Machine learning-based segmentation of aerial LiDAR point cloud data on building roof

Emon Kumar Dey, Mohammad Awrangjeb, Fayez Tarsha Kurdi & Bela Stantic

Published online: 11 May 2023.

Submit your article to this journal ⤢

Article views: 1159

View related articles ⤢

View Crossmark data ⤢

Citing articles: 1 View citing articles ⤢

Taylor & Francis
Taylor & Francis Group

Machine learning-based segmentation of aerial LiDAR point cloud data on building roof

Emon Kumar Dey [ID][a], Mohammad Awrangjeb[a], Fayez Tarsha Kurdi[b] and Bela Stantic[a]

[a]School of Information and Communication Technology, Griffith University, Nathan, QLD, Australia; [b]School of Civil Engineering & Surveying, Faculty of Health, Engineering and Sciences, University of Southern Queensland, Springfield, QLD, Australia

ABSTRACT

Three-dimensional (3D) reconstruction of a building can be facilitated by correctly segmenting different feature points (e.g. in the form of boundary, fold edge, and planar points) over the building roof, and then, establishing relationships among the constructed feature lines and planar patches using the segmented points. Present machine learning-based segmentation approaches of Light Detection and Ranging (LiDAR) point cloud data are confined only to different object classes or semantic labelling. In the context of fine-grained feature point classification over the extracted building roof, machine learning approaches have not yet been explored. In this paper, after generating the ground truth data for the extracted building roofs from three different datasets, we apply machine learning methods to segment the roof point cloud based on seven different effective geometric features. The goal is not to semantically enhance the point cloud, but rather to facilitate the application of 3D building reconstruction algorithms, making them easier to use. The calculated F1-scores for each class confirm the competitive performances over the state-of-the-art techniques, which are more than 95% almost in each area of the used datasets.

Introduction

Three-dimensional (3D) building reconstruction from Light Detection and Ranging (LiDAR) point cloud data is an emerging research topic as it has a broad range of applications, such as urban planning, solar potential estimation, building type classification, change detection, virtual tours, and gaming (Sanchez et al., 2020; Tarsha Kurdi & Awrangjeb, 2020; Dey et al. 2020; Y. Yang et al., 2021). LiDAR data consist of three independent parameters *X*, *Y*, and *Z* coordinates along with other retro-reflective properties in the form of intensities describing the topographic profile of any specific earth's surface area and/or objects in that location. Thus, it can provide more accurate geometric information than images which are more suitable to extract-specific features to describe any object accurately. In the case of 3D building reconstruction, properly extracted feature lines constructed from the calculated feature points (e.g. boundary, intersection, and planar points) can facilitate an accurate illustration of the building structure, where the feature lines can be defined as the borders of surfaces and can be categorised into the boundary and fold edge lines (Ni et al., 2016; Y. Zhang et al., 2016). Although there are various definitions of boundary and fold edges in the literature (Mérigot et al., 2010;

Y. Zhang et al., 2016), in the area of 3D building reconstruction, the boundary edge mainly represents the roof contour or facade outline (X. Chen & Yu, 2019), and the fold edge in a building roof is the line that belongs to the intersection of planes (Sampath & Shan, 2009; X. Chen & Yu, 2019). To find the proper feature lines, accurate and precise extraction of the feature points that belong to the boundary or fold area is the main challenge in this case (Dey et al., 2021; X. Chen & Yu, 2019).

Existing feature point extraction can be categorised into indirect and direct approaches. The indirect approaches first convert the point cloud data into 2D images and, then, apply the traditional image processing algorithm to extract the boundary and fold feature lines (Awrangjeb, 2016; Dai et al., 2017; R. Wang et al., 2018). The extracted feature lines are then projected back to 3D to get the corresponding feature points from the input LiDAR data. The direct approaches can be divided into two sub-categories: the segmentation-based approach and the geometric property-based approach. The former sub-category first segments or clusters the point clouds into planes and then extracts the feature outline points for each individual plane (Awrangjeb & Fraser, 2014a, 2014b; Sampath & Shan, 2009). The latter sub-category

considers the geometric properties of individual points such as angle, normal, corner, curvature, and shape to make a decision about the classes: edge, plane, or fold feature point (Ni et al., 2016; Sterri, 2021; X. Chen & Yu, 2019; Xia et al., 2020). Moreover, some authors detected buildings using the photogrammetric point cloud (Acar et al., 2019; Becker et al., 2018; Pamungkas & Suwardi, 2015; Xie et al., 2018; Xu et al., 2018). In these cases, high-density point clouds were generated by processing high-resolution images. Becker et al. (2018) used several geometric and color features for each point to classify the photogrammetric point cloud into different objects. Boundary points of the buildings were separated from the extracted roof planes based on a best-fit geometrical shape-fitting approach (Acar et al., 2019). Xie et al. (2018) used a hierarchical regularisation method to detect the boundary points from the extracted planar building structures. The extracted feature outline points using both direct and indirect approaches are then finally used for the automatic detection and reconstruction of individual buildings (Awrangjeb et al., 2010; Gilani et al., 2016, 2018).

The direct approaches of existing feature point extraction techniques based on the geometric properties are highly dependent on the selection of different parameters (e.g. distance, angle, and direction) and thresholds (Dey et al., 2021). According to literature (Bazazian et al., 2015; Dos Santos et al., 2018; X. Chen & Yu, 2019; Zhao et al., 2019), selecting a proper neighbourhood to estimate the local geometric properties is the major challenge in this case due to the unknown local geometry of the object. Most of the existing approaches use the traditional k or r neighbourhood (also known as k-nearest neighbourhood or nearest neighbourhood within radius r, respectively). Furthermore, different thresholds for the chosen geometric parameters (e.g. angle, curvature, and normal) had to be set empirically previously. For different datasets, the thresholds and parameters may vary due to the abrupt LiDAR point density and the heterogeneous point distribution (Sanchez et al., 2020). Thus, setting the thresholds globally is difficult. The wrong selection of the threshold can produce wrong outputs.

A machine learning-based classification is free from selecting different threshold values. If a system is trained with properly selected attributes of the training data, effective results can be observed on the test data.

Currently, there exist several approaches to classify the point cloud data using machine learning techniques (Gharineiat et al., 2022; Niemeyer et al., 2014; Wen et al., 2020; Y. Yang et al., 2021; Yousefhussien et al., 2018). Both handcrafted feature-based machine learning, and deep learning-based classification approaches can be found in the literature. However, all of these techniques classify the point cloud data into different objects, such as buildings, roads, trees, and ground. Thus, is also known as semantic classification, semantic labelling, or semantic segmentation (Özdemir et al., 2019). We did not find any research in the literature which specifically segments the points over a building roof for the purpose of 3D building reconstruction using machine learning. However, to establish the relationship between the extracted roof planes in the data-driven 3D reconstruction techniques, the classification of the roof point cloud to find the planar points is an essential stage. Moreover, due to the unavailability of properly labelled ground truth data, the segmentation of the extracted building roof feature points using machine learning is still an unexplored research area.

Considering the above issues and to explore the new area of point cloud segmentation over the extracted building roof using machine learning techniques for the purpose of 3D building reconstruction, we select some appropriate feature attributes and then classify the building roof point cloud into three major classes: boundary, fold, and planar points (e.g. see Figure 7). We show the effectiveness of the selected feature attributes using two different traditional machine learning classifiers. Figure 1 shows the basic workflow of this research.

The following are the highlights of the research presented in this paper:

- To segment the building roof point cloud data using machine learning techniques, we calculate and propose some effective feature attributes for each point in the extracted building roof point cloud data.
- Three major classes such as fold points, boundary points, and planar points are segmented using two traditional machine learning classifiers (Support Vector Machine, SVM, and Random Forest, RF). Additionally, a fourth class (vertical roof points) is also considered for some selected datasets (e.g. see Figure 8).
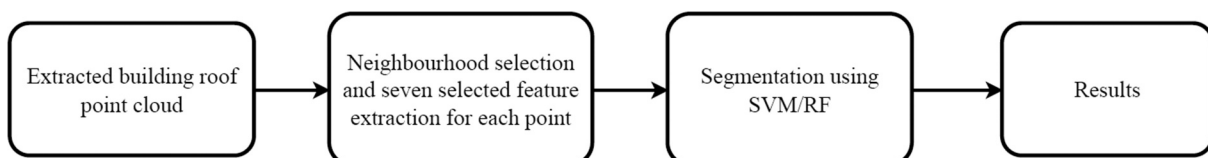

Figure 1. General workflow of the proposed research.

- To train and test the machine learning classifiers, we have manually generated labelled ground truths considering different classes (e.g. fold, boundary, planar, and vertical points) for the selected datasets we have used for the experiments.

The rest of the paper is organised as follows. Section 2 presents a review of the existing approaches to the classification of point cloud regarding building extraction and reconstruction. Proposed selected features attributes for the purpose of classification of roof point cloud along with the classifiers described in Section 3. Section 4 represents the extensive experimental results and discussion. Finally, Section 5 exposes the conclusion.

Review

To describe a building roof using feature lines, three major steps need to be followed: identifying the edges (fold and boundary), tracing the feature points and then generating 3D feature lines from the extracted feature points (Awrangjeb, 2016). The generation of accurate feature lines is highly dependent on the accurate extraction of boundary and fold feature points of a building roof. Xiong et al. (2014) considered a graph edit roof topology to describe a building roof. They used the extracted roof segments and intersecting edge feature lines to describe the graph. The edge feature lines were hypothesised for each pair of extracted nearby roof plane segments. Eigenvalue-based geometric properties (features) derived from a 3D-covariance matrix have been used by many authors to classify the edge and non-edge feature points (Dos Santos et al., 2018; Y. He et al., 2012). Other geometric properties, such as angle, normal, direction, distance, and azimuth distribution, can also be seen in literature to classify the roof points (Ni et al., 2016, 2017; X. Chen & Yu, 2019). Selecting a proper neighbourhood to calculate the geometric properties is important in all these cases (Dey et al., 2021). Although many authors used supervised machine learning, deep learning, and weakly supervised, or unsupervised machine learning to classify the LiDAR point cloud into different object classes, such as buildings, trees, roofs, facades, and roads (J. Zhang et al., 2013; Maltezos et al., 2018; Weinmann, Jutzi, et al., 2015; Y. Chen et al., 2021; Y. Lin et al., 2022; Yousefhussien et al., 2018), we did not find any machine learning-based approach in the literature able to segment the building roof point cloud into different classes such as fold, boundary, planar, or vertical points.

In this section, we first discuss the existing methods for selecting a neighbourhood in the context of features selection, and then we discuss the existing geometric features used for the classification of LiDAR point cloud data into different objects along with the classification of roof point cloud into the edge and non-edge classes.

Neighbourhood for roof feature extraction

To select neighbouring points for the purpose of extracting the feature points along with several geometric features, k-nearest neighbourhood and neighbourhood around radius r (fixed radius method) are two frequently used algorithms in literature (E. He et al., 2017). For example, several authors used the Principal Component Analysis (PCA) to calculate the geometric features by collecting k neighbourhood or neighbouring points around radius r (Rutzinger, Rottensteiner, et al., 2009; Y. He et al., 2012; Z. Wang & Prisacariu, 2020). But the performance of these methods degraded when the densities in a point cloud data varied. Moreover, selecting an appropriate value for k or r was challenging since, a smaller value of k was sensitive to the outliers, whereas, a larger value could over-smooth the sharp feature points (Ben-Shabat et al., 2019; Dey et al., 2021).

To avoid these problems, adaptive neighbourhood selection approaches had been used by many authors (E. He et al., 2017; Weinmann, Jutzi, et al., 2015; Y. He et al., 2012). Y. He et al. (2012) proposed an adaptive search range approach to consider only a limited number of points among the initially selected large k number of neighbours for point P_i. To ensure the uniformity of neighbourhood distribution and optimal search range, they adaptively calculated a fixed distance r for each point P_i, and then considered only those points as neighbours of point P_i with a distance less than r. E. He et al. (2017) used different adaptive values of k and r considering scattered and regular regions in the input point cloud. Based on the curvature value of each point they found the scattered and regular region areas.

Multiscale neighbourhood selection approaches had been used by some authors for the purpose of LiDAR point classification (Leichter et al., 2020; Weinmann et al., 2013; Weinmann, Schmidt, et al., 2015). Different scales of k and r values were selected to improve the classification accuracy in this case. For varying values of k ($k = 10$ to 100), entropy values (e.g. Shannon entropy) were calculated by various authors and value yielding minimal entropy was selected to define the optimal neighbourhood for individual points (Weinmann et al., 2013). However, it costs a high computational complexity because of different neighbouring points for each point. It also suffered from the Hughes Phenomenon, where the growing feature space

dimensionality decreased the classification accuracy (Pauly et al., 2002).

Existing features for roof point classification

As mentioned earlier, a point cloud mainly contains X, Y, and Z coordinates values for each point. Thus, a calculated covariance matrix contains three rows and three columns, and can be calculated using Eq. 1. Geometric features, based on a different combination of the calculated eigenvalues ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) and eigenvectors from the covariance matrix $\text{Cov}[P, P]$ had been widely used to classify the LiDAR point cloud in both rule-based and machine learning-based approaches (Becker et al., 2018; Belton & Lichti, 2006; Nurunnabi et al., 2015; Sampath & Shan, 2009; Xia & Wang, 2017).

$$\text{Cov}(P, P) = \frac{1}{k} \sum_{i=1}^{k} [P_i - \mu(P)][P_i - \mu(P)]^T, \quad (1)$$

where P_i is any point among k neighbours of P and $\mu(P)$ is the mean vector of its neighbours.

Belton and Lichti (2006) used the variance of curvature in a local neighbourhood using the calculated eigenvalues. They considered the corresponding eigenvectors of $\text{Cov}(P, P)$ as directions and the eigenvalues as the variance in the directions of the corresponding eigenvectors, respectively. Dos Santos et al. (2018) classified the edge and non-edge LiDAR points using different groups of measurement calculated based on the eigenvalues and eigenvectors. Points with one or two large eigenvalues among the calculated three eigenvalues are considered the edge candidates by Xia and Wang (2017). To define the threshold for large eigenvalues, they used a ratio between eigenvalues used by Lowe (2004). Azimuth angles, the direction of normal, and angular gap were also used as important features to extract the boundary and fold edge feature points by several authors (Gumhold et al., 2001; Ni et al., 2016; X. Chen & Yu, 2019). Delaunay triangulation-based approaches were used by many authors to separate the building boundary points (Awrangjeb, 2016; Boulaassal et al., 2009). For example, Awrangjeb (2016) claimed that triangles along the periphery have one side which is associated with only one triangle. Considering this fact, they divided that point cloud of a building roof into the boundary and non-boundary points. Convex hull-based approaches were used by several authors to detect the boundary points (J. Wang & Shan, 2009; Sampath & Shan, 2009).

In the supervised machine learning approach, a label is assigned to each point in the input point cloud (Becker et al., 2018). Existing literature used this approach in point cloud data to classify different objects, such as tree, building, road, car, grass, and other man-made infrastructure (Bassier et al., 2019; Becker et al., 2018; Hackel et al., 2016; Niemeyer et al., 2014; Park & Guldmann, 2019; Serna & Marcotegui, 2014; Z. Li et al., 2016). It requires some labelled data to train a classification model. The supervised learning model learns from the training data, and can predict a new unseen point. Different geometric features along with the raw point cloud of different objects can be used to train the classifiers. In the context of calculated features, we observed three major categories in the machine learning-based approach of point cloud data classification (Wen et al., 2020). The first category was the point feature-based classification, where local geometric features of each point were extracted, and a conventional machine learning classifier was used for the classification purpose (Chehata et al., 2009; J. Zhang et al., 2013; Niemeyer et al., 2014; Weinmann, Jutzi, et al., 2015). Eigenvalue-based features along with some additional features such as: point density, intensity, number of return, standard deviation, and variance of normal vector were commonly used in this case. For example, Niemeyer et al. (2014) considered seven different classes, and calculated features for each point considering a sphere of radius r. To improve the performance and to generate reliable eigen-features, C. H. Lin et al. (2014) analysed the local geometric characteristics using a weighted covariance matrix with a geometric median. Hackel et al. (2016) introduced 17 different features to classify 6 different semantic classes based on covariance, moments, height, and color of each point.

The second category was the context feature-based classification, which introduced the multi-level contextual information of the point cloud (Niemeyer et al., 2014). However, it failed to detect both large and small objects due to over-smoothing problems, thus, leading to an incorrect classification result (Zhao et al., 2018).

The third category was the deep learning-based classification approaches which could be sub-divided into feature image-based classification and direct point cloud-based classification. The feature image-based approaches firstly converted the point cloud into feature images, and then applied a convolutional neural network (CNN) to classify the objects (Z. Yang et al., 2018; Zhao et al., 2018). Considering the unordered and unstructured nature of the point cloud, the second sub-category directly applied deep learning frameworks to the unstructured data (Pohle-Fröhlich et al., 2019; Qi, Su, et al., 2017; Qi, Yi, et al., 2017; X. Li et al., 2018). PointNet architecture proposed by Qi, Su, et al. (2017) was the very first method in this category.

Feature line extraction from 3D point cloud was a sub-step of a 3D building modelling and has been a major research area for years (Ni et al., 2016). Apart from machine learning, a variety of techniques have been used to extract feature points and/or feature lines for 3D building modelling. However, the machine

learning approaches are confined to the semantic classification of LiDAR point cloud data as discussed above in this section.

In this paper, to describe the point cloud of an extracted building roof using machine learning techniques, we mainly propose seven effective machine learning features to segment three major feature points over the extracted building roof. We have used a variable point neighbourhood selection method to solve the problems of the existing fixed number of neighbouring point selection techniques. In the next section, we describe our proposed machine learning approaches for the purpose of feature point classification over a building roof.

Methodology

To segment the points over the building roof, non-building and ground points were initially separated, and building roofs were extracted using our previously developed methods (Dey & Awrangjeb, 2020; Dey et al., 2020, 2021). The separated building roof point clouds were evaluated by our robust performance evaluation metric (Dey & Awrangjeb, 2020). In this research, we considered the specific scanline pattern of aerial point cloud data over a building roof, and selected an appropriate neighbourhood for each point using our recently proposed neighbourhood selection technique (Dey et al., 2021). After that, a minimal number of geometric features to classify the point cloud over a building roof were calculated. We chose SVM and RF classifiers as representatives of conventional machine learning classifiers because of their reliable performance and extensive adoption in various applications of point cloud data classification (Liu et al., 2018).

Study sites and ground truth generation

We used three different datasets containing six different sites with different point densities and building structures to evaluate the proposed machine learning approaches. The first datasets were the high-density (12 to 40 points/m^2) Australian datasets containing three different sites (Awrangjeb & Fraser 2014a). The first (AV1) and second (AV2) sites contained 5 and 63 different residential buildings from the Aitkenvale area, respectively. The densities of these first two sites vary between 29 and 40 points/m^2. The third site (AV3) had 28 different buildings from the Hervey Bay (HB) area with a density of 12 points/m^2. The next two datasets were from the ISPRS benchmark datasets which included the buildings from Vaihingen (Germany) and Toronto (Canada). Vaihingen area contained residential buildings, historical buildings, and small-detached houses (Cramer, 2010). It had a point density of 2.5 to 3.9 points/m^2, and a total of 107 buildings larger than 2.5 m^2. The Toronto datasets contained buildings from a modern megacity in Canada with a point density of 6 to 7 points/m^2) (Cramer, 2010). It included both low- and high-story buildings with a variety of roof structures. The last datasets (Hermanni) contained large multi-story residential buildings from the Helsinki area (Finland), and belonged to the building extraction project of EuroSDR (Tarsha Kurdi et al., 2021). The point densities of these datasets were between 7 and 9 points/m^2). Figure 2 shows six different sites of our used datasets.

Our main target was to segment the building roof point cloud points into three major classes: such as planar, boundary, and fold points. To train and test the machine learning classifiers, we manually labelled the point cloud of the extracted building roofs into the planar, boundary, and fold classes for each datasets. However, most of the building roofs in the Toronto datasets were flat and did not have fold or intersection edges. Moreover, almost each building roof in this dataset contained several vertical planes as there were different planar parts of the roof on different height levels.

In this case, we considered vertical planar points instead of fold point for the Toronto buildings. Due to a large number of buildings in the AV2, Vaihingen, and Toronto sites, we chose and labelled some selected complex buildings (e.g. 25 from AV2, 30 from Vaihingen, and 30 from Toronto) from these sites, and all buildings from the AV1, HB, and Hermanni sites for generating the ground truth data. It was hard to decide the label of each point manually. To label the fold edge points of a roof, we considered the point density of the specific site and kept the points within a specific maximum distance T_f from the intersection of two different planes. T_f was calculated using equation Eq. 2 following the method used by Tarsha-Kurdi et al. (2006), where ϑ represents the point density, because, if we assume a regularly distributed point cloud data, the mean area occupied by a single LiDAR point is in a square form, and the area of the square is equal to the inverse of the point density. We can consider the side length of the square as the mean distance between two neighbouring points which satisfies Eq. 2.

$$T_f = \frac{1}{\sqrt{\vartheta}} \qquad (2)$$

The approximate distribution of the generated planar, boundary, and fold labelled points are 73%, 18%, and 9%, respectively, for the Vaihingen, Aitkenvale, and Hervey Bay areas, whereas the Toronto datasets have 77% planar, 16% boundary, and 7% vertical
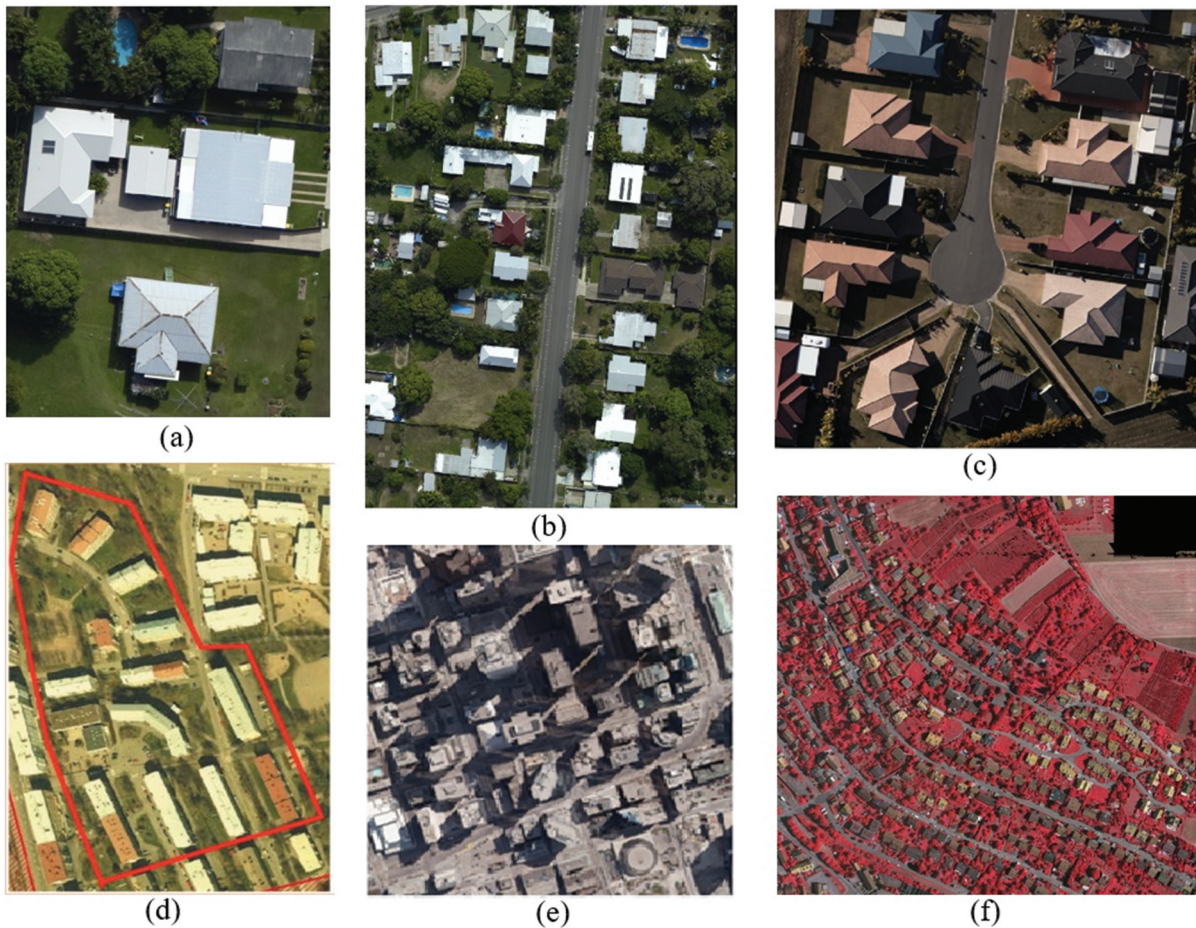
Figure 2. Datasets used in this research. (a) and (b) are two different sites from the Aitkenvale area. (c) Hervey bay area, (d) Hermanni datasets, (e) Toronto and (f) Vaihingen area from ISPRS datasets.

points. The Harmanni datasets contain 74% planar points, 10% boundary points, 12% fold points, and 4% vertical points of the generated labelled data.

Neighbourhood selection

Calculating normal vectors to find the accurate features for segmentation were necessary in our method. An inappropriate neighbourhood of a point could estimate a wrong normal direction. Instead of a fixed number of neighbouring points (k or r neighbourhood), we selected a variable point minimum neighbourhood, which led to minimise the error during the estimation of normal and other roof features for classification in our method.

The approach we selected for neighbourhood calculation was introduced by Dey et al. (2021), which considered the scanline property of aerial building point cloud data. An initial minimal number of neighbouring points (e.g. 3) were selected for each point using the k-NN algorithm. Using the selected neighbourhood and the point itself a 3D line was fitted. The standard deviation of the distance from each point of the neighbourhood to the 3D line was calculated. The value of standard deviation indicated the number of scanlines

that included the selected neighbouring points. If the value represented only one scanline, then the number of neighbouring points (k value) were increased iteratively until two or more different scanlines were observed. Because, neighbouring points selected from at least two different scanlines guarantee the accurate normal of a point in the context of the aerial roof point cloud. Figure 3 describes the scenario. The neighbouring points of P_3 were selected using the k-NN algorithm. Due to the small value of k, the neighbouring points were selected from the same scanline, thus P_3 might offer an unstable normal estimation. Using a comparatively large k value, this problem could be avoided for P_2. However, P_1 and P_4 could attract neighbouring points from other planes or objects. Thus, selecting a higher value of k could also produce a faulty normal estimation. The neighbourhood selection method we chose (Dey et al., 2021) in this paper avoided this issue by avoiding a higher number of neighbourhood from multiple scanlines. Each point chose a variable number of neighbours instead of a fixed value. In addition, the algorithm also solved the problem of selecting neighbouring points in the situation of an abrupt density variation over a roof point cloud, which was common in aerial point cloud data.
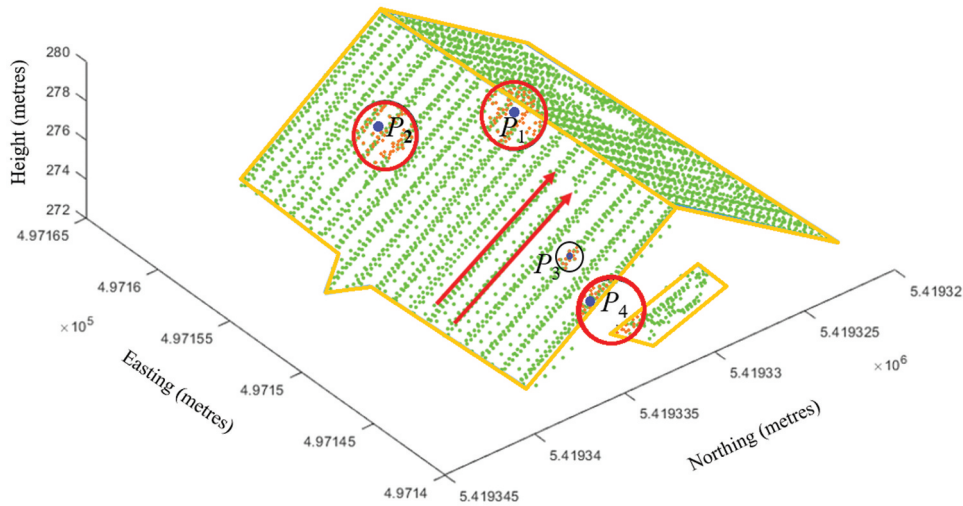
Figure 3. LiDAR points over a building roof with scanning direction (red arrows).

Selected features for machine learning

We considered seven different features based on azimuth angle, direction of the normal, distances between the points, curvature value, and eigenvalues of points to classify the roof point cloud. Below we detail these features.

Maximum Azimuths (M_τ): Considering boundary and non-boundary points have distinguishable azimuth angles, we selected the maximum of azimuth angle differences of the projected neighbouring points of a point P_i as the first feature for classification. We first estimated the neighbourhood (N_p) of each point P_i using the method of Dey et al. (2021) described in Section 3.2. After that, the azimuth angle τ_j for each point within N_p was calculated according to X. Chen and Yu (2019). In this approach, the normal vector of P_i was calculated using the weighted principal component analysis (WPCA) algorithm (Cochran & Horne, 1977), and the selected neighbouring points (N_p) were projected onto a 2D projected plane. For each projected neighbour P_j, P_i was set as an origin of a 2D coordinate system. The X-axis was formed by

extending a line segment from P_i to P_j. The Y-axis was formed by following the right-hand rule among the X-axis and the normal vector of P_i. The azimuth (τ_j) of each point P_j was calculated using the following Eq. 3, where, y_j and x_j were calculated 2D coordinates of P_j. After calculating the differences of all adjacent azimuth angles using Eq. 4, max ($\Delta\tau_j$) was taken as a feature for P_i.

Figure 4 shows the azimuth angles for non-boundary and boundary points. It is clearly visible that the boundary and non-boundary points have different maximum azimuth angles among their adjacent neighbouring points. Thus, we considered the maximum difference among the azimuth angles denoted as M_τ as a feature to classify the roof point cloud.

$$\tau_j = arctan\left(\frac{y_j}{x_j}\right), \quad (3)$$

$$\Delta\tau_j = \tau_j - \tau_{j-1}, \quad (4)$$

Maximum Normal Angle (θ_{max}): For any point P_i in a roof point cloud, angle differences between its
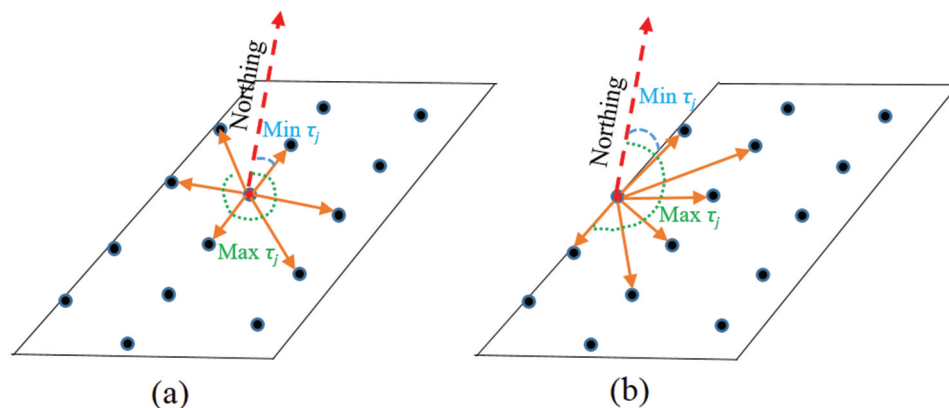

Figure 4. Azimuths (τ_j) of vectors on the 2D plane. (a) Azimuths for inside planar point, (b) Azimuths for boundary edge point.

normal and the normals of its selected neighbouring points were calculated. After estimating the neighbourhood (N_p) of each point P_i, the maximum difference θ_{max} among the normal angles was taken as the second feature for classification. Figure 5 demonstrates this feature clearly. For a point P_i in a fold edge or intersecting roof planes, the normal of its selected neighbouring points are distributed mainly into two different directions. Contrariwise, if a point belongs to a planar part, the normal directions of its neighbouring points are almost similar. Thus, if we consider the angle differences of the normals between P_i and its selected neighbouring points, and take the maximum difference value θ_{max}, a fold edge point will have a much larger value than an inside planar point.

Vertical angle (V_θ): Angle V_θ between the Z-axis and the direction of the calculated normal for each point was considered as another feature. The direction of the normal of any point P_i was calculated based on the selected neighbouring points and the WPCA algorithm (Cochran & Horne, 1977). A fold or planar roof point have smaller V_θ; however, points in a vertical plane have a larger value of V_θ. This is an important feature for classifying the vertically planar points on building roofs.

Distance (d_m): Let the set of neighbouring points including P_i be N_p. In practice, for a regularly distributed point cloud, the calculated mean point of N_p will be very close to the inner point P_i. However, if P_i is a boundary point then the mean will be away from P_i. Let the mean be $\bar{M}$. Euclidean distance d_m from P_i to $\bar{M}$ was calculated and considered as the third feature for each point in the input point cloud. Figure 6 demonstrates the distance feature d_m. Pink points in the magnified area represent the mean of the selected neighbours N_p of any point P_i.

Curvature (κ_f): Curvature of any point is the amount of deviation from being a straight line while it is a part of a curve, or it is the amount of deviation from being a plane. Thus, it is an effective feature of point cloud classification. Once the neighbouring points were determined for each point P_i, we calculated
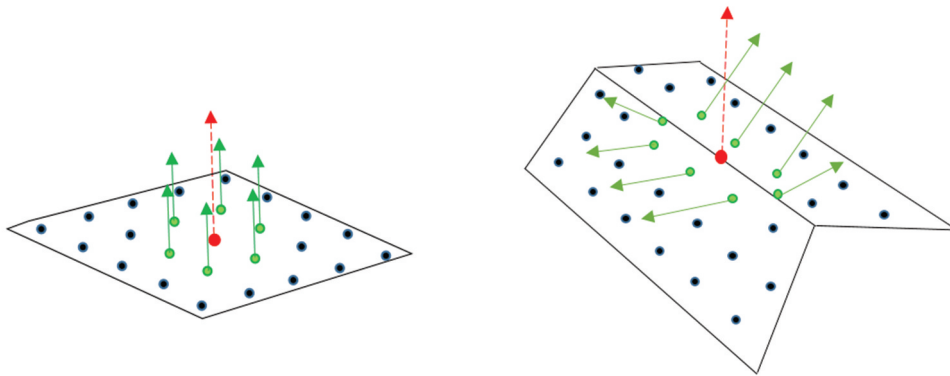

Figure 5. Direction of normals. Green points indicate the selected neighbours of red point P_i. (a) Direction of normal of a planar surface, (b) Direction of normals in a gable roof.
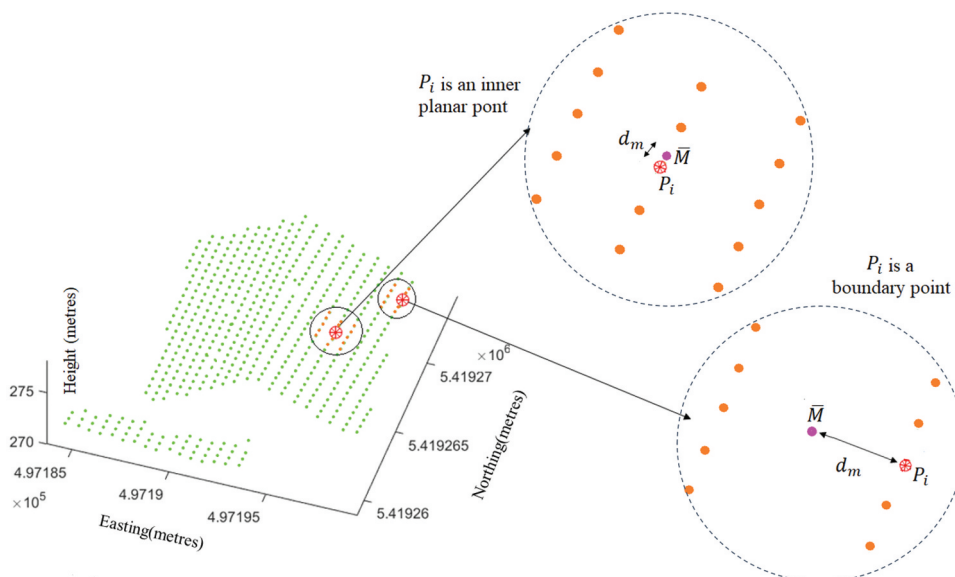

Figure 6. Distance from the mean of neighbouring points $\bar{M}$ to any point P_i.

red dots = boundary points · blue dots = fold points · cyan dots = planar points
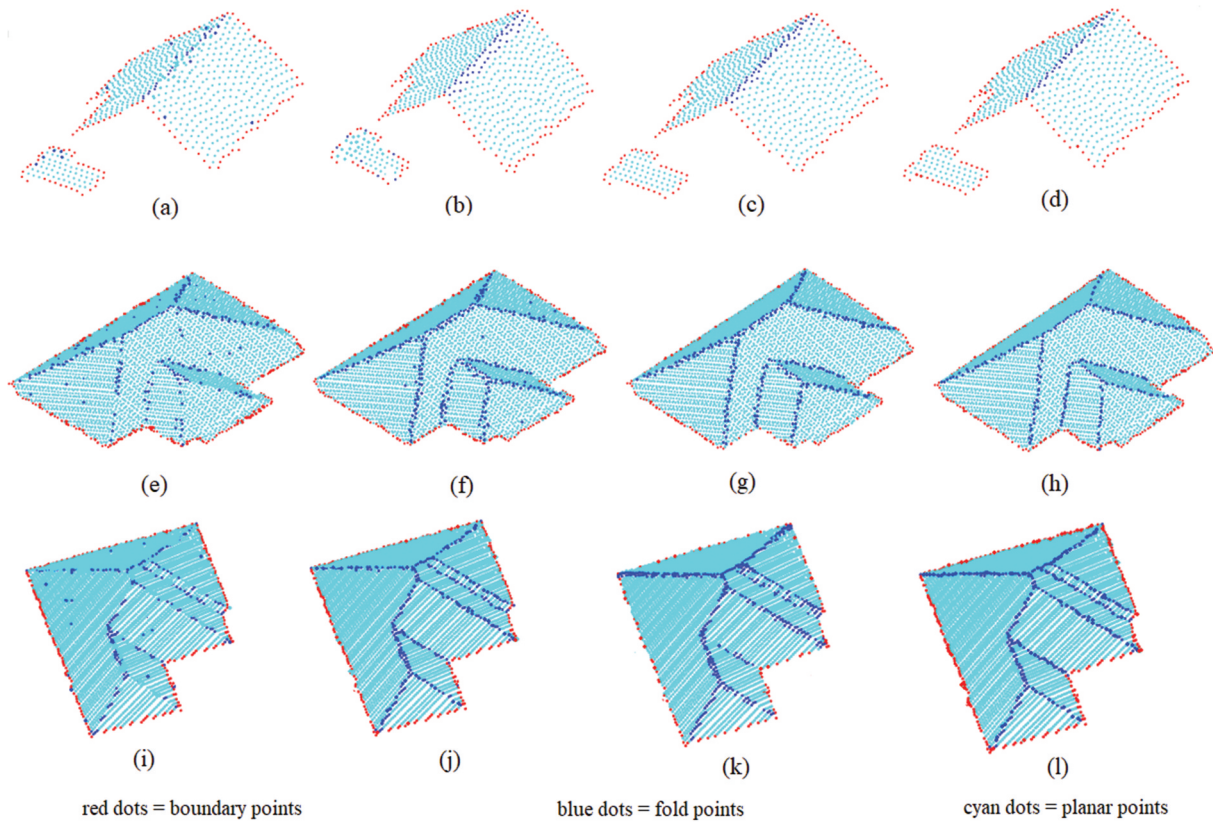
Figure 7. Qualitative performances of different methods. The first, second, and third rows indicate three representative building roofs from Vaihingen, Aitkenvale, and Hervey Bay areas, respectively. The first and second columns represent the extraction results using the methods proposed by X. Chen and Yu (2019) and Dey et al. (2021), respectively. The third and fourth columns represent the qualitative performances of the proposed approaches using the SVM and RF, respectively.

red dot = boundary point
blue dot = fold point
cyan dot = planar point
black cross = vertical point
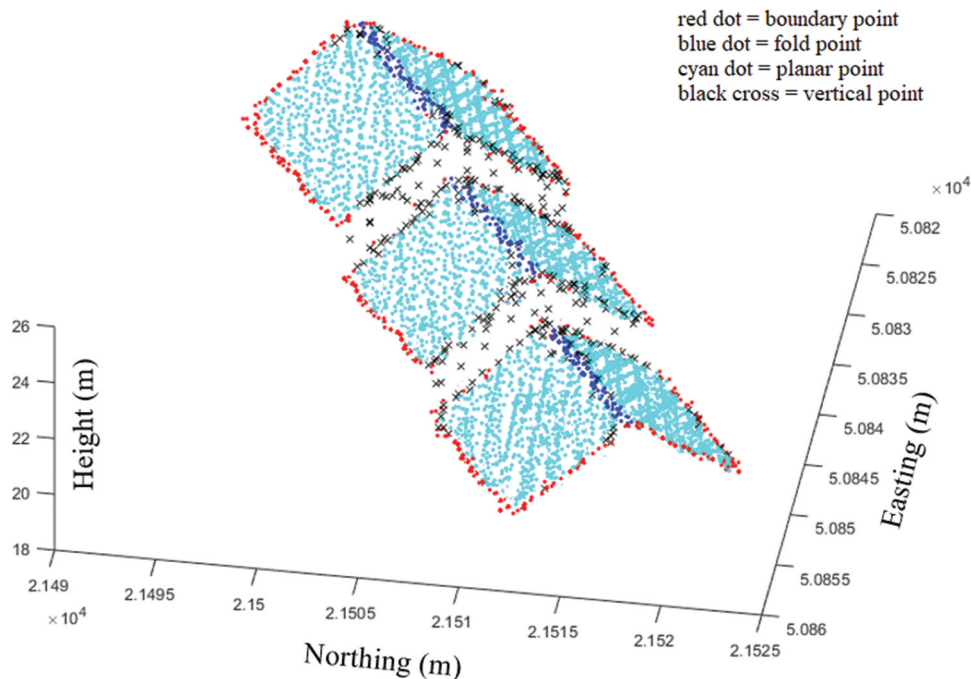
Figure 8. A sample building from Hermanni datasets with four classes of points classified using RF. Black crosses represent the vertical planar points. Red, blue and cyan dots represent the classified boundary, fold and planar points, respectively.

the covariance matrix using Eq. 1. This matrix showed how neighbourhoods of points locally disperse from their centre of gravity. Corresponding eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ were calculated from the covariance matrix,

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ (Weinmann, Jutzi, et al., 2015) and λ_3 represented the direction of the least dispersion. Thus, we calculated the change of curvature factor for each point based on the calculated

eigenvalues using Eq. 5 (Thomas et al., 2018; Weinmann, Jutzi, et al., 2015).

$$\kappa_f = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \qquad (5)$$

Linearity and Planarity: Linearity ($\mathcal{L}$) and planarity ($\mathcal{P}$) of a point were frequently used features for point cloud classification and calculated using Eqs. 6 and 7, respectively (Thomas et al., 2018). The calculated value of each of these two features was a number between 0 and 1, where a higher value indicated the higher linearity or planarity and vice versa. The highest possible measure of linearity corresponds to a perfectly linear shape (i.e. points belonging to a straight boundary line) and the highest possible measure of planarity corresponds to a perfectly planar shape (i.e. points belonging to an inside roof plane).

$$\mathcal{L} = \frac{\lambda_\infty - \lambda_\in}{\lambda_\infty}. \qquad (6)$$

$$\mathcal{P} = \frac{\lambda_2 - \lambda_3}{\lambda_1}. \qquad (7)$$

Classifiers

Using these above-selected features, we trained and tested our datasets using machine learning classifiers. Random Forest (RF) and Support Vector Machine (SVM) were selected as representatives of conventional classifiers because of their feasible and comprehensive adoption in the field of point cloud classification as mentioned earlier.

Random forest

The Random Forest (RF) is a supervised ensemble classifier (Breiman, 2001; Park & Guldmann, 2019). It grows multiple decision trees. Each individual tree in the RF can predict a class for each individual point. The calculated selected features ($M_\tau, \theta_{\max}, V_\theta, d_m, \kappa_f, \mathcal{L}, \mathcal{P}$) for each point were given as input to the RF classifier. The most popular class with a majority of votes became the final predicted class of each individual point in the input point cloud. We adopted the iterative random sampling to avoid the over and under-representation of certain classes (Belgiu & Drăguț, 2016). The manually generated ground truths (see Section 3.1) were randomly divided into two sets: training and testing. We use 80% of labelled data from each class as training and the rest of the labelled data as testing. Most of the existing semantic point cloud classification used 10 to 20 times random partitioning approaches and then took the average to find the best classification results (Park & Guldmann, 2019). We initially used 5, 10, 15 and 20 times random partitioning for our datasets and finally chose 10 as in most of the cases we found the best classification results for 10-time random partitioning. We used MATLAB 2020 to implement and used the RF classifier.

Support vector machine

The Support Vector Machine (SVM) tries to find a hyperplane in high-dimensional feature space to solve some linearly inseparable problems. It had been widely used to classify point cloud objects, such as buildings, roads and trees (Karsli et al., 2016; Lodha et al., 2006). To classify the roof feature points for the purpose of 3D building reconstruction, we used the selected features ($M_\tau, \theta_{\max}, V_\theta, d_m, \kappa_f, \mathcal{L}, \mathcal{P}$) along with the coordinates of raw point cloud to train the classifier. The modified version of LIBSVM (Chang & Lin, 2011) was used to test the performance of the probabilistic multiclass extension of the SVM classifier on our data. To avoid the over- and under-representation of certain classes, the generated training and testing sets (see Section 3.1) were randomly divided into 80% and 20%, respectively. After iterating the random process for 10 times, we took the average. To select the best iteration number of the random process, like RF, we initially tried with 5, 10, 15, and 20 random iterations and finally chose 10 because of getting the best results for our datasets.

Results and discussion

In this section, we present the extensive experimental results of the point cloud segmentation over the building roof using conventional machine learning techniques based on the selected features extracted in Section 3.

Using both SVM and RF, we tested our datasets. Moreover, we also considered state-of-the-art roof feature point extraction techniques proposed by Dey et al. (2021) and X. Chen and Yu (2019) to compare the performances. Both quantitative and qualitative performances were evaluated over the datasets. The datasets we used in this research are not balanced (see Section 3.1), hence, we abstain from the simple accuracy measure to avoid the accuracy paradox. In Tables 1-3, the quantitative classification results considering three different classes (boundary, fold, and planar points) in terms of precision, recall, and F1-scores for Vaihingen, Aikenvale, and Hervey Bay areas, respectively, are presented. To demonstrate the qualitative performances of four different methods, three sample buildings from the Vaihingen, Aitkenvale, and Hervey Bay areas are selected, respectively. The results of the methods are demonstrated in Figure 7, where, second (Figure 7b-j), third (Figure 7c-k), and fourth ((Figure 7d-i) columns represent corresponding results of Dey et al. (2021), proposed SVM, and RF classifiers, respectively. The methods proposed by X. Chen and Yu (2019) did not extract the planar point separately. They only considered boundary and

Table 1. Summary of classification using different methods for the Vaihingen area of ISPRS benchmark datasets. Results are indicated by the mean values of 10 different runs with their standard deviations.

Method	Class	Precision	Recall	F1-score
X. Chen and Yu (2019)	Boundary	0.84 ± 0.01	0.72 ± 0.02	0.78 ± 0.02
	Fold	0.74 ± 0.01	0.79 ± 0.01	0.77 ± 0.01
	Planar	-	-	-
Dey et al. (2021)	Boundary	0.83 ± 0.01	0.82 ± 0.01	0.83 ± 0.01
	Fold	0.79 ± 0.01	0.87 ± 0.01	0.83 ± 0.01
	Planar	0.91 ± 0.01	0.92 ± 0.01	0.92 ± 0.01
Proposed (SVM)	Boundary	0.92 ± 0.02	0.94 ± 0.01	0.93 ± 0.02
	Fold	0.91 ± 0.01	0.92 ± 0.01	0.91 ± 0.01
	Planar	0.95 ± 0.01	0.96 ± 0.02	0.96 ± 0.02
Proposed (RF)	Boundary	0.95 ± 0.01	0.97 ± 0.01	0.96 ± 0.01
	Fold	0.99 ± 0.01	0.96 ± 0.01	0.97 ± 0.01
	Planar	0.96 ± 0.03	0.96 ± 0.01	0.96 ± 0.02

fold points in the input data. Thus, in Figure 7a-i, we only extracted boundary (red) and fold (blue) points using their method and presented the rest of the unlabelled points using cyan color to keep consistency. We implemented the boundary and fold point extraction methods of X. Chen and Yu (2019) using Matlab 2020 platform. It is noticeable from Tables 1 and 2 , and Table 3 that machine learning approaches performed relevantly better for these datasets using the proposed selected features.

The Toronto dataset was mainly from an urban city area, the roofs of the buildings were flat, and no buildings with intersecting roof planes (e.g. gable, cross gable or hip shape roof, see Figure 2(e)). However, almost every building in this dataset contained multiple planar parts on different levels, which introduced one or more vertically planar parts in each building

roof. Also, in Hermanni datasets we noticed similar vertical planes in some building roofs. Due to the direction of the aircraft, some points could be captured from the vertical planes in a building roof point cloud data. In these cases, we considered a separate class "Vertical points" instead of "Fold point" for the Toronto datasets as there were several vertical planes on almost every roof. For the Hermanni datasets we considered four classes (fold, boundary, vertical, and planar) as some buildings contain vertical planes. We trained and tested using the SVM and RF according to the new class for the corresponding datasets. Figure 8 shows the classification results of a sample building containing four different classes from Hermanni datasets. The black cross indicates the classified vertical planar points. Red, blue, and cyan dots indicate the boundary, fold, and planar roof points, respectively.

Table 2. Summary of classification using different methods for Aitkenvale areas of Australian datasets. Results are indicated by the mean values of 10 different runs ± standard deviations.

Method	Class	Precision	Recall	F1-score
X. Chen and Yu (2019)	Boundary	0.86 ± 0.01	0.75 ± 0.01	0.80 ± 0.01
	Fold	0.75 ± 0.01	0.73 ± 0.01	0.74 ± 0.01
	Planar	-	-	-
Dey et al. (2021)	Boundary	0.90 ± 0.02	0.85 ± 0.01	0.87 ± 0.01
	Fold	0.84 ± 0.01	0.85 ± 0.01	0.84 ± 0.01
	Planar	0.93 ± 0.02	0.91 ± 0.02	0.92 ± 0.02
Proposed (SVM)	Boundary	0.83 ± 0.01	0.92 ± 0.01	0.87 ± 0.01
	Fold	0.82 ± 0.01	0.83 ± 0.02	0.83 ± 0.01
	Planar	0.98 ± 0.01	0.98 ± 0.01	0.98 ± 0.01
Proposed (RF)	Boundary	0.91 ± 0.02	0.94 ± 0.02	0.93 ± 0.02
	Fold	0.88 ± 0.02	0.86 ± 0.01	0.87 ± 0.01
	Planar	0.98 ± 0.01	0.99 ± 0.01	0.98 ± 0.01

Table 3. Summary of classification of different methods for the Hervey Bay area of Australian datasets. Results are indicated by the mean values of 10 different runs ± standard deviations.

Method	Class	Precision	Recall	F1-score
X. Chen and Yu (2019)	Boundary	0.81 ± 0.01	0.79 ± 0.01	0.79 ± 0.01
	Fold	0.78 ± 0.01	0.80 ± 0.01	0.79 ± 0.01
	Planar	-	-	-
Dey et al. (2021)	Boundary	0.87 ± 0.01	0.84 ± 0.02	0.85 ± 0.01
	Fold	0.89 ± 0.01	0.90 ± 0.01	0.89 ± 0.01
	Planar	0.91 ± 0.01	0.92 ± 0.02	0.91 ± 0.02
Proposed (SVM)	Boundary	0.94 ± 0.01	0.91 ± 0.01	0.92 ± 0.01
	Fold	0.89 ± 0.02	0.81 ± 0.02	0.85 ± 0.02
	Planar	0.95 ± 0.01	0.94 ± 0.01	0.93 ± 0.01
Proposed (RF)	Boundary	0.96 ± 0.01	0.94 ± 0.01	0.95 ± 0.01
	Fold	0.94 ± 0.01	0.92 ± 0.02	0.93 ± 0.01
	Planar	0.99 ± 0.01	0.98 ± 0.01	0.98 ± 0.01

Table 4. Summary of classification using different methods for Toronto and Hermanni datasets. Results are indicated by the mean values of 10 different runs ± standard deviations.

Method	Metric	Toronto			Hermanni			
		Vertical	Boundary	Planar	Fold	Vertical	Boundary	Planar
X. Chen and Yu (2019)	Precision	-	0.82 ± 0.02	-	0.80 ± 0.01	-	0.82 ± 0.02	-
	Recall	-	0.81 ± 0.02	-	0.83 ± 0.02	-	0.88 ± 0.02	-
	F1-score	-	0.82 ± 0.02	-	0.81 ± 0.01	-	0.85 ± 0.02	-
Dey et al. (2021)	Precision	-	0.89 ± 0.01	0.94 ± 0.01	0.93 ± 0.02	-	0.90 ± 0.010	0.94 ± 0.01
	Recall	-	0.85 ± 0.01	0.90 ± 0.01	0.90 ± 0.01	-	0.89 ± 0.01	0.94 ± 0.01
	F1-score	-	0.87 ± 0.01	0.92 ± 0.01	0.92 ± 0.02	-	0.90 ± 0.01	0.94 ± 0.01
Proposed (SVM)	Precision	0.95 ± 0.01	0.99 ± 0.01	0.98 ± 0.01	0.95 ± 0.01	1.00 ± 0.00	0.96 ± 0.01	0.99 ± 0.01
	Recall	0.94 ± 0.01	1.00 ± 0.00	0.98 ± 0.01	0.99 ± 0.01	1.00 ± 0.00	0.99 ± 0.01	0.97 ± 0.01
	F1-score	0.94 ± 0.01	0.99 ± 0.01	0.98 ± 0.01	0.97 ± 0.01	1.00 ± 0.00	0.97 ± 0.01	0.98 ± 0.01
Proposed (RF)	Precision	0.99 ± 0.01	1.00 ± 0.00	0.99 ± 0.01	0.99 ± 0.01	1.00 ± 0.00	0.96 ± 0.01	1.00 ± 0.00
	Recall	0.99 ± 0.01	1.00 ± 0.00	0.99 ± 0.01	0.99 ± 0.01	1.00 ± 0.00	0.99 ± 0.01	0.98 ± 0.01
	F1-score	0.99 ± 0.01	1.00 ± 0.00	0.99 ± 0.01	0.99 ± 0.01	1.00 ± 0.00	0.98 ± 0.01	0.99 ± 0.01

Table 4 shows the quantitative classification results for Toronto and Hermanni sites together in terms of precision, recall, and F1-score. Figure 9 shows the qualitative classification performance of two state-of-the-art techniques along with the proposed SVM, and RF classifiers, respectively, using two sample buildings from Toronto and Hermanni datasets. Blue points represent the classified vertical points in Toronto datasets and fold points in Hermanni datasets.

It is clearly noticeable from Figure 9, and also from Table 4 that, conventional classifiers SVM and RF performed very well in terms of precision, recall, and F1-scores for both Toronto and Hermanni datasets. However, the performance of RF is better than SVM. The F1-score using RF is always more than 0.98 for any class. Precision, recall, and F1-scores for boundary points in Toronto datasets and Vertical points in Hermanni

datasets are 1.00 using the RF classifier. Figures 10 and Figure 11, show the qualitative results of the classification for some selected buildings from the Toronto and Hermanni datasets, respectively, using the RF classifier.

We generated the confusion matrix for each dataset and based on the generated ground truth we found the number of correctly classified instances and calculated the accuracy based on the correctly classified instances and the total number of points in the ground truth data for each datasets. Using Figure 12, we presented the calculated accuracies to compare different methods. The accuracy of the individual method for different datasets indicates the mean values of 10 different runs. The standard deviation in each case is±0.03. It is noticeable that using the proposed features the RF classifier performs best among the other approaches.
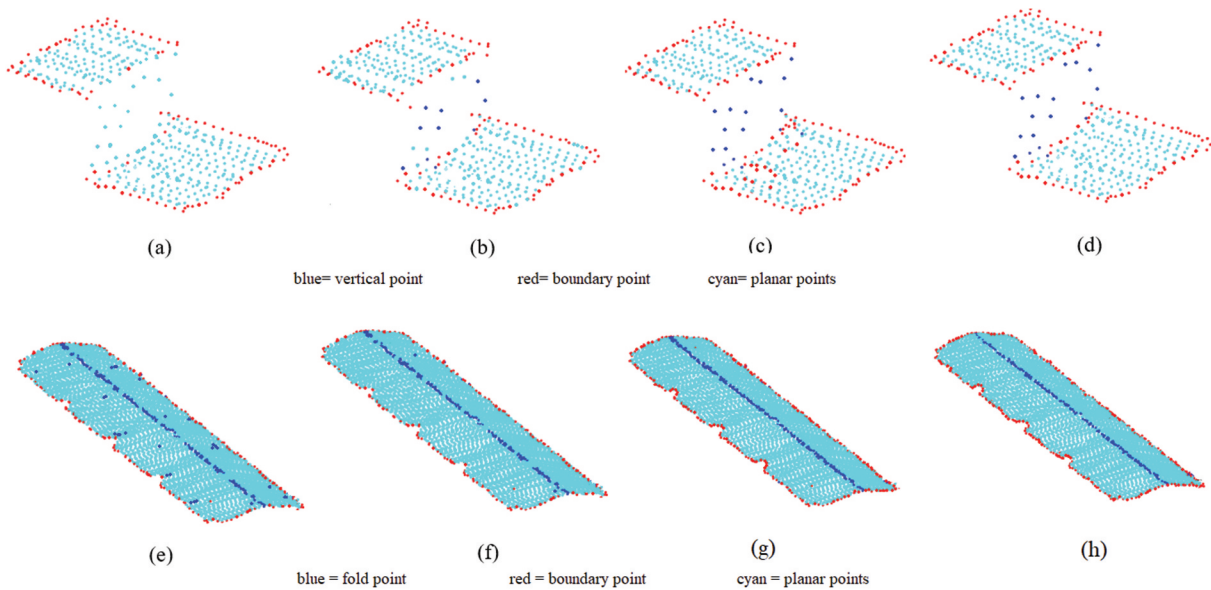

Figure 9. Comparison of the classification using different methods on two sample buildings from Toronto (first row) and Hermanni (second row) datasets. Blue points represent the vertical edge points in the Toronto building and fold edge points in the representative Hermanni building. (a) and (e) represent the results of Chen's methods where cyan points represent unclassified points, (b) and (f) represent the results of Dey et al. (2021), (c) and (g) represent results using the proposed SVM classification, (d) and (h) represent results using the proposed RF classification.
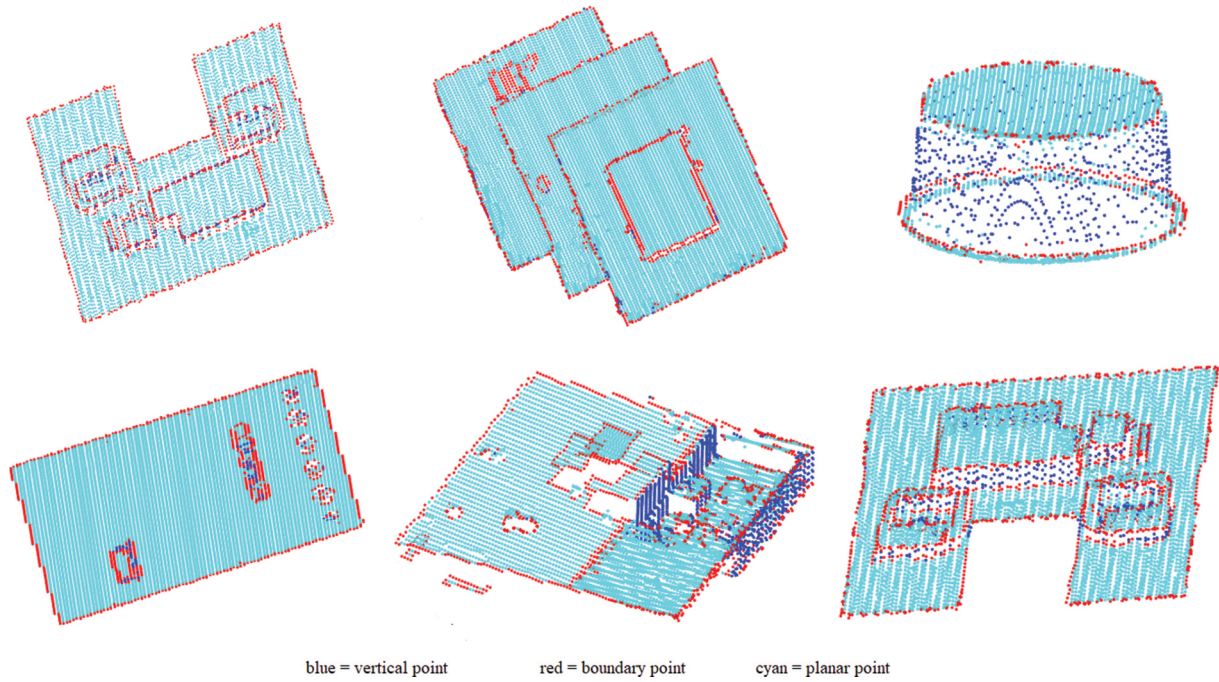
blue = vertical point   red = boundary point   cyan = planar point

Figure 10. Qualitative classification results of some selected building roofs from Toronto datasets using the RF classifier. Red points indicate classified boundary points. Cyan and blue dots indicate classified planar and vertically planar points, respectively.
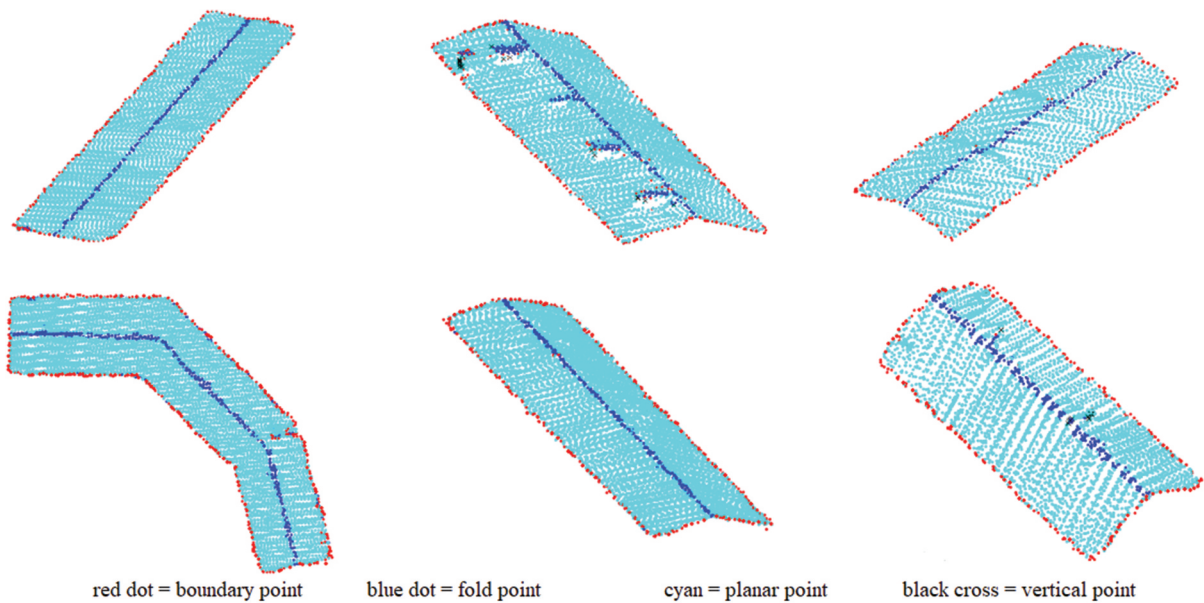

red dot = boundary point   blue dot = fold point   cyan = planar point   black cross = vertical point

Figure 11. Qualitative classification results of some selected building roofs from Hermanni datasets using the RF classifier. Red points indicate classified boundary points. Cyan and blue dots indicate classified planar and fold points, respectively.

This is due to the fact that, in most cases, the selected features have some clearly distinguishable characteristics for the points across the selected classes. For example, the boundary points have very different azimuth angles and distances from the centre point of the selected neighbourhood, which correspond to the features M_τ and d_m, respectively, we used. In addition, planar points exhibit a distinguishable normal angle than the points over the intersection line, which corresponds to the selected feature $\theta_{\max}$ we used for the classification. Moreover, the RF classifier uses decision tree partitioning which

divides the training set into small subsets until subsets are class uniform. As our training datasets are not balanced, RF performs better than SVM in this case.

To examine the universality of the extracted features irrespective of datasets in machine learning, we performed cross-database training and testing using the RF classifier. In this situation, two cases were considered. Firstly, training and testing using the features from the same datasets. Secondly, testing building roofs from datasets while training using a different one. Tables 5 and Tables 6 show the calculated F1-scores on these two cases for each class considering RF
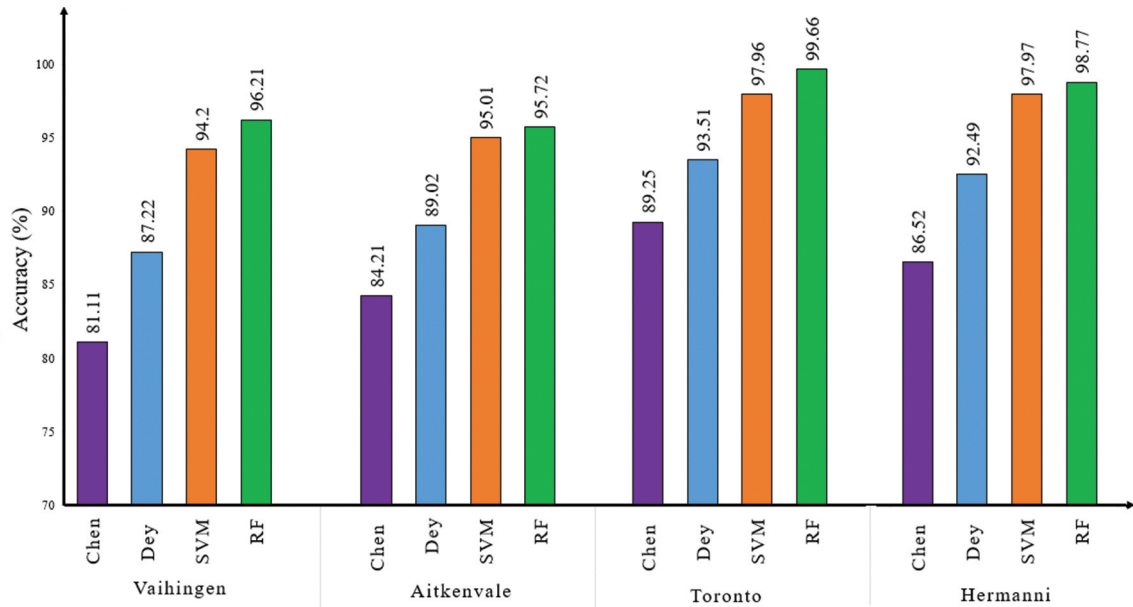
Figure 12. Summary of calculated accuracies using different methods for different datasets. Accuracies for each method indicate the mean of 10 different runs with a standard deviation of ±0.03.

Table 5. Classification performances in terms of F1-score for Vaihingen, Aitkenvale, and Hervey Bay datasets using cross-database training and testing. Results are indicated by the mean values of 10 different runs ± standard deviations.

Datasets	Vaihingen			Aitkenvale			Hervey Bay		
	Fold	Boundary	Planar	Fold	Boundary	Planar	Fold	Boundary	Planar
Vaihingen	0.97 ± 0.02	0.96 ± 0.02	0.96 ± 0.02	0.87 ± 0.01	0.69 ± 0.01	0.84 ± 0.01	0.83 ± 0.01	0.82 ± 0.01	0.86 ± 0.01
Aitkenvale	0.73 ± 0.01	0.77 ± 0.01	0.84 ± 0.01	0.87 ± 0.02	0.93 ± 0.02	0.98 ± 0.01	0.81 ± 0.01	0.86 ± 0.01	0.88 ± 0.01
Hervey Bay	0.78 ± 0.01	0.75 ± 0.01	0.89 ± 0.01	0.79 ± 0.01	0.81 ± 0.01	0.87 ± 0.01	0.88 ± 0.02	0.90 ± 0.02	0.95 ± 0.02

Table 6. Classification performance in terms of F1-score for different Toronto and Hermanni datasets using cross-database training and testing. Results are indicated by the mean values of 10 different runs ± standard deviations.

Datasets	Toronto			Hermanni		
	Boundary	Planar	Vertical	Boundary	Planar	Vertical
Toronto	1.00 ± 0.00	0.99 ± 0.01	0.98 ± 0.01	0.92 ± 0.01	0.94 ± 0.01	0.91 ± 0.01
Hermanni	0.92 ± 0.01	0.93 ± 0.01	0.89 ± 0.01	0.96 ± 0.01	1.00 ± 0.00	1.00 ± 0.00

classifiers. As we were considering similar classes (boundary, fold, and planar) for Vaihingen, Aikenvale, and Hervey Bay areas, we used Table 5 to show their performances together. For the same reason, we separated Toronto and Hermanni datasets into Table 6 as they had a different class (Vertical points). We found from these two tables that training and testing a machine learning classifier using the representative features from the same datasets provided the maximum results in terms of F1-scores. However, training and testing using different datasets also showed good results but they did not outperform the first case. This was because of different parameters in different datasets such as point density, aircraft velocity, and direction.

Conclusion

In the context of 3D building reconstruction, we have introduced the applicability of machine learning

approaches to an unexplored area of fine-grained point cloud segmentation over the extracted building roof. We have identified seven different features of the input point cloud and showed the classification results using two different conventional machine learning classifiers. The novelty and effectiveness of the selected features were demonstrated using the experimental results. Four major classes of building roof point clouds were considered and promising results have been found for each of the classes which confirmed the competitive performance over the state-of-the-art techniques. Using the RF classifier, the selected features demonstrated the maximum classification performances for each dataset. However, the performances of the machine learning classifiers are highly dependent on the training datasets.

Deep learning approaches to classify the feature points can be also applied in this area; however, the major limitation, in this case, is the absence of adequate and reliable ground truth data. We used

a manual process to generate the ground truth for our experiment. Thus, we can ensure the quality of the generated ground truth; however, the quantity of the generated ground truth may not be sufficient for a deep learning approach to be implemented. The self-supervised approaches of deep learning can be effective to generate adequate ground truth data in this case. It is a comparatively recent strategy and can be an effective alternative to supervised classification where generating ground truth is time-consuming and/or difficult.

Tracing feature lines from the classified fold and boundary feature points and construction of planar patches from the classified planar and/or vertical points are the next steps of 3D building reconstruction. In future, we will investigate the self-supervised approaches for feature point classification to avoid the manual human effort of data labelling and also an effective feature line tracing algorithm for regularisation purposes considering relationships among the constructed planar patches. Moreover, the applicability of the machine learning approaches will also be investigated in different application areas, such as 3D modelling of indoor objects from point cloud data.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCID

Emon Kumar Dey 🆔 http://orcid.org/0000-0002-2089-9557

Data availability statement

Publically available datasets can be found at https://www.isprs.org/education/benchmarks/UrbanSemLab/3d-semantic-labeling.aspx. Manually labelled and used data in this paper can be available upon a reasonable request to the corresponding author.

References

Acar, H., Karsli, F., Ozturk, M., & Dihkan, M. (2019). Automatic detection of building roofs from point clouds produced by the dense image matching technique. *International Journal of Remote Sensing*, 40(1), 138–155. https://doi.org/10.1080/01431161.2018.1508915

Awrangjeb, M. (2016). Using point cloud data to identify, trace, and regularize the outlines of buildings. *International Journal of Remote Sensing*, 37(3), 551–579. https://doi.org/10.1080/01431161.2015.1131868

Awrangjeb, M., & Fraser, C. S. (2014a). An automatic and threshold-free performance evaluation system for building extraction techniques from airborne LIDAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(10), 4184–4198. https://doi.org/10.1109/JSTARS.2014.2318694

Awrangjeb, M., & Fraser, C. S. (2014b). Automatic segmentation of raw LiDAR data for extraction of building roofs. *Remote Sensing*, 6(5), 3716–3751. https://doi.org/10.3390/rs6053716

Awrangjeb, M., Ravanbakhsh, M., & Fraser, C. S. (2010). Automatic detection of residential buildings using LIDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5), 457–467. https://doi.org/10.1016/j.isprsjprs.2010.06.001

Bassier, M., Van Genechten, B., & Vergauwen, M. (2019). Classification of sensor independent point cloud data of building objects using random forests. *Journal of Building Engineering*, 21, 468–477. https://doi.org/10.1016/j.jobe.2018.04.027

Bazazian, D., Casas, J. R., & Ruiz-Hidalgo, J., (2015), November. Fast and robust edge extraction in unorganized point clouds. In 2015 int. confe. on digital image computing: techniques and applications (DICTA), Adelaide, SA, Australia, (pp. 1–8). IEEE.

Becker, C., Rosinskaya, E., Häni, N., d'Angelo, E., & Strecha, C. (2018). Classification of aerial photogrammetric 3D point clouds. *Photogrammetric Engineering & Remote Sensing*, 84(5), 287–295. https://doi.org/10.14358/PERS.84.5.287

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Belton, D., & Lichti, D. D. (2006). Classification and segmentation of terrestrial laser scanner point clouds using local variance information. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5), 44–49.

Ben-Shabat, Y., Lindenbaum, M., & Fischer, A., (2019). Nesti-net: Normal estimation for unstructured 3d point clouds using convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, (pp. 10112–10120).

Boulaassal, H., Landes, T., & Grussenmeyer, P. (2009). Automatic extraction of planar clusters and their contours on building façades recorded by terrestrial laser scanner. *International Journal of Architectural Computing*, 7(1), 1–20. https://doi.org/10.1260/147807709788549411

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3), 1–27. https://doi.org/10.1145/1961189.1961199

Chehata, N., Guo, L., & Mallet, C. (2009). Airborne lidar feature selection for urban classification using random forests. *Laserscanning*, XXXVIII(Commission III–WG III/2), 207–212 .

Chen, Y., Liu, G., Xu, Y., Pan, P., & Xing, Y. (2021). PointNet++ Network Architecture with Individual Point Level and Global Features on Centroid for ALS Point Cloud Classification. *Remote Sensing*, *13*(3), 472. https://doi.org/10.3390/rs13030472

Chen, X., & Yu, K. (2019). Feature line generation and regularization from point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(12), 9779–9790. https://doi.org/10.1109/TGRS.2019.2929138

Cochran, R. N., & Horne, F. H. (1977). Statistically weighted principal component analysis of rapid scanning wavelength kinetics experiments. *Analytical Chemistry*, *49*(6), 846–853. https://doi.org/10.1021/ac50014a045

Cramer, M. (2010). The DGPF test on digital aerial camera evaluation – overview and test design. *Photogrammetrie – Fernerkundung – Geoinformation*, *2*(2010), 73–82. https://doi.org/10.1127/1432-8364/2010/0041

Dai, Y., Gong, J., Li, Y., & Feng, Q. (2017). Building segmentation and outline extraction from UAV image-derived point clouds by a line growing algorithm. *International Journal of Digital Earth*, *10*(11), 1077–1097. https://doi.org/10.1080/17538947.2016.1269841

Dey, E. K., & Awrangjeb, M. (2020). A robust performance evaluation metric for extracted building boundaries from remote sensing data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 4030–4043. https://doi.org/10.1109/JSTARS.2020.3006258

Dey, E. K., Awrangjeb, M., & Stantic, B. (2020). Outlier detection and robust plane fitting for building roof extraction from LiDAR data. *International Journal of Remote Sensing*, *41*(16), 6325–6354. https://doi.org/10.1080/01431161.2020.1737339

Dey, E. K., Tarsha Kurdi, F., Awrangjeb, M., & Stantic, B. (2021). Effective Selection of Variable Point Neighbourhood for Feature Point Extraction from Aerial Building Point Cloud Data. *Remote Sensing*, *13*(8), 1520. https://doi.org/10.3390/rs13081520

Dos Santos, R. C., Galo, M., & Carrilho, A. C. (2018). Building Boundary Extraction from LiDAR Data Using a Local Estimated Parameter for Alpha Shape Algorithm. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *42*(1), 127–132. https://doi.org/10.5194/isprs-archives-XLII-1-127-2018

Gharineiat, Z., Tarsha Kurdi, F., & Campbell, G. (2022). Review of Automatic Processing of Topography and Surface Feature Identification LiDAR Data Using Machine Learning Techniques. *Remote Sensing*, *14*(19), 4685. https://doi.org/10.3390/rs14194685

Gilani, S. A. N., Awrangjeb, M., & Lu, G. (2016). An automatic building extraction and regularisation technique using lidar point cloud data and orthoimage. *Remote Sensing*, *8*(3), 258. https://doi.org/10.3390/rs8030258

Gilani, S. A. N., Awrangjeb, M., & Lu, G. (2018). Segmentation of airborne point cloud data for automatic building roof extraction. *GIScience & Remote Sensing*, *55*(1), 63–89. https://doi.org/10.1080/15481603.2017.1361509

Gumhold, S., Wang, X., & MacLeod, R. S. (2001). Feature Extraction from Point Clouds. *IMR*, 293–305.

Hackel, T., Wegner, J. D., & Schindler, K. (2016). Fast semantic segmentation of 3D point clouds with strongly varying density. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *3*, 177–184. https://doi.org/10.5194/isprs-annals-III-3-177-2016

He, E., Chen, Q., Wang, H., & Liu, X. (2017). A curvature based adaptive neighbourhood for individual point cloud. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *42*, 219–225. https://doi.org/10.5194/isprs-archives-XLII-2-W7-219-2017

He, Y., Zhang, C., Awrangjeb, M., & Fraser, C. S. (2012). Automated reconstruction of walls from airborne lidar data for complete 3D building modelling. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *39*, B3. https://doi.org/10.5194/isprsarchives-XXXIX-B3-115-2012

Karsli, F., Dihkan, M., Acar, H., & Ozturk, A. (2016). Automatic building extraction from very high-resolution image and LiDAR data with SVM algorithm. *Arabian Journal of Geosciences*, *9*(14), 1–12. https://doi.org/10.1007/s12517-016-2664-7

Leichter, A., Werner, M., & Sester, M. (2020). Feature-extraction from all-scale neighborhoods with applications to semantic segmentation of point clouds. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *43*, 263–270. https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-263-2020

Lin, C. H., Chen, J. Y., Su, P. L., & Chen, C. H. (2014). Eigen-feature analysis of weighted covariance matrices for LiDAR point cloud classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *94*, 70–79. https://doi.org/10.1016/j.isprsjprs.2014.04.016

Lin, Y., Vosselman, G., & Yang, M. Y. (2022). Weakly supervised semantic segmentation of airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, *187*, 79–100. https://doi.org/10.1016/j.isprsjprs.2022.03.001

Liu, T., Abd Elrahman, A., Morton, J., & Wilhelm, V. L. (2018). Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GIScience & Remote Sensing*, *55*(2), 243–264. https://doi.org/10.1080/15481603.2018.1426091

Li, X., Yao, X., & Fang, Y. (2018). Building-a-nets: Robust building extraction from high-resolution remote sensing images with adversarial networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(10), 3680–3687. https://doi.org/10.1109/JSTARS.2018.2865187

Li, Z., Zhang, L., Zhong, R., Fang, T., Zhang, L., & Zhang, Z. (2016). Classification of urban point clouds: A robust supervised approach with automatically generating training data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(3), 1207–1220. https://doi.org/10.1109/JSTARS.2016.2628399

Lodha, S. K., Kreps, E. J., Helmbold, D. P., & Fitzpatrick, D., 2006, June. Aerial LiDAR data classification using support vector machines (SVM). In Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06), Chapel Hill, NC, USA, (pp. 567–574). IEEE.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, *60*(2), 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Maltezos, E., Doulamis, A., Doulamis, N., & Ioannidis, C. (2018). Building extraction from LiDAR data applying deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, *16*(1), 155–159. https://doi.org/10.1109/LGRS.2018.2867736

Mérigot, Q., Ovsjanikov, M., & Guibas, L. J. (2010). Voronoi-based curvature and feature estimation from point clouds. *IEEE Transactions on Visualization and Computer Graphics*, *17*(6), 743–756. https://doi.org/10.1109/TVCG.2010.261

Niemeyer, J., Rottensteiner, F., & Soergel, U. (2014). Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 152–165. https://doi.org/10.1016/j.isprsjprs.2013.11.001

Ni, H., Lin, X., Ning, X., & Zhang, J. (2016). Edge detection and feature line tracing in 3D-point clouds by analyzing geometric properties of neighborhoods. *Remote Sensing*, *8*(9), 710. https://doi.org/10.3390/rs8090710

Ni, H., Lin, X. G., & Zhang, J. X. (2017). APPLICATIONS of 3D-EDGE DETECTION for ALS POINT CLOUD. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *XLII-2/W7*, 42. https://doi.org/10.5194/isprs-archives-XLII-2-W7-277-2017

Nurunnabi, A., West, G., & Belton, D. (2015). Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data. *Pattern Recognition*, *48*(4), 1404–1419. https://doi.org/10.1016/j.patcog.2014.10.014

Özdemir, E., Remondino, F., & Golkar, A. (2019). Aerial point cloud classification with deep learning and machine learning algorithms. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 843–849. https://doi.org/10.5194/isprs-archives-XLII-4-W18-843-2019

Pamungkas, I. R., & Suwardi, I. S. (2015, March). 3D-building reconstruction approach using semi-global matching classified. In International Conference on Soft Computing, Intelligence Systems, and Information Technology (pp. 382–391). Springer, Berlin, Heidelberg.

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

Pauly, M., Gross, M., & Kobbelt, L. P., (2002). Efficient simplification of point-sampled surfaces. In IEEE Visualization, 2002. VIS 2002, Boston, MA, USA, (pp. 163–170). IEEE.

Pohle-Fröhlich, R., Bohm, A., Ueberholz, P., Korb, M., & Goebbels, S. (2019). Roof Segmentation based on Deep Neural Networks. *VISIGRAPP (4: VISAPP)*, 326–333.

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, Honolulu, HI, USA, (pp. 652–660).

Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, *30*.

Rutzinger, M., Elberink, S. O., Pu, S., & Vosselman, G. (2009). Automatic extraction of vertical walls from mobile and airborne laser scanning data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *38*(W8), 7–11.

Rutzinger, M., Rottensteiner, F., & Pfeifer, N. (2009). A comparison of evaluation techniques for building extraction from airborne laser scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *2*(1), 11–20. https://doi.org/10.1109/JSTARS.2009.2012488

Sampath, A., & Shan, J. (2009). Segmentation and reconstruction of polyhedral building roofs from aerial lidar point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, *48*(3), 1554–1567. https://doi.org/10.1109/TGRS.2009.2030180

Sanchez, J., Denis, F., Dupont, F., Trassoudaine, L., & Checchin, P. (2020). Data-driven modeling of building interiors from lidar point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*, 395–402. https://doi.org/10.5194/isprs-annals-V-2-2020-395-2020

Serna, A., & Marcotegui, B. (2014). Detection, segmentation and classification of 3D urban objects using mathematical morphology and supervised learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *93*, 243–255. https://doi.org/10.1016/j.isprsjprs.2014.03.015

Sterri, A. J. E. (2021). Building Boundary Extracting from Pointcloud with a Generative Adversarial Network (Master's thesis, NTNU).

Tarsha Kurdi, F., & Awrangjeb, M. (2020). Automatic evaluation and improvement of roof segments for modelling missing details using Lidar data. *International Journal of Remote Sensing*, *41*(12), 4702–4725. https://doi.org/10.1080/01431161.2020.1723180

Tarsha Kurdi, F., Awrangjeb, M., & Munir, N. (2021). Automatic filtering and 2D modeling of airborne laser scanning building point cloud. *Transactions in GIS*, *25*(1), 164–188. https://doi.org/10.1111/tgis.12685

Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P., & Smigiel, E., (2006), September. New approach for automatic detection of buildings in airborne laser scanner data using first echo only. In ISPRS Comm. III Symposium, Photogrammetric Comp. Vision, Bonn, Germany, (pp. 25–30).

Thomas, H., Goulette, F., Deschaud, J. E., Marcotegui, B., & LeGall, Y. (2018). Semantic classification of 3D point clouds with multiscale spherical neighborhoods. 2018 International Conference on 3D Vision (3DV), Verona, Italy, 390–398.

Wang, R., Peethambaran, J., & Chen, D. (2018). Lidar point clouds to 3-D urban models: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(2), 606–627. https://doi.org/10.1109/JSTARS.2017.2781132

Wang, Z., & Prisacariu, V. A. (2020). Neighbourhood-insensitive point cloud normal estimation network. 31st British Machine Vision Conference, 2020, UK.

Wang, J., & Shan, J. (2009, March). Segmentation of LiDAR point clouds for building extraction. In *American Society for Photogram. Remote Sens. Annual Conf* (pp. 9–13).

Weinmann, M., Jutzi, B., Hinz, S., & Mallet, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 286–304. https://doi.org/10.1016/j.isprsjprs.2015.01.016

Weinmann, M., Jutzi, B., & Mallet, C. (2013). Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *5*(W2), 1. https://doi.org/10.5194/isprsannals-II-5-W2-313-2013

Weinmann, M., Schmidt, A., Mallet, C., Hinz, S., Rottensteiner, F., & Jutzi, B. (2015). Contextual classification of point cloud data by exploiting individual 3D neigbourhoods. *ISPRS Annals of the Photogrammetry,*

Remote Sensing and Spatial Information Sciences II-3 (2015), *II-3/W4*, 271–278. Nr. W4, 2(W4). https://doi.org/10.5194/isprsannals-II-3-W4-271-2015

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

Xia, S., Chen, D., Wang, R., Li, J., & Zhang, X. (2020). Geometric primitives in LiDAR point clouds: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 685–707. https://doi.org/10.1109/JSTARS.2020.2969119

Xia, S., & Wang, R. (2017). A fast edge extraction method for mobile LiDAR point clouds. *IEEE Geoscience and Remote Sensing Letters*, *14*(8), 1288–1292. https://doi.org/10.1109/LGRS.2017.2707467

Xie, L., Zhu, Q., Hu, H., Wu, B., Li, Y., Zhang, Y., & Zhong, R. (2018). Hierarchical regularization of building boundaries in noisy aerial laser scanning and photogrammetric point clouds. *Remote Sensing*, *10*(12), 1996. https://doi.org/10.3390/rs10121996

Xiong, B., Elberink, S. O., & Vosselman, G. (2014). A graph edit dictionary for correcting errors in roof topology graphs reconstructed from point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, *93*, 227–242. https://doi.org/10.1016/j.isprsjprs.2014.01.007

Xu, Y., Tuttas, S., Hoegner, L., & Stilla, U. (2018). Reconstruction of scaffolds from a photogrammetric point cloud of construction sites using a novel 3D local feature descriptor. *Automation in Construction*, *85*, 76–95. https://doi.org/10.1016/j.autcon.2017.09.014

Yang, Y., Tang, R., Wang, J., & Xia, M. (2021). A hierarchical deep neural network with iterative features for semantic labeling of airborne LiDAR point clouds. *Computers & Geosciences*, *157*, 104932. https://doi.org/10.1016/j.cageo.2021.104932

Yang, Z., Tan, B., Pei, H., & Jiang, W. (2018). Segmentation and multi-scale convolutional neural network-based classification of airborne laser scanner data. *Sensors*, *18*(10), 3347. https://doi.org/10.3390/s18103347

Yousefhussien, 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. https://doi.org/10.1016/j.isprsjprs.2018.03.018

Zhang, Y., Geng, G., Wei, X., Zhang, S., & Li, S. (2016). A statistical approach for extraction of feature lines from point clouds. *Computers & Graphics*, *56*, 31–45. https://doi.org/10.1016/j.cag.2016.01.004

Zhang, J., Lin, X., & Ning, X. (2013). SVM-based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sensing*, *5*(8), 3749–3775. https://doi.org/10.3390/rs5083749

Zhao, R., Pang, M., Liu, C., & Zhang, Y. (2019). Robust normal estimation for 3D LiDAR point clouds in urban environments. *Sensors*, *19*(5), 1248. https://doi.org/10.3390/s19051248

Zhao, R., Pang, M., & Wang, J. (2018). Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network. *International Journal of Geographical Information Science*, *32*(5), 960–979. https://doi.org/10.1080/13658816.2018.1431840