



3D-MFDNN: Three-dimensional multi-feature descriptors combined deep neural network for vegetation segmentation from airborne laser scanning data

Dheerendra Pratap Singh, Manohar Yadav*

Geographic Information System (GIS) Cell, Motilal Nehru National Institute of Technology Allahabad, Prayagraj 211004, India

ARTICLE INFO

Keywords:

Airborne laser scanning (ALS)
Point cloud
Feature descriptor
Deep neural network (DNN)
Vegetation segmentation

ABSTRACT

Airborne laser scanning (ALS) is a state-of-the-art technique for fast and accurate three-dimensional information acquisition of land cover including vegetation. This paper presents a three-dimensional multi-feature descriptors combined deep neural network-based methodology (3D-MFDNN) for ALS data-based vegetation segmentation with three well-designed steps namely generation of feature descriptors, 3D-MFDNN method's training, testing, and performance comparison using ALS datasets. The proposed 3D-MFDNN method is straightforward to implement, where accurate segmentation of tree points are effectively dealt in several complex cases, such as tree branches connected with other objects, tree with understory low vegetation, low-lying plants on the sloping surface, large tree with volumetric shape and branches hanging near sloping ground surface, etc. The method performance was evaluated using six datasets having different levels of scene complexity, and vegetation segmentation was performed at F1-score and accuracy of 83.94 % and 92.13 %, respectively. The method achieves significant improvement in comparison with several state-of-the-art methods.

1. Introduction

To maintain and improve the quality of life and ensure sustainable urban and semi-urban environment, it is essential to map and analyse the vegetation cover [1]. The urban and semi-urban forests can provide several benefits to the human health [2] and play a significant role in boosting environmental quality [3] by reducing air pollution [4], maintaining the city environment through several ecosystem services [5], climate change mitigation, and habitat [6], decentralized green infrastructure to protect life [7], etc. The urban and semi-urban vegetation cover including forests can also contribute to ensure the local food and nutrition security, urban biodiversity, reducing carbon emissions, cooling air, filter urban pollutants and fine particulates [8]. The overall benefits and impacts of urban forests in the sustainable growth of cities and communities are derived through the analysis at individual tree level, therefore trees mapping and their identification are very crucial.

The vegetation covers information in urban and semi-urban areas was traditionally extracted using field inventory in-situ methods however; these methods are labour-intensive, time-consuming, costly, and not scaled to larger areas [9]). Optical remote sensing data, i.e., satellite imageries, are extensively used to extract vegetation information from

distinct texture and spectral information [10]. However, it is limited to 2D information and is vulnerable to weather conditions, In the image-based data, the object's three-dimensional (3D) geometry and vertical structural information are not available; however they are very crucial for urban trees inventory [11]. With the advent of light detection and ranging (LiDAR), an active remote sensing technology, it has been possible to map the large area in 3D collecting detailed object's geometry with high accuracy and point density [12]. The LiDAR-derived point cloud data and its processing have brought feature extraction to a new era [13]. The aerial platform mounted LiDAR system is a good choice for the laser scanning of large urban areas, where due to its top-down scanning method and extensive flight coverage, highly accurate and dense point clouds are captured in a shorter time [14]. Many methods were developed for the airborne laser scanning (ALS) data processing for trees segmentation [15] and their species identification [16].

In the recent years, the machine learning (ML) methods developed for the pattern recognition and imagery data processing were employed for the 3D point cloud data processing. The conventional ML methods have been mostly used to for ALS data classification to identify the urban tree species [17]. Now, the growing availability of multi-source and

* Corresponding author.

E-mail addresses: dheeru.dp@gmail.com (D.P. Singh), ssmyadav@mnnit.ac.in (M. Yadav).

multi-temporal data makes the classification difficult for conventional ML methods due to the increase in data size, objects information, and scene complexities [18]. To address these issues, deep learning (DL) methods, a well-known family of artificial intelligence approaches, have been developed that handle extensive data effectively, solve the complex problems and efficiently automate the data processing [19].

1.1. Related studies

Many efforts have been made to address the problem of vegetation segmentation in urban, semi-urban and forested areas using ALS point cloud. The existing ALS-based vegetation segmentation studies can be categorized into knowledge-based and machine/deep learning-based (ML/DL) methods. Knowledge-based methods (KBM) use the knowledge of human experts to support decision-making in urban tree segmentation [20], while ML/DL methods partially or fully automate the segmentation process [21]. In these studies, several LiDAR-derived key variables and models were generated before the segmentation, such as digital terrain model (DTM) [22], digital elevation model (DEM) [23], canopy height model (CHM) [24], local maxima [25]. A summary of recent vegetation segmentation and detection methods for ALS point cloud is summarized in Table 1.

1.1.1. Knowledge-based methods

The knowledge-based method (i.e. rule-based) uses a set of rules defined on the basis of geometric, radiometric, colorimetric and radiometric properties of various objects for their identification and classification [39]. Several knowledge-based methods were developed for the vegetation segmentation using ALS datasets, which are summarized separately in Table 1.

Liu et al. [22] proposed a *k*-NN and random forest based method for tree segmentation using ALS point cloud data, where first DTM was

generated from the point cloud. In the method of Ma et al. [26] an urban site ALS point cloud data was processed for tree segmentation. The canopy morphological features (CMF) were extracted using height normalized data, and trees were segmented employing *k*-means clustering and region growing method. Torresan et al. [27] proposed point cloud and raster-based local maximum region growth-based methods, in which individual tree was extracted using height, crown area, diameter at breast height (DBH), and above-ground biomass. Fekete and Cserep [28] detected the change in trees cover of large urban areas using LiDAR data of two different epochs. Each epoch data was processed to generate CHM and non-trees were eliminated. Further tree crown was segmented and clustered datasets of two epochs were paired to calculate height and volume differences. Xu et al. [29,30] proposed a crown morphology knowledge-based tree detection approach in which DTM and CHM were used as input features. Yan et al. [17] proposed a marker-controlled watershed algorithm-based individual tree segmentation technique focused on 3D spatial distribution recognition from ALS point clouds. Based on the recognition of point cloud distribution in 3D space, the multidirectional spatial distribution analysis was performed to refine the probable tree apex sites. Finally, the 3D coarse-to-fine segmentation of individual trees was achieved using the *k*-means clustering algorithm.

1.1.2. ML/DL-based methods

The several ML/DL-based methods were developed including the conventional ML methods such as the support vector machine (SVM), decision tree, random forest (RF), *k*-nearest neighbor (*k*-NN), and artificial neural network (ANN) technique for vegetation covers segmentation and identification using ALS data [17]. These methods are summarized in Table 1. Zhang et al. [40] evaluated the performance of LiDAR-derived geometry and intensity parameters in tree species classification using SVM approach. Dian et al. [24] fused ALS data and hyperspectral image, where fused dataset was used to generate vertical,

Table 1
Summary of vegetation segmentation methods for airborne laser scanning data.

Knowledge-based methods					
Number and type of test sites (Area/Pts)	Point density (Pts/m ²)	Input data format	Pre-processed LiDAR-derived variables and models	Key steps	References
1:Forest 1:Semi-urban	160 25	Point cloud Point cloud	DTM Normalized height, CMF	<i>k</i> -NN, Random forest Region Growing & <i>k</i> -means clustering Region growing	[22] [26]
1: Forest	193	Point cloud, Raster image	CHM	Region growing	[27]
2:Urban	8, 24	Point cloud	CHM, local maxima	Centroid & Hausdorff distance based tree segmentation	[28]
2:Urban 7:Forest	160 40.57	Point cloud Point cloud	DTM, CHM, Local maxima CHM, Local maxima	Crown morphology Watershed segmentation, <i>k</i> -means clustering	[29,30] [17]
ML/DL-based methods					
1: Urban	12	Point cloud, Hyperspectral image	DEM, DSM, CHM	SVM	[24]
1:Urban	NA	Point cloud	CHM, Geometric features (47) and radiometric features (100).	Random Forest	[31]
2:Forest	70	Point cloud	Geometric features (16), radiometric features (Intensity-based: 11, echo-based:10)	Random forest	[32]
1:Forest 2:Forest	9 2, 50	Point cloud Point cloud	Voxelized point cloud DSM, 2D-CNN	3D-CNN FCN, Softmax	[33] [34]
1:Urban	5–80	Point cloud, Voxelized point cloud	K-D Tree, DEM, F-RNN	3D-FCN, PointNet	[23]
4:Urban	NA	Voxelized grid point cloud	Multiple local maxima	CSF, PointNet	[35]
3:Urban	8	Point cloud	CHM, Local maxima	Watershed clustering; Mean shift segmentation, PointNet	[36]
1:Forest 1: Urban	NA 16	Point cloud Point cloud	Watershed, point distance Geometric features (12), Intensity- based features (12)	LayerNet SVM, Random forest, MLP	[37] [38]
1:Urban(4 km ²)	16.6	Voxelized point cloud	Learned features per voxel (32)	SSH(PNP/ KPConv)	[16]

Pts: Points; CHM: Canopy height model; CMF: Canopy morphology features; CNN: Convolutional neural network; CSF: Cloth simulator filters; DEM: Digital elevation model; DSM: Digital surface model; DTM: Digital terrain model; FCN: Fully connected network; MLP: Multi-layer perceptron; PNP: PointNet+; SCN: Sparse convolutional network; SSH: Semantic segmentation head; SVM: Support vector machine.

spatial, and spectral features to train the SVM model for urban trees identification. The tree point cloud segmentation was performed by Koma et al. [31] using RF-based method, which was trained by geometric and radiometric features derived from the full-waveform ALS data. Shi et al. [32] discussed the importance of using 37 different LiDAR features extracted using ALS data of leaf-on and leaf-off environments to train the RF method for tree species segmentation.

In recent years, the DL approaches have been preferably used for object identification [41], segmentation and classification [42] using ALS data. Ayrey and Hayes [33] processed the vowelized ALS point cloud using 3D-CNN to generate the forest metrics such as above-ground biomass, tree count, percent leaf, etc. The 2D-CNN was used to categorize trees as coniferous or deciduous using a 2D representation of ALS point cloud data [34]. Windrim and Bryson [23] identified trees in 2D raster representation of high-resolution aerial LiDAR from a bird's-eye view using Faster RCNN. Further, 3D-FCN and k-d tree were used to segment trees in 3D point cloud into their stem and leaf components. Chen et al. [35] first classified the point cloud data as aboveground points and ground points using the cloth simulation filtering. The aboveground points were voxelized, and each voxel was analyzed by the trained PointNet framework to label them as tree or non-tree. Kippers et al [36] created a tree map using the PointNet deep learning model and performed watershed algorithm-based tree segmentation. Liu and Han [37] discussed a point-based deep neural network called Layer Net for tree species classification using ALS data, where local 3D structural features of trees computed for each of the layers were combined and global feature was generated via convolution for the trees classification. Cetin and Yastikli [38] used ML-based 3D LiDAR point cloud classification approaches, namely SVM, RF and DL-based MLP to categorise urban tree species as deciduous or coniferous. Schmohl et al. [16] used sparse convolutional network for three-dimensional feature extraction of ALS point clouds, wherein deep 3D single-shot detection network was used for urban tree detection.

1.2. Challenges with ALS data-based vegetation segmentation

Some of the existing methods exhibit good overall performance for the vegetation segmentation including urban trees using ALS data, but some significant limitations are still there, such as: (i) However, several vegetation segmentation methods perform well in various scene environment, still the accurate vegetation segmentation becomes challenging in one or many specific cases, such as inter- and intra-class variability, diverse spatial arrangements of objects, objects of varying shapes and sizes, occlusions, etc. [16] (ii) Partial objects geometry information and data gaps in the ALS point cloud, which cause the limited performance of the vegetation segmentation methods (iii) Several methods performed ground filtering first before the segmentation, but accurate ground filtering becomes challenging in one or many cases, such as low vegetation, objects on steep slopes, small and low objects, large building with flat rooftop, attached objects, scene border and discontinuity [43] (iv) Vegetation segmentation becomes more challenging with varying tree height and canopy diameter which are very common among the urban and semi-urban vegetation covers compared with the forest (v) The spatial heterogeneity specially in urban vegetation covers, where different tree species with varying shapes and sizes are clustered together, creates a complex scene to be dealt with.

A three-dimensional multi-feature descriptors combined deep neural network-based methodology (3D-MFDNN) is proposed to overcome the abovementioned challenges. This 3D-MFDNN method has several merits, such as: (i) Ground filtering is not required as an initial step before the vegetation segmentation; (ii) It can be widely used to accurately segment the vegetation points from any kind of ALS datasets irrespective of data complexity, objects heterogeneity as well as the original data structure; (iii) It is a data-driven methodology based on point-wise feature descriptors that works accurately for all types of vegetation and does not require point cloud-derived models, like DEM,

DSM, CHM; (iv) In the proposed methodology, it is straightforward to combine surface-based, moment-based, PCA-based, and density-based feature descriptors to train DNN model and differentiate vegetation and non-vegetation objects, where the accurate vegetation segmentation is effectively achieved in generalized scene as well as in the presence of scene complexities, such as tree branches connected with other objects, tree with understory low vegetation, low-lying plants on the sloping surface, large tree with volumetric shape and branches hanging near sloping ground surface, etc. Further, in experiment section, the proposed 3D-MFDNN method's performance is discussed in detail.

2. Study area and datasets

The performance evaluation of proposed 3D-MFDNN method and its comparison with state-of-the-art ML methods: SVM and RF were performed on benchmark ALS datasets: Dayton Annotated Laser Earth Scan (DALES) dataset. It is well-known open source data [44]. The DALES datasets are large-scale ALS data, generally used in the methods developed for point cloud segmentation and classification. The DALES ALS data were collected over the city of Surrey in British Columbia, Canada using a Riegl Q1560 dual-channel ALS system. DALES datasets contain 40 tiles with each tile area of $500 \times 500 \text{ m}^2$ and average point density of 50 pts/m². This dataset spans a 10 km² area with 12 million points and 40 tiles. In this dataset, the mean error was determined to be $\pm 8.5 \text{ cm}$ at a 95 % confidence level for vertical accuracy. The data is already labelled as ground truth data into eight categories: ground, car, truck, fence, pole, building, power line, and vegetation, with seven multiple return levels.

The six test sites were chosen from different tiles of DALES datasets for conducting the experiment and performance assessment of proposed method. The datasets of these sites are grouped into two different categories: urban and semi-urban. These datasets differ in terms of types and density of trees and buildings, intra-class variability, attached and overlapped objects, occlusions and data gaps, and spatial heterogeneity of high vegetation (Table 2).

2.1. Urban test site

The dataset #1 was from the DALES datasets tile number: 5100_54490 (Fig. 1(a)). The dataset #1 represented urban scene with gable roof houses (i.e. small buildings) constructed in multiple rows and trees located along each row of the houses. At many places tree branches were hanging over the house roof. There were decorated trees (i.e. small vegetation) planted along residential area and urban sports open playground. Dataset #2 was an urban industrial area with large warehouses and gable roof residential buildings (Fig. 1(b)). It was chosen from tile number: 5100_54440. The low objects like vehicle were parked in front of warehouses. There were trees, and power lines along the roads and buildings, where at some places power lines passed very close to the tree. Trees with varying height and crown diameter created heterogeneity in their spatial distribution. Dataset #3 was from tile number: 5145_54340, which was having variations in tree types (Fig. 1(c)). The low vegetations, that is decorated trees and bushes were located throughout the test site. The power lines and high-voltage transmission lines both were present and hanging in the air over the vegetation at varying heights. The buildings in the test site were multi-storey and large. The dataset #4 was from residential urban site belongs to tile number: 5145_54340, where houses of varying size gable roof were present throughout the site along many rows (Fig. 1(d)). It was densely vegetated site with low and high vegetation of different shapes and sizes. At several places trees were overlapped with other tree or connected with building.

2.2. Semi-urban test site

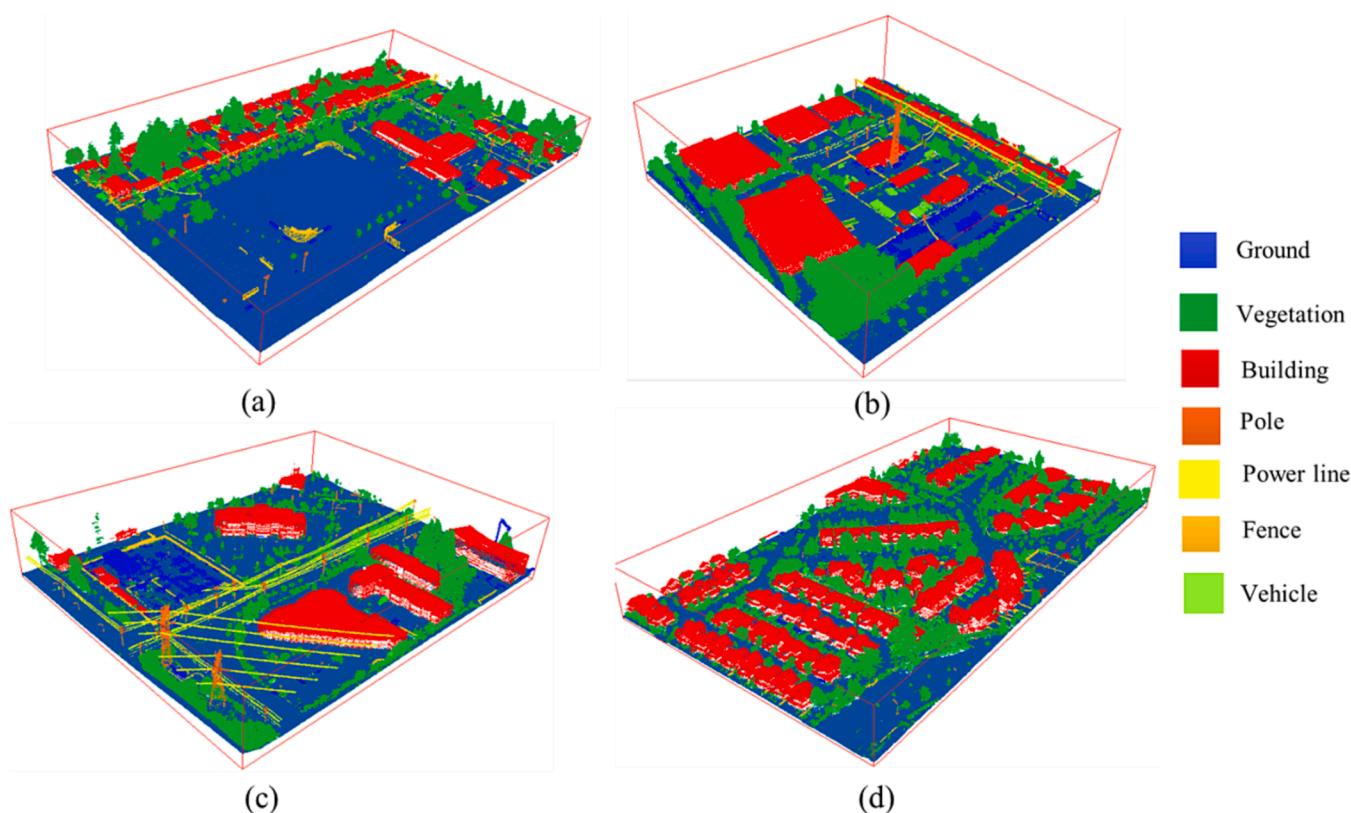
The two datasets #5 and #6 were from semi-urban areas captured

Table 2

Overall characteristics of input ALS datasets acquired from the test sites of varying geospatial scene.

Data set	Test site	Number of points (Area)	Terrain slope	Low vegetation	Objects on sloping terrain	High vegetation	Small buildings	Attached and overlapped objects	Occlusions and data gaps	Intra-class variability	Spatial heterogeneity of high vegetation
#1	Urban	4,393,404 (89,470 m ²)	Relatively flat	***	*	***	**	*	*	*	*
#2	Urban with mixed types of trees and buildings	3,643,834 (73,959 m ²)	Flat	*	**	**	**	*	**	**	**
#3	Urban with mixed types of trees and PL/TL lines	3,296,716 (74,046 m ²)	Relatively flat	***	*	**	**	*	*	*	*
#4	Urban residential	3,217,416 (63,112 m ²)	Flat	**	***	***	***	**	**	*	**
#5	Semi-urban	3,019,309 (64,464 m ²)	Gentle	*	**	**	***	**	**	**	**
#6	Semi-urban with water body	2,755,738 (52,875 m ²)	Gentle	*	**	**	**	*	*	*	**

The relative density of objects in a dataset is specified by symbols: * indicates low, ** indicates medium, and *** indicates high.

**Fig. 1.** Three-dimensional perspective view of ALS dataset (a) #1 from urban test site, (b) #2 from urban site having mixed types of trees and buildings, (c) #3 from urban site with mixed types of trees and PL/TL lines, and (d) #4 from urban residential site; where colours are assigned to the points object class-wise.

and recorded in tile numbers 5135_54435, and 5135_54435, respectively of DALES datasets. In dataset #5, varying shape and size of trees were present in dense cluster at several locations and they were partially overlapped at crown level (Fig. 2(a)). Some trees were present on gentle sloping ground terrain. The houses in the test site were mainly double-storey with gable roof, where small objects, such as vehicles were parked in front of the houses. In dataset #6, mainly trees were present in

several dense clusters, where transmission lines were hanging over the trees (Fig. 2(b)). In the test site, the connected and partially overlapped trees cluster encloses a water body and they were grown on sloping bank of water body.

In the proposed method a binary classification is performed to segment the vegetation and non-vegetation objects, therefore the labelled DALES test datasets #1 to #6 of six test sites and training

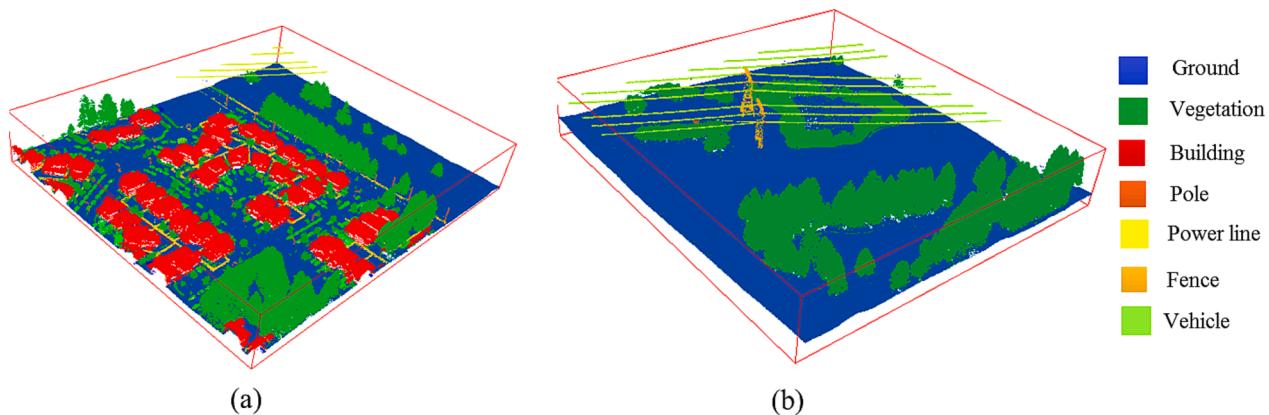


Fig. 2. Three-dimensional perspective view of ALS dataset (a) #5 from semi-urban test site, (b) #6 from semi-urban site having water body; where colours are assigned to the points object class-wise.

dataset are grouped into two classes only, vegetation and non-vegetation. As discussed previously, the labelled DALES dataset has eight categories, with vegetation as one of the categories. All non-vegetation categories are merged into a single class as non-vegetation. The training dataset is imbalance as non-vegetation points are larger than vegetation points. When a class imbalance exists within training data, model learning will typically over-classify the majority group due to its increased prior probability. As a result, the instances belonging to the minority group are misclassified more often than those belonging to the majority group [45]. Hence for resolving the data imbalance, equal number of random samples of vegetation and non-vegetation are provided to the model as training data.

3. Methodology

The proposed vegetation cover segmentation 3D-MFDNN methodology consists of three main steps: (1) feature descriptors generation (2) model training (3) vegetation segmentation, where in first step, the feature descriptors are selected for vegetation segregation from non-vegetation objects in ALS dataset (Fig. 3). Surface-based, moment-based, PCA-based and density-based feature descriptors are generated for each ALS points of training and testing datasets. The feature descriptors derived from training data are used to train proposed 3D-

MFDNN method and its parameter setting and same set of descriptors are also used for SVM and RF-based methods training. Further using the six different ALS datasets having several complex cases (Table 2) to be dealt with, the testing of proposed 3D-MFDNN method and its performance comparison with the SVM and RF-based methods in terms of evaluation metrics: Precision, Recall, F1-score, and Accuracy are performed. The following sections explain proposed methodology's steps in detail.

3.1. Point-based feature descriptors generation

The objects present in the geospatial scene are distinct in terms of their geometrical traits, such as shape, size and spatial arrangements and these geometrical traits are used to differentiate the object types, such as vegetation, building, ground, etc. [46]. These geometrical traits define object's geometry are acquired in the ALS point cloud datasets and they are important feature for the objects segmentation [47–49]. The geometrical traits are defined comprehensively using various feature descriptors derived point-wise in ALS data and they are very effective in ML and DL models training and objects segmentation, such as vegetation [32]. In this section, these feature descriptors are discussed in detail with their properties for distinguishing various object types present in the point cloud data.

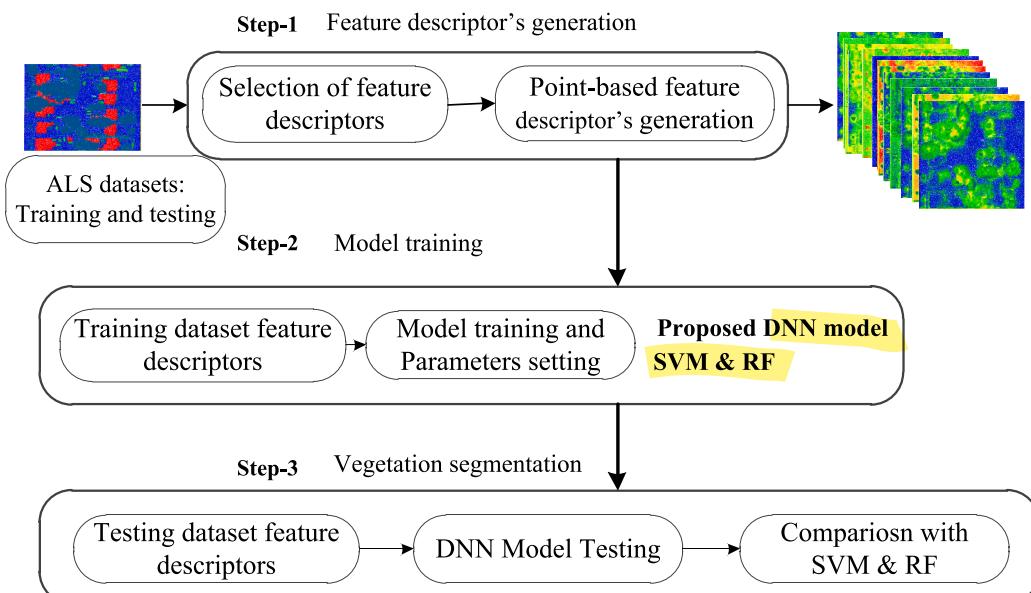


Fig. 3. Workflow of proposed methodology for vegetation cover segmentation.

The ALS points are represented by set \mathbf{P} , where each point is $p_i = (x_i, y_i, z_i) \in \mathbb{R}^3$, $i = 1, \dots, n$. The i and n represent an index of an ALS point and total points in the input data, respectively. These points are embedded in 3D Euclidean space and they are representing the actual object's geometry that is shape and size. Therefore each ALS point as seed point and its neighbourhood are considered for the computation of feature descriptors [48]. The spherical neighbourhood is selected around the seed point, where the radius of the sphere called as scale r_j is a critical parameter for the characterization of object geometry and its variation among the several scene objects [50]. A single scale r_j for $j = 1$ is selected for defining the generalized local neighbourhood, and handling the variations in point density for the geometric characterization of major object classes, such as vegetation [49]. As the objective of proposed methodology is to segment the point cloud of semi-urban and urban fabric in vegetation and non-vegetation objects, the characteristics of local neighbourhood is analysed using multiple scales. The multiple scales r_j for $j = 1 \dots m$ are needed to define the local neighbourhood to characterize the varying shapes and sizes of vegetation present in the geospatial scene where variation in vegetation species within small pocket trees is quite common [50]. In multiple scales selection, the number of scales m and scale intervals $|r_{j+1} - r_j|$ are important and they are decided heuristically [47].

There are many feature descriptors, which have been used to characterize the object's geometry and differentiate the objects based on their geometry. In the proposed methodology, nineteen descriptors are selected through comprehensive literature review and preliminary experiments considering their performance for segmentation of vegetation and non-vegetation points in ALS point cloud. The selected descriptors can be grouped into four categories: (i) surface-based, (ii) moment-based, (iii) PCA-based, and (iv) density-based feature descriptors. The Table 3 describes the importance of each feature descriptors in detail for the vegetation segmentation in ALS datasets. The computation of these descriptors is performed using each ALS point as seed and its local neighbourhood defined at various scales.

The surface-based feature descriptors include surface roughness and normal change rate, where surface roughness (S_{rgh}) discriminates smooth and rough surfaces, while normal change rate (N_r) evaluate surface curvature. The S_{rgh} is estimated by calculating the distance between the seed ALS point and the best fitting plane of its k neighbouring points within the spherical neighbourhood [50].

The second category that is moment-based, the first order moment (μ) is computed as a feature descriptor, where $\mu = \sum_{i=1}^k (p_i - \bar{p})/k$, the k is total points within the spherical neighbourhood around a ALS point p . The third category is PCA-based descriptors, which are generated by first computing the variance covariance matrix $\sum_{3 \times 3}$ using the points p_i for $i = 1 \dots k$ of local spherical neighbourhood of seed point and mean \bar{p} of k number of points (Equation (1) [54]. The characteristic equation $\det(\sum_{3 \times 3} - \lambda I_{3 \times 3}) = 0$ is solved to compute three eigenvalues λ_1, λ_2 and λ_3 ($\lambda_1 \geq \lambda_2 \geq \lambda_3$).

$$\sum_{3 \times 3} = \frac{1}{k} \sum_{i=1}^k (p_i - \bar{p})(p_i - \bar{p})^T \quad (1)$$

The equation $(\sum_{3 \times 3} - \lambda I_{3 \times 3})v_j = 0$ is solved for λ_j where $j = 1, 2, 3$ to compute three eigenvectors, which are termed as PCA1 (v_1), PCA2 (v_2), and PCA3 (v_3) corresponding to λ_1, λ_2 and λ_3 , respectively. Further, the normalized eigenvalues $\beta_i = \lambda_i / \sum_{i=1}^3 \lambda_i$ for $i = 1, 2$ and 3 are computed. For ALS point cloud vegetation segmentation, the effective PCA-based geometric feature descriptors are eigenvalues ($\lambda_1, \lambda_2, \lambda_3$) [47], eigenvectors (v_2, v_3) [53], and eigenvalues-derived descriptors. The normalized eigenvalues are used to compute eigenvalues sum ($E_{\text{sum}} = \lambda_1 + \lambda_2 + \lambda_3$ [48] and eigenentropy ($E_{\text{entropy}} = -\sum_{i=1}^3 \beta_i \times \ln(\beta_i)$ [49] as feature descriptors. The other eigenvalues-derived descriptors are Linearity (L_λ), Planarity (P_λ), Sphericity (S_λ), Anisotropy (A_λ), and Omnivariance (O_λ) [48,49,51], which are recorded into a vector F_{eigen}

Table 3

The geometric feature descriptors and their importance for objects differentiation in vegetation segmentation using ALS datasets.

Surface-based feature descriptors			
Feature descriptor	Symbol	Description	Reference
Surface Roughness	S_{rgh}	The surface roughness discriminates smooth surface objects such as road, building rooftop from rough surface objects such as vegetation.	[50]
Normal change rate	N_r	It discriminates the objects based on surface curvature such as curvature of vegetation is higher than the building rooftop and roads.	[46]
Moment-based feature descriptor			
1 st order moment	μ	It computes the information of mean location of point's distribution.	[48]
PCA-based feature descriptors			
1 st Eigen value	λ_1	It discriminates the linear structure from non-linear structure. For linear objects λ_1 is greater than the λ_2 and λ_3 .	[47]
II nd Eigen value	λ_2	It discriminates the planer structure from non-planer structure. For planer structure, λ_1 and λ_2 are nearly equal and more than λ_3 .	[47]
III rd Eigen value	λ_3	It discriminates the volumetric structure from non-volumetric structure. For volumetric structure: $\lambda_1 \approx \lambda_2 \approx \lambda_3$.	[47]
Eigen values sum	E_{sum}	The sum of λ_1, λ_2 and λ_3 is the Trace of $\sum_{3 \times 3}$ which reflects the invariant point number based on the eigenvectors of $\sum_{3 \times 3}$.	[48]
Eigenentropy	E_{entropy}	The eigenentropy represents a measure describing the order/disorder of 3D points within the local 3D neighborhood.	[49]
Linearity	L_λ	It discriminates the linear feature objects such as poles, tree trunks, etc. from other linear-linear objects such as buildings, low vegetation, etc. If in a neighbourhood, points are distributed along a line, then L_λ gets close to 1.	[48]
Planarity	P_λ	It discriminates the planner feature objects such as buildings top, roads from other objects such as poles, power lines, trees, etc. For perfect planar point's distribution the planarity P_λ is close to 1.	[48]
Sphericity	S_λ	It discriminates trees with variable trunk size, poles, vehicle, etc. from other non-volumetric objects. For perfect volumetric distribution, the Sphericity S_λ is close to 1.	[51]
Anisotropy	A_λ	It reflects linearity as Anisotropy A_λ is 1 when β_1 is 1 and β_3 is 0. Anisotropy A_λ is 0 for volumetric objects, where $\beta_1 = \beta_2 = \beta_3$.	[48]
Omnivariance	O_λ	It describes the local 3D structure around a seed ALS point. Omnivariance O_λ reaches its highest value: 1, when all eigen values are same.	[49]
Surface variation (change of curvature)	C_λ	It refers to the change of surface curvature that is high in the vegetation compared with non-vegetation objects.	[52]
PCA2	v_2	It provides the direction of points spread along which λ_2 is measured; like in case of building rooftop it represents the rooftop orientation.	[53]

(continued on next page)

Table 3 (continued)

Surface-based feature descriptors			
Feature descriptor	Symbol	Description	Reference
PCA3	v_3	It represents normal of planar point's distribution and used to distinguish the orientation of plane approximating the point's distribution.	[53]
Verticality	θ	It distinguishes the ground objects such as roads, bare land, parking lots, and building rooftop from the non-ground objects like poles, trees, and other vertical surfaces.	[47]
Density-based feature descriptors			
Volume density	V_d	It discriminates the volumetric and compact objects, such as trees from rest of the objects.	NA
Surface density	S_d	It discriminates the planner surface objects such as roads, parking lots, and building rooftops from the non-planar surface objects.	NA

(Equation (2)), where first, second, third, fourth, and fifth column represent L_λ , P_λ , S_λ , A_λ , and O_λ , respectively.

$$F_{\text{eigen}} = \left[\frac{\beta_1 - \beta_2}{\beta_1}, \frac{\beta_2 - \beta_3}{\beta_1}, \frac{\beta_1 - \beta_3}{\beta_1}, \sqrt[3]{\beta_1 \beta_2 \beta_3} \right] \quad (2)$$

The surface variation (C_λ) [52] and verticality (θ) are recorded into vector a vector $\mathbf{F}_{\text{structure}}$ (Equation (3), where first, and second column represent C_λ , and θ , respectively.

$$F_{\text{structure}} = \left[\frac{\beta_3}{\beta_1 + \beta_2 + \beta_3}, \arccos|v_3 \cdot \hat{a}_z| \right] \quad (3)$$

Where \hat{a}_z is unit vector along z -axis. These PCA-based geometric feature descriptors are summarized in Table 3, they play important roles in differentiating vegetation from non-vegetation objects [53].

In fourth category, the density-based feature descriptors: volume density (V_d) and surface density (S_d) are recorded into vector a vector $\mathbf{F}_{\text{density}}$ (Equation (4), where first, and second column represent V_d , and S_d , respectively. The $\mathbf{F}_{\text{density}}$ is used to differentiate the planar and non-planar volumetric objects, such as vegetation.

$$F_{\text{density}} = \left[\frac{k}{\left(\frac{4\pi R^3}{3} \right)}, \frac{k}{\pi R^2} \right] \quad (4)$$

The feature descriptors as summarized in Table 3 are generated using a sample ALS point cloud and shown in Fig. 4. Now in set \mathbf{P} for each point, the feature descriptor's value is assigned in place of z value that is calculated using the point as seed and its neighbourhood points. Now \mathbf{P} is updated to new set, where each point is represented as $(x, y, \text{feature descriptor's value})$, therefore total nineteen sets are formed. These feature descriptors distinctly define the scene objects point-wise in their local neighbourhood, and differentiate the objects types from each other.

The feature descriptors, which define the geometrical traits of the objects in the ALS dataset, are effectively utilized by the well-known ML methods: SVM and RF [38], and proposed 3D-MFDNN method for vegetation segmentation and these descriptors are used for the methods training and testing. The proposed 3D-MFDNN method's performance is compared with the SVM and RF for quantitative and qualitative performance assessment. In the following sections, the working of SVM and RF is summarized and then the proposed 3D-MFDNN architecture is explained for vegetation points' segmentation.

3.2. Machine learning methods

There are two well-known ML methods: SVM [40] and RF [55], which have been used for objects identification and classification using point cloud data. They are used for the tree point's segmentation in the ALS data using the selected feature descriptors as discussed in the previous section. SVM is a supervised classification technique based on a statistical theory to separate the class levels using the best line or best decision boundary (hyperplane) so that class labels are easily assigned to new data points. SVM algorithm finds the closest point of the lines from both the classes in binary separation. These points are called support vectors, where the distance between the vectors and the hyperplane is measured as margin. The optimal hyperplane is measured as hyperplane with maximum margin and optimal decision boundary (hyperplane) in a multi-dimensional space with maximum margin between marginal lines

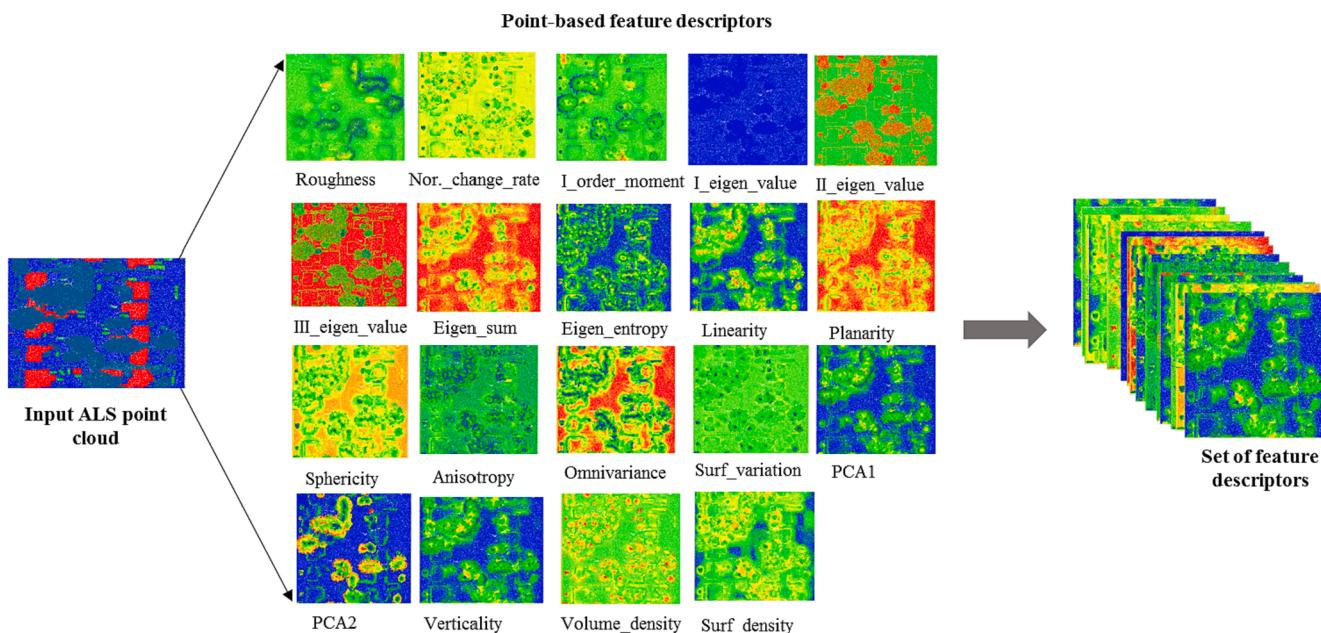


Fig. 4. The feature descriptors computed point-wise using input ALS point cloud dataset and for each input ALS point a color is assigned based on the corresponding descriptor value.

is searched. The marginal lines are the parallel hyperplanes that are optimal hyperplanes separating the data of different object classes (Fig. 5(a)). The orientation and position of the hyperplanes are controlled by support vectors which are the points close to or on the marginal lines. The hyperplane dimensionality is decided by the number of object classes, such as vegetation and non-vegetation. For two object features, the hyperplane is linear, and for three object features, a hyperplane is a plane [56].

The RF is an ensemble learning-based method, which uses forest of decision trees for the segmentation of vegetation and non-vegetation points based on the class votes from all decision trees (Fig. 5(b)). The technique uses the attribute bagging concept, which provides the model with a random subset of ALS training datasets and feature descriptors (Table 3). The concept of a large number of trees and attribute bagging of RF technique provide high classification gain and less sensitive to overfitting. The minimum number of trees, the number of feature descriptors at each split, the minimum number of nodes, sample size, etc., are essential parameters which are decided while RF-based method training.

The SVM and RF-based methods perform well, but their performances are limited in various cases, such as RF and SVM find challenging: (i) to handle missing data and data gaps, (ii) to segregate isolated objects and shrubs or ground-attached vegetation as vegetation class. SVM is more sensitive to noisy data, where object classes are overlapping [57] and it can be affected by the curse of dimensionality [58]. RF has certain notable drawbacks, such as difficulty in visualizing the trees as it uses several trees to make predictions [59], and split rules for classification are mysterious/unclear. Hence, RF is considered as a black-box type classifier [59]. Further, SVM has limited performance for large ALS datasets.

3.3. Proposed 3D-MFDNN architecture

A methodology framework: 3D-MFDNN based on deep neural network (DNN) using feature descriptors is proposed for vegetation segmentation from ALS data. The limitations of conventional ML methods are addressed in the proposed method. The DNN is characterized as an artificial neural network (ANN) with several hidden layers between the input and output. As shown in Fig. 6, the DNN architecture contains fully connected layers, where neurons in each layer are forwardly connected to all the neurons of the successive layers. A standard neural network model comprises three types of layers: one input, multiple hidden layers, and an output layer. A hidden layer consists of a set of neurons called nodes, and each neuron has a

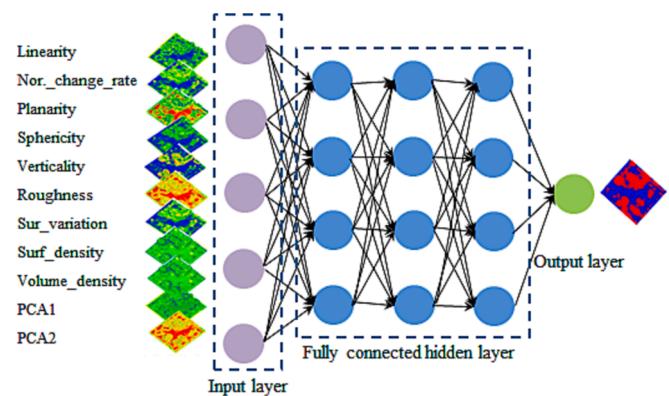


Fig. 6. A schematic diagram of 3D-MFDNN comprised of feature descriptors, input layer, hidden layers and output layer.

summation and activation function. The summation function integrates the received input with the help of correlated weights. Based on the integrated input, the activation function is applied to activate the neuron to pass the processed information to the next layer. The neurons used to maintain the operations are connected to the nodes of each layer by a set of correlational weights. The number of neurons in the input and output layer depends on the inputs given to the input layer and outputs taken from the output layer.

At an arbitrary neuron (p) of a hidden layer, the inputs are weighted output of previous layer neurons. The output (Y_p) of the neuron is calculated (Fig. 6) using the equation as follows:

$$Y_p = \sum_{i=1}^m W_{ip} \cdot O_i \quad (5)$$

Where m is total neurons in the previous layer, O_i and W_{ip} are the output, and weight function, respectively of previous layer i^{th} neuron.

The net output Y_p generated by summation function is passed through the activation function f that transforms the information in Y_p and gives controlled output $O_p = f(Y_p)$ to the subsequent layer neurons as an input. The activation function f determines whether or not a neuron should be stimulated based on weighted sums of inputs with added bias.

The main objective of the activation function is to introduce non-linearity in the output of neurons. Several activation functions are available; where some of them are ReLU, sigmoid, Tanh, Leaky ReLU, Parametric ReLU. In the proposed method, the ReLU is used as an activation function to maintain the non-linearity in output of neurons in

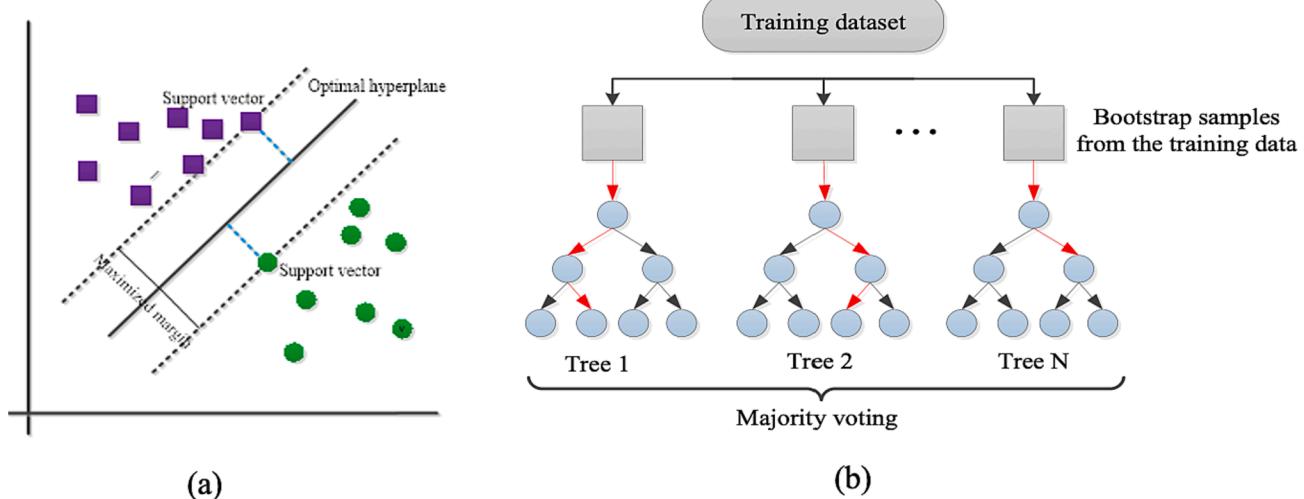


Fig. 5. Pictorial representation of how (a) SVM and (b) RF work to segment the data into meaningful object's segments, such as vegetation.

hidden layers and a sigmoid activation function is used in the output layer for their binary nature. ReLU activation function (Equation (6)) and Fig. 7(a)) in hidden layers induces non-linearity in each neuron's processed output and controls the passing information to next layer neurons.

$$f(x) = \max(0, x)f(x) \quad (6)$$

The sigmoid activation function is selected in the output layer and it is generally preferred for the binary classification problems [34] such as segmentation of vegetation and non-vegetation points. The sigmoid activation function (Equation (7)) is shown in Fig. 7(b).

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

The output layer of 3D-MFDNN architecture produces the predicted output which is compared with the true value that is labelled points as vegetation and non-vegetation. Based on the difference between the predicted and actual results, the loss value is generated with the help of the loss function. In the proposed method, the binary cross entropy loss function is used that compares the predicted probabilities to actual class output, which can be either 0 or 1 [34]. On the other hand, this binary cross entropy function can be seen as reducing the difference between the distribution caused by the 3D-MFDNN method and the empirical distribution of training data [60]. An optimizer is used to update the weights and bias values to minimize the loss value. During back propagation, Adam optimizer updates the weight and bias by the iterative training and tuned the weight and bias by an optimized value. In forward propagation, the initial value of weight and bias value are assigned by an initializer that is normal initializer, which is used in the proposed method.

In the proposed 3D-MFDNN method, the DNN model training and testing are the two important steps for the vegetation point's segmentation. In this DNN-based model, the ALS data processing strategy for vegetation segmentation is completed in two phases, forward and backward propagation. In the forwarding propagation, a loss value is generated, while to minimize this loss value, weights and bias of the hidden layer are updated in backward propagation. Training of the model is completed with the help of backpropagation. The Adam optimizer is used while training to optimize the loss value and update the weights w_n (Equation (8)) and biases.

$$w_n = w_{n-1} - \eta \frac{\hat{m}_n}{\sqrt{\hat{v}_n + \epsilon}} \quad (8)$$

Where η is the learning rate and ϵ is very small (10^{-8}) constant value used to avoid vanishing gradient descent. Mean and uncertainty variance of stochastic gradient, are estimated by $\hat{m}_n = m_n / \sqrt{1 - \beta_1^n}$, $\hat{v}_n = v_n / \sqrt{1 - \beta_2^n}$.

The proposed 3D-MFDNN method has several hyper parameters, which are finalized while training for the vegetation points' segmentation. The main hyper parameters, which are tuned are number of hidden layers, number of neurons in each hidden layers, and epoch. After completing the 3D-MFDNN model training for vegetation point's segmentation, the model testing is conducted to evaluate the performance of 3D-MFDNN model.

3.4. Model's performance evaluation:

The performance evaluation of ML methods and proposed 3D-MFDNN method is conducted using the evaluation metrics, such as precision, recall, accuracy and F1-score [26]. Proposed method is also compared with another popular DL model: PointNet++ [61] that is based on classic U-Net structure: the encoding and symmetric decoding parts. To compute the evaluation metrics for ALS data segmentation into vegetation and non-vegetation classes, that is binary classification a confusion matrix is generated. The confusion matrix is formed using the parameters: true positive (TP), true negative (TN), false positive (FP), and false negative (FN); which are computed comparing the segmented ALS data into vegetation and non-vegetation points with their ground truth value that is pre-labelled ALS data. The parameter TP is the points correctly segmented as tree points, while FN is the non-tree points, which are correctly segmented. FP is the number of non-tree points incorrectly segmented as tree points, while FN is the tree points, but segmented as non-tree points.

The confusion matrix parameters: TP, TN, FP, and FN are used to compute the evaluation metrics; precision (Equation (9)), recall (Equation (10)), accuracy (Equation (11)) and F1-score (Equation (12)). These metrics assist to evaluate the performance of the proposed 3D-MFDNN method and its comparison with the ML methods. The precision delivers the percentage of correctly segmented vegetation points (i.e. target object's points) among the segmented vegetation points as method's result, while the recall represents the percentage of correctly segmented vegetation points with reference to ground truth of

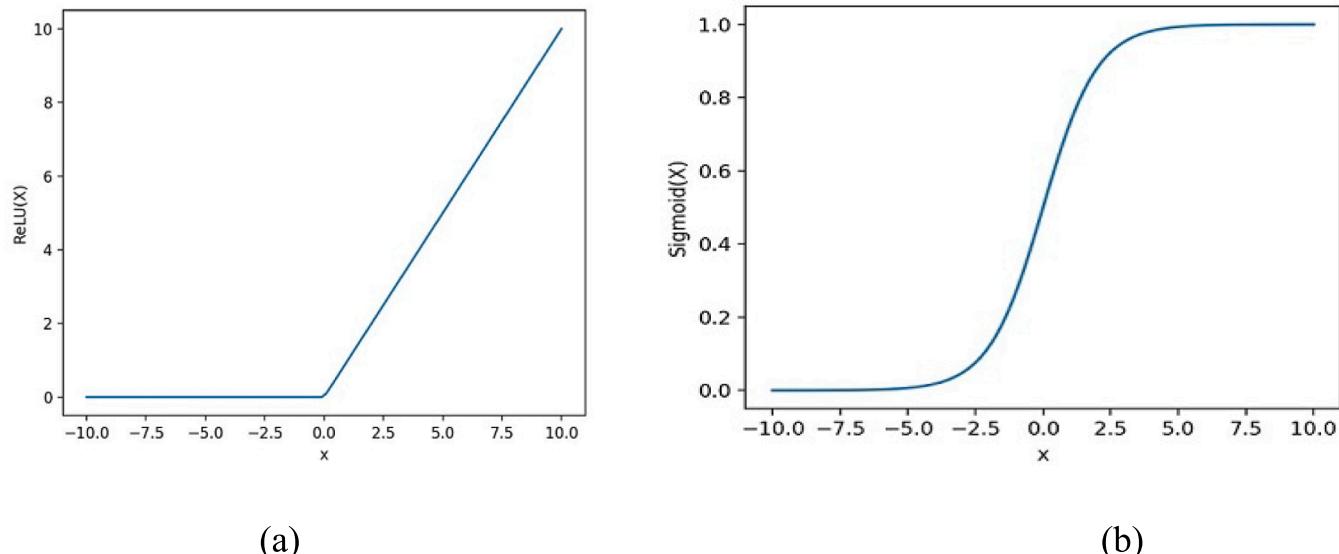


Fig. 7. Activation function graph: (a) ReLU, and (b) Sigmoid.

vegetation. The accuracy delivers overall performance in terms of correctly segmented vegetation and non-vegetation points among their ground truth. F1-score gives a better measure of the incorrectly segmented points than the accuracy measure. Accuracy is used when the true positive and true negative are more important while F1-score is used when the false negative and false positive are crucial.

$$\text{Precision (\%)} = \frac{TP}{TP + FP} \times 100 \quad (9)$$

$$\text{Recall (\%)} = \frac{TP}{TP + FN} \times 100 \quad (10)$$

$$\text{Accuracy (\%)} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (11)$$

$$\text{F1-score (\%)} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \quad (12)$$

4. Experiments

4.1. Parameter setting and DNN-based model training

In this section, the selection of various parameters required for the implementation of proposed 3D-MFDNN method and other models used for the comparison is discussed. In the proposed 3D-MFDNN method, the point-based feature descriptors are generated for each point of ALS dataset considering its spherical neighbourhood. The spherical neighbourhood in terms of scale parameter (i.e. radius of spherical neighbourhood) needs to be optimally selected. The several scale parameters: R3 = 3 m, R5 = 5 m, and R3R5 = both 3 m & 5 m were chosen considering the canopies diameters and optimal value was selected based on the experiments (Fig. 8). The performance of proposed 3D-MFDNN method was compared with RF, SVM, where their best performances were achieved in terms of precision, recall, F1-score, and accuracy using both scale parameters R3 and R5 together as shown in

Fig. 8. Further, the proposed method was compared with another DL model: PointNet++. The two parameters: (i) radius (R) around centroids, which are identified as most dominant points to represent 3D structural shape in their neighbourhood; and (ii) associated hyperparameter K that is number of points around the centroids govern the PointNet++ model's performance. Therefore, their tuning was carried out through several iterative experiments and model's best performance was achieved at $R = 0.35\text{m}$ and $K = 16$. The other hyperparameters: learning rate, epochs, batch-size, optimizer, and loss function were optimally chosen 0.001, 32, 4096, Adam, Null loss respectively as a set of best performing hyperparameters after several experiments.

While the process of 3D-MFDNN model training, the model parameters, such as weights and bias of neurons of each layer are tuned, but the important hyper parameters are adjusted based of several experiments on training datasets of the study area and evaluating the trained 3D-MFDNN model performance using the evaluation metric that is accuracy (Equation (11)). The hyper parameters to be fine-tuned for the best performance of 3D-MFDNN model are number of hidden layers (L), number of neurons for hidden layers (N), and number of epochs (ep). The several experiments are performed using the various combinations of L , N , ep , where L is chosen 1–10, N between 10 and 100, and three epochs: 100, 150, and 200. In the experiments, two scale parameters: R3 and R5 are selected for the computation of feature descriptors. Three cases are formed using the scale parameter as discussed before, in the first case only R3 is used, in second case R5, and in the third case using both R3 and R5 scales, the feature descriptors are generated. The selection of scale parameter is limited to the generation of feature descriptors (Table 3), which are the best possible geometric representors of the scene objects. Through the comprehensive experiments using the labelled sample testing dataset from the study area, the optimal hidden layers L are decided and they are 5–7, and number of neurons N for each hidden are 50–70. In Fig. 9, the accuracies for the several combinations of R3, R5, L, N, and ep are presented using radar diagram, where it is depicted that for R3R5L5N70ep150 combination, the accuracy of the

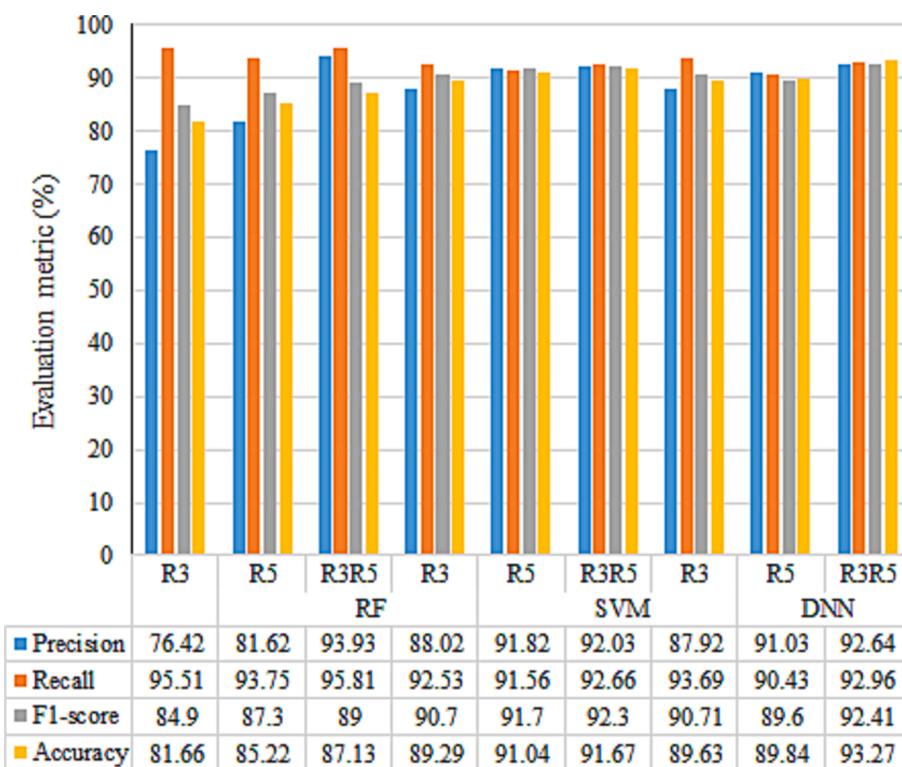


Fig. 8. Graph and table showing variations in the evaluation metrics, which are generated by the processing of training datasets using RF, SVM and proposed 3D-MFDNN method at scale parameters: R3 = 3 m, R5 = 5 m, and R3R5 = both 3 m & 5 m.

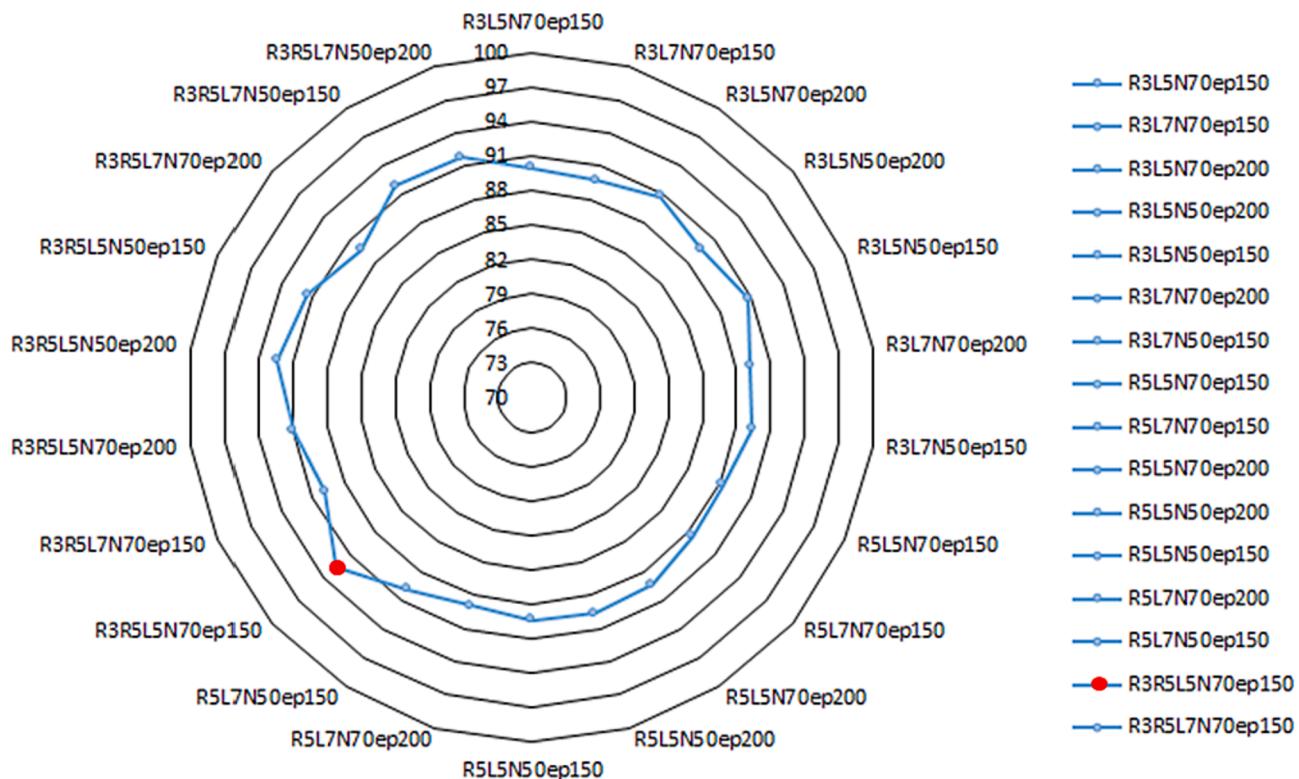


Fig. 9. 3D-MFDNN model performance (i.e. radar diagram), which is evaluated in terms of accuracy (i.e. varies from 88.27 % to 92.58 %) of tree points segmentation using several combination of scales and hyper parameters for their tuning and selection of best performing combination that is R3R5L5N70ep150.

3D-MFDNN model for tree points segmentation is highest, that is 92.58 %. The coding scheme followed for the scale and hyper parameters combination, such as R3R5L5N70ep150 states that both R3 and R5 scales are used and L5, N70, ep150 means $L = 5$, $N = 70$, $ep = 150$, respectively. The uses of two scales R3 and R5 lead to the computation of 2×19 feature descriptors corresponding to each point and its spherical neighbourhood. In Fig. 10, the proposed 3D-MFDNN model performance is presented in terms of accuracy and loss with change in number of epochs.

4.2. Results

The proposed methodology was implemented using Python programming language and its associated libraries. The six testing datasets selected from various tiles of ALS DALES dataset were processed using the proposed trained 3D-MFDNN method and the method's overall performance was reported in terms of precision, recall, F1-score and

accuracy. The data processing was performed on personal computer with configuration: Xeon Processors @2.3 GHz, A100 Tensor Core GPU, 16 GB RAM. The evaluation metrics: precision, recall, and F1-score were used for emphasizing method's performance in tree point's segmentation, while the metric: accuracy presented the method's performance in segmentation of both tree and non-tree points. The Table 4 summarizes these accuracy metrics for all datasets #1 to #6.

4.2.1. Urban test sites

The four urban datasets #1, #2, #3 and #4 were processed using the proposed 3D-MFDNN method and tree points were accurately segmented and rest of the objects were segmented as non-tree (Fig. 11). In dataset #1, low vegetation and small buildings were predominant among all the non-ground objects, and they were accurately segmented as tree and non-tree, respectively (Fig. 11(a)). The buildings had gable roofs, which had close geometrical association to some of the low vegetation with wide tree crown, though they were accurately labeled as

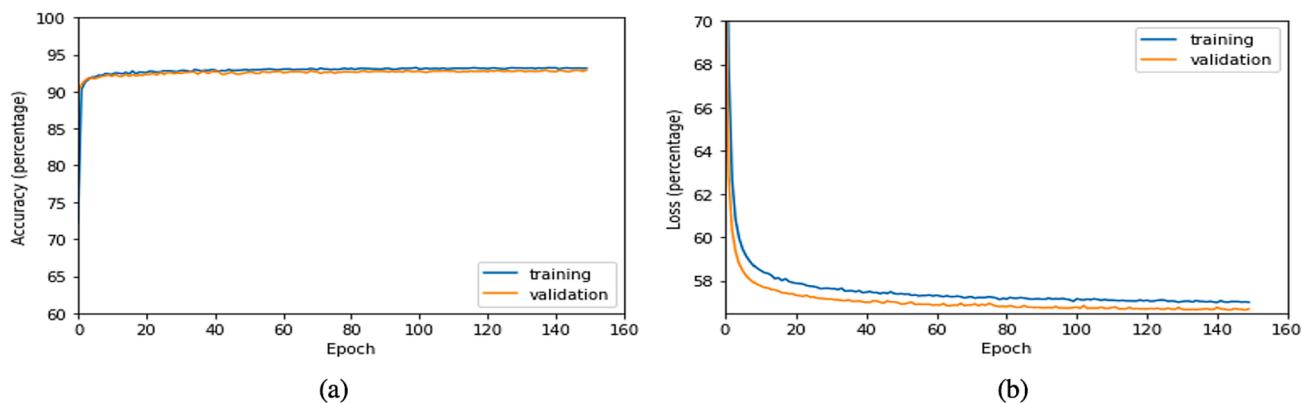


Fig. 10. 3D-MFDNN method's (a) accuracy and (b) loss, variations with change in epochs.

Table 4

Evaluation metrics, that is precision, recall, F1-score and accuracy in tree points segmentation of datasets #1 to #6.

Dataset →	#1	#2	#3	#4	#5	#6	Average (#1 to #6)
Precision (%)	74.73	79.83	71.45	80.14	83.31	96.61	81.01
Recall (%)	92.11	82.98	77.81	88.87	89.57	92.84	87.36
F1-score (%)	82.52	81.34	74.50	84.28	86.32	94.65	83.94
Accuracy (%)	91.22	90.22	90.42	91.24	93.50	96.16	92.13

non-tree. The trees points' segmentation was achieved at F1-score of 82.52 %, while the accuracy was 91.22 %. The few geometric parts of trees and buildings at their boundaries cannot segmented accurately. The industrial area with large warehouses, gable roof residential buildings and trees with varying height and crown diameter were acquired in dataset #2, the trees were accurately segmented, and warehouses and gable roofs were separated from tree as non-tree points (Fig. 11(b)). Trees points very close to the power line and house were readily segmented and isolated from non-tree objects. The F1-score and accuracy were 81.34 %, and 90.22 %, respectively. At the overlapping boundaries of vegetation and non-vegetations objects, some parts of trees are falsely segmented as non-tree and vice versa.

Dataset #3 represented highly variable and partially overlapped vegetation cover having trees with various shape and size, which were accurately segmented as tree (Fig. 11(c)). The power lines crossed over the small trees and located very close to the high vegetation were accurately separated as non-tree points. 74.50 % and 90.42 % were achieved as F1-score and accuracy, respectively of tree point's segmentation. In dataset #4, dense houses were surrounded by different shapes and sizes of trees, where tree and non-tree points were accurately segmented with accuracy of 91.24 % (Fig. 11(d)). The objects overlapping was predominant in this test site, such as tree branches were connected and hanging over the gable roof; trees were overlapping,

though proposed method distinguished tree and non-tree points in such cases. The F1-score achieved 84.28 %.

4.2.2. Semi-urban test sites

The datasets #5 and #6 were from the semi-urban environment, where the trees were overlapped clustered together along the road and at several locations, these clustered trees were accurately segmented. In dataset #5, the cluster of high vegetation covered the understory low vegetation, though the vegetation was segmented accurately as tree from rest of the objects (i.e. non-tree) (Fig. 12(a)). Vegetation at sloping surface was differentiated from ground surface as tree and low lying objects, such as vehicles under the trees were segmented as non-tree. In this dataset, the accuracy was 93.50 % and F1-score achieved 86.32 %. In dataset #6, a trees cluster encircle a water body and trees were on the sloping bank and two other trees clusters were densely populated, where trees were accurately segmented (Fig. 12(b)) and F1-score and accuracy were reported 94.65 % and 96.16 %, respectively.

4.3. Discussion

The proposed 3D-MFDNN method combines point-wise computed multiple feature descriptors as summarized in Table 3. These feature descriptors are generated efficiently on different scales R3 and R5, which are used to define the 3D neighbourhood and 3D-MFDNN method is trained for vegetation points segmentation. This method is straightforward to implement, thus it has the potential for broader use in the urban vegetation mapping and forestry application. It is a data-driven methodology based on point-wise feature descriptors that works accurately for all types of vegetation and does not require point cloud-derived models, like DEM, DSM, CHM. Ground filtering is not required as an initial step before the vegetation segmentation in the proposed methodology unlike many published works, where accurate ground filtering is a key challenge. In this section, the performance of the proposed methodology is analysed quantitatively and qualitatively. The parameter sensitivity analysis and model stability in 3D-MFDNN methodology are also discussed in this section.

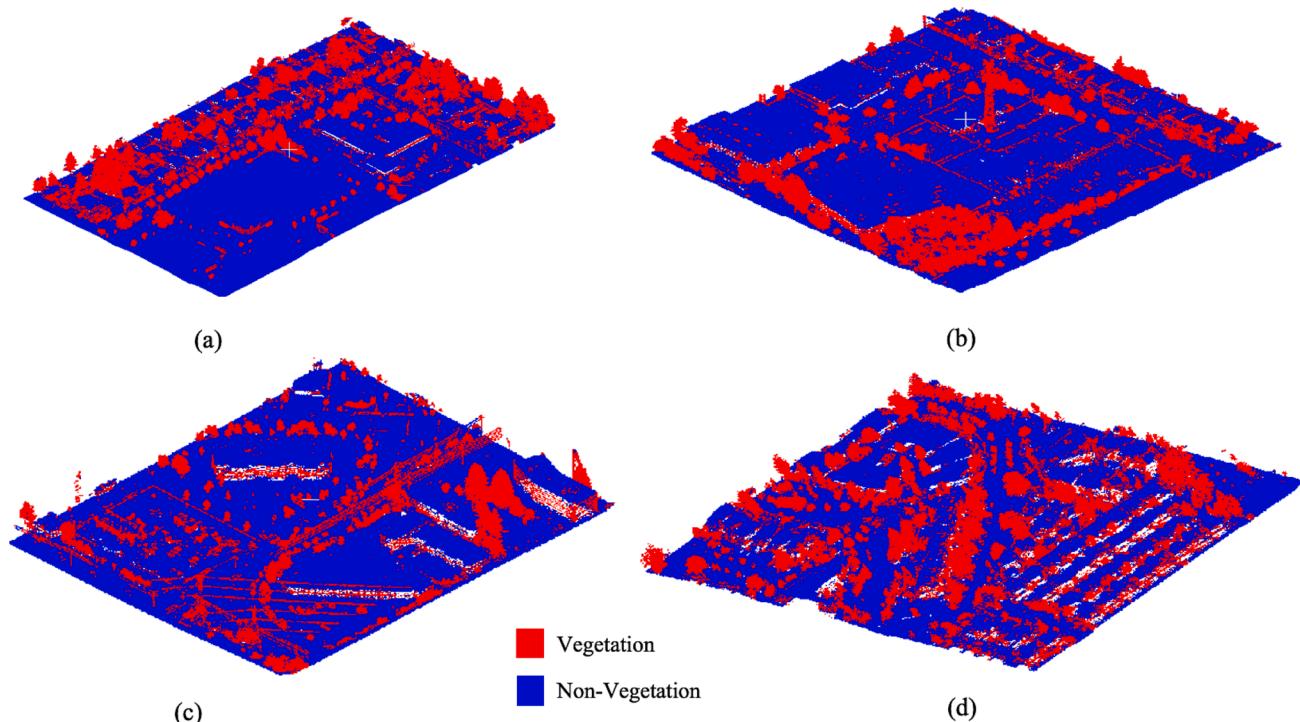


Fig. 11. Perspective three-dimensional views of vegetation (red color) and non-vegetation objects (blue color) in ALS dataset (a) #1 from urban test site, (b) #2 from urban site having mixed types of trees and buildings, (c) #3 from urban site with mixed types of trees and PL/TL lines, and (d) #4 from urban residential site.

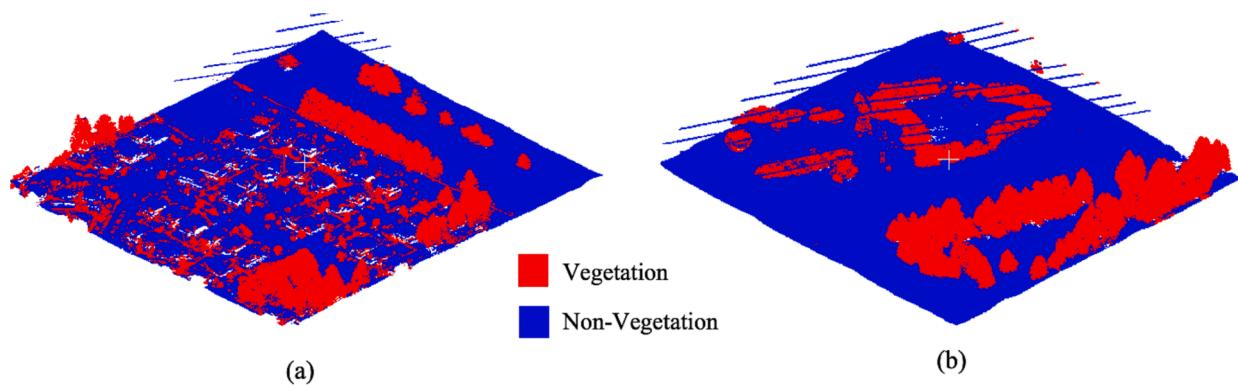


Fig. 12. Perspective three-dimensional views of vegetation (red color) and non-vegetation objects (blue color) in ALS dataset (a) #5 from semi-urban test site, (b) #6 from semi-urban site having water body.

4.3.1. Quantitative performance analysis and comparison

The quantitative evaluation of vegetation point's segmentation in six different datasets is shown in Table 4. The average, F1-score of 83.94 % and accuracy of 92.13 % were achieved in these six datasets. This quantitative evaluation emphasizes, the proposed 3D-MFDNN method adopts the test sites variations (Table 2) and delivers the vegetation and non-vegetation points' segmentation accuracy over 90 %. Datasets #5 and #6 had vegetation on gentle slope with common special features, such as dense and overlapped vegetation on sloping surface and hanging objects (i.e. wires). In these datasets, the average F1-score and accuracy were over 90 %.

The proposed 3D-MFDNN method utilizes the important feature descriptors (Table 3) generated for each ALS point for the vegetation point's segmentation. These descriptors, which are categorized as surface-based, moment-based, PCA-based, and density-based feature, distinctly define and differentiate the various objects' points, such as vegetation. The proposed method is compared with two popular machine learning methods: SVM and random forest, which are also, trained using same set of feature descriptors as used in proposed 3D-MFDNN method. A popular DL model: PointNet++ was also used for the comparison with the proposed method. The performance comparison is done using the evaluation metrics: precision, recall, F1-score and accuracy

computed on datasets #4 and #5, which are representing a complex scene environment to be dealt with among all six datasets (Table 2). The dataset #4 represents an urban environment (Fig. 13), while dataset #5 is semi-urban (Fig. 14) and both datasets have several complex and special cases to be dealt for vegetation point's segmentation.

The overall performance of proposed 3D-MFDNN method outperforms the SVM, RF and PointNet++ methods in these two datasets (Table 5). In terms of F1-score, the proposed method perform better than SVM, RF and PointNet++ methods with F1-score equal to 81.70 %, 81.30 % and 81.70 % respectively in dataset #4. In dataset #5, the F1-score is 85.20 %, 85.90 %, and 85.68 % for SVM, RF and PointNet++ methods, while it is better (i.e. 86.32 %) in case of proposed method. The low vegetation, that is shrubs-like decorated small tree attached with ground successfully segmented as vegetation by the proposed method, while SVM and RF cannot segregate them from the ground surface and segment them as non-vegetation. The small trees connected with the houses are segregated by the proposed method, while SVM and RF segment them including small tree as non-vegetation. The several inner points of large tree are labelled as non-vegetation by RF, while proposed method performs better and segment high vegetation points accurately. Similarly in case of vegetation segmentation from sloping surfaces, the proposed method performs better than SVM, RF and PointNet++. The

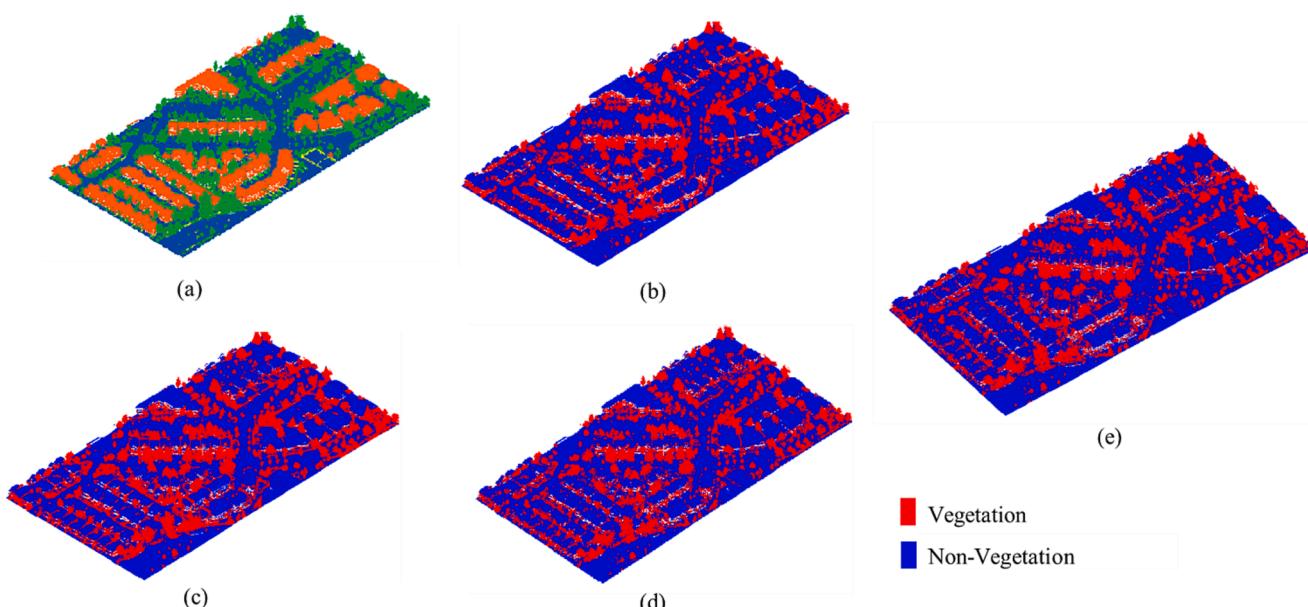


Fig. 13. Three-dimensional perspective view of (a) input ALS dataset #4, where colours are assigned to the points object class-wise; and corresponding outputs generated by: (b) proposed 3D-MFDNN method, (c) SVM, (d) RF, and (e) PointNet++ models, where vegetation and non-vegetation objects are shown in red and blue color, respectively.

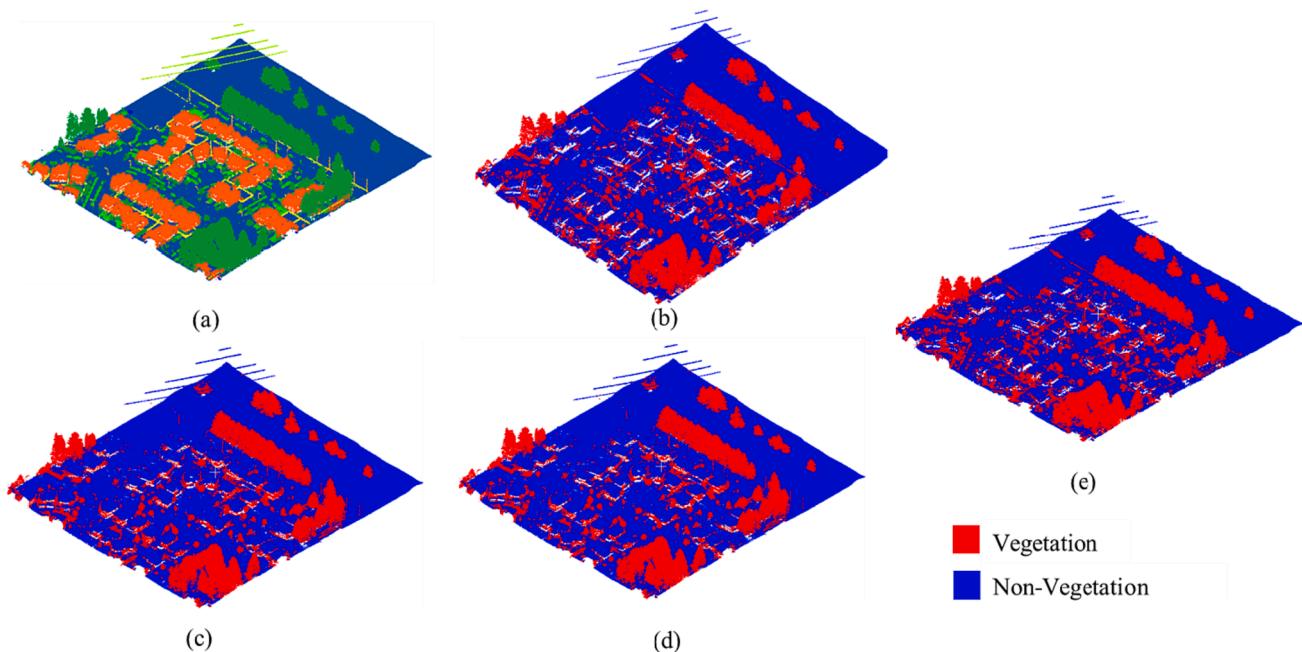


Fig. 14. Three-dimensional perspective view of (a) input ALS dataset #5, where colours are assigned to the points object class-wise; and corresponding outputs generated by: (b) proposed 3D-MFDNN method, (c) SVM, (d) RF, and (e) PointNet++ models, where vegetation and non-vegetation objects are shown in red and blue color, respectively.

Table 5

Accuracy metrics computed on the ALS DALES dataset for vegetation point's segmentation using well-known ML methods: RF & SVM, a popular DL method: PointNet++, and proposed 3D-MFDNN method.

Site	Indices →	TP	TN	FP	FN	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Dataset#4	RF	765,258	2,100,720	266,878	84,546	74.14	90.05	81.30	89.07
	SVM	696,760	2,208,641	158,957	153,044	81.42	81.99	81.70	90.30
	PointNet++	760,345	2,117,384	250,214	89,459	75.24	89.38	81.70	89.44
	3D-MFDNN	755,204	2,180,460	187,151	94,600	80.14	88.87	84.28	91.24
	RF	616,721	2,204,995	126,714	76,166	82.95	89.00	85.90	93.28
	SVM	607,582	2,206,222	125,487	85,305	82.88	87.69	85.20	93.03
	PointNet++	618,835	2,205,588	126,121	74,052	83.07	88.45	85.68	93.38
	3D-MFDNN	620,604	2,207,351	124,358	72,283	83.31	89.57	86.32	93.50

several points from linear feature objects like pole and power line are labelled as vegetation by SVM, while they are accurately assigned as non-vegetation by the proposed method. The accurate segregation of tree and building points, where they are overlapped and accuracy of low vegetation segmentation are better in case of proposed method compared with the PointNet++. One of reasons for proposed 3D-MFDNN method outperforms the PointNet++ is the important feature descriptors selected by the user govern the training of 3D-MFDNN method unlike PointNet++, where feature generation is internal and automatic process and user cannot control it. The selection of important feature descriptors, such as surface-based, moment-based, PCA-based, density-based descriptors is based on their importance for differentiating vegetation from other non-vegetation objects point-wise in ALS datasets (Table 3). These feature descriptors comprehensively train the 3D-MFDNN model for vegetation segmentation by defining the object classes' geometrical and statistical differences in 1D, 2D and 3D. Further, the 3D-MFDNN method effectively handles the complex and challenging scenes and performs well compared with the PointNet++ by selecting the specific feature descriptors, which have capability to accurately discriminate the object class types in case of objects overlap, occlusions, and data gaps. The training of 3D-MFDNN model is less complex compared with the PointNet++ as the large and versatile training datasets are required in case of PointNet++ model training,

where internal features are generated from the point coordinate (x , y , z) only.

4.3.2. Qualitative performance analysis

The proposed 3D-MFDNN method is robust to the generalized as well as complex ALS datasets with connected and overlapped objects having non-uniform and irregular distribution. It works successfully for various types of scene environment (Figs. 11, 12, 15 and Tables 2, 4). The proposed method performance in several challenging cases is summarized in this section.

The vegetation cover with overlapped trees becomes a challenging case for accurate vegetation point's segmentation, where low and high vegetation are mixed together. As shown in Fig. 15(a), the proposed method works satisfactorily to identify the tree point's cluster having both low and high vegetation. It is shown in this figure, a small hut roof elevated from the ground and touches the tree branches is successfully differentiated from vegetation and segmented as non-vegetation. The low-lying plants on the sloping surface, where the plant's lower part is very close to ground, form the volumetric shape including the ground surface (Fig. 15 (b)). In this geometric shape, the plant's height gradually decreases and ground surface height gradually increases away from the geometric centre, which resulted into the volumetric shape. From this geometrical structure, the vegetation points are accurately

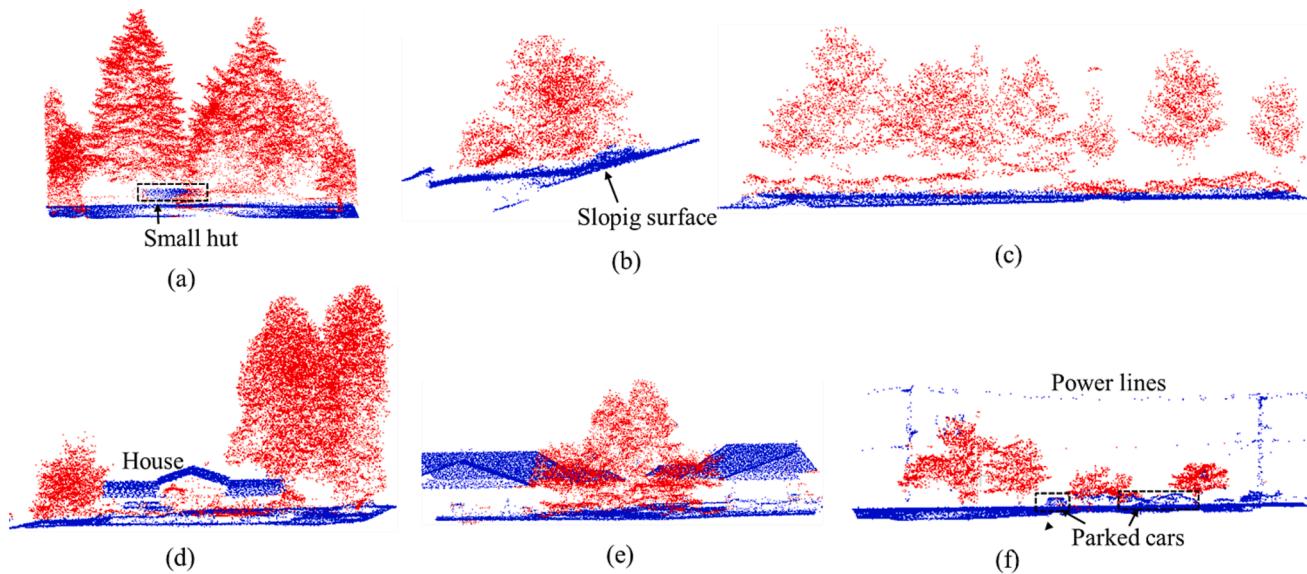


Fig. 15. Specific terrain cases: (a) tree point's cluster having both low and high vegetation, (b) low-lying plants on the sloping surface, (c) low vegetation surrounds the trees, (d) house surrounded by trees and low vegetation, (e) tree with understory low vegetation sandwiches between two gable roof houses, and (f) power lines pass very close to or over the tree's branches and low lying objects.

segmented out from the ground points. The low vegetation surrounds the trees and spreads along particular direction is an example of decorated trees used as boundary, where layer of low vegetation is just below the trees (Fig. 15 (c)). Proposed method successfully segments the low vegetation and trees from the ground surface. A case of large tree having volumetric shape and its lower crown hanging close to the sloping ground surface is present in dataset #1, where tree points accurately segmented and removed from ground.

The specific cases, where vegetation, such as trees are connected or overlapped with other objects, such as buildings, powers lines, etc. pose challenges in vegetation points' segmentation. In Fig. 15 (d) a house is surrounded by trees and low vegetation. At many locations tree's branches are touching the house's roof, through the proposed method segregated them and accurately segmented the vegetation's points. A case of tree with understory low vegetation sandwiches between two gable roof houses is shown in Fig. 15 (e), where tree branches are touching the sloping downward gable roof. In this case the vegetation's points are successfully segmented despite the geometrical complexity. The power lines pass very close to or over the tree's branches and low lying objects, such as cars parked under tree are considered as a complex case to deal with, though the proposed successfully segments the

vegetation's points and isolated the understory low lying objects and power lines as non-vegetation Fig. 15 (f).

4.3.3. Performance evaluation of scale parameter and feature descriptors

The scale parameter and its effect in the proposed 3D-MFDNN method's performance are evaluated. The scale parameter is the radius of spherical neighbourhood that is used to select the neighbouring points around each point in the ALS datasets. Since the feature descriptors of each point are generated with the help of its neighbouring points, the optimal selection of scale parameter is very crucial for the method's performance. Considering the method's objective, that is tree points segmentation, and variations in the canopy diameters, the scale parameter is chosen as single scale: R3, R5, and multi-scale: R3R5, where R3 = 3 m, and R5 = 5 m for the experiments. The F-score and accuracy values are highest, when both R3 & R5 are used for the multi-scale feature descriptors generation (Fig. 16(a)). Though the maximum deviations in the F-score and accuracy are 2.33 % and 3.04 %, respectively from their peak values at multi-scale parameters: R3R5, therefore the 3D-MFDNN method's performance is stable to the variations in the scale parameter.

Further, the performance of feature descriptors are evaluated by

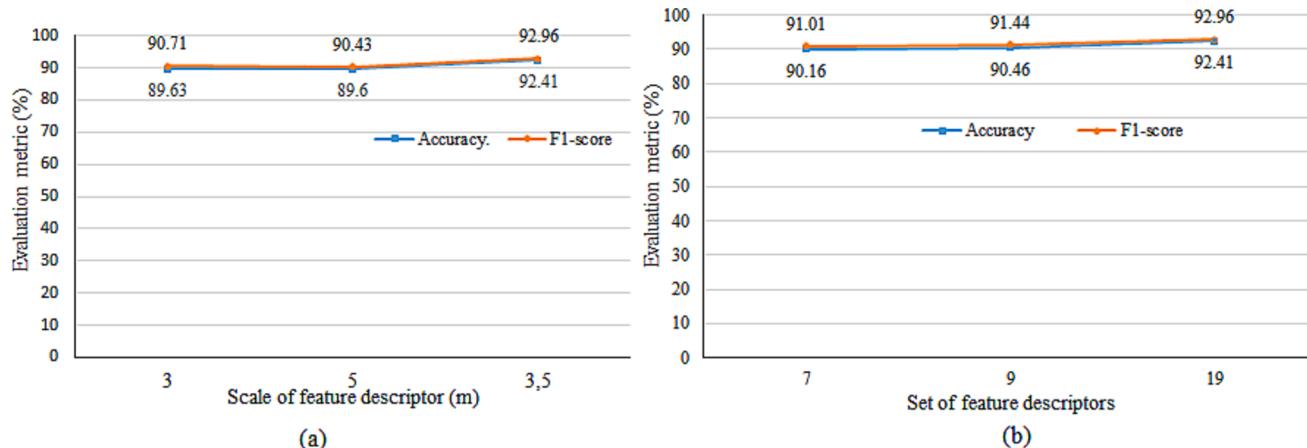


Fig. 16. Graphs showing variations in accuracy and F1-score with: (a) scale of feature descriptor (b) set of feature descriptors.

conducting the experiments using sets of feature descriptors comprised of 7, 9 and 19 descriptors (Fig. 16(b)). These sets are formed based on aggregating the best performing descriptors. The F1-score and accuracy are 91.01 % and 90.16 %, respectively when set of 7 feature descriptors (i.e. roughness, surface density, planarity, Ist order moment, linearity, sphericity, verticality) are used for proposed method implementation and testing. When two more descriptors are added (i.e., PCA2, PCA3) to make 9 feature descriptors set, the F1-score and accuracy are increased by 0.47 % and 0.33 %, respectively, while considering 19 feature descriptors set, where 10 more descriptors are added (i.e. normal change rate, volume density, eigen values sum, omnivariance, eigenentropy, anisotropy, surface variation, Ist eigen value, IInd eigen value, IIIrd eigen value), the F1-score and accuracy are increased to 92.96 % and 92.41 %, respectively. The improvement in F1-score and accuracy are 2.14 % and 2.5 %, respectively when feature descriptors are changed from 7 to 19 in feature sets.

4.3.4. Sensitivity analysis

The deep learning models, such as proposed 3D-MFDNN method needs to be tested for the sensitivity analysis of their hyper parameters, which regulate the model's learning process and directly affect model's parameters such as weights and biases. The sensitivity analysis of the hyper parameters, such as hidden layers, and neurons decides the overall method's architecture stability and performs fine-tuning of these hyper parameters (Fig. 17).

In the proposed 3D-MFDNN method, the hyper parameters: hidden layers and neurons are varied from 1 to 10 & 10–100, respectively and proposed method's evaluation metrics: F1-score and accuracy are computed using the validation datasets. In the first case (Fig. 17(a)), keeping all parameters constant, the hidden layers are varied from 1 to 10 and the corresponding accuracy obtained by the proposed method's testing increases from 89.12 % to 94.35 % (at hidden layers = 5), which is highest and further declines to 90.38 %. Similarly, the accuracy is highest at number hidden layers equal to 5. The maximum deviations in F1-score and accuracy are 7.03 % and 5.54 %, respectively from their peak values, which states that the proposed 3D-MFDNN method is not very sensitive to the change in the number of hidden layers and it is considered as a stable method. Further, this sensitivity analysis approach helps to fine tune the number of hidden layers, which optimal value is 5 and it delivers the best tree point's segmentation results. In the second case (Fig. 17(b)), the numbers of neurons are varied from 10 to 100 and results are evaluated in terms of F1-score and accuracy keeping all other parameters constant (i.e. hidden layers = 5). The highest F1-score and accuracy are achieved at numbers of neurons is equal to 70. The maximum deviation in the accuracy from its peak value is not too high; it is less than 10 % even in wide range of changes in numbers of

neurons from 10 to 100.

5. Conclusions and future work

In this paper, the proposed methodology effectively segments vegetation points from input ALS datasets having different levels of scene complexity and heterogeneity. The overall methodology consists of three main steps. The first step is generation of feature descriptors, which are derived point-wise using training datasets. The chosen feature descriptors characterize and differentiate the scene objects comprehensively and they are categorized as surface-based, moment-based, PCA-based and density-based descriptors. In second step, the training data-derived feature descriptors are used to train proposed 3D-MFDNN method and its parameter setting and same set of descriptors are also used for SVM and RF-based methods training. In third step, the proposed 3D-MFDNN method is tested on ALS dataset and its performance is compared with SVM, RF and PointNet++ methods in terms of evaluation metrics: Precision, Recall, F1-score, and Accuracy. The proposed methodology was tested on six ALS datasets from different test sites, such as urban, and semi-urban having various levels of scene complexities. In these test sites, the vegetation segmentation performance of 3D-MFDNN method is reported as average F1-score and accuracy of 83.94 % and 92.13 %, respectively.

The proposed 3D-MFDNN method's results in these six different test sites along with thorough analysis of existing methods are used to conclude that the proposed method successfully works to segment accurately the vegetation points from ALS data of generalized roadway scene as well as in presence of challenging cases of scene features in vegetation segmentation, such as: (i) trees overlapped or connected with other objects, (ii) tree with understory low vegetation, (iii) low-lying plants on the sloping surface, (iv) power lines pass very close to or over the tree's branches, (v) objects like cars parked under the tree, (vi) large tree with volumetric shape and branches hanging near sloping ground surface, etc. Future work will focus on classification of vegetation cover types and automatic dendrometry of tree cover using ALS datasets.

CRediT authorship contribution statement

Dheerendra Pratap Singh: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Manohar Yadav:** Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision.

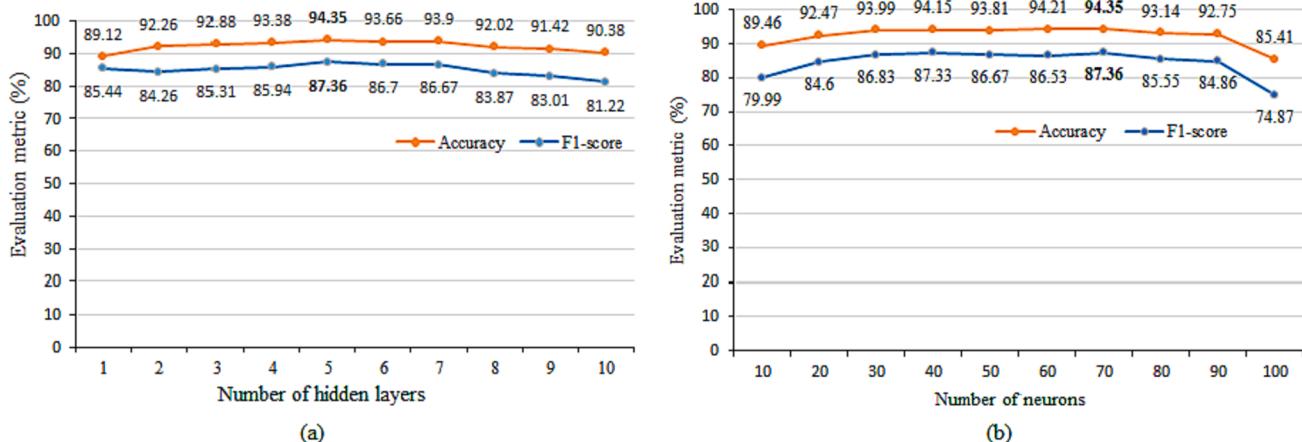


Fig. 17. Graphs showing variations in evaluation metrics: F1-score and accuracy in case of (a) change in number of hidden layers from 1 to 10, while number of neurons = 70; (b) change in number of neurons from 10 to 100, while hidden layers = 5.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] M. Yang, Y. Mou, S. Liu, Y. Meng, Z. Liu, P. Li, W. Xiang, X. Zhou, C. Peng, Detecting and mapping tree crowns based on convolutional neural network and Google Earth images, *Int. J. Appl. Earth Obs. Geoinf.* 108 (2022), 102764.
- [2] E. Banzhaf, T. Arndt, J. Ladiges, Potentials of Urban Brownfields for Improving the Quality of Urban Space, 2018, pp. 221–232.
- [3] L.D. Philip, F. Emir, A.A. Alola, The asymmetric nexus of entrepreneurship and environmental quality in a developing economy, *Int. J. Environ. Sci. Technol.* 19 (8) (2022) 7625–7636.
- [4] D.J. Nowak, S. Hirabayashi, M. Doyle, M. McGovern, J. Pasher, Air pollution removal by urban forests in Canada and its effect on air quality and human health, *Urban For. Urban Green.* 29 (2018) 40–48.
- [5] D.J. Nowak, E.J. Greenfield, Declining urban and community tree cover in the United States, *Urban For. Urban Green.* 32 (2018) 32–55.
- [6] Sawyer (FIMMM), Gervais, Applied tree biology, *Int. Wood Prod. J.* 11(3) (2020) 162.
- [7] H.P. Borland, T.A. Schlacher, B.L. Gilby, R.M. Connolly, N.A. Yabsley, A.D. Olds, Habitat type and beach exposure shape fish assemblages in the surf zones of ocean beaches, *Mar. Ecol. Prog. Ser.* 570 (2017) 203–211.
- [8] Pucci, Flo, The Significance of the Urban Forest in the Urban Environment, 2020, <https://ucanr.edu/blogs/blogcore/postdetail.cfm?postnum=43772>.
- [9] Y. You, T. Chen, Y. Shen, Z. Wang, Graph Contrastive Learning Automated, 2021, <http://arxiv.org/abs/2106.07594>.
- [10] Y. Yang, H. Guan, M. Shen, W. Liang, L. Jiang, Changes in autumn vegetation dormancy onset date and the climate controls across temperate ecosystems in China from 1982 to 2010, *Glob. Chang. Biol.* 21 (2) (2015) 652–665.
- [11] A.A. Plowright, N.C. Coops, B.N.I. Eskelson, S.R.J. Sheppard, N.W. Aven, Assessing urban tree condition using airborne light detection and ranging, *Urban For. Urban Green.* 19 (2016) 140–150.
- [12] M. Liu, Z. Han, Y. Chen, Z. Liu, Y. Han, Tree species classification of LiDAR data based on 3D deep learning, *Measur.: J. Int. Measur. Confed.* 177 (2021), 109301.
- [13] X. Wang, X. Zhang, X. Ren, L. Li, H. Feng, Y. He, H. Chen, X. Chen, Point cloud 3D parent surface reconstruction and weld seam feature extraction for robotic grinding path planning, *Int. J. Adv. Manuf. Technol.* 107 (1-2) (2020) 827–841.
- [14] Z. Szabó, A. Schlosser, Z. Túri, S. Szabó, A review of climatic and vegetation surveys in urban environment with laser scanning: a literature-based analysis, *Geographica Pannonica* 23 (4) (2019) 411–421.
- [15] T. Yun, K. Jiang, G. Li, M.P. Eichhorn, J. Fan, F. Liu, B. Chen, F. An, L. Cao, Individual tree crown segmentation from airborne LiDAR data using a novel Gaussian filter and energy function minimization-based approach, *Remote Sens. Environ.* 256 (2021), 112307.
- [16] S. Schmohl, A. Narváez Vallejo, U. Soergel, Individual tree detection in urban ALS point clouds with 3D convolutional networks, *Remote Sens. (Basel)* 14 (6) (2022), 1317.
- [17] S. Yan, L. Jing, H. Wang, A new individual tree species recognition method based on a convolutional neural network and high-spatial resolution remote sensing imagery, *Remote Sens. (Basel)* 13 (3) (2021) 1–21.
- [18] S.A. Bello, S. Yu, C. Wang, J.M. Adam, J. Li, Review: Deep learning on 3D point clouds, *Remote Sens. (Basel)* 12 (11) (2020) 1729.
- [19] J. Serey, L. Quezada, M. Alfaro, G. Fuertes, M. Vargas, R. Ternero, J. Sabattin, C. Duran, S. Gutierrez, Artificial intelligence methodologies for data management, *Symmetry* 13 (11) (2021), 2040.
- [20] M. Bohanec, V. Rajkovic, Knowledge Acquisition and Explanation for Multi-Attribute Decision Making, 1988.
- [21] F. Matrone, E. Grilli, M. Martini, M. Paolanti, R. Pierdicca, F. Remondino, Comparing machine and deep learning methods for large 3D heritage semantic segmentation, *ISPRS Int. J. Geo Inf.* 9 (9) (2020), 535.
- [22] K. Liu, X. Shen, L. Cao, G. Wang, F. Cao, Estimating forest structural attributes using UAV-LiDAR data in Ginkgo plantations, *ISPRS J. Photogramm. Remote Sens.* 146 (2018) 465–482.
- [23] L. Windrim, M. Bryson, Detection, segmentation, and model fitting of individual tree stems from airborne laser scanning of forests using deep learning, *Remote Sens. (Basel)* 12 (9) (2020), 1469.
- [24] Y. Dian, Y. Pang, Y. Dong, Z. Li, Urban tree species mapping using airborne LiDAR and hyperspectral data, *J. Indian Soc. Remote Sens.* 44 (4) (2016) 595–603.
- [25] Z. Yan, T. Duckett, N. Bellotto, Online learning for 3D LiDAR-based human detection: experimental analysis of point cloud clustering and classification methods, *Auton. Robot.* 44 (2) (2020) 147–164.
- [26] Z. Ma, Y. Pang, D.i. Wang, X. Liang, B. Chen, H. Lu, H. Weinacker, B. Koch, Individual tree crown segmentation of a larch plantation using airborne laser scanning data based on region growing and canopy morphology features, *Remote Sens. (Basel)* 12 (7) (2020), 1078.
- [27] C. Torresan, F. Carotenuto, U. Chiavetta, F. Miglietta, A. Zaldei, B. Gioli, Individual tree crown segmentation in two-layered dense mixed forests from UAV Lidar data, *Drones* 4 (2) (2020), 10.
- [28] A. Fekete, M. Cserep, Tree segmentation and change detection of large urban areas based on airborne LiDAR, *Comput. Geosci.* 156 (2021), 104900.
- [29] W. Xu, S. Deng, D. Liang, X. Cheng, A crown morphology-based approach to individual tree detection in subtropical mixed broadleaf urban forests using UAV Lidar data, *Remote Sens. (Basel)* 13 (7) (2021), 1278.
- [30] X. Xu, Z. Zhou, Y. Tang, Y. Qu, Individual tree crown detection from high spatial resolution imagery using a revised local maximum filtering, *Remote Sens. Environ.* 258 (2021), 112397.
- [31] Z.s. Koma, K. Koenig, B. Höfle, Urban tree classification using full-waveform airborne laser scanning, *ISPRS Ann. Photogr. Remote Sens. Spat. Inf. Sci.* III-3 (2016) 185–192.
- [32] Y. Shi, T. Wang, A.K. Skidmore, M. Heurich, Important LiDAR metrics for discriminating forest tree species in central Europe, *ISPRS J. Photogramm. Remote Sens.* 137 (2018) 163–174.
- [33] E. Ayrey, D. Hayes, The use of three-dimensional convolutional neural networks to interpret LiDAR for forest inventory, *Remote Sens. (Basel)* 10 (4) (2018) 649.
- [34] H. Hamraz, N.B. Jacobs, M.A. Contreras, C.H. Clark, Deep learning for conifer/deciduous classification of airborne LiDAR 3D point clouds representing individual trees, *ISPRS J. Photogramm. Remote Sens.* 158 (2019) 219–230.
- [35] X. Chen, K. Jiang, Y. Zhu, X. Wang, T. Yun, Individual tree crown segmentation directly from UAV-borne Lidar data using the Pointnet of deep learning, *Forests* 12 (2) (2021) 131.
- [36] R.G. Kippers, L. Moth, S.J. Oude Elberink, Automatic modelling of 3D trees using aerial Lidar point cloud data and deep learning, *Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci.* - ISPRS Arch. 43 (B2-2021) (2021) 179–184.
- [37] J. Liu, G. Han, Tracing riverine sulfate source in an agricultural watershed: constraints from stable isotopes, *Environ. Pollut.* 288 (2021), 117740.
- [38] Z. Cetin, N. Yastikli, The use of machine learning algorithms in urban tree species classification, *ISPRS Int. J. Geo Inf.* 11 (4) (2022) 226.
- [39] S. Barnea, S. Filin, Extraction of objects from terrestrial laser scans by integrating geometry image and intensity data with demonstration on trees, *Remote Sens.* (Basel) 4 (1) (2012) 88–110.
- [40] J. Zhang, X. Lin, X. Ning, SVM-based classification of segmented airborne LiDAR point clouds in urban areas, *Remote Sens. (Basel)* 5 (8) (2013) 3749–3775.
- [41] J. Hu, et al., MAT-Net: medial axis transform network for 3D object recognition, in: IJCAI International Joint Conference on Artificial Intelligence 2019-August, 2019, pp. 774–781.
- [42] Y. Zhou, H. Chen, Y. Li, Q. Liu, X. Xu, S. Wang, P.-T. Yap, D. Shen, Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images, *Med. Image Anal.* 70 (2021), 101918.
- [43] B. Kumar, M. Yadav, B. Lohani, A.K. Singh, A two-stage algorithm for ground filtering of airborne laser scanning data, *Int. J. Remote Sens.* 39 (20) (2018) 6757–6783.
- [44] N. Varney, V.K. Asari, Q. Graehling, DALES: a large-scale aerial LiDAR data set for semantic segmentation, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops 2020-June, 2020, pp. 717–726.
- [45] J.M. Johnson, T.M. Khoshgoftaar, Survey on deep learning with class imbalance, *J. Big Data* 6 (1) (2019).
- [46] C.C. Feng, Z. Guo, A hierarchical approach for point cloud classification with 3D contextual features, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14 (2021) 5036–5048.
- [47] A. Nurunnabi, et al., An efficient deep learning approach for ground point filtering in aerial laser scanning point clouds, *Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci.* - ISPRS Arch. 43 (B1-2021) (2021) 31–38.
- [48] L. Wang, W. Meng, R. Xi, Y. Zhang, C. Ma, L. Lu, X. Zhang, 3D point cloud analysis and classification in large-scale scene based on deep learning, *IEEE Access* 7 (2019) 55649–55658.
- [49] M. Weinmann, B. Jutzi, S. Hinz, C. Mallet, Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers, *ISPRS J. Photogramm. Remote Sens.* 105 (2015) 286–304.
- [50] R. Zhao, M. Pang, J. Wang, Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network, *Int. J. Geogr. Inf. Sci.* 32 (5) (2018) 960–979.
- [51] S.A. Yang, J. Yoon, K. Kim, Y.K. Park, Measurements of morphological and biophysical alterations in individual neuron cells associated with early neurotoxic effects in Parkinson's disease, *Cytometry A* 91 (5) (2017) 510–558.
- [52] C. Catandaslar, M. Zeybek, Extraction of forest inventory parameters using handheld mobile laser scanning: a case study from Trabzon, Turkey, *Measurement* 177 (2021) 109328 (1–16).
- [53] S.C. Sevgen, Airborne Lidar data classification in complex urban area using random forest: a case study of Bergama, Turkey, *Int. J. Eng. Geosci.* 4 (1) (2019) 45–51.
- [54] M. Yadav, B. Lohani, Identification of trees and their trunks from mobile laser scanning data of roadway scenes, *Int. J. Remote Sens.* 41 (4) (2020) 1233–1258.
- [55] H. Ni, X. Lin, J. Zhang, Classification of ALS point cloud with improved point cloud segmentation and random forests, *Remote Sens. (Basel)* 9 (3) (2017), 288.
- [56] M. Kanevski, et al., Machine learning models for geospatial data, in: Handbook of Theoretical and Quantitative Geography (April 2018), 2009, pp. 175–227.
- [57] G. Mountrakis, J. Im, C. Ogole, Support vector machines in remote sensing: a review, *ISPRS J. Photogramm. Remote Sens.* 66 (3) (2011) 247–259.

- [58] B. Waske, S. van der Linden, J.A. Benediktsson, A. Rabe, P. Hostert, Sensitivity of support vector machines to random feature selection in classification of hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 48 (7) (2010) 2880–2889.
- [59] V.F. Rodriguez-Galiano, M. Chica-Olmo, F. Abarca-Hernandez, P.M. Atkinson, C. Jeganathan, Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture, *Remote Sens. Environ.* 121 (2012) 93–107.
- [60] S.I. Amari, *Information geometry and its applications*, *Appl. Math. Sci. (Switzerland)* 194 (2016).
- [61] C.R. Qi, L. Yi, H. Su, L.J. Guibas, “PointNet++: deep hierarchical feature learning on point sets in a metric space, in: *Advances in Neural Information Processing Systems*, 2017-Decem, 2017, pp. 5100–5109.