Semantic Segmentation of Coastal Zone on Airborne Lidar Bathymetry Point Clouds

Sajjad Roshandel, *Member, IEEE*, Weiquan Liu, Cheng Wang, *Senior Member, IEEE*, and Jonathan Li, *Senior Member, IEEE*

***Abstract*— Large-scale semantic segmentation point cloud is an ongoing research topic for on-land environments. However, there is a rare deep learning research study for the sub-surface environment. Although, PointNet and its successor PointNet++ have become the cornerstone of point cloud segmentation. However, these techniques handle a relatively small number of points. This poses a natural difficulty in a large spatial scene with millions of possible points. In particular, for shallow water of coastal zone, the small number of points where the seabed and water surface meet, close points may belong to different classes. In our work, we present the semantic segmentation on a large-scale airborne Lidar bathymetry (ALB) point cloud containing millions of sample points into two classes of water surface and seabed with the voxel sampling pre-processing (VSP) approach. The proposed approach will allow us to capture the complicated outdoor natural scene components of water surface and seabed more accurately and more realistic through nonuniform voxelization in the mixture of dense and sparse points of the ALB point cloud. The performance of validation results show improvement in a per-point accuracy of 72.45% compared with other state-of-the-art deep learning-based methods.**

***Index Terms*— Airborne Lidar bathymetric (ALB) point cloud, coastal zone, deep learning, seabed, semantic segmentation.**

I. Introduction

SCENE understanding is a key problem and fundamental issue for various applications such as augmented reality, robotics, industrial reconstruction, and geospatial modeling. Semantic segmentation refers to the ability to distinguish accurate object borders within its surroundings. This is useful where the detailed information is very important such as medical scans, criminal issues, bio-chemical, scene modeling, etc. 3-D scene semantic labeling can generally be grouped into two categories; on-surface (or on-land) environments, which mainly considered the human-made features, and sub-surface and underwater environments. For on-surface semantic segmentation, convolutional neural networks (CNNs) architectures, which is generally processed 2-D images, is widely used by many researchers to process 3-D point cloud by transforming the point cloud into an image [1], [2].

For instance, Boulch *et al.* [1] performs the CNN approach by converting 3-D points to 2-D snapshots. They applied pixel-wise labeling for 2-D snapshots using fully convolutional networks. However, this approach results in some missing data during the transformation, and the network may have difficulties extracting whole features and distinguishing them. To this endpoint-base methods widely have been proposed based on PointNet [3]. Other studies have investigated 3-D semantic segmentation tasks in indoor and outdoor scenes [4], [5]. For outdoor scenes, Winiwarter *et al.* [4] proposed a PointNet model on multiple scales of 3-D point clouds acquired by airborne laser scanning (ALS) to automatically learn a representation of local neighborhoods in an end-to-end approach. For indoor scenes, [5] employed deep learning techniques via building volumetric data representation and their methodology can be adapted to different kinds of datasets.

Although, PointNet [3] and PointNet++ [6] have become the cornerstone of point cloud segmentation. However, these techniques handle a relatively small number of points. This poses a difficulty in scenes that cover a large spatial area with possibly millions of points. While large-scale point cloud segmentation is an ongoing research topic, major efforts in this direction are inspired by superpixel-based algorithms in computer vision [7]. One major approach in this direction is the Superpoint graph (SPG) for 3-D point clouds [8] which creates a downsampled version of the original point cloud.

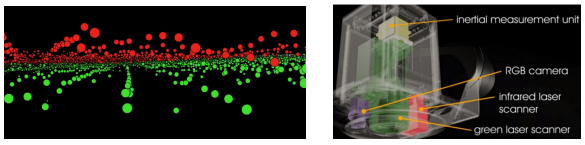
Most current efforts focused on human-made features such as roads, cars, buildings [8], [9] for the on-surface environment. Although sub-surface and seabed scene understanding is one the most challenging environments for on-time monitoring due to geomorphological changes, sediment transportation, etc. However, there is a rare research study using deep neural network (DNN) approaches for 3-D semantic segmentation applied to underwater geospatial 3-D point clouds data. This is mainly due to lack of enough labeled datasets. Moreover, most of recent underwater deep learning approaches applied to 2-D data such as Synthetic Aperture Sonar images [10], [11]. The main challenges of their research are related to the noise and low quality of underwater imaging for recognizing objects. Other research used multibeam echo-sounder (MBES) point cloud data for underwater environment [12], [13]. Daniel and Dupont [13] investigated a DNN to classify sea-bottom

Manuscript received November 16, 2021; revised February 12, 2022 and March 7, 2022; accepted March 14, 2022. Date of publication March 21, 2022; date of current version April 4, 2022. This work was supported in part by the China Postdoctoral Science Foundation under Grant 2021M690094 and in part by the China Scholarship Council (CSC). *(Corresponding author: Weiquan Liu.)*

Sajjad Roshandel, Weiquan Liu, and Cheng Wang are with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Informatics, Xiamen University, Xiamen 361005, China (e-mail: roshandel.hydraulic@gmail.com; wqliu@xmu.edu.cn; cwang@xmu.edu.cn).

Jonathan Li is with the GeoSTARS Laboratory, Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: junli@uwaterloo.ca).

Digital Object Identifier 10.1109/LGRS.2022.3161191



(a) (b)

Fig. 1. (a) Illustration of annotated of water surface and seabed points at breaking wave zone of very shallow water. Seabed points are in green color and water surface points are in red color. (b) Airborne laser bathymetry (ALB) system for point cloud acquisition (RIEGL VQ-880-G).

morphology. They obtained accuracy results no better than 65% for classifying a steep seabed into two classes. Likewise, [14] used geometric and full-waveform features and applied artificial neural network (ANN) for seabed classification on ALB point cloud.

Almost no efforts focused on the seabed 3-D semantic segmentation using airborne Lidar bathymetry (ALB) point cloud. In this letter, we present the semantic segmentation of water surface and seabed on the large-scale ALB point cloud. The main contributions are summarized as follows.

1) We present a pipeline for semantic labeling of ALB point clouds, which combines the voxel-sampling preprocessing (VSP) approach with state-of-the-art SPGs. This innovative study improved the performance by 72.45% per-point accuracy.
2) Present the modified SPG approach on the underwater shallow water of coastal zones for more accurate monitoring, evaluation, and management of shore and near-shore of the ocean area.

II. Methodology

A. Voxel-Sampling Pre-Processing (VSP)

An ALS sensor is used for capturing a wide range of point clouds data. In data collection, due to the sensor characteristic, the broom part of the laser makes dense quality data in the central scope of the sensor. Whilst the opposite is seen when further away from the sensor center, the data will be sparser. In conclusion, we encountered less quality data in the far-end part of data.

In semantic segmentation, the uniform representation is vital in the case of detailed information. Tchapmi *et al.* [15] used 3-D convolutions on a regular voxel grid. However, the voxel gross need to be created carefully in our dataset because of the sparse data parts. Since voxelization is the main problem in this step, we decided to implement different sampling strategies for segmentation. Previously we were sampling sparse and dense data at the same rate. In another effort, we sampled them at different rates. We sample the sparse data at a higher rate and the dense data at a lower rate. The hierarchical approaches were implemented where the segmentation is done at different scales for the model to learn the global picture of the profile.

In this case, we design the VSP strategy (see Fig. 2). To conduct nonuniform sampling, we follow the voxelization procedure. In the voxelization procedure, we define a small voxel volume R corresponding to an $R \times R \times R$ cube. All points inside this voxel are represented by the voxel. For

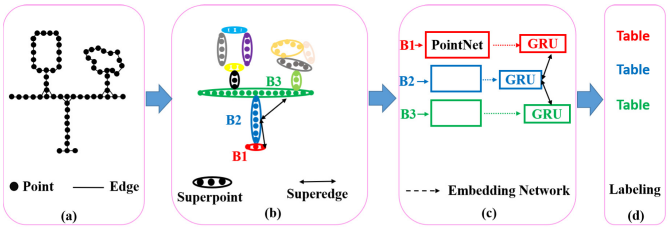


Fig. 2. General framework of semantic labeling with SPG for small indoor objects. The above shows the desk, mirror, and flowerpot. (a) Coherent set of points which provide a simple and meaningful shape of the object. (b) SPG, connecting each related Superpoint (node) by Superedges. Each node corresponds to a small part of the object. (c) PointNet embedded each Superpoint. GRU then updates itself by incoming messages through Superedges. (d) Final labeling for whole table object.

sparse parts of the data, we set $R = 5$ and for dense part of the data, we set $R = 16$, except where the water and seabed meet, where we set $R = 2$. By doing this, we reduce the size of the point cloud set. The subsampled set in this way reflects the geometric profile of our data. We directly render the point cloud as nonuniform voxelization by using sub-voxel in the dense location and large-voxel at the sparse location. The aforementioned approach is an alternative to geometric partitioning which is proposed in the Superpoint [8]. We iterate, however, that this does not prevent us from using the Voronoi tessellation approach of the Superpoint method.

B. Geometric Separation

Generally, for point-set segmentation, the objective is the small single compact objects such as guitar, bed or chair, and small room, etc. The basic idea in point-set segmentation is providing a coherent set of points to a model such that the collection of points provide an idea of the shape of the profile. However, in this dataset, we have natural outdoor objects. The first object is noncompact, formable, and fluid water, the other is an impressible, moveable and erodible object as soil.

Due to breaking waves happening at the surfing zone of the shoreline, the water surface and seabed points are so near each other and the created turbulences result in a mixture of points of both classes. Therefore, the mixture points will make difficult to identify the exact geometry of each class clearly. Fig. 1(a) shows the water surface and seabed points at the breaking wave zone of very shallow water. Likewise, the underwater points acquired with ALB are without RGB values since the reflected green laser wavelength just captures the geometry of each point, i.e., the spatial coordinates of the features (x, y, z). This is one of the reasons why we chose to use geometric separation. Fig. 2 shows the general framework of semantic labeling for small indoor objects.

By breaking down the objects into simple meaningful parts through simple geometric clusters, the model can separate the object-parts more accurately and not cover other objects which are part of different classes. Fig. 3 represents the semantic labeling of seabed and water surface through ALB point cloud. It can be seen the simple profile of a water surface and seabed points at the shoreline where these points are very close to each other. Referring to step (c) of Fig. 3, $B1$, $B2$, and $B3$ simply

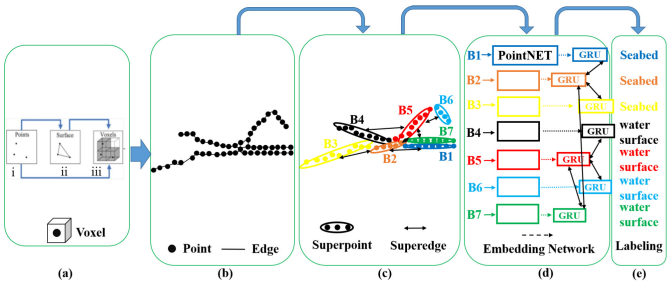


Fig. 3. Illustration of semantic labeling of seabed and water surface through ALB point cloud. (a) VSP approach: (i) Superpoint takes a point cloud set; (ii) creates a geometric set by using Voronoi tessellation; and (iii) this geometric is then further directly processed as non-uniformly voxelization. (b) Geometric division of the water surface and seabed points in shoreline. (c) SPG and Superedge. (d) Embedding Network. (e) Final labeling of water surface and seabed points.

illustrate the seabed points and $B4$, $B5$, $B6$, and $B7$ represent the fluctuation of water surface points. Assume the input point cloud C as a set of n 3-D point set: $P_i, P_i, P_i, P_n \in R_d$ and each point $i \in C$ are defined by its 3-D position p_i, and, if available, other observations o_i such as color or intensity. For every point, we compute a set of d_g geometric features $f_i \in R^{dg}$ describing the shape of its local neighborhood. We also compute the elevation of each point, defined as the z coordinate of p_i normalized over the whole input cloud. The geometrically homogeneous partition is defined as the following (as optimization problem):

$$\arg\min_{g \in R^{dg}} \sum_{i \in C} \|g_i - f_i\|^2 + \mu \sum_{(i,j) \in E_{nn}} w_{i,j} [g_i - g_j \neq 0] \quad (1)$$

where μ: regularization strength determines the coarseness of the resulting partition. W: The edge weight; i, j: points (local nodes); (g_i, f_i): geometric features vector; E_{nn}: adjacency relationship between segments; $[\cdot]$: the iverson bracket [8]. Equation (1) simplifies the geometrical components and is defined as Superpoint in this letter.

C. Superpoint Structure

Referring to Section II-B for making relevant neighborhood graphs and their relatives, the network creates a downsampled version of the original point cloud. Note that our data are simple geometric components so this allows us to utilize the embedding networks as an abbreviated PointNet [8]. The system architecture then learns embedding for each node of the graph followed by recurrent modules which model the relationships between different nodes. This amounts to modeling the structural context of the scene.

The point cloud is represented via each big point defined by oriented attributed graph $G = (B, \Gamma, F)$. B is the set of big points (Superpoints) or nodes and Γ is edges proximity between Superpoints. The Superedges need to be annotated by a set of d_f features: $F \in R^{E \times df}$ characterizing the proximity relationship between Superpoints [8]. In our dataset, we have a mixture of dense and sparse components. One disadvantage of Voronoi tessellation on large point clouds is that having been based entirely on the Euclidean metric, Voronoi tessellation fails to capture the structure of the point cloud. This is especially true in our application where the seabed and water surface points are very close to each other and maybe grouped entirely together by the Voronoi tessellation, leading to improper learning [16]. Indeed, geometric distinction plays a key role in the success of segmentation. We, therefore, adopted a different approach. We begin by subsampling the point cloud based on the following two principles: 1) points close to each other likely belong to the same class and 2) for a small number of points where the seabed and the water surface points meet each other, close points may belong to different classes. We note that the above observations are based on our dataset. Hence, we favor a nonuniform sampling. Nonuniform sampling means the following steps: 1) sparse parts of the data, which are in the far-end scope, are sampled more densely and 2) dense parts of the data are sampled more sparsely.

In large point clouds, (1) or (minimized partition functional) cannot properly act well [8]. In this case, it can be a good solution to find proximity with a few graph-cut iterations. Therefore, we will have more constant big points (Superpoint) in simple geometric components.

Superedge: To make Superpoint B and T adjacent, either B or T must be at one end; with one edge in E_{vor}, (2). The E_{vor} and C are called symmetric Voronoi adjacency graph for the whole input point cloud [8], [17]

$$\Gamma = \{(B, T) \in B^2 \mid \exists (i, j) \in E_{\text{vor}} \cap (B \times T)\} \quad (2)$$

(B, T): the set of Superpoints, E_{vor}: edge of Voronoi adjacency graph. Equation (3), distinguishes the important features of Superedge B and T as a set of offsets $\delta(B, T)$

$$\delta(B, T) = \{(p_i - p_j) \mid (i, j) \in E_{\text{vor}} \cap (B \times T)\}. \quad (3)$$

Every Superpoint B_i is embedded into a vector z_i of fixed-size dimensionality d_z. Then contextual information is provided by the graph convolutions.

D. Segmentation

Classifying each big point with its meaningful surrounding is the final step of the segmentation task in this section. In image segmentation, classification acts based on informative context and finds the relationship of the nearby pixels (neighborhood). For 3-D big points with three coordination, graph convolution networks (GCNs), is used. The general idea is to update the embedding of each Superpoints according to pieces of information transmitted through Superedges. In this regard, gated recurrent unit (GRU) [8] is used as retaining embedded information for each Superpoint. Then over iteration GRU updates itself by incoming messages $m_i{}^t$.

III. Experiments and Results

A. Dataset

Dataset was recorded by a 3-D Bathymetry LiDAR sensor mounted on a moving air-vehicle passing through various coastal-urban scenes of Tampa Bay, Florida, USA, on October 15, 2015. The maximum altitude flight was about 504.617 m (latitude 27°53′35″N, longitude 82°51′06″W). Figs. 4 and 5 show the study area and original bathymetry point clouds

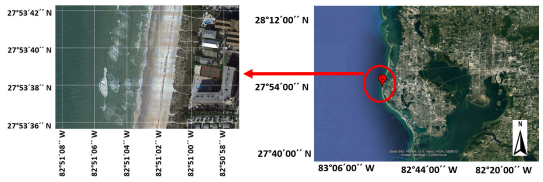


Fig. 4. Study area - The coastal urban area, Tampa Bay (Florida, USA).

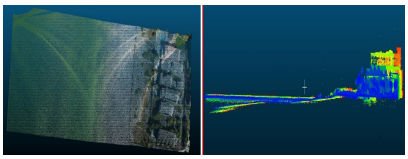


Fig. 5. Some bathymetry point clouds. Cross section of point cloud shows on the right side. Buildings, trees, and beachside can be seen. Points with different reflection intensities are rendered in different colors.

respectively. As shown in Fig. 1(b), a RIEGL VQ-880-G scanning system which integrates green laser scanners, infrared laser scanner, RGB camera, and inertial navigation devices are installed on traveling aircraft to capture 3-D point clouds. The total number of ALB point cloud at shallow water are around five million points by 1.2 km × 1.8 km. The average point cloud density is 1–3 points/m^2 for the sparse part of the data, in particular for high depth seabed (>4 m) and 5–8 points/m^2 for dense part of the data.

B. Georeference and Complexity Computation

For geo-referencing of the captured data, we adjust the dataset with the national oceanic and atmospheric administration (NOAA) benchmark, which is around 33 ground control points. These ground control points also include the bathymetric points in the shallow part of the dataset. All experiments were performed on CentOS 8 with Intel Xeon Gold CPU and 192-GB RAM with four NVIDIA V100 GPU cards. The training process for 241 780 points consumed 31 308 s with over 50 epochs. This part of data is including very shallow water. This time excludes time spent on data pre-processing and loading.

C. Experimental Results

Referring to the addressed challenges, the comparisons are conducted with five state-of-the-art deep-learning-based methods; PointNet [3], PointNet++ [6], Dynamic Graph CNN (DGCNN) [18], RandLA-Net [19], SPG [8], and our proposed method on the ALB dataset. Note that, still lack of enough ALB datasets, limited us for further comparisons with other similar datasets. The performance comparisons and validation accuracy of different methods of our dataset are reported in Table I. Fig. 6 compares semantic labeling results of water surface and seabed on very shallow water. The water depth is 0.8 to 1 m. Since the ALB system captures the geometry of underwater features (seabed features) in unstructured space, point cloud doesn't have additional information such as RGB. Moreover, superimposed echoes in very shallow water results in unreliable extraction of the sea-bottom part

TABLE I

PERFORMANCE COMPARISONS ON THE VALIDATION SET OF OUR DATA. INTERSECTION OVER UNION METRIC FOR THE CLASSES OF OUR DATASET. MIOU REFERS TO THE UNWEIGHTED AVERAGE OF IOU OF EACH CLASS

Method	Validation accuracy	mIoU
PointNet [3]	0.4405	0.547
PointNet ++ [6]	0.4920	0.552
Dynamic Graph CNN (DGCNN) [18]	0.6250	0.611
RandLA-Net [19]	0.6790	0.631
Superpoint Graph (SPG) [8]	0.6918	0.742
SPG + VSP (Ours)	**0.724**	**0.813**

from waveform analysis. Therefore, we conducted the related point-based methods. PointNet operates on each point independently through permutation invariant and then applies a symmetric function to accumulate features. However, this point-independency approach leads to inattention of the geometric relationships among points.

In SPG, the validation accuracy after 60K iterations is a 69.89% which is substantially higher than the 49% of accuracy with PointNet++. This shows the less effectiveness of using the Pointnet++ method for the large-scale point cloud data. RandLA-Net use the random sampling approach to decrease point density of large-scale datasets. Random sampling selects K points from the whole original points. Although RandLA-Net when compared with other point sampling approaches such as farthest point sampling (FPS) has low memory cost. However, in our dataset, because of the mixture of dense and sparse components, particularly in very shallow water, this approach cannot capture the water surface and seabed points accurately. Therefore, we propose an approach to overcome the bathymetry dataset which is a mixture of dense and sparse components without RGB values.

Voxel-sampling preprocessing (VSP) allows us to use effective deep learning tools which would not be able to handle the large-scale data. Fig 7 shows some visual experimental results of sparse and dense parts of the dataset with DGCNN [18], RandLA-Net [19], and SPG with VSP (Ours).

Water surface and seabed points are very close to each other in the very shallow water of shoreline and Euclidean metric is not able to capture the structure of the point cloud. We solve this by nonuniform sampling. We carry out nonuniform sampling on 70% of the data selected as a training set and train the system as per normal loss functions of Superpoint. We then validate the performance of the remaining 30% of the data. Fig. 8 illustrates the classified semantic labels of water surface and seabed in different water depths with SPG + VSP approach. Fig. 9 shows some comparison of semantic labeling results as a cross section of very shallow water. The water depth range is 0.8 to 6 m. Our method significantly improves accuracy comparing PointNet, PointNet++, DGCNN, RandLA-Net, and SPG on our dataset. We obtain a per-point accuracy of 72.45%.

IV. CONCLUSION

In this letter, we presented the application of deep learning for seabed semantic segmentation of large-scale ALB point

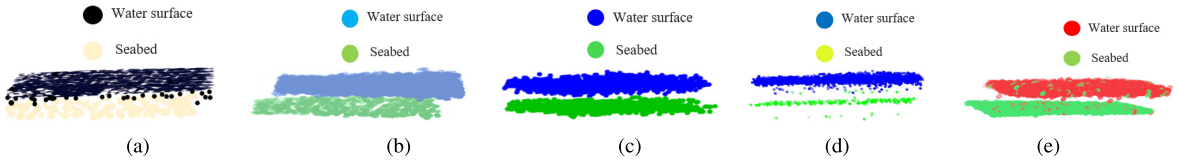


Fig. 6. Comparison of semantic labeling results of water surface and seabed on very shallow water. The water depth is 0.8 to 1 m. (a) PointNet++, (b) Dynamic Graph CNN (DGCNN), (c) RandLA-Net, (d) SPG, and (e) SPG with VSP on the ALB point cloud.

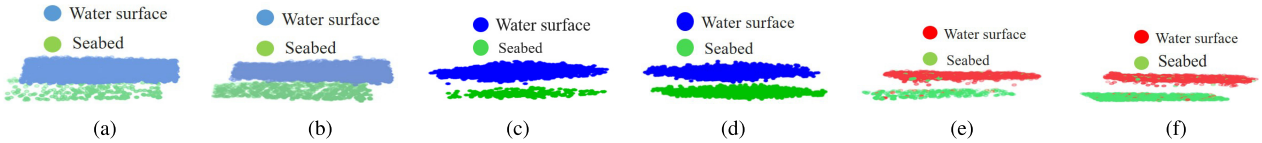


Fig. 7. Examples of semantic labels of water surface and seabed on the ALB dataset at dense and sparse parts. For each set, Left: Sparse part of data. Right: Dense part of data. (a) and (b) dynamic Graph CNN (DGCNN) results, (c) and (d) RandLA-Net results, and (e) and (f) SPG with VSP results (ours).

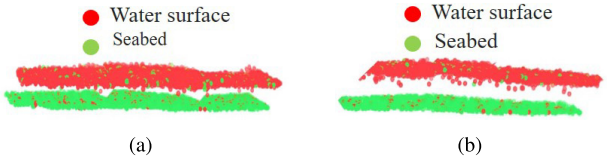


Fig. 8. Classified semantic labels of water surface and seabed in different water depth via SPG with VSP. (a) Shows the semantic label with depth of 0.8 to 1.5 m in shoreline. (b) Shows the semantic label with depth of more than 2.5 m.

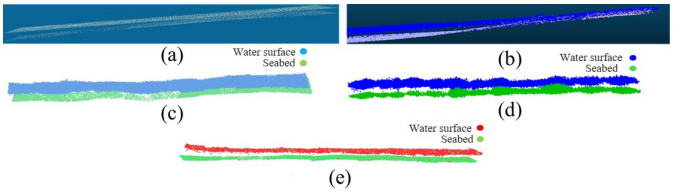


Fig. 9. Compare semantic labeling results of cross section of shallow water in the shoreline on ALB point cloud. (a) Raw point cloud of water surface and seabed. The water depth is between 0.8 and 6 m. (b) Groundtruth. Water surface is in blue color and seabed is in light blue color. Semantic label with (c) DGCNN, (d) RandLA-Net, and (e) SPG + VSP (ours).

cloud based on VSP approach. We proposed an approach to overcome the challenges of 3-D large-scale ALB point cloud dataset which is a mixture of dense and sparse components. We showed that VSP allows us to use effective deep learning tools and able us to handle the large-scale ALB point cloud data. Our method significantly improves accuracy when comparing PointNet, PointNet++, DGCNN, RandLA-Net, and SPG on the ALB dataset. We obtained a per-point accuracy of 72.45%. Our work infers implicit relationships to improve regimentation results. Indeed, additional information such as underwater images, more labeled data and data integration with 3-D LiDAR point clouds will improve the performance of DNN approaches on sub-surface semantic segmentation task.

REFERENCES

[1] A. Boulch, B. Le Saux, and N. Audebert, "Unstructured point cloud semantic labeling using deep segmentation networks," in *Proc. 3DOR@ Eurograph.*, vol. 3, Apr. 2017, pp. 17–24.

[2] F. J. Lawin, M. Danelljan, and P. Tosteberg, "Deep projective 3D semantic segmentation," in *Proc. Int. Conf. Comput. Anal. Images Patterns*. Ystad, Sweden: Springer, 2017, pp. 95–107.

[3] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "PointNet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 652–660.

[4] L. Winiwarter, G. Mandlburger, S. Schmohl, and N. Pfeifer, "Classification of ALS point clouds using end-to-end deep learning," *PFG J. Photogramm., Remote Sens. Geoinf. Sci.*, vol. 87, no. 3, pp. 75–90, Sep. 2019.

[5] K. Babacan, L. Chen, and G. Sohn, "Semantic segmentation of indoor point clouds using convolutional neural network," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 4, pp. 101–108, Nov. 2017.

[6] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in *Proc. NIPS*, 2017, p. 30.

[7] Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering," in *Proc. IEEE CVPR*, Jun. 2015, pp. 1356–1363.

[8] L. Landrieu and M. Simonovsky, "Large-scale point cloud semantic segmentation with superpoint graphs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4558–4567.

[9] X. Yang, X. Li, Y. Ye, R. Y. K. Lau, X. Zhang, and X. Huang, "Road detection and centerline extraction via deep recurrent convolutional neural network U-Net," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7209–7220, Sep. 2019.

[10] X. Shang, J. Zhao, and H. Zhang, "Automatic overlapping area determination and segmentation for multiple side scan sonar images mosaic," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2886–2900, 2021.

[11] Q. Wang *et al.*, "RT-SEG: A real-time semantic segmentation network for side-scan sonar images," *Sensors*, vol. 19, no. 9, p. 1985, Apr. 2019.

[12] D. Stephens, A. Smith, T. Redfern, A. Talbot, A. Lessnoff, and K. Dempsey, "Using three dimensional convolutional neural networks for denoising echosounder point cloud data," *Appl. Comput. Geosci.*, vol. 5, Mar. 2020, Art. no. 100016.

[13] S. Daniel and V. Dupont, "Investigating fully convolutional network to semantic labelling of bathymetric point cloud," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. V-2-2020, pp. 657–663, Aug. 2020.

[14] T. Kogut and A. Slowik, "Classification of airborne laser bathymetry data using artificial neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1959–1966, 2021.

[15] 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.

[16] S. M. Manson, P. A. Burrough, and R. A. McDonnell, "Principles of geographical information systems: Spatial information systems and geostatistics," *Econ. Geography*, vol. 75, no. 4, p. 422, Oct. 1999.

[17] J. W. Jaromczyk and G. T. Toussaint, "Relative neighborhood graphs and their relatives," *Proc. IEEE*, vol. 80, no. 9, pp. 1502–1517, Sep. 1992.

[18] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph CNN for learning on point clouds," *ACM Trans. Graph.*, vol. 38, no. 5, pp. 1–12, Nov. 2019.

[19] Q. Hu *et al.*, "RandLA-Net: Efficient semantic segmentation of large-scale point clouds," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11108–11117.