

# INVESTIGATING FULLY CONVOLUTIONAL NETWORK TO SEMANTIC LABELLING OF BATHYMETRIC POINT CLOUD

S. Daniel<sup>1\*</sup>, V. Dupont<sup>1</sup>

<sup>1</sup> Dept of Geomatic Science, Université Laval, 1055 avenue du séminaire, Québec (QC), Canada – [Sylvie.daniel@scg.ulaval.ca](mailto:Sylvie.daniel@scg.ulaval.ca);  
[vincent.dupont.1@ulaval.ca](mailto:vincent.dupont.1@ulaval.ca)

**KEY WORDS:** Point cloud, hydrographic survey, classification, fully convolutional neural network, sea-bottom morphology

## ABSTRACT:

The benefit of autonomous vehicles in hydrography is largely based on the ability of these platforms to carry out survey campaigns in a fully autonomous manner. One solution is to have real-time processing onboard the survey vessel. To meet this real-time processing goal, deep learning based-models are favored. Although Artificial Intelligence (AI) is booming, the main studies have been devoted to optical images and more recently, to LIDAR point clouds. However, little attention has been paid to the underwater environment. In this paper, we present an investigation into the adaptation of deep neural network to multi-beam echo-sounder (MBES) point cloud in order to classify sea-bottom morphology. More precisely, the paper investigates whether fully convolutional network can be trained while using the native 3D structure of the point cloud. A preprocessing approach is provided in order to overcome the lack of adequate training data. The results reported from the test data sets show the level of complexity related to natural, underwater terrain features where a classification accuracy no better than 65% can be reached when 2 micro topographic classes are used. Point density and resolution have a strong impact on the seabed morphology thereby affecting the classification scheme.

## 1. INTRODUCTION

### 1.1 Introduction

The exploration and use of ocean resources is one of the major challenges of the 21st century. Although 71% of the Earth's surface is covered with water, many areas are not yet mapped (e.g. Arctic), or the available cartographic representations are not sufficiently up-to-date or do not provide an adequate level of detail. One of the obstacles to the intensification of hydrographic data production is their very high cost, which in the vast majority of cases must be carried out from hydrographic vessels requiring significant logistics and qualified personnel that are often difficult to recruit. One of the ways envisioned to overcome these difficulties is the use of autonomous surface vehicles (ASV) (Desa et al., 2006).

The benefit of autonomous vehicles in hydrography is largely based on the ability of these platforms to carry out survey campaigns in a fully autonomous manner. With the survey systems and processes currently available, the platform is programmed for a predefined mission during which hydrographic data are collected and recorded on board the vehicle for further post-processing (i.e. offline). Such an approach has two major disadvantages. The first concerns the bottleneck between the increasing volume of data acquired from these autonomous platforms and the processing capacity available to process quickly and efficiently the data. The second concerns the quality of the data acquired. Indeed, using ASV, the hydrographer has no way of verifying before receiving the data that the survey meets the expected specifications in terms of accuracy, conformity and uncertainty of the soundings. One of the solutions to overcome these disadvantages is to have real-time processing onboard the ASV dedicated to denoising the

data and computing error estimators in order to assess if the survey meets the required quality. Such estimators are generally applied on flat or slope sea-bottom. Therefore, conducting a morphological analysis of the seabed is often a prior to the error estimator computation.

To meet the real-time processing goal, deep learning based-models are favored. Although Artificial Intelligence (AI) is booming, particularly techniques based on deep neural networks, the main studies have been devoted to optical images (LeCun et al., 2015). More recently, research efforts have been turned to LiDAR (Light Detection And Ranging) point clouds, collected either from airborne or terrestrial platforms (Liu et al., 2019). However, little attention has been paid to the underwater environment regarding the design and implementation of deep neural networks (DNN).

In this paper, we present an investigation into the adaptation of deep neural network to multi-beam echo-sounder (MBES) point cloud in order to classify sea-bottom morphology. More precisely, the paper investigates whether fully convolutional network can be trained while using the native 3D structure of the point cloud. To our knowledge, this is the first attempt at applying deep neural network to native bathymetric point cloud for classification purpose. As such, this work aims at designing and conducting a series of experiments to better understand the behavior of deep neural network when applied to underwater natural terrain and the sea-bottom features that impact it. The remainder of this paper is organized as follows. Section 2 is devoted to related work and main issues when applying deep neural network to bathymetric data. Then Section 3 focuses on the methodology including the data preprocessing and the network architecture we propose to classify bathymetric data morphology. Experimental results are presented in Section 4.

---

\* Corresponding author

Finally, Section 5 provides conclusions and perspectives of this work.

## 2. CHALLENGE OF BATHYMETRIC DATA AND RELATED WORK

### 2.1 Challenge of bathymetric data for deep neural network

Multi-beam systems record measurements of water depth (i.e. bathymetry), from which a number of secondary layers that provide information of seafloor morphology can be generated (e.g. seafloor slope, terrain variability) (Brown et al., 2019). The closest analogous instrument used for terrestrial surveying is an airborne laser scanner (ALS). MBES is based on the same principle as LiDAR namely it measures the angle and two way travel time of a transmitted pulse. In both cases, the returned signals are stored as raw, ungridded point clouds. The two systems have significant geometric differences. One noteworthy is the ratio of the beam footprint fluctuations with relation to the surface elevation. For the ALS, these fluctuations are in the order of 3-4% while for the MBES they are in the order of 40% (Hughes Clark, 2018). As a result, variations in the effective resolution are going to be much more pronounced across a multi-beam swath than an ALS swath. This will impact the achievable terrain discrimination using geomorphic techniques.

Echo-sounder uses acoustic wave while LiDAR uses electromagnetic wave. The significantly slower speed of the acoustic wave (ex. average speed 1500m/s compare to  $3.10^8$  m/s for the light) makes it more sensitive to the sensor platform movements. In addition, the sound celerity varies through the water column according to the temperature, salinity and pressure of the environment, which generates a refraction of the wave. The combination of such effects is the source of many spurious soundings and inconsistencies in the point cloud.

As the MBES platform does not follow a straight line path and rotates on three axes (roll, pitch, yaw), the point cloud density is uneven. In addition, the ensonified area is not of uniform size and the incidence angle of acoustic energy is highly variable. As a result, both the point cloud density and resolution vary strongly with both elevation and resolution across a single swath. These factors will impact the resolved terrain roughness thereby affecting any classification scheme based on surface characteristics (Hughes Clark, 2018).

Although progress is being made in applying DNN to object detection and semantic classification, most current efforts center on extracting man-made features such as buildings, roads, cars (Cheng et al., 2017, Deng et al., 2017; Yao et al., 2017, Roh, Lee, 2017). Three major challenges prevent advancement in natural, underwater terrain feature classification. The first challenge is the lack of a properly labelled seafloor database for training the DNN models adapted to bathymetric point clouds. Many existing open data sets exist in computer vision (ex. Imagenet (Deng et al., 2009)) and even in LiDAR (ex. KITTI (Geiger et al., 2013), Semantic3D (Heickel et al., 2017)). Similar data sets related to MBES point cloud classification are still missing. Although the performance of DNN has been promising, achieving high accuracy requires good-quality labeled data. To overcome this issue it requires creating in-house seafloor database and applying various data augmentation techniques. The second challenge is the vague boundary or ambiguous edge of seafloor terrain features such as slope, dunes. The third challenge is the lack of a thorough understanding of the DNN model's performance and the factors

that impact that performance, such as capability of the convolution modules and the hyper-parameter settings.

### 2.2 Related work

In the literature, there is seldom research study about DNN applied to bathymetric point cloud. Seafloor classification is usually done using machine learning or statistical methods. The methods that use MBES data for seafloor classification are primarily based on SVM, random forest, learning vector quantization (LVQ), self-organizing feature map (SOM) classifiers, and cluster analysis methods (Kaski, 1998, Li et al., 2011). These techniques often require parameters to be tuned for different areas, and offer limited scalability. In addition, automated tools often cannot provide consistent level of accuracy and analyst revert to more semi-manual processing (Stephens et al., 2020).

Most of the current deep learning solutions applied to underwater data concern sonar images (Denos et al., 2017). Even if there are challenges when processing such images (e.g. noisy underwater environment, lack of geometric features, ...), the network structure remains similar to conventional computer vision problems given the similarity of the image structure. Regarding MBES point cloud, only one DNN solution has been found in the literature (Stephens et al., 2020). In this paper, the authors present a first attempt at applying 3D Convolutional Neural Network (CNN) to the problem of denoising bathymetric point cloud. The results reported from the test sets show a promising performance with kappa scores of 0.94 and accuracy of 0.977. The data is structured as voxels before being fed to the network. Such voxelisation would not be appropriate in all circumstances. As underlined in Section 2.1, strong variations are expected in both point density and resolution across a multi-beam swath and according to the elevation. Thus, the voxel size may impact strongly the solution performance. In addition, in Stephens et al.'s approach, the sounding densities are scaled inside the voxels, which may not be a reasonable choices given the expected point density variation across a data set.

Given LiDAR and LiDAR point cloud similarity with MBES system and bathymetric point cloud respectively, we conducted a literature review in this field. Currently, there are few survey and review papers providing state-of-the-art deep learning models directly addressing point clouds (Liu et al., 2019, Griffiths, Boehm, 2019). LiDAR point cloud labeling methods can generally be grouped into two main categories: direct methods that operate immediately on the point clouds; indirect methods, which transform the input point cloud into an image or a volume as a preconditioning step. Given the discontinuous nature of point clouds, 3D Point cloud to 2D image conversion has the consequence of rendering adjacent in the 2D representation characteristics that are not necessarily close in 3D space. A straightforward volume representation of the point cloud is voxels. However, voxelized data imply many underlying challenges (ex. setting the voxel size; memory occupancy). Designing an efficient conversion of the point cloud into a dense voxel structure is doable as described in (Zhou, Tuzel, 2018). In order to directly operate on point clouds and avoid transforming the data to a different representation, Qi and al. proposed Point Net++, which achieved satisfactory results while enabling the network to learn local structures at different scales (Qi et al., 2017). The input of the network is an  $N \times M$  array of unordered data points where  $N$  is the number of points and  $M$  is the number of features of each points, i.e. the spatial coordinates of the sounding ( $X, Y,$

Z). Point Net++ demonstrates that unlike in 2D CNNs where small kernels are preferred, when point density is sparse, larger point samples are required for robust pattern extraction. Several extensions of Point Net++ have been proposed to achieve increased performances (Engelmann et al., 2018). Other significant works aim to incorporate a spatial convolution operator within the network. SplatNet (Su et al., 2018) is an example of such research effort.

The previous references concern methods that directly process unordered point clouds. A variety of literature is focused on ordering point cloud for processing for classification and segmentation. In some studies, the point cloud data is represented and indexed as shallow octrees (Riegler et al., 2017) or using kd-tree structure (Klokov, Lempitsky, 2017). New data structures can also be introduced as the superpoint graph (SPG) (Landrieu, Simonovsky, 2017). SPGs are derived by partitioning the point cloud into geometrically homogeneous elements.

A prominent issue with point cloud remains the lack of adequate quality training data. Even if labelled point cloud data sets are available, their size does not compare with the size of 2D image data sets.

### 3. METHODOLOGY

#### 3.1 Building the training data sets

As underlined, a key issue when applying deep neural network to bathymetric point cloud is the availability of relevant quality training data, and the amount of such training data required by DNN. To overcome this issue, two steps have been designed in the proposed methodology. The first step (cf. section 3.1.1) focuses on labelling bathymetric point cloud using morphological classes. The second step (cf. section 3.1.2) focuses on augmenting the training labelled data sets. The partitioning of the point cloud in order to structure the training data into batches is also addressed in this second step.

**3.1.1 Data classification using geomorphons:** To meet the DNN quality requirement, a dedicated morphological classification algorithm has been selected, namely the geomorphon algorithm (Jasiewicz, Stepinski, 2013). The geomorphon method is an efficient solution to process dedicated seafloor. However, like machine learning techniques, it presents limited scalability and variable accuracy according to the point cloud density and terrain topography. Furthermore, its processing time may prevent such solution from real-time usage. For all these reasons, geomorphon is a relevant classification method to build a quality training data set for DNN, but not a feasible solution to be embedded onboard ASV. The geomorphon is a raster-based algorithm. It analyses terrain surface (i.e. DTM) in order to detect and extract the 10 most relevant micro topographic structures such as flat, peak, spur, slope and so on. The technique uses texture analysis tools from the field of computer vision adapted to topographic forms rather than differential geometry tools.

When running, the geomorphon algorithm analyzes each cell of the DTM. For each of them, it looks in the direction of its 8 neighbours and evaluates whether the surface is going up, down or stays at the same height using the principle of line of sight. From a maximum search distance  $S$ , it looks in the direction of the 8 neighbours and calculates the zenith  $\phi$  and nadir  $\psi$  angles between the analyzed cell and the intercepted surface. A flatness

threshold  $F$  is applied to the zenith and nadir angle difference in order to assess if the neighbouring direction is descending (-1), ascending (1) or if it is at the same level (0) with respect to the analysed cell. Each group of 8 direction values can be classified as a dedicated micro topographic structures using a lookup table. The geomorphon algorithm provides bounding polygons as output of the terrain classification, where the raster cells inside a polygon have the same micro topography label. These polygons are then used to classify the MBES point cloud. The polygon label is transferred to all the points inside the polygon convex hull.

In order to assess the sensitivity of the approach to the sea-bottom topography and morphology, experiments have been conducted with various datasets using different values for the algorithm main parameters, namely the DTM resolution, the search radius  $S$  and the flatness threshold  $F$ . The relevance of the result has been, first, visually assessed. Then, the appropriateness of the classification was assessed by comparison with the result of a region growing method based on this classification. In (Dupont et al., 2019), authors described in more details the region growing approach. Figure 1 provides two classified sea-bottoms, a steep slope on one hand, a dune field on the other hand, for which the same geomorphon parameters have been used, namely: DTM resolution = twice the Canadian Hydrographic Services requirement (Canadian Hydrographic Services, 2020);  $S = 10$  pixels;  $F = 1^\circ$ . These experiments confirmed the quality of the classification results met the expectation for the DNN training data sets.

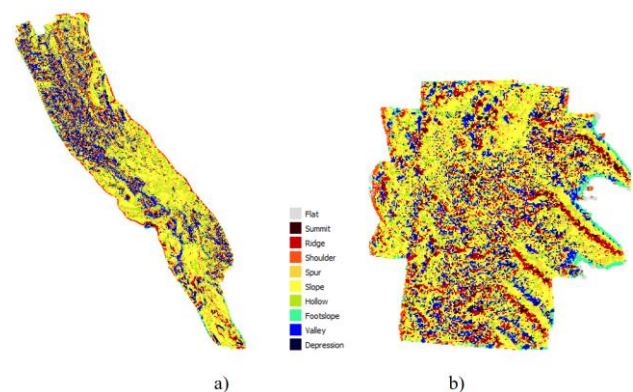


Figure 1. Sea-bottom classification using geomorphon approach: a) steep slope river bed with boulders; b) field of dunes on a flat seabed

**3.1.2 Data partitioning and augmentation:** similarly as PointNet++ (Qi et al., 2017), the bathymetric point clouds used as training data sets are partitioned in blocks of 8192 points. A sampling of each data set is carried out in order to regularly distribute positions that will constitute the centres of a predefined neighbourhood of points. All points in the MBES point cloud located within a radius  $R$  of the centre will be selected. The radius  $R$  has been determined as a function of the point density, namely 6.0 m. The centres are spaced at a distance equal to  $4R/3$ , which ensures complete coverage of the area by point blocks (there is a partial overlap of the point blocks). The data partitioning step aims at selecting 8192 points per block. If the block does not contain enough points, some points are randomly duplicated. If the block contains too many points, the 8192 points are randomly drawn.

In order to take into account the representativeness of each class within the blocks of points, the weight of each point according to the class is calculated. The weight is equal to the number of points of the class out of the total number of points, ensuring that the sum of the weights equals 1. This weight will intervene during the training of the network so that the network is not biased towards the dominant class in the point cloud.

While dividing the MBES point clouds into blocks of points increases the number of training samples, robustness to orientation is further improved by augmenting the training data with transformed versions of the original data. The transformation consists in randomly rotating points around the z-axis.

### 3.2 Network architecture

The proposed network architecture is based on recent works where 1D-fully convolutional networks are used to generate point-wise labeling of an airborne LiDAR point cloud (Yousefhusien et al., 2017). While PointNet-like networks are designed to deal with CAD-model or indoor point clouds, the proposed 1D-fully convolutional network was designed to overcome typical obstacles related to airborne LiDAR, namely noise, occlusions, scene clutter, and terrain variation. Since airborne laser scanner is the closest analogous instrument to MBES, bathymetric point cloud share significant similarities with ALS point cloud. However, as underlined in section 2.1, there are still dedicated issues with relation to the underwater environment the proposed network was not engineered to solve. Also, the fully convolutional network proposed by Yousefhusien et al. takes advantage of three spatial coordinates and three corresponding spectral values for each point. Bathymetric data consists only of three spatial coordinates limiting the features that can be learnt from the point cloud.

The network architecture is structured in order to learn local and global features. The input of the network is an  $N \times M$  array of unordered data points where  $N$  is the number of points and  $M$  is the number of features of each points, i.e. the spatial coordinates of the sounding ( $X, Y, Z$ ). The feature learning part of the network aims at extracting both local features (at the point level) and global features (at the block-level stemming from the preprocessing step), while providing the required invariances for the labelling task. It consists of a Fully Connected Network (FCN) combining a series (five in total) of  $1 \times 1$  convolutional layer, batch normalization (BN) layer and a rectified linear unit (ReLU). The convolutional layers involve from 64 to 2048 output features. Global features can be extracted using a pooling layer, which simultaneously provides permutation-invariance. Local features are obtained from an intermediary convolutional layer. The second part of the network consists of a convolutional neural network used to label the points. The input consists of the global feature vector concatenated with the point level feature vector. The convolutional layers output 512 features. The final output is the point label provided by a softmax classifier present at the end of the network. Figure 2 provides a general overview of the network.

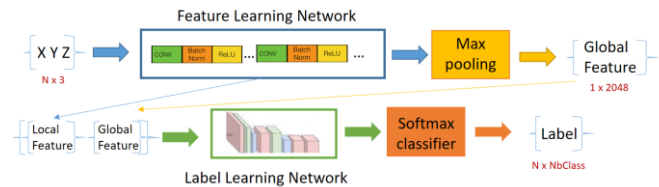


Figure 2. Fully convolutional neural network for bathymetric point cloud labelling

## 4. EXPERIMENTS

### 4.1 Data set

Two data sets were used in the experiments. The first data set is a control zone used for patch test purpose. It is a flat area located about 3km outside Rimouski's harbor (Quebec, Canada). The bathymetric point cloud was recorded by CIDCO (Research and Development Center in coastal and ocean mapping) in September 2015 on board of the FJ Saucier vessel. The Reson Seabat T20P was used to record the data. The area is approximately 325m long by 200 m wide. The depth varies between 5 to 12m. The point cloud consists of about 4 millions points and the average point density is 73pts/m<sup>2</sup>.

The second bathymetric point cloud was recorded during a survey carried out by Groupe Océan in 2016 on board of the Korok vessel. The study area is situated in the Saint-Lawrence river near the Chaudière river estuary (Quebec, Canada). It is approximately 200 metres long by 50 metres wide. The depth varies between few meters to 13m. The vessel carried out 8 parallel survey lines and 2 additional perpendicular lines for validation purpose. The R2Sonic 2022 MBES was used to record the data. The point cloud consists of about 3 millions points and the point density is approximately 250pts/m<sup>2</sup>. Each study area is divided into two regions, one for training and one for testing. Figure 3 presents the two study sites represented as bathymetric surfaces.

The geomorphon method enable the labelling of the bathymetric point cloud into 10 micro topographic structures. As explained in the paper introduction, to fully take advantage of hydrographic survey using ASV, onboard processing dedicated to computing error estimators to assess if the survey meets the required quality is required. Since such estimators are generally applied on flat or slope sea-bottom, the 10 classes provided by the geomorphon algorithm have been merged into two classes as follows: classes Flat and Slope have been merged into a single class; the rest of the 10 classes have been merged into a second class.



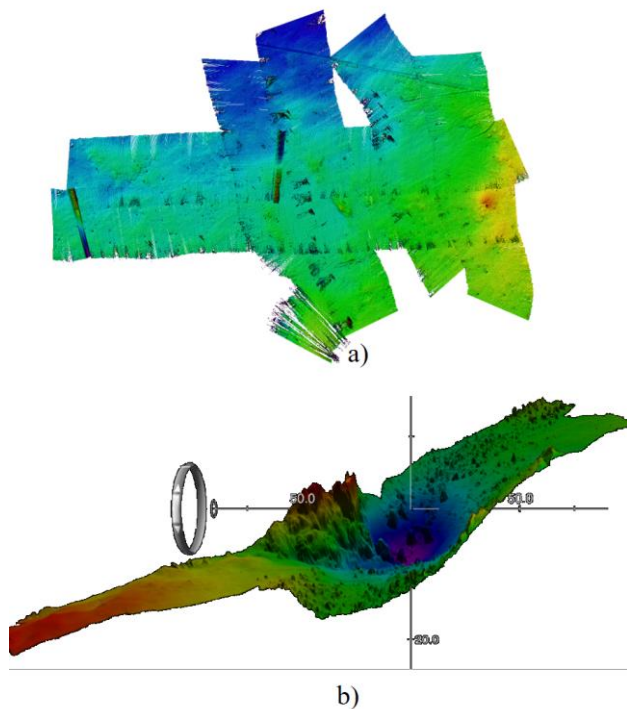


Figure 3. Sea-bottom surface representation of the two study sites: a) Rimouski control zone; b) Chaudière river estuary

#### 4.2 Training parameters

We used the Adam optimizer (Kingma, Ba, 2014) with an initial learning rate 0.001, a momentum of 0.9 and a batch size of 16. The learning rate is iteratively lowered as the learning progresses. An exponential decay function is applied to the initial learning rate. The network has been implemented using Keras (Chollet, 2015) with the Tensorflow backend. The training proceeds for a total of 60 epochs using a Tesla V100 GPU. We monitor the loss and overall accuracy progress during training and validation.

#### 4.3 Labelling results

Several experiments have been conducted in order to increase the understanding of the fully connected network architecture when applied to bathymetric data related to natural sea-bottom. Such network has already demonstrated significant performance when applied to airborne LiDAR data represented as a native 3D point cloud (Yousefhusien et al., 2017). The impact of the terrain complexity as well as the number of classes used to label the point cloud were assessed through the experiments. The first experiment concerned the labelling of the control zone near Rimouski using 2 classes namely Flat seabed / Other. The goal was to verify the ability of the network to classify a sea-bottom with an obvious and explicit topography. The second experiment concerned the Chaudière river estuary where the global topography is a steep slope but the local topography includes various micro topographic structures. As such, the seabed displays some geometrical features that could be learnt by the network. The network has been trained in order to label the point cloud using, first, the 10 micro topographic structures provided by the geomorphon, and second, two classes namely Flat/Slope and Other. Table 1 provides the percentage of points in the two most represented classes of the training data set. For each experiment, the following performance measures have

been computed: the training accuracy per epoch; the test accuracy per epoch. The class omission and commission were computed only for the labelling into 2 classes. Table 2 provides an overview of these measures for each experiment.

	Rimouski seabed (2 classes labelling)	Chaudière river seabed (10 classes labelling)	Chaudière river seabed (2 classes labelling)
Main class	99,88%	37,51%	62,48
Second class	0,12%	19,71%	37,52%

Table 1. Percentage of points in the two main classes of the data sets used in the experiments

	Rimouski seabed (2 classes labelling)	Chaudière river seabed (10 classes labelling)	Chaudière river seabed (2 classes labelling)
Max Accuracy Training	99,96%	40%	61%
Max Accuracy Test	99,87%	37%	65%
% omission	0%	NA	0%
% commission	0,13%	NA	36%

Table 2. Performance measures provided by the three experiments (NA: not applicable)

For the first experiment, the convergence of the network is reached after few epochs. The training accuracy starts already at 93% and is equal to 99,9% after convergence. The test accuracy is not progressing during the training phase. It remains equal to 99,8%. This is due to the large number of points labelled as flat/slope sea-bottom in comparison to the other class. Labelling all the points as flat seabed without any learning, the chance to mistake is less than 1% given the class proportion in the training data set (cf. Table 1).

To assess the network capacity to label different seabed morphology, it is trained using the second data set where the 10 classes from the geomorphon approach have been retrieved. After 60 epochs, the best training accuracy is about 40% and the test accuracy reached 37%. This is in the order of the main class proportion in the training data set (cf. Table 2). As a result, there is no learning in the network. There may be several reasons to explain such a result. First, there may be a lack of data to learn features that will discriminate up to 10 classes. In machine learning, the more classes to discriminate, the more data samples required. Furthermore, even if the network architecture involves feature learning at the local and global level, the point cloud is analyzed using a single neighborhood size. Identifying different micro topographic structures may require different neighborhood sizes. When dealing with airborne LiDAR data, it is usual to have network fed with block of points extracted using various sizes in order to be able to extract objects at different scale like buildings and trees versus cars. In addition, one can notice the various shapes displayed by micro topographic structures belonging to the same class. Unlike the urban environment, the classification of the natural terrain involves forms that are not very repeatable for groups of points associated with similar morphologies.

Given the lack of data to discriminate 10 classes, the same data set is used to train the network but using only 2 classes. The

training accuracy is about 61% and the final test accuracy is about 65%. These results are again in the order of the class distribution in the training data set. The slight increase in training accuracy tends to indicate that the network is learning. However, when looking at the data set labelled in two classes, one can notice the lack of spatial patterns allowing discriminating the two classes. Again, there is a lack of repeatable features in the point cloud. Areas related to the same class will have different point distribution.

These results demonstrate the network difficulty to learn from the proposed data sets. As explained in section 2, the point cloud density and resolution vary strongly with both elevation and resolution across a single swath. There is a lack of repeatability of the seabed morphology appearance all along the point cloud, which could make it difficult for the network to identify pattern of points to associate to a dedicated class. Fully connected network learning approach relies on a certain form of logic in the training data sets. Such consistency tends to be lacking in underwater terrain.

## 5. CONCLUSIONS AND PERSPECTIVES

This paper has introduced a first investigation of applying deep neural network to MBES point cloud. There is currently a gap in deep neural network applications and natural terrain analysis. This gap is even larger when it comes to underwater environment. We addressed the challenges of such a context and of MBES point cloud. We proposed an approach to overcome the lack of relevant labelled data sets to train deep neural network. We implemented a fully connected network adapted from an architecture displaying significant performances with airborne LiDAR data recorded in an urban context. The proposed network provided accuracy results no better than 65% when classifying a steep seabed into two classes. The classification rate was slightly better than the class distribution in the point clouds. Results underlined the network difficulty to learn from the geometric features in the point cloud. Point density and resolution have a strong impact on the seabed morphology thereby affecting the classification scheme.

Further investigation is required to better adapt the network to the sea-bottom features and bathymetric point distribution. Multiple block size may be required to adapt to the various coverages of the micro topographic structures. Additional information, like the acoustic backscattered signal, could be used to help the labelling process. Indeed, works on airborne LiDAR data show the performances of deep neural network significantly increase when multispectral images are used jointly with the LiDAR point cloud. Also, further efforts need to be invested in designing a network architecture specifically dedicated to point cloud.

## ACKNOWLEDGEMENTS

This work is funded in part by the Discovery Program of the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by the Fonds de Recherche du Québec – Nature et Technologie, Team program. The authors would like to thank CIDCO and Groupe Océan for the study data sets.

## REFERENCES

Brown, C. J., Beaudoin, J., Brissette, M., & Gazzola, V., 2019. Multispectral multibeam echo sounder backscatter as a tool for improved seafloor characterization. *Geosciences*, 9(3), 126.

Canadian Hydrographic Services, 2012. Recommended base surface resolution, Fisheries and Oceans Canada - Canadian Hydrographic Service (CHS). [www.charts.gc.ca/data-gestion/bathymetric-traitement/index-fra.html](http://www.charts.gc.ca/data-gestion/bathymetric-traitement/index-fra.html) (28 January 2020).

Cheng, G., Wang, Y., Xu, S., Wang, H., Xiang, S., & Pan, C., 2017. Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 55 (6), 3322–3337.

Chollet, F. et al. 2015. Keras [github.com/fchollet/keras](https://github.com/fchollet/keras) (28 January 2020).

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.

Deng, Z., Sun, H., Zhou, S., Zhao, J., & Zou, H., 2017. Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10 (8), 3652–3664.

Denos, K., Ravaut, M., Fagette, A., & Lim, H. S. (2017, June). Deep learning applied to underwater mine warfare. In *OCEANS 2017-Aberdeen* (pp. 1-7). IEEE.

Desa, E., Madhan, R., & Maurya, P., 2006. Potential of autonomous underwater vehicles as new generation ocean data platforms. Indian Academy of Sciences.

Dupont, V., Daniel, S., & Larouche, C., 2019. A region growing algorithm adapted to bathymetric point clouds. In *OCEANS 2019 MTS/IEEE SEATTLE* (pp. 1-6). IEEE.

Engelmann, F., Kontogianni, T., Schult, J., & Leibe, B., 2018. Know what your neighbors do: 3D semantic segmentation of point clouds. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 0-0).

Geiger, A., Lenz, P., Stiller, C., & Urtasun, R., 2013. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11), 1231-1237.

Griffiths, D., & Boehm, J., 2019. A Review on deep learning techniques for 3D sensed data classification. *Remote Sensing*, 11(12), 1499.

Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., & Pollefeys, M., 2017. Semantic3d. net: A new large-scale point cloud classification benchmark. *arXiv preprint arXiv:1704.03847*.

Hughes Clarke, J. E., 2018. The Impact of acoustic imaging geometry on the fidelity of seabed bathymetric models. *Geosciences*, 8(4), 109.

Jasiewicz, J., & Stepinski, T. F., 2013. Geomorphons—a pattern recognition approach to classification and mapping of landforms. *Geomorphology*, 182, 147-156.

Kaski, S., 1998. Dimensionality reduction by random mapping: Fast similarity computation for clustering; Proceedings of the

- 1998 IEEE International Joint Conference on Neural Networks, IEEE World Congress on Computational Intelligence; Anchorage, AK, USA. 4–9 May 1998; pp. 413–418. *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4490-4499).
- Kingma, D. P., Ba, J., 2014. Adam: A method for stochastic optimization. CoRR abs/1412.6980.
- Klokov, R. and Lempitsky, V., 2017. Escape from cells: Deep kdnetworks for the recognition of 3D point cloud models, in ICCV, 2017, pp. 863–872.
- Landrieu, L., & Simonovsky, M., 2018. Large-scale point cloud semantic segmentation with superpoint graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4558-4567).
- LeCun, Y., Bengio, Y., & Hinton, G., 2015. Deep learning. *Nature*, 521(7553), 436.
- Li J., Heap A.D., Potter A., Huang Z., Daniel J.J., 2011. Can we improve the spatial predictions of seabed sediments? A case study of spatial interpolation of mud content across the southwest Australian margin. *Cont. Shelf Res.* 2011;31:1365–1376. doi: 10.1016/j.csr.2011.05.015.
- Liu, W., Sun, J., Li, W., Hu, T., & Wang, P., 2019. Deep learning on point clouds and its application: A survey. *Sensors*, 19(19), 4188.
- Qi, C. R., Su, H., Mo, K., & Guibas, L. J., 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 652-660).
- Riegler, G., Osman Ulusoy, A. and Geiger, A., 2017. OctNet: Learning deep 3D representations at high resolutions, in CVPR, 2017, pp. 3577–3586.
- Roh, M. C. and Lee, J. Y., 2017. Refining faster-RCNN for accurate object detection. In *Machine Vision Applications (MVA), 2017 Fifteenth IAPR International Conference on IEEE*, Hamar, Norway, 514-517.
- Stephens, D., Smith, A., Redfern, T., Talbot, A., Lessnoff, A., & Dempsey, K., 2020. Using three dimensional convolutional neural networks for denoising echosounder point cloud data. *Applied Computing and Geosciences*, 5, 100016.
- Su, H., Jampani, V., Sun, D., Maji, S., Kalogerakis, E., Yang, M. H., & Kautz, J. (2018). Splatnet: Sparse lattice networks for point cloud processing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2530-2539).
- Yao, Y., Jiang, Z., Zhang, H., Cai, B., Meng, G., & Zuo, D., 2017. Chimney and condensing tower detection based on faster R-CNN in high resolution remote sensing images. *Geoscience and Remote Sensing Symposium*, IEEE, Fort Worth, Texas, 3329–3332
- Yousefhusien, M., Kelbe, D. J., Ientilucci, E. J., & Salvaggio, C., 2018. A multi-scale fully convolutional network for semantic labeling of 3D point clouds. *ISPRS journal of photogrammetry and remote sensing*, 143, 191-204.
- Zhou, Y., & Tuzel, O., 2018. Voxnet: End-to-end learning for point cloud based 3d object detection. In *Proceedings of the*