Full-Waveform Airborne LiDAR Data Classification using Convolutional Neural Networks

Stefano Zorzi, Eleonora Maset, Andrea Fusiello and Fabio Crosilla

***Abstract*—Point-cloud classification is one of the most important and time consuming stages of airborne LiDAR data processing, playing a key role in the generation of cartographic products. This paper describes an innovative algorithm to perform LiDAR point-cloud classification, that relies on Convolutional Neural Networks and takes advantage of full-waveform data registered by modern laser scanners. Our method accurately identifies even challenging classes such as power line and transmission tower.**

***Index Terms*—LiDAR · Full-waveform · Classification · Deep learning · Convolutional Neural Network**

I. Introduction

AIRBORNE laser scanning (ALS) relies on the LiDAR (Light Detection and Ranging) principle, namely to measure the time of flight of a short laser pulse travelling to the target and back, that allows to compute the distance between the sensor and the target. Ranges are then converted to discrete 3D points exploiting GNSS and IMU (Inertial Measurement Unit) data. During its path, the laser ray can be reflected by more than one surface placed at different heights, e.g. part of the laser beam can be reflected from the top of a tree and some part within the tree or the ground surface. The first commercial laser scanners detected only the first and last echo per emitted pulse. Nowadays, most instruments have the ability to record up to six reflections for each emitted pulse and, since 2004, these multi-echo laser scanners have been joined by a new category, the so called *full-waveform* laser scanners, that are finally able to record the entire waveform of the reflected signal. Several studies have shown that these instruments provide a higher spatial point density as well as additional information on the characteristics of the target [1], [2], [3]. In fact, the shape and size of the backscattered waveform is related to the geometry and the reflectance properties of the hit surface.

ALS is currently being employed in a variety of applications, including urban planning, natural hazard management, forestry and facilities monitoring. In almost all the applications, the classification of LiDAR point-cloud is required, being a necessary processing step, e.g., to create Digital Terrain Models (DTMs), to perform analyses on data belonging to particular classes (e.g., to evaluate the vegetation density) and to automatically determine the relationships between different classes (e.g., to calculate the distance between power line conductors and vegetation or buildings).

The aim of this paper is to propose a new classification method for full-waveform airborne LiDAR data using Convolutional Neural Networks (CNNs) and exploiting both full-waveform and spatial information. Thanks to this combination the proposed network is able to distinguish among (e.g.) six classes with an overall accuracy of 92.6%.

II. Related Work

As shown in several studies [4], [5], [6], [7], the LiDAR point-cloud classification process can significantly benefit from the data collected by full-waveform laser scanners. In fact, the waveform registered by these instruments offers the possibility to extract additional features related to the reflectivity characteristics of the target. Over the last years, several classification methods have been proposed in the literature using full-waveform data and the features derived from them [8]. Among these, we mention simple thresholds, both set up manually [9] and automatically [10]. The first method distinguishes between vegetation and non-vegetation points with an accuracy of 89.9% for a dense natural forest and 93.7% for a baroque garden area, while the latter exploits the backscatter coefficient derived from the waveform to classify urban areas into vegetation, roads and building roofs. Amplitude, pulse width and number of pulses are the features used in [4] to perform a binary classification and extract vegetation points via a decision tree. Other methods are based on statistical learning classifiers like Support Vector Machines (SVMs, [11]), which belong to non-parametric methods and perform non-linear classification. This algorithm is well suited for high dimensional problems with limited training set and proved to reach high accuracy (around 95%) when distinguishing between three classes, namely ground, vegetation and building. For urban vegetation detection Höfle et al. [12] use instead geometric and radiometric features that are fed to an artificial neural network classifier consisting of a single hidden layer of neurons and trained by back propagation. Finally, Wang and Glennie [13] apply a ”voxelization” method that divides the waveform data into voxels, merging the ones falling in the same voxel into a synthesized waveform. Features are then extracted and fused with the information derived from hyperspectral images, constituting the input of a SVM that is able to discriminate between 9 classes with an overall accuracy of 92.6%.

All these algorithms rely on hand-crafted features, that are subsequently fed to statistical classifiers or simple machine learning algorithms. An alternative approach is the one

Stefano Zorzi is with the Institute of Computer Graphics and Vision (ICG), Graz University, Inffeldgasse 16, 8010 Graz, Austria. E-mail: stefano.zorzi@icg.tugraz.at

Eleonora Maset, Andrea Fusiello and Fabio Crosilla are with the DPIA, University of Udine, Via Delle Scienze, 206, 33100 Udine, Italy. E-mail: maset.eleonora@spes.uniud.it; (fabio.crosilla; andrea.fusiello)@uniud.it

proposed by Maset et al. [14], that exploits a Kohonen's Self Organizing Maps (SOMs) to perform the unsupervised classification of raw full-waveform data without the need of extracting features from them. The method proved to reach an accuracy of 93.1% over three different classes: grass, trees and road.

In the last years disciplines such as computer vision, speech and audio processing, robotics and bioinformatics have pushed forward and exploited the potential of Deep Learning [15]. Approaches based on hand-engineered features can nowadays be effectively replaced by methods that learn both features and classifier from the data end-to-end. In particular, Convolutional Neural Networks (CNNs) represent the most powerful and reliable tool for classification and segmentation [16], [17].

While many researchers are focused on the development of new architectures for image and video processing, the application of deep learning to LiDAR data – and, notably, to full-waveform data – is still almost unexplored.

In the case of conventional LiDAR data, the recent works of Hu et al. [18] and Yang et al. [19] can be recalled, in which the potential of CNNs for the classification of LiDAR data is demonstrated. More specifically, in [18] a CNN is used to detect ground points, exploiting a point-to-image framework. For each point in the dataset, context information are computed from the neighbouring points in a window and subsequently transformed into an image that is fed to a CNN. In this way, point classification is treated as the binary classification of an image. Similarly, Yang et al. [19] perform a multi-class segmentation of the point-cloud by first transforming the 3D neighbourhood features of a point into a 2D image that is then classified by a CNN. The method reaches an overall accuracy of 82.3% when distinguishing between nine classes, showing however poor performances in the identification of points belonging to small and thin objects such as power line and fences.

Our method is the first that applies CNNs to full-waveform data.

III. Experiments and Results

A. Dataset

Our networks have been trained and validated using a dataset acquired by Helica s.r.l. with a Riegl LMS-Q780 full-waveform airborne laser scanner. The surveyed area contains both natural surfaces such as ground and vegetation, as well as artificial objects such as buildings, power lines and transmission towers.

Three different information are associated to every measured point contained in the dataset, namely the waveform registered by the LiDAR full-waveform sensor, described by a vector of 160 values, the 3D coordinates of the point and the label that shows the class to which the point belongs. These labels have been assigned manually among six classes that were identified: *ground*, *vegetation*, *building*, *power line*, *transmission tower* and *street path*.

The point-cloud is composed by more than 9.8 million points, unevenly distributed over the classes. The dataset is indeed very imbalanced due to the different shape of the scanned objects and the occupied area: e.g., the number of points belonging to vegetation and ground is much higher than the number of points belonging to power line and transmission tower classes. Table I shows in detail the points distribution over the classes.

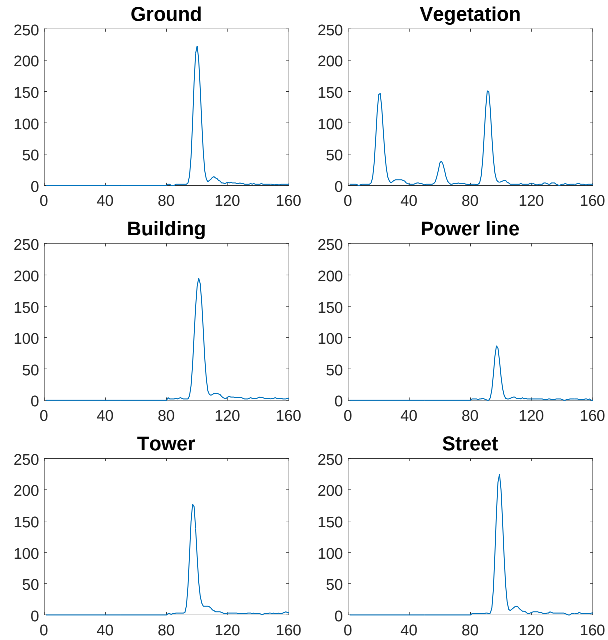


Fig. 1: Waveform samples.

TABLE I: Points distribution over the six classes, divided into training and test sets.

		Training		Test	
Label	**Class**	**# Points**	**%**	**# Points**	**%**
1	*ground*	1787352	20.4	193070	18.1
2	*vegetation*	4719634	53.9	765327	71.7
3	*building*	1514486	17.3	49138	4.6
4	*power line*	71978	0.8	8151	0.8
5	*tower*	32008	0.4	1829	0.2
6	*street path*	633606	7.2	49580	4.6

To handle the entire point-cloud, the dataset is divided into tiles, each containing a different number of points. In the experiments, one tile is used as test dataset (corresponding to approximately 10% of the total number of points), while the remaining tiles are exploited to train the models.

B. Testing

In order to report results that are independent from the training stage, to some extent, five trainings were performed independently, each time reinitializing the weights from scratch and randomly extracting the training dataset from the entire point-cloud. The resulting overall accuracy, computed on the test set, is equal to $92.6(\pm 0.7)\%$, while the average per class accuracy is $87.0(\pm 0.3)\%$.

As can be noticed from the confusion matrix represented in Fig. 2, that reports the results for one out of the five trainings,

	ground	*vegetation*	*building*	*power line*	*tower*	*street*
ground	0.84	0.07	0.00	0.00	0.00	0.09
vegetation	0.03	0.97	0.00	0.00	0.00	0.00
building	0.01	0.06	0.93	0.00	0.00	0.00
power line	0.00	0.05	0.00	0.91	0.04	0.00
tower	0.01	0.07	0.00	0.04	0.88	0.00
street	0.30	0.01	0.00	0.00	0.00	0.69

Fig. 2: Confusion matrix: each row of the matrix represents the instances in an actual class while each column represents the instances in a predicted class. Values are normalized so that the sum of every row is equal to 1.

the network performs very well for the classes *vegetation*, *building*, *power line* and *transmission tower*. Instead, points belonging to the class *street path* are often confused with the class *ground*. This is probably due to the fact that the shape of the waveforms belonging to these two classes are often indistinguishable (see Fig. 1) and also the geometric characteristics of *ground* and *street path* points can be very similar. In practical applications (e.g. for the creation of DTMs) these two classes are usually merged together. If we consider *ground* and *street path* as a unique class, the overall accuracy increases to $96.1(\pm 0.2)\%$ and the average per class accuracy to $92.5(\pm 0.5)\%$.

TABLE II: Synopsis of state-of-the-art methods.

Ref	# classes	Method	Accuracy
Ours	6	CNN	92.6
Ours	5	CNN	96.1
[9]	2	threshold	89.9 - 93.7
[11]	3	SVM	95
[13]	9	SVM	92.6 (+ hyperspectral)
[14]	3	SOM	93.1

Although a direct comparison with other methods using full-waveform LiDAR is not possible, Tab. II suggest that our method compares favourably with the state of the art (the table refers to the methods described in Sec. II).

Examples of the input provided to our network and of the obtained results for the test set are shown in Figs. 3, 4 and 5
.

This confirms that our approach reaches high accuracy in the point-cloud classification thanks to the combination of full-waveform data and spatial support.

IV. Conclusion

In this paper we presented an innovative algorithm based on CNNs to perform full-waveform LiDAR point-cloud classification. The proposed network employs directly the raw full-waveform data, learning both features and classifier end-to-end, unlike other methods that require preliminary extraction of features. It can be applied to the classification of points belonging to any kind of area and no prior knowledge on the data characteristics is required.

Experiments reports an overall accuracy of 92.6%, on six classes including challenging instances such as *power line* and *transmission tower*. Although a direct comparison with other methods using full-waveform LiDAR is not possible, experiments suggest that our method compares favourably with the state of the art.

References

[1] W. Wagner, A. Ullrich, T. Melzer, C. Briese, and K. Kraus, "From single-pulse to full-waveform airborne laser scanners: potential and practical challenges," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 35, pp. 201–206, 2004.

[2] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS journal of Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 100–112, 2006.

[3] C. Mallet and F. Bretar, "Full-waveform topographic lidar: State-of-the-art," *ISPRS Journal of photogrammetry and remote sensing*, vol. 64, no. 1, pp. 1–16, 2009.

[4] V. Ducic, M. Hollaus, A. Ullrich, W. Wagner, and T. Melzer, "3d vegetation mapping and classification using full-waveform laser scanning," in *Proc. Workshop on 3D Remote Sensing in Forestry. EARSeL/ISPRS, Vienna, Austria, 1415 February 2006*, 2006, pp. 211–217.

[5] C. Mallet, F. Bretar, and U. Soergel, "Analysis of full-waveform lidar data for classification of urban areas," *Photogrammetrie Fernerkundung Geoinformation*, vol. 5, pp. 337–349, 2008.

[6] A. L. Neuenschwander, L. A. Magruder, and M. Tyler, "Landcover classification of small-footprint, full-waveform lidar data," *Journal of applied remote sensing*, vol. 3, no. 1, p. 033544, 2009.

[7] J. Reitberger, P. Krzystek, and U. Stilla, "Benefit of airborne full waveform lidar for 3d segmentation and classification of single trees," in *ASPRS 2009 Annual Conference*, 2009, pp. 1–9.

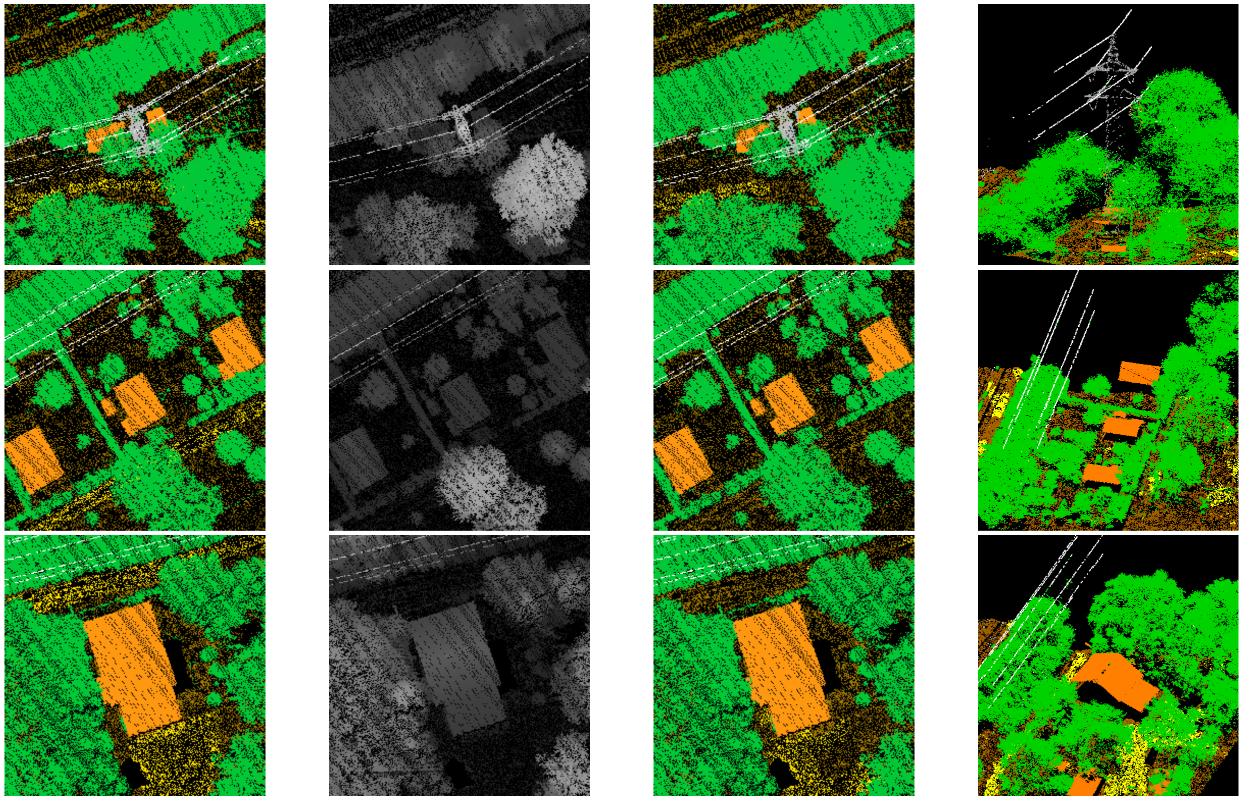
[8] K. D. Fieber, I. J. Davenport, J. M. Ferryman, R. J. Gurney, J. P. Walker, and J. M. Hacker, "Analysis of full-waveform lidar data for classification of an orange orchard scene," *ISPRS journal of photogrammetry and remote sensing*, vol. 82, pp. 63–82, 2013.

[9] W. Wagner, M. Hollaus, C. Briese, and V. Ducic, "3d vegetation mapping using small-footprint full-waveform airborne laser scanners," *International Journal of Remote Sensing*, vol. 29, no. 5, pp. 1433–1452, 2008.

[10] C. Alexander, K. Tansey, J. Kaduk, D. Holland, and N. J. Tate, "Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 5, pp. 423–432, 2010.

[11] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, "Relevance assessment of full-waveform lidar data for urban area classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, pp. S71–S84, 2011.

[12] B. Höfle, M. Hollaus, and J. Hagenauer, "Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne lidar data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 67, pp. 134–147, 2012.



(a) (b) (c) (d)

Fig. 3: Sample images and results (best viewed in colour). (a) Ground truth used for training and validation; (b) Height channel; (c) Labels produced by our network; d) 3D views of the classified point-cloud, coloured with the predicted labels. Classes: *ground* (brown), *vegetation* (green), *building* (orange), *power line* (white), *transmission tower* (grey), *street path* (yellow).

[13] H. Wang and C. Glennie, "Fusion of waveform lidar data and hyperspectral imagery for land cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 108, pp. 1–11, 2015.

[14] E. Maset, R. Carniel, and F. Crosilla, "Unsupervised classification of raw full-waveform airborne lidar data by self organizing maps," in *International Conference on Image Analysis and Processing*. Springer, 2015, pp. 62–72.

[15] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press Cambridge, 2016, vol. 1.

[16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

[18] X. Hu and Y. Yuan, "Deep-learning-based classification for dtm extraction from als point cloud," *Remote Sensing*, vol. 8, no. 9, p. 730, 2016.

[19] Z. Yang, W. Jiang, B. Xu, Q. Zhu, S. Jiang, and W. Huang, "A convolutional neural network-based 3d semantic labeling method for als point clouds," *Remote Sensing*, vol. 9, no. 9, p. 936, 2017.

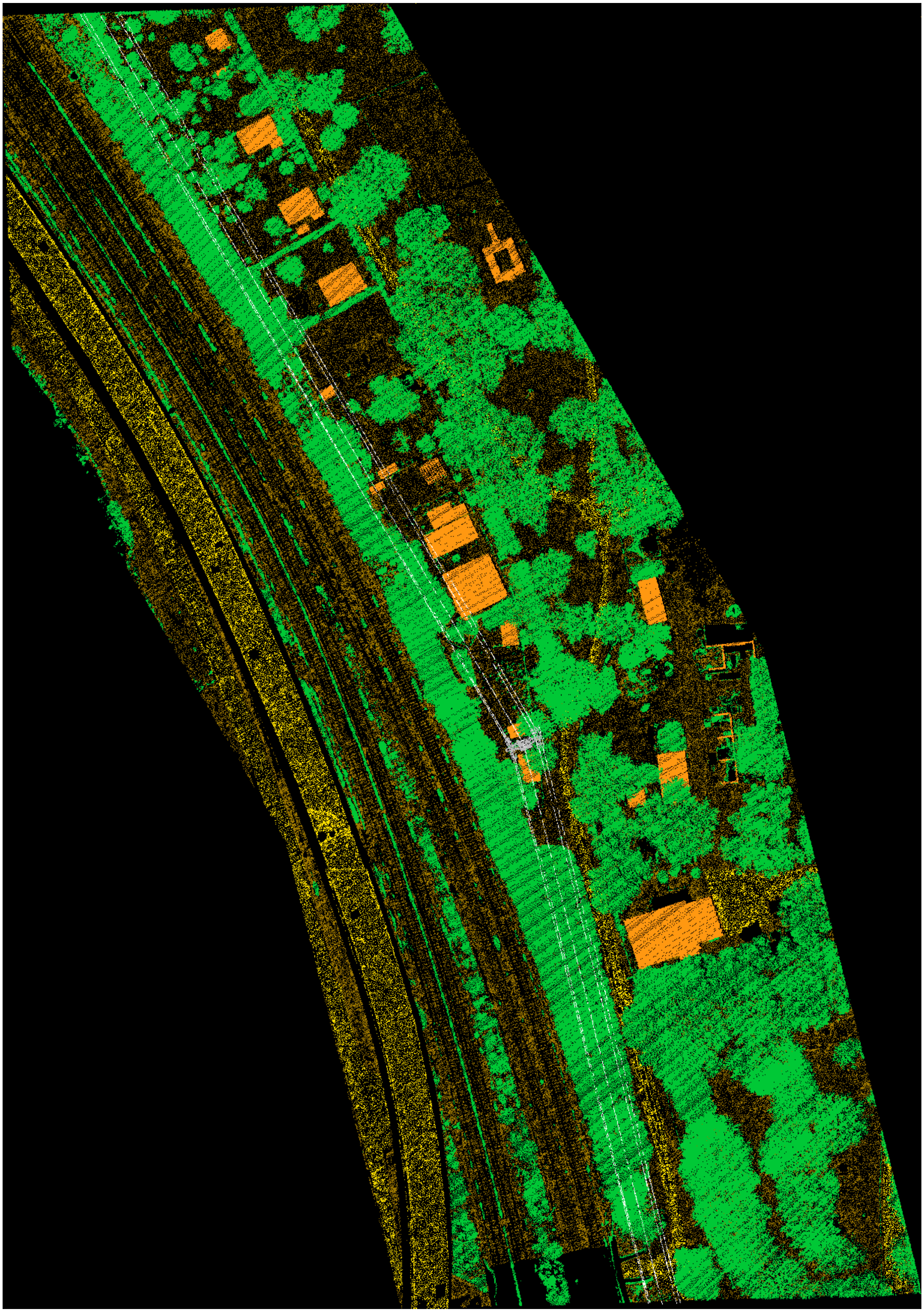


Fig. 4: Ground truth.

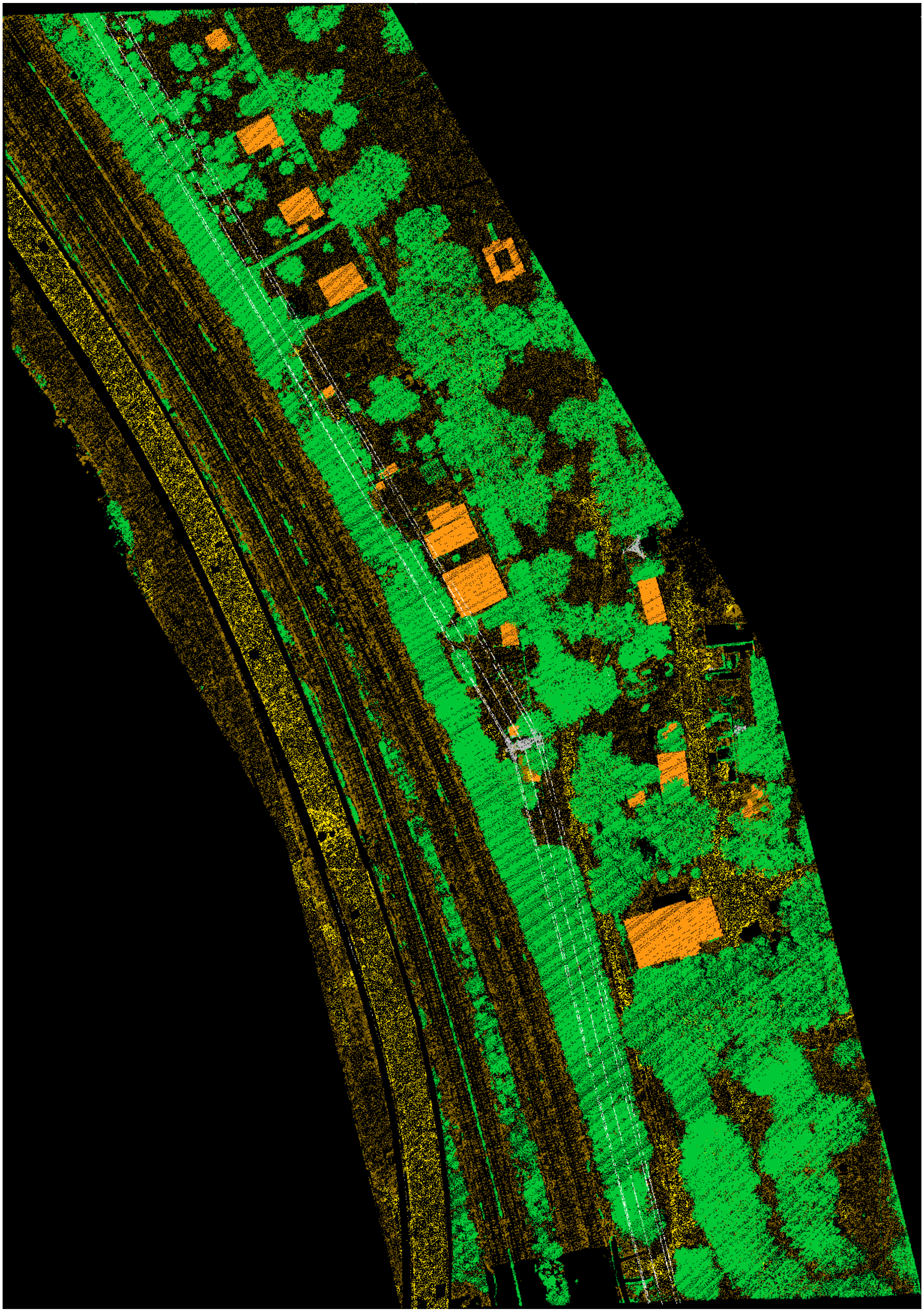


Fig. 5: Point cloud classified by our method.