Integrating LiDAR Point Cloud Classification and Building Footprints for Enhanced 3D LOD Building Modeling: A Deep Learning Approach

Lasithasree Lakshmanan[1], Sudhagar Nagarajan[2]

[1]llakshmanan2023@fau.edu
[2]snagarajan@fau.edu
[1], [2] Department of Civil, Environmental and Geomatics Engineering, Florida Atlantic University, 777 Glades Road, FL 33431.

Keywords: LiDAR Point cloud classification, Building Footprint Extraction, 3D Building Modeling, Open3D Library.

Abstract

3D Building modeling is crucial for urban planning, helping stakeholders make informed decisions on critical issues such as flood risk assessment, urban heat island effect, and sustainable infrastructure development. In this research, we use RandLA-Net, a cutting-edge deep learning algorithm to classify LiDAR point cloud data to distinguish building structures. This identification of building points is essential for the subsequent creation of reliable 3D Building LOD models. To enhance the classification accuracy, this research utilizes building footprint vector data as a reference layer, which aids in refining the detection of building points and ensures validation of the results. Once the building points are classified and improvised with Building Footprint vector layer, they are utilized to reconstruct detailed 3D geometric models. This study employs the Open3D library to generate Levels of Detail (LOD) models for buildings. By combining advanced LiDAR point cloud classification with Building Footprint Extraction and 3D Building modeling techniques, this approach maximizes the utility of publicly accessible LiDAR point cloud data, delivering detailed 3D models that support a wide range of spatial decision-making processes.

1. Introduction

LiDAR has become a robust technology for capturing highly detailed 3D representations of the Earth's surface. LiDAR sensors collect 3D points that define the geometry of objects and surfaces, making it indispensable for applications like urban planning, infrastructure development, disaster management, and environmental monitoring (Edson & Wing, 2011).

LiDAR data significantly differs from traditional satellite imagery in several ways. While multispectral satellite imagery typically captures surface reflectance, providing a two-dimensional view, LiDAR offers three-dimensional spatial information with specifications like coordinates, intensity and return number. (Urbazaev et al., 2018). This capability allows for more detailed modelling of building architecture, vegetation canopy structure, and terrain features, which are often difficult to detect from 2D imagery alone. Furthermore, integrating LiDAR point cloud data with multispectral imagery enhances the understanding of complex environments (Parmehr et al., 2016). However, this fusion is with challenges, such as misalignment between LiDAR point cloud and multispectral satellite imagery caused by relief displacement in tall buildings (Gopalakrishnan et al., 2020).

Some of the key outputs of LiDAR point cloud data includes Digital Terrain Models (DTMs), Digital Elevation Models (DEMs), Canopy Height models (CHMs), and building footprints, all of which are crucial for a broad field of applications (Sharma et al., 2021). Despite the utility of these outputs, extracting useful information from LiDAR data requires effective classification and extraction techniques.

Traditional classification methods, such as Random Forest classifiers, have been used to categorize buildings into rooftops, walls, ground, and high outliers (Park & Guldmann, 2019).

However, these methods often rely on manual feature extraction and extensive preprocessing, limiting their efficiency. Methods like region growing, clustering, and surface fitting have provided the foundation for LiDAR point cloud classification, but they struggle with the complexity and irregularity of the points.

In contrast, deep learning models like RandLA-Net offer significant advantages by providing better automation, flexibility, and performance (Sarker et al., 2024). Additionally, the integration of building footprint vector data, which are from local cadastral offices and sources like Microsoft's building footprint, has been shown to enhance classification accuracy. By overlaying building boundaries, detection and classification of building points is improved. (Park & Guldmann, 2019) (Lloyd et al., 2020) (Esch et al., 2020).

Recent progress in deep learning has enhanced the LiDAR point cloud classification (Garcia-Garcia et al., 2017). While early deep learning models focused primarily on multispectral imagery (Zhao et al., 2019) (Hoeser & Kuenzer, 2020). The advancements in deep learning for 2D imagery has prompted its application to point cloud data, which is gaining increasing attention (Jiang et al., 2021). The availability of publicly accessible benchmark datasets for point cloud data, such as ISPRS Vaihingen 3-D dataset (Li et al., 2020; Wen et al., 2020), Toronto 3D (Tan et al., n.d.), Semantic3D (Hackel et al., 2017) and the SemanticKITTI benchmark dataset (Behley et al., n.d.), has contributed to growth in the LiDAR data classification research.

Deep learning models for LiDAR data classification fall into three main categories: projection-based, point-based, and hybrid-based methods. Projection-based techniques apply 2D convolutional networks to extract features by converting the data into 2D space, while point-based models like PointNet,

PointNet++, RandLA-Net and DGCNN directly process raw point cloud data without transformation. Hybrid methods combine both approaches to improve performance (Liu et al., 2019).

Among the deep learning models, RandLA-Net stands out due to its efficient random sampling technique, which allows it to handle large-scale point clouds effectively than many of the other deep learning models (Hu et al., n.d., 2022). It mitigates the loss of essential features caused by random sampling by employing a local feature aggregation module and performs competitively and even surpasses other leading models such as PointNet++, DGCNN, and KPConv on benchmark datasets like S3DIS, Semantic3D, and SemanticKITTI (Hu et al., n.d.).

For the analysis, processing, and visualization of LiDAR data, Open3D library plays a major role in the research, offering essential functionalities. It also facilitates implementing advanced algorithms, such as DBSCAN enabling clustering of urban structures (Zhou et al., 2018). DBSCAN clustering method is effective in segmenting LiDAR data, particularly in identifying clusters within noisy datasets (Wang et al., 2019). Combining DBSCAN with advanced deep learning models like RandLA-Net could offer a comprehensive approach to 3D modeling from LiDAR data (Zeng & Gevers, n.d.).

In summary, this research addresses the use of LiDAR data for urban planning, such as the integration of building footprint vector data with LiDAR point cloud classification. By combining this method with the 3D building modeling, this research emphasizes an integrated approach that produces detailed Levels of Detail (LOD) 3D models for urban planning.

2. Data

LiDAR data for the research was sourced from the U.S. Geological Survey (USGS) and it is publicly accessible. The region of interest for this research is a small portion of the residential area in the Boca Raton, Florida.
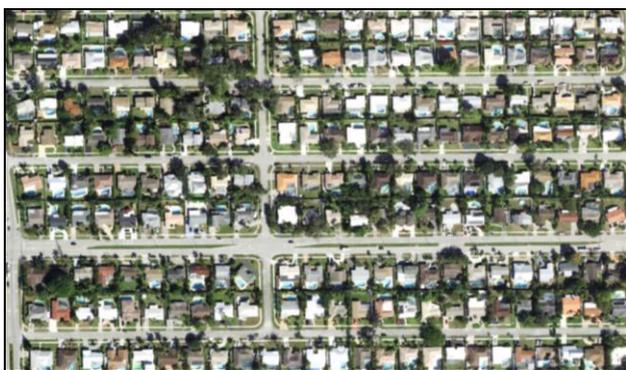

Figure 1. Region of study – A small portion of Residential Area in Boca Raton, Florida. (Source: NAIP Imagery)

The data collection took place between May 8, 2021, and October 29, 2021, and was later published on March 3, 2022. The data is available in LAS format, compressed as LAZ. This LiDAR dataset has a Quality Level 2 (QL2) classification, which indicates a moderate resolution.

3. Methodology

The methodology for this research includes five main steps, LiDAR Data Preprocessing, LiDAR point cloud classification,

Building Footprint extraction, Cross validation and improvement and 3D Building (LOD- Level 1) Model Extraction, and each of them are explained detailly below.
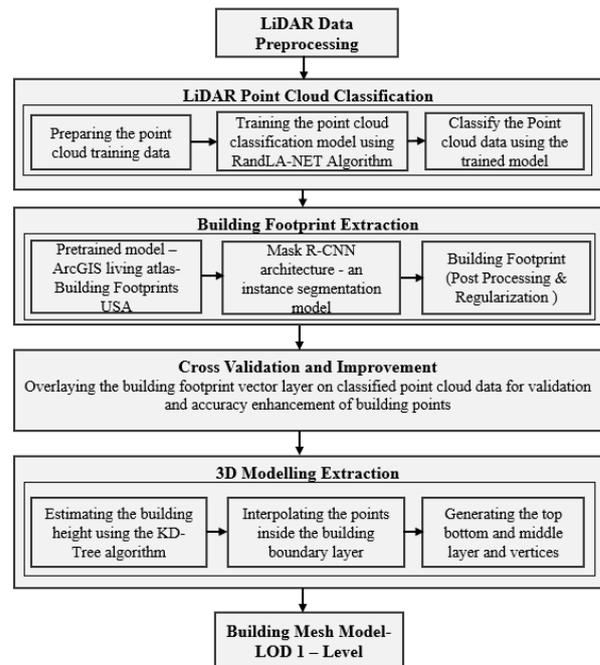

Figure 2. Overall Methodology – Flow Chart

3.1 LiDAR Data Preprocessing and Classification

The raw LiDAR data, sourced from the U.S. Geological Survey (USGS), was pre-processed to remove high outliers, which could interfere with the classification further.

The classification of the LiDAR point cloud data was performed using the RandLA-Net algorithm, a deep learning model for large-scale 3D point cloud classification. RandLA-Net was chosen for its better accuracy and its effectiveness in handling the large-scale point cloud data. The classification process began with the preparation of 70% of training and 30% of validation datasets and the RandLA-Net model was trained over 10 epochs, with 100 iterations per epoch, ensuring that the model sufficiently learned to distinguish the features of building points.

Figure 3. The classified Building points using RandLA-Net Algorithm (Building points are visualised in red colour)

Once the building points were classified, the DBSCAN clustering technique was used to cluster the building points into distinct segments, each representing a different building segment. The clustering parameters, such as epsilon and minimum cluster points, were fine-tuned to identify the building clusters. Specifically, an epsilon value of 10 and a minimum of 200 points per cluster were set, resulting in the identification of 227 unique building clusters.
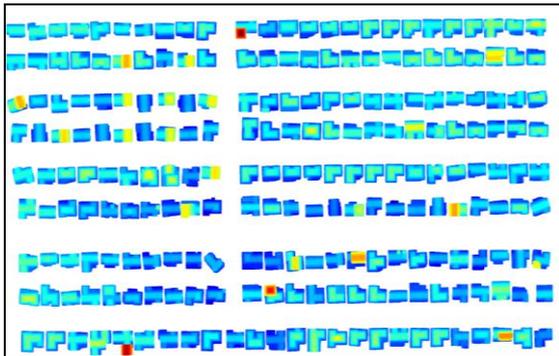

Figure 4. 227 Building clusters that were obtained by using DBSCAN clustering technique over the classified building points.

3.2 Building Footprint Extraction and Cross-Validation

Building footprints represent the precise outlines of buildings, capturing the spatial boundaries of building segments within a given area. The building footprint vector layer was extracted from NAIP imagery, using the pre-trained model ArcGIS Living Atlas – Building Footprints USA to detect buildings within the region of interest. The Mask R-CNN architecture was employed for segmentation, enabling the detection of individual buildings. Post-processing steps includes building footprint regularization to correct the distortions and unwanted artifacts, ensuring accuracy in the building geometries.
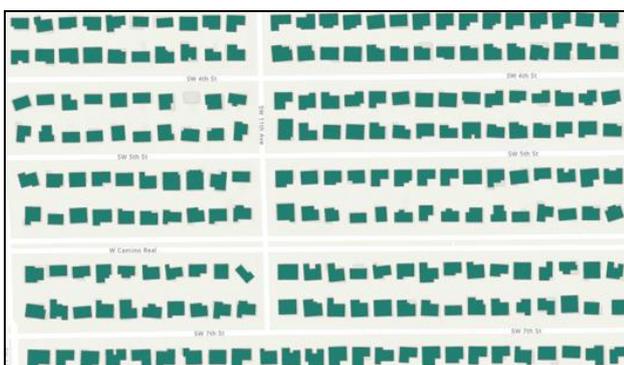

Figure 5. 226 Buildings were detected using Building Footprint Extraction from NAIP imagery.

As a part of the next step, cross-validation was done between the extracted building footprints and the classified LiDAR point cloud data. In the cross-validation and improvement method, the building footprint vector layer was overlaid onto the classified LiDAR point cloud data to enhance the accuracy of the building detection process. This overlay facilitated the alignment and validation of classified building points against known building footprints, enabling the identification and rectification of any misclassified areas. The comparison between the detected building points and the building footprint vector layer provided a reference for confirming the presence and boundaries of buildings.

3.3 3D Building LOD Model Extraction

After the refinement of LiDAR building points with the building footprint vector layer, the next step is to derive Levels of Detail (LOD) building models from the classified points. The initial step of this LOD model extraction involves extracting the base outline from the LiDAR point cloud data, for every building cluster.

Additionally, 3D building model requires the heights of their respective buildings to form the extrusion for the LOD Level 1 Building. To estimate the extrusion height, the minimum height values of nearby ground points were identified using a KD-Tree search algorithm. This query located the 100 nearest ground points around the building segment, to establish the local ground level. After the nearby ground points were extracted, the true building height was calculated by determining the difference between the rooftop points and the identified ground level.

Figure 6. Side view of the point cloud of a single segment and vector layer generated after the simplification process.

The process of generating a 3D building model includes creating the top, bottom and the middle layers from the base outline. This outline serves as a foundation for generating the 3D model. To construct the top, bottom and middle layers, the base outline will define the boundaries at different heights corresponding to the rooftop and ground level. These layers are then stacked to form the building's vertical structure. By defining these layers, a closed volume is created, representing the 3D LOD Level 1 of the building.
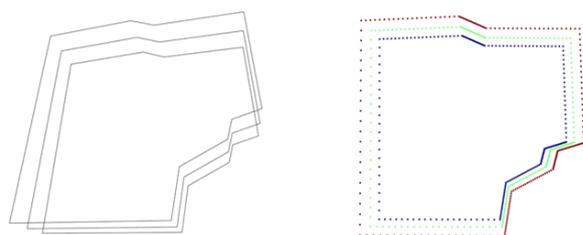

Figure 7. Top, middle and bottom base and vertices layer with the height extrusion.

The visualization above indicates the distinction between the top, middle and bottom layers, showing how the building's vertical extents are defined. This layered approach ensures that the 3D building model is constructed, capturing the height differences. Following the clustering process, all the clustered building segments after going through the process of extraction of top, bottom and middle layers, points within the boundary

layer were then interpolated to fill gaps in the original sparse point cloud, by creating additional points within the bounding box of the layers. Mesh generation was conducted using alpha shape method to achieve a detailed and precise fit that captures structural nuances of buildings. The result is a Level of Detail (LOD-1) 3D mesh, effectively representing the building's geometry.
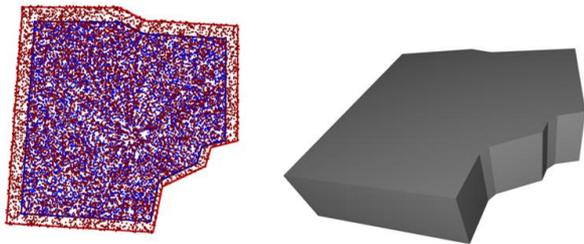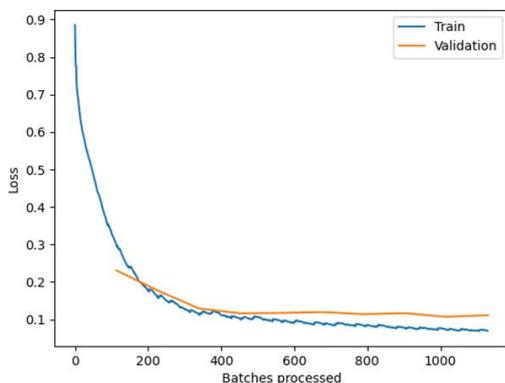
EPOCH	TRAINING_LOSS	VALIDATION_LOSS	ACCURACY	PRECISION	RECALL	F1_SCORE
0	0.19	0.23	0.94	0.92	0.82	0.86
1	0.08	0.18	0.95	0.95	0.86	0.89
2	0.08	0.13	0.96	0.95	0.9	0.92
3	0.08	0.12	0.97	0.96	0.9	0.92
4	0.08	0.12	0.97	0.96	0.9	0.93
5	0.08	0.12	0.97	0.97	0.9	0.93
6	0.08	0.11	0.97	0.96	0.9	0.93
7	0.09	0.12	0.97	0.97	0.9	0.93
8	0.07	0.11	0.97	0.96	0.91	0.93
9	0.08	0.11	0.97	0.97	0.91	0.93

Figure 10. A summary of model's performance metrics across 10 epochs.

This table summarizes the model's performance over 10 epochs, showing steady improvements in key metrics. The training and validation losses stabilize around 0.08 and 0.11, indicating low error. Accuracy reaches 97%, with precision and recall stabilizing around 97% and 91%, respectively, with an F1 score of 93%. These metrics demonstrate the model's reliability in classifying LiDAR point cloud data, with a balance between precision and recall metrics.

Figure 8. Mesh Model Generation from the top and bottom layers.

4. Results and Discussion

The classification of the LiDAR point cloud data was performed using the RandLA-Net algorithm, achieving an accuracy of 96.9%. This trained model was applied to classify the LiDAR point cloud data for the region of interest.
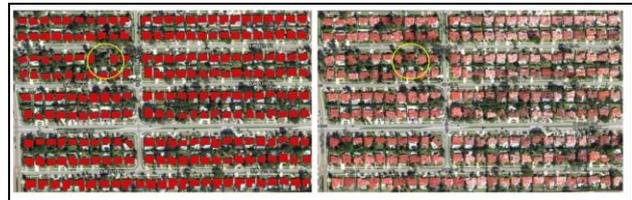

Figure 11. Cross-validation between the Building Footprints vector data and building clusters.

The building footprint extraction process from NAIP imagery, identified 226 buildings out of the 227 known structures within the region. In contrast, the LiDAR point cloud classification using RandLA-Net and DBSCAN clustering detected all 227 buildings. This discrepancy reveals that one building was missed in the NAIP-derived footprint extraction. This is due to the limitation of the model to detect the building, when it was overcrowded by the trees around. Consequently, the missing building boundary was added to the building footprint vector layer. This layer was then overlaid on the classified LiDAR point cloud data to enhance classification accuracy, allowing outliers to be filtered out.

Training and Validation loss

Analysis of the model

Accuracy: 0.9695125818252563

Figure 9. Training and Validation loss curves of the model over processed batches

Initially, both the training and validation losses start at high ranges, with a steep decline as the model processes more batches. This rapid decrease suggests that the model quickly learns essential features in the initial stages. After about 200 batches, the training loss stabilizes and continues to decrease slowly, indicating that the model is learning and improving its accuracy with small changes to the loss. The validation loss also shows a consistent decline initially and stabilizes at a lower loss value, which suggests that the model is generalizing well to the validation data.

Figure 12. The improvised building points classification overlaid with the building footprint vector layer.

The 3D Level of Detail (LOD) building model was generated for each of the 227 clusters. These models closely resemble the

LOD Level 1 models produced by the LAS Building Multipatch tool in ArcGIS Pro. The mesh creation, performed using Open3D library's alpha shape mesh modeling, provided a detailed representation of each building capturing their spatial geometry.
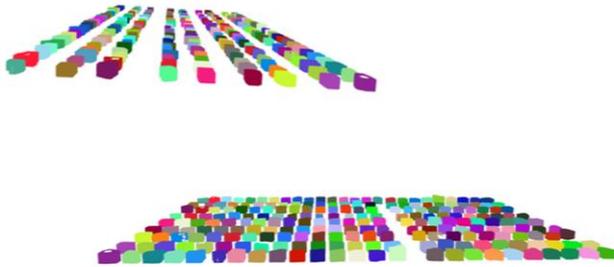

Figure 13. 3D LOD Level-1 Mesh models

5. Conclusion

5.1 Conclusion

This study demonstrates the use of publicly accessible, moderate-resolution LiDAR point cloud data to produce low-cost 3D Level of Detail (LOD) models for urban planning and development. Using the RandLA-Net model for building classification achieved a high accuracy of 96.9%, and the resulting LOD Level 1 mesh models closely resembled those generated by the LAS Building Multipatch Tool in ArcGIS Pro, However, some of the generated building models provided sparse areas for the mesh representation of certain buildings, indicating a limitation in mesh completeness this may impact applications requiring highly detailed LOD Level 1 models. While this method was applied to buildings in a residential environment, it can be adapted to other urban environments, though the level of detail may vary with different building types and areas.

A key contribution of this research is the integration of the building footprint vector layer with LiDAR point cloud classification, providing an improvement to the classification and as a cross-validation step to refine building segment detection.

5.2 Future Research

Future research will focus on implementing the extraction workflow with higher-resolution LiDAR point cloud data to yield more detailed results. This enhancement is expected to improve mesh completeness and reduce sparse areas within building mesh models, addressing limitations. Additionally, future work could also expand to generate more detailed building models, such as LOD3 and LOD4 to capture even finer details of the building.

References

Ballouch, Z., Hajji, R., Poux, F., Kharroubi, A., & Billen, R. (2022). A Prior Level Fusion Approach for the Semantic Segmentation of 3D Point Clouds Using Deep Learning. *Remote Sensing*, *14*(14). https://doi.org/10.3390/rs14143415

Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., & Gall, J. (n.d.). *SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences*. www.semantic-kitti.org

Edson, C., & Wing, M. G. (2011). Airborne light detection and ranging (LiDAR) for individual tree stem location, height, and biomass measurements. *Remote Sensing*, *3*(11), 2494–2528. https://doi.org/10.3390/rs3112494

Esch, T., Zeidler, J., Palacios-Lopez, D., Marconcini, M., Roth, A., Mönks, M., Leutner, B., Brzoska, E., Metz-Marconcini, A., Bachofer, F., Loekken, S., & Dech, S. (2020). Towards a large-scale 3D modeling of the built environment-joint analysis of tanDEM-X, sentinel-2 and open street map data. *Remote Sensing*, *12*(15). https://doi.org/10.3390/RS12152391

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). *A Review on Deep Learning Techniques Applied to Semantic Segmentation*. http://arxiv.org/abs/1704.06857

Gopalakrishnan, R., Seppänen, A., Kukkonen, M., & Packalen, P. (2020). Utility of image point cloud data towards generating enhanced multitemporal multisensor land cover maps. *International Journal of Applied Earth Observation and Geoinformation*, *86*. https://doi.org/10.1016/j.jag.2019.102012

Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., & Pollefeys, M. (2017). *Semantic3D.net: A new Large-scale Point Cloud Classification Benchmark*. http://arxiv.org/abs/1704.03847

Hoeser, T., & Kuenzer, C. (2020). Object detection and image segmentation with deep learning on Earth observation data: A review-part I: Evolution and recent trends. In *Remote Sensing* (Vol. 12, Issue 10). MDPI AG. https://doi.org/10.3390/rs12101667

Hu, Q., Ang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., & Markham, A. (n.d.). *RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds*.

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., & Markham, A. (2022). Learning Semantic Segmentation of Large-Scale Point Clouds With Random Sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(11), 8338–8354. https://doi.org/10.1109/TPAMI.2021.3083288

Jiang, D., Li, G., Tan, C., Huang, L., Sun, Y., & Kong, J. (2021). Semantic segmentation for multiscale target based on object recognition using the improved Faster-RCNN model. *Future Generation Computer Systems*, *123*, 94–104. https://doi.org/10.1016/j.future.2021.04.019

Li, W., Wang, F. D., & Xia, G. S. (2020). A geometry-attentional network for ALS point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *164*, 26–40. https://doi.org/10.1016/j.isprsjprs.2020.03.016

Lin, Y. C., Shao, J., Shin, S. Y., Saka, Z., Joseph, M., Manish, R., Fei, S., & Habib, A. (2022). Comparative Analysis of Multi-Platform, Multi-Resolution, Multi-Temporal LiDAR Data for

Forest Inventory. *Remote Sensing*, *14*(3). https://doi.org/10.3390/rs14030649

Liu, W., Sun, J., Li, W., Hu, T., & Wang, P. (2019). Deep learning on point clouds and its application: A survey. In *Sensors (Switzerland)* (Vol. 19, Issue 19). MDPI AG. https://doi.org/10.3390/s19194188

Lloyd, C. T., Sturrock, H. J. W., Leasure, D. R., Jochem, W. C., Lázár, A. N., & Tatem, A. J. (2020). Using GIS and machine learning to classify residential status of urban buildings in low and middle income settings. *Remote Sensing*, *12*(23), 1–20. https://doi.org/10.3390/rs12233847

Park, Y., & Guldmann, J. M. (2019). Creating 3D city models with building footprints and LIDAR point cloud classification: A machine learning approach. *Computers, Environment and Urban Systems*, *75*, 76–89. https://doi.org/10.1016/j.compenvurbsys.2019.01.004

Parmehr, E. G., Amati, M., & Fraser, C. S. (2016). MAPPING URBAN TREE CANOPY COVER USING FUSED AIRBORNE LIDAR AND SATELLITE IMAGERY DATA. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *III–7*, 181–186. https://doi.org/10.5194/isprsannals-iii-7-181-2016

Sarker, S., Sarker, P., Stone, G., Gorman, R., Tavakkoli, A., Bebis, G., & Sattarvand, J. (2024). A comprehensive overview of deep learning techniques for 3D point cloud classification and semantic segmentation. *Machine Vision and Applications*, *35*(4). https://doi.org/10.1007/s00138-024-01543-1

Sharma, M., Garg, R. D., Badenko, V., Fedotov, A., Min, L., & Yao, A. (2021). Potential of airborne LiDAR data for terrain parameters extraction. *Quaternary International*, *575–576*, 317–327. https://doi.org/10.1016/j.quaint.2020.07.039

Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., Yang, K., & Li, J. (n.d.). *Toronto-3D: A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways*. https://www.cloudcompare.org

Urbazaev, M., Thiel, C., Cremer, F., Dubayah, R., Migliavacca, M., Reichstein, M., & Schmullius, C. (2018). Estimation of forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. *Carbon Balance and Management*, *13*(1). https://doi.org/10.1186/s13021-018-0093-5

Van den Broeck, W. A. J., Terryn, L., Cherlet, W., Cooper, Z. T., & Calders, K. (2023). THREE-DIMENSIONAL DEEP LEARNING FOR LEAF-WOOD SEGMENTATION OF TROPICAL TREE POINT CLOUDS. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *48*(1/W2-2023), 765–770. https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-765-2023

Wang, C., Ji, M., Wang, J., Wen, W., Li, T., & Sun, Y. (2019). An improved DBSCAN method for LiDAR data segmentation with automatic Eps estimation. *Sensors (Switzerland)*, *19*(1). https://doi.org/10.3390/s19010172

Wen, C., Yang, L., Li, X., Peng, L., & Chi, T. (2020). Directionally constrained fully convolutional neural network for airborne LiDAR point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 50–62. https://doi.org/10.1016/j.isprsjprs.2020.02.004

Zeng, W., & Gevers, T. (n.d.). *3DContextNet: K-d Tree Guided Hierarchical Learning of Point Clouds Using Local and Global Contextual Cues*.

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object Detection with Deep Learning: A Review. In *IEEE Transactions on Neural Networks and Learning Systems* (Vol. 30, Issue 11, pp. 3212–3232). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/TNNLS.2018.2876865

Zhou, Q., Park, J., & Koltun, V. (2018). Open3d: a modern library for 3d data processing.. https://doi.org/10.48550/arxiv.1801.09847