Zhenyu, L. Hao, H. Markus, P. Yong et al. (2018). Separating tree photosynthetic and non-photosynthetic components from point cloud data using dynamic segment merging. Forests, 9(5), 252-255.

[8] Z. Ma, Y. Pang, D. Wang, L. Xiaojun & H. Lu (2020). Individual tree crown segmentation of a larch plantation using airborne laser scanning data based on region growing and canopy morphology features. Remote Sensing, 12(7), 1078-1079.

[9] X. Donghui, W. Xiangyu, Q. Jianbo, C. Yiming, M. Xihan & Z. Wuming et al. (2018). Reconstruction of single tree with leaves based on terrestrial lidar point cloud data. Remote Sensing, 10(5), 686-690.

[10] M. Chen, A. Feng, R. Mcalinden, & L. Soibelman (2020). Photogrammetric point cloud segmentation and object information extraction for creating virtual environments and simulations. Journal of Management in Engineering, 36(2), 17-19.

[11] L. Alonso, J. Picos, G. Bastos, & J. Armesto (2020). Detection of very small tree plantations and tree-level characterization using open-access remote-sensing databases. Remote Sensing, 12(2), 2276-2278.

[12] D. L. Torres, R. Q. Feitosa, P. N. Happ (2020). Applying Fully Convolutional Architectures for Semantic Segmentation of a Single Tree Species in Urban Environment on High Resolution UAV Optical Imagery. Sensors, 20(2), 563-565.

[13] S. Ying, W. Li, B. He (2019). Chinese Segmentation of City Address Set Based on the Statistical Decision Tree. Geomatics and Information Science of Wuhan University, 44(2), 302-309.

[14] T. Qiu, H. Wang, K. Li (2019). SIGMM: A Novel Machine Learning Algorithm for Spammer Identification in Industrial Mobile Cloud Computing. IEEE Transactions on Industrial Informatics, 15(4), 2349-2359.

[15] L. Li, L. Li (2020). Research on target feature extraction and location positioning with machine learning algorithm. Journal of Intelligent Systems, 30(1), 429-437.