IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 October 2023, accepted 26 October 2023, date of publication 14 November 2023, date of current version 21 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3333280

RESEARCH ARTICLE

Point Sampling Net: Revolutionizing Instance Segmentation in Point Cloud Data

NANDHAGOPAL GOMATHI[1], KRISHNAMOORTHI RAJATHI[1], MIROSLAV MAHDAL[2], AND MUNIYANDY ELANGOVAN[3,4]

[1]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai 600062, India
[2]Department of Control Systems and Instrumentation, Faculty of Mechanical Engineering, VSB—Technical University of Ostrava, 70800 Ostrava, Czech Republic
[3]Department of Research and Development, Bond Marine Consultancy, EC1V 2NX London, U.K.
[4]Department of Biosciences, Saveetha Institute of Medical and Technical Sciences, Chennai 602105, India

Corresponding author: Miroslav Mahdal (miroslav.mahdal@vsb.cz)

This work was supported by the Application of Machine and Process Control Advanced Methods through the Ministry of Education, Youth and Sports, Czech Republic, under Project SP2023/074.

ABSTRACT Today, there is a great need for 3D instance segmentation, which has several uses in robotics and augmented reality. Unlike projective observations like 2D photographs, 3D models offer a metric reconstruction of the sceneries without occlusion or scale ambiguity of the environment. In agriculture, understanding Plant growth phenotyping enhances comprehension of complex genetic features and accelerates the advancement of contemporary breeding and smart farming. A reduction in crop production quality is caused by leaf diseases in agriculture. In order to increase productivity in the agricultural industry, it is therefore possible to automate the recognition of leaf diseases. Diverse leaf disease patterns affect the detection's accuracy in the majority of systems. During phenotyping, 3D PCs (PC) of components of plants like the stems and leaves are segmented in order to follow autonomous growth and estimate the level of stress the crop has experienced. This research proposed a Point Sampling Method with occupancy grid representation for segmenting PCs of different plant species, which was developed. To handle unordered input sets, this approach mainly relies on the application of the single symmetric function max pooling. In reality, a set of optimization functions are used by the network to choose points which is more curious or instructive from the PC and encapsulate the selection reason, and the fully connected layers, used for shape classification or shape segmentation, integrate these learned ideal significances hooked on a global descriptor regarding the overall shape. After being trained on the Point Sampling Network-created plant dataset, the network can simultaneously realize semantic and leaf instance segmentation.

INDEX TERMS Instance segmentation, PC data, point sampling net, point clustering, semantic segmentation, leaf segmentation.

I. INTRODUCTION

A new field of study called plant phenotyping links genetics to agriculture, ecology, and plant physiology [1]. In order to intuitively depict plant growth, it investigates a group of markers produced by the ever-changing interaction of the environment and genes throughout development. Study [2] investigates the correlation between phenotypes and genotypes by computerized digitalization. This research endeavor seeks to enhance our comprehension of intricate genetic features and, ultimately, expedite the advancement of contemporary breeding techniques and precision agriculture [3], [4]. In general, the investigation of plant phenotype is primarily concerned with the organs, covering aspects like the shape and structure of the root system and the characteristics of the fruits and leaves. The primary site of photosynthesis and respiration is found in leaves because they are that organ's largest surface area [5]. Thus, some of the most significant phenotypic characteristics of a leaf are its width, length, and inclination [6]. The stem system acts as the solid foundation of the plant's structure and connects all other organs geographically, including the plant's leaves,

The associate editor coordinating the review of this manuscript and approving it for publication was Abdel-Hamid Soliman.

flowers, and fruits. How much stress the plant has been exposed to can be determined by the phenotyping of the stems [7], [8]. Plant phenotyping success depends on the efficient and precise segmentation of plant parts. The task of segmenting plant organs has been the subject of various studies since the 1990s, particularly segmenting leaves to identify diseases. The basis of 2D picture-based phenotyping is typically conventional algorithms for pattern recognition, image processing, and machine learning for segmentation based on a threshold value [9], edge detection method [10], region growing approach [11], and clustering algorithm [12]. Convolutional neural network (CNN)–based deep learning algorithms have recently advanced to state-of-the-art levels in the classification of images and image segmentation [13]. Picture networks with deep learning were used in references [14] to separate both leaves and fruits in pictures of plants.

In spite of this, the few monocotyledonous plants with fewer leaves that the 2D phenotyping approaches often operate with are plant 1 and plant 2. The primary cause for this is that depth information cannot be included in a 2D image because it can only be taken from one viewing angle. The blocking and overlap in the canopy between the leaves also substantially hinder segmentation techniques based on 2D images [15], [16]. Analysis of the observed behavioral features' statistical significance is also weakened because pictures do not adequately depict the whole location distribution of the plant's structure. In addition to color, texture, and other visual signals, 3D models can convey the most important information depth as compared to pictures. The cornerstone for very accurate phenotypic assessment is depth directly resolves the problems of occlusion and overlaps. Plant phenotyping techniques based on depth pictures or PC data have been quickly developing in recent years as a result of the advent of inexpensive and high-precision 3-dimensional imaging technologies [17], [18]. The Lidar approach, which utilizes laser-based sensors, has become a widely employed tool for ''3D reconstruction'' and phenotyping. This is mostly due to its exceptional precision in capturing detailed information about large trees in three-dimensional imaging [19], monocot plant [20], cotton plant [21], and several other cash crops [22]. 3D sensors based on time-of-flight (ToF) and structured light & are essential tools for Plant 3D phenotyping due to their exceptional real-time capabilities [23]. In the reference [24], a range of greenhouse and cash crops were recreated and analyzed using binocular stereo vision. References [25] performed Crop phenotypic analysis and 3D reconstruction utilizing the multiple view stereo (MVS) method. Yang et al. [26] created a toolset called Label 3D maize plant that manually identifies 3D PC data of maize plant branches so that it may be used for testing and training machine learning models.

- This research acknowledges the challenge posed by diverse leaf disease patterns in the majority of systems. By using 3D instance segmentation, it seeks to address this issue by enhancing the accuracy of disease detection.
- The proposed Point Sampling Method with occupancy grid representation is a novel approach to segmenting 3D point clouds of different plant species.
- The network's integration of shape classification and shape segmentation into a global descriptor for overall shape is a valuable contribution.
- This approach allows the network to simultaneously realize semantic and leaf instance segmentation, providing a comprehensive understanding of the plant's structure and health.

Despite having good 3D data, the separation of plant entities after the developing chunk and the subsequent division of Phenotyping processes are challenging when determining the phenotypic parameters for each plant's organs. The use of 3D PCs for unsupervised leaf segmentation has already gained prominence. To distinguish leaves from stems, enhanced the PC net identification technique and created a mixed segmentation algorithm that could account for the physical differences between different individuals of cotton. He et al. [27] presented an approach that was data-driven, dynamic, and proposal-free and generated the proper convolution kernels to use depending on the circumstances. They investigated a vast context by assembling identical points that had within votes for the geometric centroids and shared the same semantic categories in order to make the kernels discriminative. Poux et al. [28] presented a method for the automated extraction of objects from 3D point clouds. Their approach involved an unsupervised segmentation technique and a classification process based on an ontology, which was further enhanced by self-learning mechanisms. Wang et al. [29] investigated a technique for accomplishing organ-level 3D instance segmentation in lettuce. Their objective was to train a neural network capable of accurately segmenting point clouds of various plant leaves into distinct instances. Zong et al. [7] employed a novel approach to enhance the performance of a 3D point cloud segmentation model, with the aim of guiding the segmentation process for cells and accurately fitting the structure of the container ship cell guide. As a result, it facilitated the implementation of the container simulation test box functionality. Following the segmentation of the cell guide, a precision inspection was carried out concurrently to assess its practicality. Wang et al. [30] introduced a novel unsupervised pipeline that effectively carried out the tasks of single-tree separation and leaf-wood classification concurrently. The approach suggested in their study is founded upon a unique superpoint graph that incorporates comprehensive node and edge properties. The utilization of a joint dual-task network facilitated the acquisition of class labels and the grouping of data points through the incorporation of graph operators. Vu et al. [31] introduced SoftGroup, a method that demonstrates simplicity and effectiveness in the task of instance segmentation on 3D point clouds. The company SoftGroup created a method of grouping soft semantic scores in order to resolve the issue

that arises from the practice of hard grouping on objects that are locally ambiguous. The suggestions derived from the grouping stage were allocated to either positive or negative samples. Next, a refining stage was implemented using a top-down approach to enhance the positive aspects and mitigate the unfavorable aspects. Yi et al. [32] introduced GSPN, an innovative object suggestion network designed for the purpose of instance segmentation in 3D point cloud data. The GSPN algorithm produced suggestions for objects of excellent quality and objectness, resulting in a significant enhancement of performance for an instance segmentation framework. Zanjani et al. [33] introduced a novel comprehensive learning framework, referred to as Mask-MCNet, designed for the purpose of tooth instance segmentation within a three-dimensional point cloud of intra-oral scan data. The precise segmentation of teeth is a crucial component in the development of automated computational dentistry.

Point cloud instance segmentation is essential for comprehending 3D scenes. There are still a few issues that need to be resolved, including (i) the absence of a common down sampling technique for point clouds that are specifically prepared for deep learning and (ii) the network architecture for multipurpose segmenting a point cloud is difficult; for example, a network is difficult to maintain equilibrium between the instance segmentation task and the organ semantic segmentation task. A deep learning network called PointNet was created to perform 2 Plant Phenomics plant body semantic segmentation and plant instance segmentation simultaneously using a fully annotated point cloud dataset comprising several species in order to overcome the aforementioned difficulties. Deep learning based Down sampled Point-wise instance labels algorithm proposed to segmentation on point cloud data.

Further, this research work is organized; section II provides an explanation of the materials and related processes. Comparative tests and outcomes are provided in Section III. Section IV offers more comments and analyses. The final portion is when the conclusion is reached.

II. MATERIALS AND METHODS

A. DATA PREPROCESSING

Data preprocessing, for instance, segmentation in point cloud data, involves cleaning, transforming, and organizing 3D point clouds for better model input, often including noise removal, downsampling, and feature extraction. Here, a median filter is chosen to preprocess the data.

Utilizing digital methods, image noise reduction minimizes undesirable brightness and color fluctuations caused by low light or heightened sensor sensitivity. Algorithms discern and mitigate noise while preserving vital details, elevating image quality. Employing filters and statistical approaches enhances clarity, rendering it apt for photography, medical imaging, and computer vision, ensuring accuracy in depiction.

$$Median(C) = Med\ \{\mathrm{Ci}\} \tag{1}$$

$$= \mathrm{Ci}(M+1)/2, M \text{ is Odd} \tag{2}$$

$$= 1/2[\mathrm{Cj}(M/2) + Cj(M/2)+1],\ Mis \tag{3}$$

B. PC SAMPLING

Efficiently managing complex point clouds can pose significant computational challenges due to their inherent high precision, frequently comprising tens of thousands of data points. Furthermore, a majority of current 3D neural networks demand standardized inputs featuring a fixed number of points, thus mandating the need for subsampling prior to in-depth processing and modeling [34]. In this research, we introduce novel enhancements to the well-established techniques of ''Voxel-Based Sampling (VBS) and Farthest Point Sampling (FPS)''.

The suggested technique demonstrates efficient preservation of local point cloud (PC) density during the sampling process, requiring low computational effort. While the existing FPS (Farthest Point Sampling) sections are inherently concise and sparse, there is a potential risk of rapid information loss, which constitutes a notable drawback. Within the three-dimensional (3D) space of the PC, a voxel is characterized by three parameters: height, width, and length, with Voxel-based Sampling (VBS) employed as a down-sampling method. VBS replaces the original coordinates within each voxel with the voxel's center of gravity to achieve this down-sampling

VBS handles data quickly, but it faces two significant limitations: (i) The voxelization scale relies on three variables that fluctuate based on the PC's size, density, and the total pixel count post-VBS sampling, making it unsuitable for processing large batches of PCs and datasets due to unpredictability. (ii) VBS produces PCs with an evenly distributed data distribution, which is not conducive to efficient deep learning. After carefully considering the pros and cons of both FPS and VBS, we introduce a novel PC downsampling technique known as Point Sampling (PS). Figure 1 shows the Point Sampling Technique.

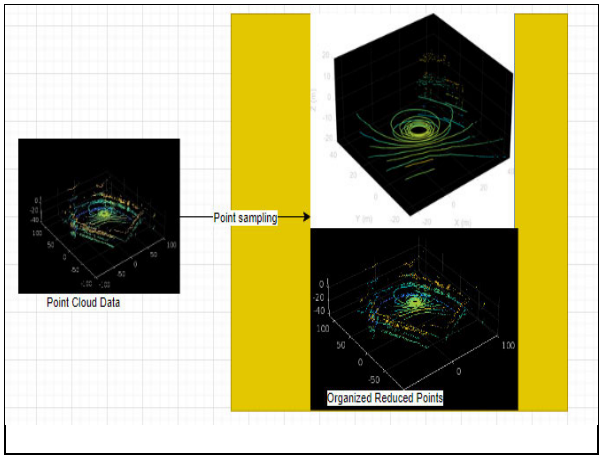


FIGURE 1. Point sampling technique.

N, the quantity of down-sampled points, comes first and must be established. In order to run Point Sampling on the initial PC and create a PC that contains a few extra points than N, the second stage includes changing the pixel settings. The temporary pixels of the PC are subjected to the percentage of points to be reduced, and the result is a down-sampled N-point image. For all of the tests in this investigation, N was set to 4096.

C. NETWORK ARCHITECTURE

The entire design of the backbone architecture of Point Net is displayed in Figure 2. The labeling network receives n data points as input, transforms the input and features, and then pools all of the point's features together. The results are my labeling scores. The labeling net is expanded upon by the segmentation network [35], [36]. It combines local and global elements, as well as outcomes measured in points. All layers using ReLU employ the batch normalization method. In a label classification net, dropout layers are employed for the final mlp. The three main components of an end-to-end network. The front half contains an encoder-like structure typical of deep neural networks and four consecutive Neighborhood Point Attribute Extraction for Modules (APEMs) for feature calculation [37], [38]. Prior to each APEM, the prominent area is down-sampled, compressing the features and decreasing the number of points. Deep Fusion Net [36] created a model of the network's core. The lower decoder, which is more complex than the more efficient decoder and indicates a more fine-grained process, is initially executed by this component, which we refer to as the Point Attribute Fusion Module (PAFM).

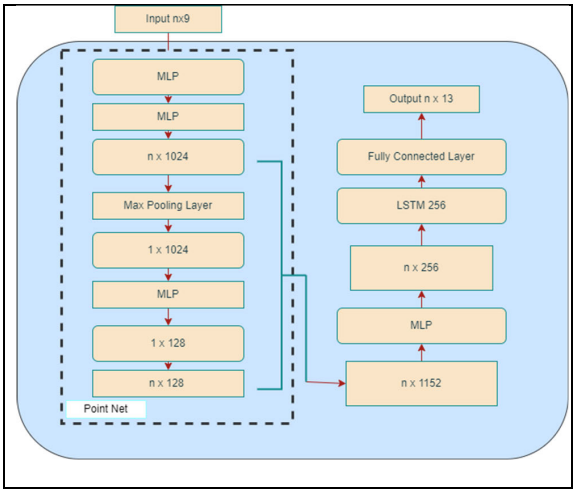


FIGURE 2. Point net architecture.

At the end of the PAFM, feature fusion is realized using convolutions and a concatenation operation [39]. The feature flows of the top and bottom branches are utilized to depict the instance segmentation problem and the semantic segmentation challenge correspondingly. In addition, in order to complete the process of semantic segmentation, the semantic flow within the output module of PointNet applies the "argmax operation" to the last feature layer, therefore extracting the predicted semantic labels for each individual point. Meanwhile, Mean Shift clustering is effectively employed within the instance flow's final feature layer to achieve precise instance segmentation. The fine-grained 3D instance segmentation is quite difficult. Assigning part category is the task given a 3D scan data. Label each point or face (e.g., leaf). We test using the Velodyne lidar. Most categories of objects have labels that a pair or components. Annotations pertaining to ground truth are tagged on the forms' sampled points. Part segmentation is formulated as a per-point classification problem.

Algorithm 1 Point Clustering

Input: G, P_{ins} (ij)
Output: Point-wise instance labels for the entire PC 1 set $G = -1$;
2 ClusterCount $\leftarrow 0$;
3 ***Repeat for*** each block i ***do*** 4 check i equals to block1, **then**
5 Repeat ***for*** every point P_{ins} (ij) in block I ***do***
6 Define x where P_{ins} (ij) is located in the xth cell of G;
7 $G_x \leftarrow P_{ins}$ (ij);
8 ***end*** *for*
9 ***else***
10 Repeat ***for*** every instance y_j in block y do
11 Define G_x points in y_j are located in cells G_i;
12 $G_i \leftarrow$ the cells in G_i that do not have value -1; 13 check ***if*** the frequency of the mode < 30, ***then***
14 $G_i \leftarrow$ ClusterCount;
15 ClusterCount $\leftarrow$ ClusterCount $+ 1$;
16 ***else***
17 $G_i \leftarrow$ the mode of G_i;
18 ***endif***
19 ***end for***
20 ***end*** if
21 ***end for***
22 ***Repeat for*** every point P_{ins} (ij) in the whole scene ***do***
23 Define k where P_{ins} (ij) is located in the kth cell of G;
24 $G_i \leftarrow G_i$;
25 ***end for***

D. ADJACENT POINT EXTRACTION FOR MODULE (APEM)

To improve PointNet's front-end feature extraction, APEM was developed. In the front section, there are four APEMs that follow one another. APEM is always followed by a feature space-down sampling in this front part that resembles an encoder. APEM possesses the following two characteristics: (i) It concentrates on simultaneously extracting high-level and low-level information, and (ii) The successful utilization of the encoder facilitates the extraction of information that is transported from a local scale to a global scale. This is achieved by consistently "aggregating the neighborhood information" inside the feature space.

E. POINT ATTRIBUTE FUSION MODULE

According to Deep Fusion Net, both fine-grained point-level features and coarse-grained voxel-level data are very beneficial for feature learning. Combining the data at two different granularities enhances this network's capacity to extract characteristics [40], [41]. We incorporate a ''Point Attribute Fusion Module (PAFM)'' into the core of our PointNet as an homage to Deep Fusion Net. The PAFM model employs two decoders, one for coarse-grained learning and another for fine-grained learning, to generate feature layers that exhibit different levels of learning granularity [42], [43]. The module first obtains the coarse-grained and fine-grained feature maps, denoted as Fc and Ff, respectively. These feature maps are then combined using techniques such as average pooling and feature concatenation. The resulting aggregated feature denoted as FDGF, is obtained after applying a 1D convolution with the rectified linear unit (ReLU) activation function. The results of PAFM have the potential to enhance the performance of the following semantic and instance segmentation tasks [44], [45].

F. INTERNET OF INTENSITY AND LOCATION MODULE

The study of attention is now a viable subject despite being widely used in 2D image recognition and deep learning. It is still early in the development of 3D PC learning networks to apply attention mechanisms. Based on this, an Intensity and Location Attention Module was developed with two sub-attentional mechanisms: Location Attention (LA) and Intensity Attention (IA) [46], [47], [48]. In order to emphasize the significance of crucial points in the PC representation, Location Attention (LA) emphasizes the data points that may successfully express essential attributes with Neighborhood location and give greater priority. The feature vectors with the same intensity that include the more intriguing data are more likely to be the focus of intensity attention (IA). By aggregating along the attribute channel direction, the significance of each feature distribution is automatically determined [38]. Then, for every channel based on its significance, the significant feature dimensions are amplified while the unrelated ones are simultaneously muted.

In the attention module, the top layer point is responsible for instance segmentation, and the bottom layer point is responsible for semantic segmentation. The specifications of concurrent modifications are not the same, even though the processing parameters for both flows are the same. A vector of size 4096 is constructed by performing mean pooling on the 4096 output characteristic produced by the geographical attentiveness module of the PAFM. The resulting vector is then subjected to the sigmoid function to obtain the position weight vector. To create the feature that is reinforced by SA, whose size is 128, the weight vector and PAFM feature are multiplied [49], [50].

Finally, a 128-channel frequency vector is produced using the sigmoid method. The reinforcing attribute vector of the precise location process, which is also an outcome of the attention module, is created by multiplying the aggregated feature vector by the weight vector. After applying an additional 1D convolution for the calculation of the attention unit, the instance flow produces ''a 4096 5 feature map'' for the segmentation of individual PC instances. The instance segmentation result is then produced using the characteristic map and Mean Shift clustering technique. After the clustering process, the loss function is computed to evaluate the segmentation accuracy of the examples during the training phase. By integrating an additional one-dimensional convolution and an Argmax algorithm into the output feature of the attention mechanism (AM) for semantic flow, where C denotes the number of semantic classes, we can generate a 4096. On this page, the semantic segmentation loss is determined.

Each scene is divided into 1 × 1 meters. Also, the entire scene divided into a 400 × 400 × 400 grid G is. Gi serves as a visual cue for the cell's instance label. i ∈ (0, 400 × 400 × 400). A point clustering algorithm is created to combine object instances from distinct blocks given G, and point-wise instance the names of each block Pins anywhere Pins (ij) denotes the block I instance label for the jth point. In Algorithm 1, the process of clustering the points from different blocks after labeling.

G. LOSS FUNCTIONS

Our PointNet controls numerous operations simultaneously using carefully built various loss functions. Deep learning networks must be trained using loss functions. The standard cross-entropy loss function, or Lsem, is used for the semantic segmentation position and is defined as follows.

$$\text{loSEG} = -\sum_{\text{i=0}}^{n}\sum_{\text{j=1}}^{c} \text{p}\,(\text{X})_j^N \text{logp}\,(X)_J(i) \tag{4}$$

where

xj - i is the expected likelihood that ''the current point pi belongs to class j.'' xj is the predicted probability that the current point pi belongs to class j (example leaf) and the one hot encoding of the point's actual semantic class is represented by xj. If the argument really falls within category j, with xj 'as its value If equals 1, then 0. In the task of instance segmentation.

If the point actually falls within category j, it returns 1 else. It is 0. The quantity of PC input instances for the instance segmentation task can vary. Hence, in order to supervise the training process, ''a comprehensive loss function'' is utilized, consisting of three sub-losses with varying weights across an indeterminate number of instances. The instance loss loINS equation is presented as follows:

$$\text{loINS} = \text{l}\alpha * \text{sl}\,(\text{SEG}) + \text{l}\beta * \text{sl}\,(\text{INSLBL}) + \text{l}\gamma * \text{sl}(\text{REGU}) \tag{5}$$

The weights of sub-loss for segments from PC repel different labels from clusters, and regularization losses aid in forming a regular ordered and effective feature limit for computation. Here, l and sl represent loss and sub-loss in instance segmentation; the symbols α, β, Υ, respectively, show total

losses. The center of ith insatance is taken from feature space and the feature vector created. In this research, n of instance is taken for computing the loss function. Parameter d is taken as a threshold for two neighborhood clusters. Finally, Features that are both coarse and fine-grained are integrated in order to get the output goal.

A Double-hinge Loss (DHL) point level was suggested in reference to taking into account the limitations of the instance task on the semantic task on the network's mid-level characteristics. Since DHL was immediately translated Based on the feature pattern (Lo_{fm}), the PointNet total loss function is depicted as follows.:

$$TotalLoss = Lo_{seg} + Lo_{ins} + Lo_{fm} \tag{6}$$

III. EVALUATION MEASURES

To assess the effectiveness of PointNet in the semantic segmentation of the PC, four quantitative metrics for each semantic class: Precision (Pre), Recall (Re), Intersection over Union (IoU), and F1. It is better if each of the four semantic metrics has a higher score. The percentage of all points predicted by the network that were correctly categorized into this semantic class is the measure of precision. IoU measures the amount of overlap between each semantic category's expected and actual areas, whereas F1 is a full indication produced by harmonically averaging Pre and Re.

The instance segmentation outcomes were evaluated using the metrics mCov, mWCov, mPrec, and mRec. The initial knowledge point set for the mth case inside a given semantic group and the anticipated points set for the nth event within the same semantic class are collectively referred to as IGm. The highest possible value of the entire assessment is denoted by max 12. The network's predicted point set and the ground truth point set are the two inputs that the binary function IoU uses to build the semantic IoU equation. The parameter C determines how many semantic classes are utilized to calculate Mean Recall (mRec) and Mean Precision (mPrec).

IV. EXPERIMENTS AND RESULTS

The data utilized in this investigation was sourced from the plant PC, which was generated through laser-scanned point cloud data collection. The dataset includes observations of three distinct crops thriving in various environmental conditions, with each crop undergoing repeated scanning over a 30-day period. Show the crops in Figure 3 at various growth phases (Gaidon, Wang, Cabon, & Vig, 2016). The dataset has a 25 m maximum scanning error. The collection has a total of 546 unique PCs, including 129 sorghum PCs and 312 tomato PCs. While the lowest PC only has about 10,000, the largest has over 100,000. In semantic segmentation, the training and validation process begins with a labeled dataset, where images are annotated with pixel-level class labels. This dataset is divided into training and validation sets, with the former used for model training and the latter for monitoring and fine-tuning. The accuracy score reflects the ratio of correctly classified pixels to the total number of pixels in the validation set, helping gauge the model's performance in segmenting objects and scenes.

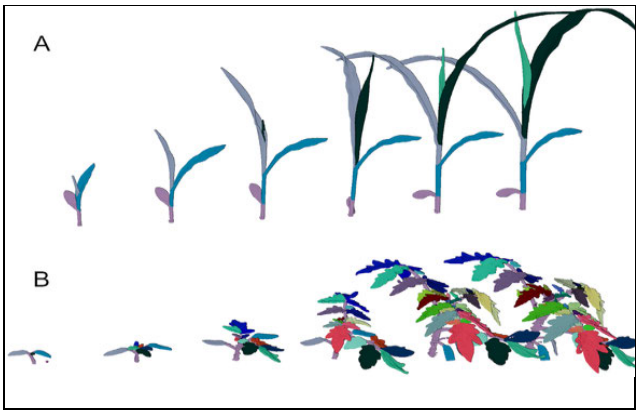


FIGURE 3. Shows sample PC data from plant 1 and plant 2.

V. DATA PREPARATION AND TRAINING DETAILS

This proposed research work is carried out in a hardware platform consisting of an NVIDIA, an ENVI-supported CPU with 128 GB of RAM, 16 cores, and 32 threads. Tensor Flow 1.13.1 is the deep learning framework used. The batch size remains constant at 8 during training, and the process of learning rate is initiated at 0.002., and after 10 epochs per iteration, the gaining knowledge rate is 30% lower. Since the dataset is not open source, the data sets exhibit solve time vs. number of pose graph nodes on a huge dataset but suffice it to state that the parameters suggested are effective. 7th generation benchmark on a low-power computer. It has a moderate CPU and memory footprint and can map very vast regions. Default settings increase the number of pose graph members by O(N). Ceres and the SCHUR_JACOBI preconditioner with the SPARSE_NORMAL_CHOLESKY solver, based on extensive testing. Comparable to the dog leg subspace technique, using LM at the trust region is a superior option because it is well-supported.

A. SEMANTIC SEGMENTATION OUTCOMES

Figure 4 displays the qualitative PointNet semantic crop species findings. To highlight the actual performance, we prominently exhibit test samples from various growing settings and phases of PointNet that appear to be successful in differentiating between the stem and the leaves because The line dividing the two semantically driven classes only contains a few infrequent incorrect segmentations. Table 1 displays the results of PointNet's quantitative semantic division of the entire test set. The majority of metrics are greater than 85 %, suggesting satisfactory semantic performance. From the metrics shown in Table 1, it is clear that PointNet's leaf segmentation results are superior to its stem results since there are far fewer points connected with stems in the training data than there are with leaves. Out of the three species, plant 2 offers the best semantic segmentation results, which might be accounted for by the training dataset's predomi-

TABLE 1. Quantitative measure for semantic segmentation of plant.

	Samples	PREC%	REC%	FI%	IOU%
Plant 1	STEM	93.71	86.42	88.59	82.79
	LEAF	95.12	97.73	97.24	93.63
Plant 2	STEM	95.63	96.02	95.68	93.73
	LEAF	97.98	97.59	98.29	95.63
Plant 3	STEM	90.14	86.69	87.57	78.89
	LEAF	98.04	97.63	97.33	95.72

nance of plant 2 PCs. Additional information from the two distinct species can be added to this unbalanced training to make it better results.

B. INSTANCE SEGMENTATION OUTCOMES

Sample PCs from various growth stages together with the comparative examination of three crops' PointNet instance results [51], [52]. All three species have effectively segmented leaf specimens [53], [54], [55]. Significant differences exist between the three species' leaf structures. Large, broad leaves are found on tobacco plants, complex leaves with at least one large leaflet and two smaller leaflets are found on tomato plants, and long, thin leaves are found on sorghum plants [56], [57]. PointNet does well at segmenting leaf instances across all three leaf varieties. The quantitative assessments of PointNet's leaf segmentation by instance for the test set are shown in Table 2. The bulk of metrics show average performance is 85%, which is appropriate for instance, segmentation.

TABLE 2. Quantitative measure, for instance, segmentation of plant.

Samples	MPREC %	MREC %	MCOV %	MWCOV %
Plant 1	88.1	81.09	81.18	91.27
Plant 2	88.92	79.62	85.19	90.45
Plant 3	86.68	81.13	83.97	88.46

C. EVALUATION OF ALTERNATIVE METHODS

On the same plant dataset, the Python tool is used, and our PointNet is compared with a number of well-used PC segmentation networks in this subsection. The only ones that can do anything other than semantic segmentation are PointNet++ and DGCNN. ASIS and PlantNet can simultaneously do both the semantic and instance division tasks, much like our network can. Using the same set of conceptual and instance labels, we develop all three dual-function networks [58], [59], [60]. From each of their original articles, we adopted the suggested parameter combinations for the comparing networks. In the majority of cases, PointNet outperformed the competition and outperformed them on all four averaged quantitative criteria. Instance segmentation performance comparison between PointNet and two dual-function segmentation networks is shown in Table 2.

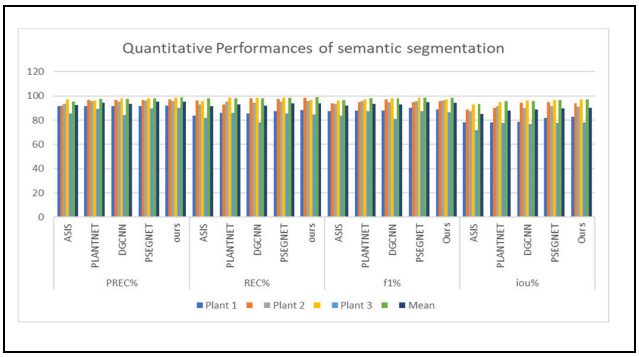


FIGURE 4. Performances of semantic segmentation are compared to quantitative.

The evaluation metrics considered for this comparison include Mean Precision (MPREC), Mean Recall (MREC), Mean Coverage (MCOV), and Mean Weighted Coverage (MWCOV) for four distinct methods: ASIS, PLANTNET, PSEGNET, and our proposed method.

Our proposed PointSamplingNet has obtained the best performance, taking into account all four averaged metrics, with the exception of the mCov values, mRec values, and mPrec values of leaves.

We qualitatively compared PointNet to the state-of-the-art on challenges requiring semantic division and instance segmentation. Figure 5 provides a detailed evaluation of the qualitative performance of instance segmentation and semantic splitting applied to various plant species. The assessment is based on several key metrics, including Mean Precision (MPREC), Mean Recall (MREC), Mean Coverage (MCOV), and Mean Weighted Coverage (MWCOV), for four distinct methods: ASIS, PLANTNET, PSEGNET, and our novel proposed method. The examples in the two images demonstrate that PointNet outperforms networks designed primarily for semantic segmentation as well as ''dual-function segmentation networks, such as ASIS and PlantNet.''

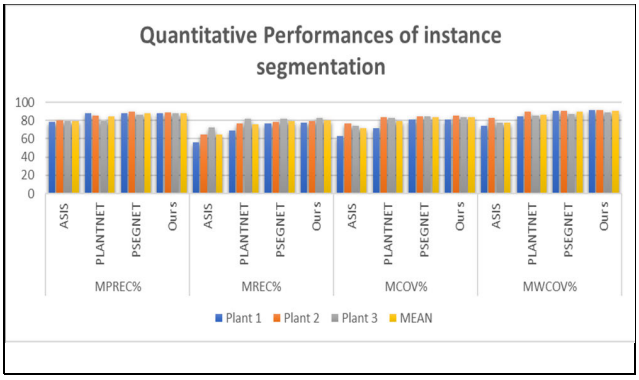


FIGURE 5. Quantitative performances of instance segmentation.

D. COMPARATIVE ANALYSIS

In this section, we conduct an analysis of the proposed method with respect to mPrec and mRec metrics and provide a comparative assessment against existing methods, namely

DKnet and ISBNet. Figure 6 shows the expected point group of the n-th instance in the same semantic class, which reflects the ground truth points group of the m-th instance. The highest value of every parameter tested is shown by max. The binary function IoU is computed precisely as the semantic IoU equation. It takes two inputs: the ground truth point set and the anticipated point set from the network. The number of semantic classes used to calculate the Mean Precision (mPrec) and Mean Recall (mRec) is denoted by the parameter. The results of this analysis are presented in Figure 7.

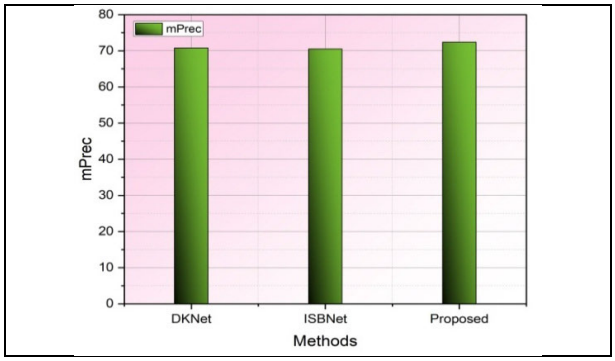


FIGURE 6. mPrec comparison.

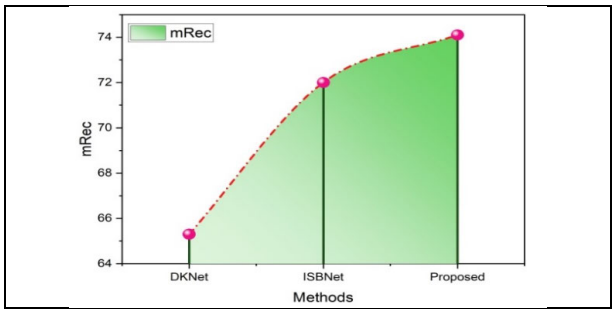


FIGURE 7. mRec comparison.

VI. DISCUSSION

A. GENERALIZATION ABILITY OF POINTNET

Here, we demonstrate that the suggested PointNet is adaptable and adequate to be used with further kinds of PC data. We used the Stanford Large-Scale 3D interior Spaces (S3DIS) for practicing and evaluating our PointNet in order to confirm its suitability for PCs of interior sections, which differ greatly from 3D crops. The S3DIS dataset's points are broken down into 13 semantic classifications, including floor, table, window, and more. Each block served as a single PC input and was downscaled to 4096 points.

S3DIS Area 5's PCs were employed in screening, while the other S3DIS regions were employed in training. This work trained PointNet using the same hyperparameter settings used for training the plant dataset. Displayed are the outcomes of the acute semantic and instance segmentation in Figure 8, in turn. Comparing PointNet to the GT, the majority of the points were correctly categorized, and our network appears to be particularly adept at identifying furniture like tables and chairs with a variety of forms and orientations. The four separate rooms exhibit good instance segmentation from our PointNet, despite the fact that instance-based segmentation on the dataset S3DIS is thought to be a difficult process. The loss of instance segmentation is computed during training just after clustering. Using a for the semantic segmentation task an Argmax operation and an additional 1D convolution on the using AM's semantic flow output function, we arrive to a result of one-hot encoded feature of $4096 \times C$, where C stands for how many semantic classes there are. Here, the semantic segmentation loss is computed.

Instance segmentation on small things appears to be more effective for PointNet than on large objects.

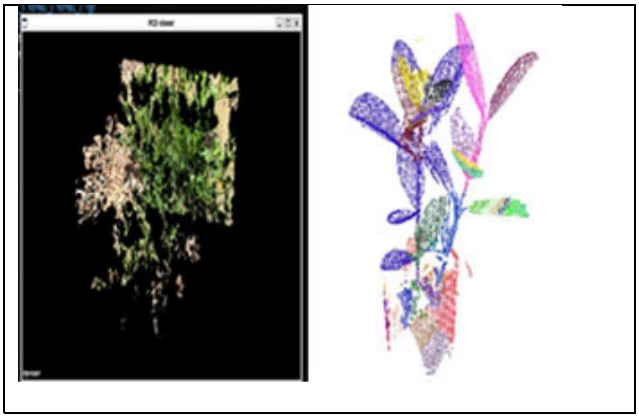


FIGURE 8. Leaf segment from PC data (Left), Leaf instance segment (Right).

B. DISCUSSION OF THE EFFECTIVENESS

This part of the article shows more detail about how this study addresses the three problems that exist in the state-of-the-art deep learning models for PC segmentation. Building a network for Segmenting multipurpose PCs is challenging. Each segmentation task necessitates a dedicated network segment that is governed by a unique loss function, resulting in two causes. In PointNet, Lsem and Lins, respectively, constrain both the instance and semantic segmentation pathways. Additionally, we introduced the point-level losses to the feature attribute map in order to more effectively resolve the training on the principal network. As a result, PointNet's entire network is constrained by three losses combined. Modifying An additional design concern is the potential impact of the weight assigned to the loss of one branch on the overall loss, which may result in an imbalance in the load of the other branch. For instance, it is projected that instance performance will suffer once the level of semantic degradation in PointNet is increased. By altering the weights given to each segment's losses, it is possible to achieve an appropriate ratio between semantic segmentation and instance segmentation. Poor generalization of species frequently leads to two unfavorable outcomes for a PC-acquiring model. The first is that it is possible to wrongly label parts of a plant 2's stem as ''the stem of a plant 1'' when using semantic segmentation; for instance,

it is possible to identify elements of a plant species "1" on a PC of a plant species "2". Another phenomenon is the possibility of a significant segmentation performance differential between monocotyledons and dicotyledons as a result of the substantial variations in 3D structure. Figure 9 shows the Training and Validation Accuracy for semantic segmentation on the Left side, and the right side shows the Training and Validation Accuracy, for instance, segmentation.

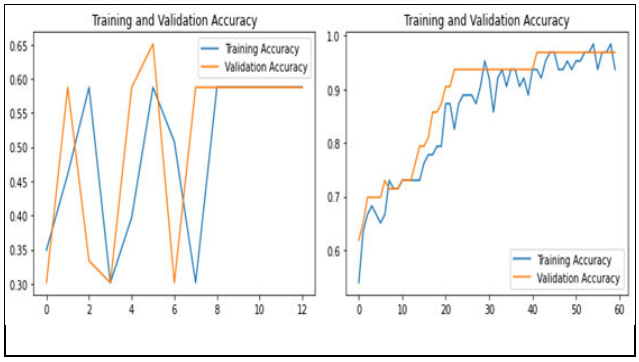


FIGURE 9. Training and validation accuracy for semantic segmentation (Left) and instance segmentation (Right).

PointNet outperformed well-known networks on tasks requiring semantic and instance segmentation, and results that are both qualitative and quantitative in nature showed There were two undesired phenomena detected infrequently. Furthermore, we found that for the numerous object categories seen in PCs, PointNet has good generalization capabilities. The network's applicability to other disciplines, such as "indoor Simultaneous Localization and Mapping and self-driving cars," was shown by the PointNet test on the indoor S3DIS dataset. The proposed algorithm is also tested with real-time data, which was recorded using Velodyne lidar. The PAEM independently gathers both high and low-level traits from two nearby locations in order to improve learning. Using two concurrent decoders, First, PAFM builds feature layers using two distinct learning granularities. Then, to create thorough feature learning, PAFM integrates the two feature granularities. Using two types of interest (location and channel), the focused attention component of PointNet increases the efficiency of network training for both segmentation tasks. The negative effects on classification training and procedure are shown in Figure 10. Numerous training sessions aid in increasing accuracy and minimizing losses. The quantitative performance comparison of PSegNet on instance segmentation against the function segmentation networks. With the exception of the plant leaf mPrec, mRec, and mCov, our PSegNet has demonstrated the best performance across all cases, encompassing all four averaged measurements.

Additionally, we qualitatively contrasted PSegNet with the state of the art on tasks related to semantic and instance segmentation. The PointNet's quantitative semantic segmentation findings for the whole test set show satisfactory

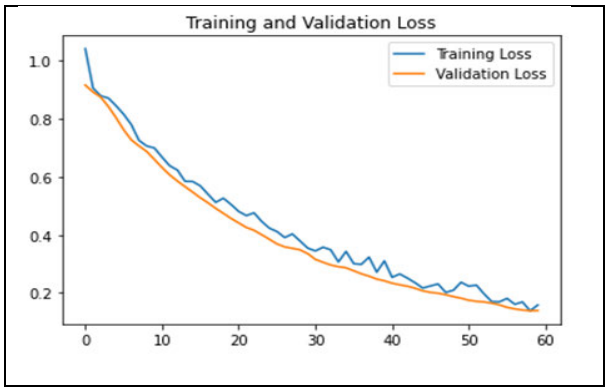


FIGURE 10. Training and validation loss on segmentation process.

semantic segmentation performance, with most measures reaching 75.0%.

C. FAILINGS

PointNet still performs badly on seedlings even though it can do a pair of distinct kinds of categorizations on three different species with various growth phases. The stage of seeds is significantly smaller than the earlier phases of development and has a distinctive 3D shape for many plant species. Consequently, because of the unique characteristics of the seedling stage, it may be difficult to train and evaluate deep learning networks. Sadly, this is true of all such nets. PointNet's segmentation performance will decline as the degree of detail of the plant structure increases. There are two causes. The network cannot function immediately on examples of trees because there aren't any leafy plant specimens in the available dataset. Second, the system can only receive 4096 data as one sample source owing to hardware limitations. The plant PC with dense foliage is expected to have a limited number of points on each organ, resulting in inadequate feature learning and hence generating suboptimal segmentation outcomes.

VII. CONCLUSION

In this study, the Point Down Sampling approach is introduced. The novel PC down-sampling method has advantages over both the conventional distance point sampling technique and the voxel-based sampling strategy. Given that it can adjust the number of points after downsampling and add randomization to the dataset, it is especially highly suited for deep learning network testing and training. Additionally, using data from laser scanning of various species, this research develops PointNet, an innovative dual-function classification network appropriate for crop PCs.

The entire PointNet consists of the Point Fusion Module, the middle element, and the front part, which is often an encoder in deep neural networks. The Point Fusion Module decodes and combines two characteristics with different granularities. The instance segmentation task is represented by the upper branch and the lower branch of the attention modules of the third PointNet component, respectively. Our PointNet outperformed a variety of well-known networks, namely

PointNet++ and ASIS, with respect to both organ semantic and leaf occurrence segmentation tasks in qualitative and quantitative comparisons. Future work will concentrate on two areas. First, we will add further high-precision crop PC data and new species to the collection, particularly plants that are monocotyledonous with thin organs. Second, we'll develop fresh deep neural network architectures that are better at comprehending and evaluating 3D plant structures. Further study will ultimately suggest compressed nets with high accuracy in segmentation to satisfy real-time needs in a few cases in the agriculture sector.

REFERENCES

[1] A.-V. Vo, L. Truong-Hong, D. F. Laefer, and M. Bertolotto, "Octree-based region growing for point cloud segmentation," *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 88–100, Jun. 2015.

[2] B. Yang, J. Wang, R. Clark, Q. Hu, S. Wang, A. Markham, and N. Trigoni, "Learning object bounding boxes for 3D instance segmentation on point clouds," in *Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–10.

[3] N. Ganesh, R. Shankar, R. Cep, S. Chakraborty, and K. Kalita, "Efficient feature selection using weighted superposition attraction optimization algorithm," *Appl. Sci.*, vol. 13, no. 5, p. 3223, Mar. 2023.

[4] J. Priyadarshini, M. Premalatha, R. Cep, M. Jayasudha, and K. Kalita, "Analyzing physics-inspired metaheuristic algorithms in feature selection with K-nearest-neighbor," *Appl. Sci.*, vol. 13, no. 2, p. 906, Jan. 2023.

[5] L. Jiang, H. Zhao, S. Shi, S. Liu, C.-W. Fu, and J. Jia, "PointGroup: Dual-set point grouping for 3D instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 4866–4875.

[6] L. Han, T. Zheng, L. Xu, and L. Fang, "OccuSeg: Occupancy-aware 3D instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 2937–2946.

[7] C. Zong and Z. Wan, "Container ship cell guide accuracy check technology based on improved 3D point cloud instance segmentation," *Brodogradnja*, vol. 73, no. 1, pp. 23–35, Mar. 2022.

[8] H. Sheng, R. Cong, D. Yang, R. Chen, S. Wang, and Z. Cui, "UrbanLF: A comprehensive light field dataset for semantic segmentation of urban scenes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 11, pp. 7880–7893, Nov. 2022.

[9] M. A. Uy, Q.-H. Pham, B.-S. Hua, T. Nguyen, and S.-K. Yeung, "Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1588–1597.

[10] F. Zhang, J. Fang, B. Wah, and P. Torr, "Deep FusionNet for point cloud semantic segmentation," in *Computer Vision—ECCV*. Cham, Switzerland: Springer, 2020, pp. 644–663.

[11] M. Weinmann, M. Weinmann, C. Mallet, and M. Brédif, "A classification-segmentation framework for the detection of individual trees in dense MMS point cloud data acquired in urban areas," *Remote Sens.*, vol. 9, no. 3, p. 277, Mar. 2017.

[12] L. Tchapmi, C. Choy, I. Armeni, J. Gwak, and S. Savarese, "SEGCloud: Semantic segmentation of 3D point clouds," in *Proc. Int. Conf. 3D Vis. (3DV)*, Oct. 2017, pp. 537–547.

[13] T. Cortinhal, G. Tzelepis, and E. E. Aksoy, "SalsaNext: Fast, uncertainty-aware semantic segmentation of LiDAR point clouds," in *Advances in Visual Computing*. Cham, Switzerland: Springer, 2020, pp. 207–222.

[14] Y. Xiang, C. Xie, A. Mousavian, and D. Fox, "Learning RGB-D feature embeddings for unseen object instance segmentation," in *Proc. Conf. Robot Learn.*, 2021, pp. 461–470.

[15] J. Sun, C. Qing, J. Tan, and X. Xu, "Superpoint transformer for 3D scene instance segmentation," in *Proc. Conf. AAAI Artif. Intell.*, vol. 37, Jun. 2023, pp. 2393–2401.

[16] K. Shaik, J. V. N. Ramesh, M. Mahdal, M. Z. U. Rahman, S. Khasim, and K. Kalita, "Big data analytics framework using squirrel search optimized gradient boosted decision tree for heart disease diagnosis," *Appl. Sci.*, vol. 13, no. 9, p. 5236, Apr. 2023.

[17] S. Agrawal, S. Pandya, P. Jangir, K. Kalita, and S. Chakraborty, "A multi-objective thermal exchange optimization model for solving optimal power flow problems in hybrid power systems," *Decis. Anal. J.*, vol. 8, Sep. 2023, Art. no. 100299.

[18] R. C. Narayanan, N. Ganesh, R. Cep, P. Jangir, J. S. Chohan, and K. Kalita, "A novel many-objective Sine–Cosine algorithm (MaOSCA) for engineering applications," *Mathematics*, vol. 11, no. 10, p. 2301, May 2023.

[19] S. Qiu, S. Anwar, and N. Barnes, "Semantic segmentation for real point cloud scenes via bilateral augmentation and adaptive fusion," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 1757–1767.

[20] J. Hou, A. Dai, and M. Nießner, "3D-SIS: 3D semantic instance segmentation of RGB-D scans," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 4416–4425.

[21] Z. Zhang, R. Girdhar, A. Joulin, and I. Misra, "Self-supervised pretraining of 3D features on any point-cloud," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 10232–10243.

[22] H.-Y. Meng, L. Gao, Y.-K. Lai, and D. Manocha, "VV-Net: Voxel VAE net with group convolutions for point cloud segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 8499–8507.

[23] B. Wu, A. Wan, X. Yue, and K. Keutzer, "SqueezeSeg: Convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 1887–1893.

[24] S. A. Bello, S. Yu, C. Wang, J. M. Adam, and J. Li, "Review: Deep learning on 3D point clouds," *Remote Sens.*, vol. 12, no. 11, p. 1729, May 2020.

[25] Y. Cui, R. Chen, W. Chu, L. Chen, D. Tian, Y. Li, and D. Cao, "Deep learning for image and point cloud fusion in autonomous driving: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 722–739, Feb. 2022.

[26] Z. Yang, Y. Sun, S. Liu, X. Shen, and J. Jia, "STD: Sparse-to-dense 3D object detector for point cloud," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Seoul, Korea (South), Oct. 2019, pp. 1951–1960.

[27] T. He, C. Shen, and A. van den Hengel, "DyCo3D: Robust instance segmentation of 3D point clouds through dynamic convolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Nashville, Tennessee, Jun. 2021, pp. 354–363.

[28] F. Poux and J. J. Ponciano, "Self-learning ontology for instance segmentation of 3D indoor point cloud," *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 43, pp. 309–316, Aug. 2020.

[29] L. Wang, L. Zheng, and M. Wang, "3D point cloud instance segmentation of lettuce based on PartNet," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, New Orleans, LA, USA, Jun. 2022, pp. 1646–1654.

[30] D. Wang, "Unsupervised semantic and instance segmentation of forest point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 165, pp. 86–97, Jul. 2020.

[31] T. Vu, K. Kim, T. M. Luu, T. Nguyen, and C. D. Yoo, "SoftGroup for 3D instance segmentation on point clouds," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, USA, Jun. 2022, pp. 2698–2707.

[32] L. Yi, W. Zhao, H. Wang, M. Sung, and L. J. Guibas, "GSPN: Generative shape proposal network for 3D instance segmentation in point cloud," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 3942–3951.

[33] F. G. Zanjani, A. Pourtaherian, S. Zinger, D. A. Moin, F. Claessen, T. Cherici, S. Parinussa, and P. H. N. de With, "Mask-MCNet: Tooth instance segmentation in 3D point clouds of intra-oral scans," *Neurocomputing*, vol. 453, pp. 286–298, Sep. 2021.

[34] H. Liu, H. Yuan, J. Hou, R. Hamzaoui, and W. Gao, "PUFA-GAN: A frequency-aware generative adversarial network for 3D point cloud upsampling," *IEEE Trans. Image Process.*, vol. 31, pp. 7389–7402, 2022.

[35] Q. Ni, J. Guo, W. Wu, and H. Wang, "Influence-based community partition with sandwich method for social networks," *IEEE Trans. Computat. Social Syst.*, vol. 10, no. 2, pp. 819–830, Apr. 2023.

[36] Q. Ni, J. Guo, W. Wu, H. Wang, and J. Wu, "Continuous influence-based community partition for social networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 3, pp. 1187–1197, May 2022.

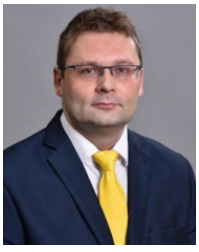
[37] Q. Liu, H. Yuan, R. Hamzaoui, H. Su, J. Hou, and H. Yang, "Reduced reference perceptual quality model with application to rate control for video-based point cloud compression," *IEEE Trans. Image Process.*, vol. 30, pp. 6623–6636, 2021.

[38] Y. Song, R. Xin, P. Chen, R. Zhang, J. Chen, and Z. Zhao, "Identifying performance anomalies in fluctuating cloud environments: A robust correlative-GNN-based explainable approach," *Future Gener. Comput. Syst.*, vol. 145, pp. 77–86, Aug. 2023.

[39] J. Zhang, Y. Liu, Z. Li, and Y. Lu, "Forecast-assisted service function chain dynamic deployment for SDN/NFV-enabled cloud management systems," *IEEE Syst. J.*, vol. 17, no. 3, pp. 4371–4382, Sep. 2023.
[40] Y. Zheng, X. Lv, L. Qian, and X. Liu, "An optimal BP neural network track prediction method based on a GA–ACO hybrid algorithm," *J. Mar. Sci. Eng.*, vol. 10, no. 10, p. 1399, Sep. 2022.
[41] X. Liu, G. Zhou, M. Kong, Z. Yin, X. Li, L. Yin, and W. Zheng, "Developing multi-labelled corpus of Twitter short texts: A semi-automatic method," *Systems*, vol. 11, no. 8, p. 390, Aug. 2023.
[42] J. Gong and A. Rezaeipanah, "A fuzzy delay-bandwidth guaranteed routing algorithm for video conferencing services over SDN networks," *Multimedia Tools Appl.*, vol. 82, no. 17, pp. 25585–25614, Jan. 2023.
[43] Y. Wang, N. Xu, A.-A. Liu, W. Li, and Y. Zhang, "High-order interaction learning for image captioning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 7, pp. 4417–4430, Jul. 2022.
[44] Y. Shen, N. Ding, H.-T. Zheng, Y. Li, and M. Yang, "Modeling relation paths for knowledge graph completion," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 11, pp. 3607–3617, Nov. 2021.
[45] L. Li, P. Wang, X. Zheng, Q. Xie, X. Tao, and J. D. Velásquez, "Dual-interactive fusion for code-mixed deep representation learning in tag recommendation," *Inf. Fusion*, vol. 99, Nov. 2023, Art. no. 101862.
[46] Y. Chen, Z. Chen, D. Guo, Z. Zhao, T. Lin, and C. Zhang, "Underground space use of urban built-up areas in the central city of Nanjing: Insight based on a dynamic population distribution," *Undergr. Space*, vol. 7, pp. 748–766, Oct. 2022.
[47] Y. Zhuang, N. Jiang, and Y. Xu, "Progressive distributed and parallel similarity retrieval of large CT image sequences in mobile telemedicine networks," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–13, Jul. 2022.
[48] B. Cheng, M. Wang, S. Zhao, Z. Zhai, D. Zhu, and J. Chen, "Situation-aware dynamic service coordination in an IoT environment," *IEEE/ACM Trans. Netw.*, vol. 25, no. 4, pp. 2082–2095, Aug. 2017.
[49] Z. Dai, Z. Ma, X. Zhang, J. Chen, R. Ershadnia, X. Luan, and M. R. Soltanian, "An integrated experimental design framework for optimizing solute transport monitoring locations in heterogeneous sedimentary media," *J. Hydrol.*, vol. 614, Nov. 2022, Art. no. 128541.
[50] W. Fan, L. Yang, and N. Bouguila, "Unsupervised grouped axial data modeling via hierarchical Bayesian nonparametric models with Watson distributions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 9654–9668, Dec. 2022.
[51] H. Wang, X. Zhang, and S. Jiang, "A laboratory and field universal estimation method for tire–pavement interaction noise (TPIN) based on 3D image technology," *Sustainability*, vol. 14, no. 19, p. 12066, Sep. 2022.
[52] A.-A. Liu, Y. Zhai, N. Xu, W. Nie, W. Li, and Y. Zhang, "Region-aware image captioning via interaction learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 6, pp. 3685–3696, Jun. 2022.
[53] T. Guo, H. Yuan, L. Wang, and T. Wang, "Rate-distortion optimized quantization for geometry-based point cloud compression," *J. Electron. Imag.*, vol. 32, no. 1, Feb. 2023, Art. no. 013047.
[54] H. Zhang, Y. Mi, X. Liu, Y. Zhang, J. Wang, and J. Tan, "A differential game approach for real-time security defense decision in scale-free networks," *Comput. Netw.*, vol. 224, Apr. 2023, Art. no. 109635.
[55] T. Li, Y. Fan, Y. Li, S. Tarkoma, and P. Hui, "Understanding the long-term evolution of mobile app usage," *IEEE Trans. Mobile Comput.*, vol. 22, no. 2, pp. 1213–1230, Feb. 2023.
[56] H. Liu, H. Yuan, Q. Liu, J. Hou, H. Zeng, and S. Kwong, "A hybrid compression framework for color attributes of static 3D point clouds," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 3, pp. 1564–1577, Mar. 2022.
[57] C. Mi, S. Huang, Y. Zhang, Z. Zhang, and O. Postolache, "Design and implementation of 3-D measurement method for container handling target," *J. Mar. Sci. Eng.*, vol. 10, no. 12, p. 1961, Dec. 2022.
[58] S. Lu, M. Liu, L. Yin, Z. Yin, X. Liu, and W. Zheng, "The multi-modal fusion in visual question answering: A review of attention mechanisms," *PeerJ Comput. Sci.*, vol. 9, p. e1400, May 2023.
[59] Y. Liu, K. Wang, L. Liu, H. Lan, and L. Lin, "TCGL: Temporal contrastive graph for self-supervised video representation learning," *IEEE Trans. Image Process.*, vol. 31, pp. 1978–1993, 2022.
[60] X. Liu, T. Shi, G. Zhou, M. Liu, Z. Yin, L. Yin, and W. Zheng, "Emotion classification for short texts: An improved multi-label method," *Humanities Social Sci. Commun.*, vol. 10, no. 1, pp. 1–9, Jun. 2023.

NANDHAGOPAL GOMATHI received the P.D.F. degree in Germany, in 2018. She is currently a Professor with the School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai. She has nearly 25 years of experience in teaching and research. She has published nearly 50 articles to her credit. Her research interests include information retrieval, image processing, computational techniques, optimization techniques, artificial intelligence, machine learning, networking, and the IoT.

KRISHNAMOORTHI RAJATHI is currently an Associate Professor with the School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai. She has nearly 15 years of teaching experience. She has published nearly 20 articles to her credit. Her research interests include AI, simultaneous localization and mapping, point cloud data processing, and autonomous navigation.



MIROSLAV MAHDAL is currently the Vice-Dean of Science, Research, and Doctoral Studies with the Faculty of Mechanical Engineering, VSB—Technical University of Ostrava, and as an Associate Professor with the Department of Control Systems and Instrumentation. He has published nearly more than 80 articles to his credit. His research interests include control of mechatronic systems, control systems, automatic control theory, wireless technologies, artificial intelligence, cloud computing, optimization methods, and programming of control systems.

MUNIYANDY ELANGOVAN received the Doctor of Engineering (Dr.-Eng.) degree in naval architecture from Hiroshima University, Japan, and the master's (M.Eng.) degree in engineering design from the Government College Technology, Coimbatore, Tamil Nadu, India.

He is currently the Director (Research and Development and Consulting) with DTC Corporation, Chennai. He provides high-level software development services in Python, SQL, AI, and ML. He started his career as a Ship Designer with the National Ship Design and Research Centre, Visakhapatnam, which is a Central Government Organization, and later, he worked for well-known companies, such as Jagson International Ltd., New Delhi, Allseas Engineering, Dubai, IRS, Mumbai, NAPA, Finland, and GreenSHIP, Bengaluru. He has visited many countries like Japan, China, South Korea, France, Brazil, Malaysia, and Singapore for technical discussions and conferences. Altogether, he has 22 years of industrial experience, which includes five years of teaching at a university. Altogether, he has completed research and consultancy project worth 40 crores and published more than 100 papers in journals and conferences.

Dr. Elangovan was awarded a funded project from the Naval Science and Technological Laboratory (NSTL), Visakhapatnam, India, and completed it with successful output. He is an editor of two journals and a reviewer of four journals. He has made remarkable contributions to marine hydrodynamics, design, CFD, underwater marine vehicles, robotics, energies, sensors, and industrial robots.