ISPRS Journal of Photogrammetry and Remote Sensing 209 (2024) 324–343

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Supervised terrestrial to airborne laser scanner model calibration for 3D individual-tree attribute mapping using deep neural networks

Zhouxin Xi, Chris Hopkinson *, Laura Chasmer

Department of Geography and Environment, University of Lethbridge, 4401 University Dr W., Lethbridge, AB T1K 3M4, Canada

ARTICLE INFO

Keywords:
3D deep learning
Tree classification
Individual-tree segmentation
Tree attribute scaling
Terrestrial laser scanning
Airborne laser scanning
Biomass mapping

ABSTRACT

Capturing subcanopy forest information from airborne laser scanning (ALS) is constrained by signal occlusion. This study demonstrates the potential of close-range terrestrial laser scanner (TLS) scanning to mitigate the constraints of ALS in acquiring stem-level forest attributes. A transformer-based neural network was adapted to classify and segment 3D individual trees from ALS data. A deep neural network combined with a gaussian process layer was proposed to estimate tree diameter-at-breast-height (DBH) from ALS data. The performance of these methods was compared to other benchmarked methods using the same dataset, including a total of seven classifiers, five segmentors, and six attribute regressors. The study was conducted across four ALS sample areas and ten combined TLS/ALS plots, primarily in montane forests. Manual delineation of TLS trees provided a precise validation reference. The proposed methods demonstrated high accuracies, with a mean intersection-over-union (mIoU) of 0.92 for ALS tree classification, 0.70 for tree segmentation, and a RMSE of 4.2 cm or 18.9 % for DBH estimation on average of the ten plots. Tree detection accuracy was strongly associated with the final segmentation accuracy. Factors such as tree height, overlapping, inclination, and neighboring conditions impacted segmentation accuracy. Our segmentation method effectively mitigated accuracy loss for short and occluded trees. Overall, this study presents scalable and cost-effective solutions for TLS calibration of ALS scans over two *meso*-scale montane valleys. Leveraging deep neural networks enables scaling of stem attributes to landscape scales, thereby linking fine-scale forest inventory with sustainable management of expansive forest resources. Our codes are available at https://github.com/truebelief/artemis_treescaling.

1. Introduction

1.1. Calibrating ALS with TLS for individual trees

Among existing 3D forest survey tools, airborne laser scanning (ALS) is renowned for rapidly capturing forest structural variation over large areas, yet with constraints of signal penetration for sub-canopy details (Chasmer et al., 2004; Hilker et al., 2012). Modern ALS acquisitions demonstrate point densities up to and exceeding 1–30 points/m^2 (Okyay et al., 2019) and can occlude over 25 % of canopy volume (Kükenbrink et al., 2017). Such ALS data occlusions become increasingly problematic in stands with dense crown and canopy cover. One solution could be to integrate data captured by high-resolution terrestrial laser scanning (TLS) systems with flexible scanning angles from the ground (Bang et al., 2008; Hopkinson et al., 2013a; van Leeuwen et al., 2011). The degree of TLS integration for ALS-level mapping, also referred to as scaling, can be categorized into three levels. The most general and routine integration is based on inventory metrics such as stem diameters, stem volume, and crown dimension (Chasmer et al., 2006; Hilker et al., 2010; Lovell et al., 2003). The individual-tree metrics derived from TLS are used to enable or enhance ALS predictions of forest indicators such as crown area (Jung et al., 2011), basal area (Alonso-Rego et al., 2021), and biomass (Greaves et al., 2017; Hauglin et al., 2014; Rocha et al., 2023). A finer integration level is to combine detailed point cloud metrics from both ALS and TLS including leaf area profiles (Hopkinson et al., 2013b) and percentile distributions of laser returns or voxel volume (Liu et al., 2017). These metrics, often height-specific 2D summaries of 3D TLS point clouds, are used to calibrate the corresponding ALS metrics, typically utilizing machine learning regression models (Corte et al., 2020; Gleason and Im, 2012). The finest level, relying directly on point cloud properties without manually designated features, has yet to be thoroughly examined in the academic literature. Deep neural network models, capable of direct mapping from point cloud data to forest variables of interest, present a viable metric-free option. Note that fusion-

* Corresponding author.
E-mail address: c.hopkinson@uleth.ca (C. Hopkinson).

https://doi.org/10.1016/j.isprsjprs.2024.02.010
Received 26 May 2023; Received in revised form 12 February 2024; Accepted 12 February 2024
Available online 19 February 2024

Table 1
Literature using deep learning for ALS tree detection, segmentation, and attribute scaling.

Task	Reference and method	Sites and model input	Sampling	Best accuracy
Tree point detection	Multi-Scale FCN (Rizaldy et al., 2018)	Forests in NetherlandsLayered image	Training: 10 patches Testing: 3 patches Patch size: 500 × 500 m^2	F1-score = 0.93
	PFCN (Jin et al., 2020)	Mixed conifer3D points	Training: 14 patches Testing: 6 patches Patch size: 500 × 500 m^2	Classification accuracy = 0.96
Individual-tree isolation	Faster RCNN (Windrim and Bryson, 2020)	25 plots of pine plantation3-channel image	Training: 1527 treesTesting:140 trees	F1-score = 0.87
	Local maxima assisted PointNet++ (Wang et al., 2023)	14 alpine plots3D points	Training: three folds Testing: one-fold	RMS* detection rate = 61 %
	Offset point transformer (Zhang et al., 2023)	Four conifer and mixed plots 3D points	Training: 1558 trees Testing: 382 trees	F1-score = 0.90
Scaling	Deep-RBN and FCN regression (Liu et al., 2020)	Eucalyptus and fir standsLidar metrics	Training: 168 plotsTesting: 72 plots	Volume: RMSE%=15 % R^2 = 0.67
	OCNN-HRNet and DGCNN (Seely et al., 2023)	Canadian temperate forest3D voxel and points	Training: 1635 plotsTesting: 351 plots	Biomass: MAPE = 45.4 % R^2 = 0.80
	PointNet, KPConv, MSENet (Oehmcke et al., 2024)	Danish NFI forests3D points	Training: 4270 subplotsTesting: 911subplots	Biomass: MAPE = 180 % R^2 = 0.83

based integration at point level only merges point clouds from the two scales (Huo et al., 2022; Paris et al., 2017), while the question of how to extend this relationship beyond the TLS-covered area remains unaddressed.

Another variable of integration is the unit scale, which is transitioning from area-based to individual-tree level, driven by advancements in sensor resolution and modeling techniques over the past decades (Vastaranta et al., 2012). The tree-wise integration is influenced not only by the calibration method but also by the individual tree crown (ITC) extraction technique. Successful integration requires both accurate tree detection and individual-tree isolation from ALS data. Detecting ALS tree points is well-established and includes tools such as LAStools (Isenburg, 2014), Terrascan (Soininen, 2020), lidR (Roussel et al., 2020), AMS3D (Ferraz et al., 2012), and TreeVaW (Popescu, 2007). Individual tree extraction algorithms have also been explored in numerous studies (Gupta et al., 2010; Holmgren and Lindberg, 2019; Strîmbu and Strîmbu, 2015). Kaartinen et al. (2012) suggest that the most effective automation methods may surpass manual processing in reliability and accuracy. However, many studies employing hand-labeled trees as reference without ground validation (Weinstein et al., 2020) may yield overly optimistic results. Majority of studies have concentrated on well-defined trees of interest for attribute integration (Jung et al., 2011), resulting in sampling bias, particularly in co-dominant and suppressed layers. Validated using TLS, Kaartinen et al. (2012) demonstrate that most methods have low tree detection rates from ALS, a finding corroborated by Sackov et al. (2017). Coomes et al. (2017) noted that the accuracy of tree-wise biomass analysis from ALS is compromised by up to 20.5 % due to the inadequate detection of individual understory trees, rendering it less favorable than traditional area-based solutions. A hybrid use of ITC and area-based methods, proposed by Kelley et al. (2022), provides a more comprehensive solution. Huo et al. (2022) highlighted that most 2D raster-based algorithms focus solely on detecting dominant forest trees, leaving low vegetation in need of separate treatment to prevent underestimation. By incorporating TLS data, they reduced the root-mean-square error (RMSE) for understory tree count estimates from 39 % to 19 %.

1.2. Advancing ALS scaling with deep learning

The issue of tree detectability from ALS is closely associated with point density and scan occlusion (Tao et al., 2014). Despite sufficient point density, the effectiveness of many existing tree detection methodologies—predominantly unsupervised and reliant on a predefined parameter set—varies across different forest types and regions, especially in diverse topographical conditions. The adoption of deep learning as a supervised, non-linear classification approach introduces enhanced flexibility and robustness to the challenge of tree detection, as highlighted by Jin et al. (2020) and summarized in Table 1. A range of deep learning architectures, including Convolutional Neural Networks (CNN), U-Net (Ronneberger et al., 2015), PointNet++ (Qi et al., 2017), PointCNN (Li et al., 2018), and more recently, transformers (Vaswani et al., 2017), have been successfully used for ALS vegetation detection (Chen et al., 2023; Diab et al., 2022; Hamedianfar et al., 2022; Hu et al., 2023; Rizaldy et al., 2018; Windrim and Bryson, 2020). Compared to the machine learning counterpart, deep learning could typically increase vegetation detection accuracy from ~ 80 % to ~ 90 % (Li et al., 2021). Conversely, isolating individual trees is a more complex issue that has yet to find a universally effective solution. Traditional tree extraction methods often falter when faced with challenges such as irregular tree shapes, overlapping crowns, obscured treetops, or varying tree size (Coomes et al., 2017). While deep neural models experimented on Unmanned Aircraft Systems (UAS) LiDAR show promising potential to address the intricacies of scale, noise, and occlusion, as per Straker et al. (2023) and Xiang et al. (2023), many models require further validation across large-scale natural forest landscapes. In contrast to tree detection, adapting popular deep learning frameworks like the R-CNN series (Girshick et al., 2014) and the YOLO series (Redmon et al., 2016) for 3D point cloud analysis requires model optimization to manage computational resources effectively (Wang et al., 2023; Windrim and Bryson, 2020). Converting 3D data to 2D attributed rasters, while a common workaround, fails to leverage the inherent 3D spatial variability (Sun et al., 2022; Windrim and Bryson, 2019), leading to challenges in accurately delineating complex boundaries, especially in cases of overlapping crowns (Osco et al., 2021; Schmohl et al., 2022; Zhang et al., 2023). There is a significant demand for a lidar-specific tree segmentation model that transcends the limitations of the R-CNN and YOLO frameworks. A promising approach by Zhang et al. (2023) utilizes an offset point transformer to simultaneously determine tree tops and their cluster affiliations, marking a notable advancement in the field. The method's reliance on single-point tree tops and the post refinement of noise clusters represent space for further development. Until recently, the application of deep neural networks for mapping forest attributes from ALS data was scarce. Seely et al. (2023) demonstrated an enhancement in biomass estimation accuracy using two point-based

Table 2
Characteristics of the joint plots with both ALS and TLS scan collections. Field plot size is given either as circular radius (rad) or square side length (sq). *Std: standard deviation.

Plot	TLS Date	ALS Date	Location	Plot size (m)	Tree height (std) (m)	Stem density (ha^{-1})	Subcanopy height (std)(m)	Slope (°)
LPine#1	2016/8/7–8	2016/8/7	49.67°,-109.51°	20 (rad)	18.92(4.46)	1203	0.84 (0.55)	4.1
LPine#2	2016/8/29–30	2016/8/7	49.68°,-109.52°	20 (rad)	14.63(4.65)	2164	0.64 (0.44)	5.0
LPine#3	2021/7/21	2021/8/9	52.79°,-118.01°	25 (rad)	12.73(6.34)	450	1.19 (0.81)	3.5
LPine#4	2021/7/21	2021/8/9	52.79°,-118°	25 (rad)	9.32(5.28)	300	1.39 (1.17)	2.5
LPine#5	2021/7/23	2021/8/3	52.87°,-118.25°	25 (rad)	16.44(5.63)	1501	2.34 (1.69)	20.8
LPine#6	2021/7/24	2021/8/9	52.72°,-117.93°	25 (rad)	14.17(5.43)	1101	2.63 (2.48)	7.2
LPine#7	2021/7/24	2021/8/9	52.72°,-117.93°	25 (rad)	16.94(5.90)	1101	2.41 (2.17)	6.9
TAspen	2018/5/2	2018/7/16	49.35°,-114.41°	20 (rad)	12.14(2.90)	627	1.52 (1.26)	6.7
RPine	2015/7/8–10	2015/7/2	44.08°,-79.32°	20 (sq)	24.82(2.26)	462	6.80 (4.00)	10.5
SMaple	2015/7/8–10	2015/7/2	44.08°,-79.32°	20 (sq)	20.83(5.29)	344	5.71 (3.42)	8.9

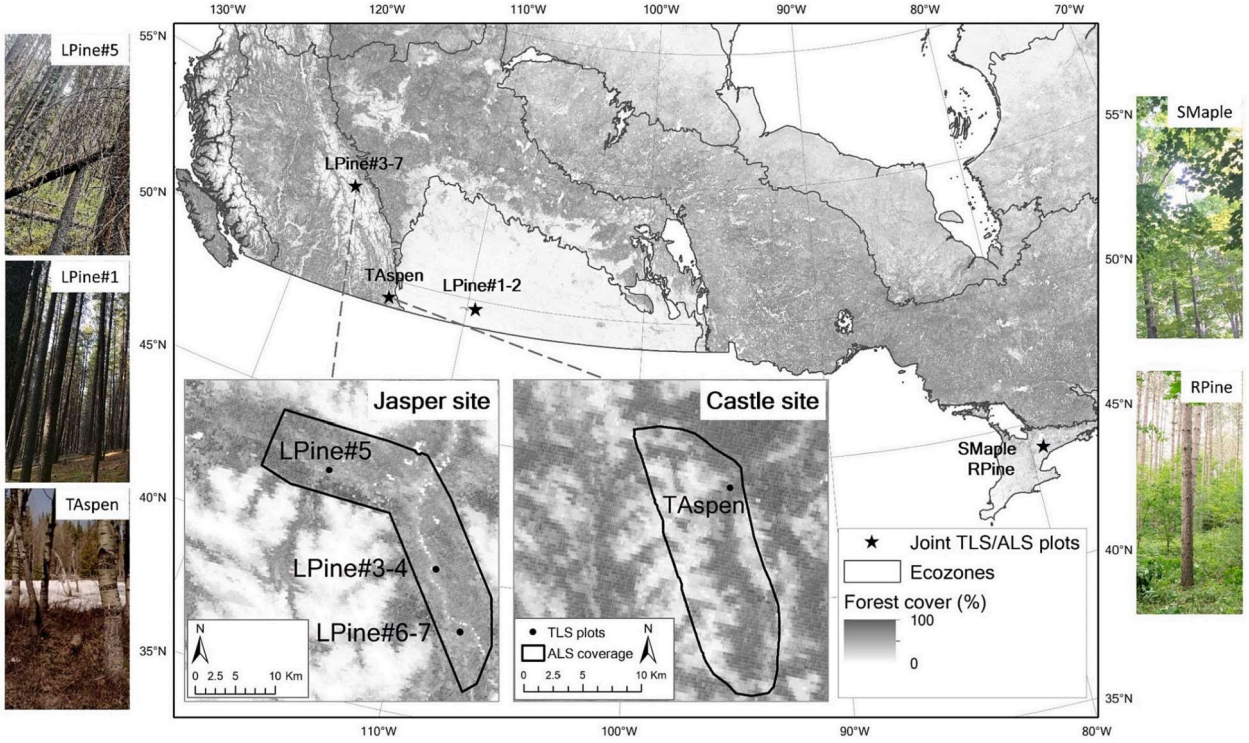

Fig. 1. Locations of the TLS/ALS joint plots and in-situ photos from five plot examples. Black dots represent TLS plots and black polygon outlines the ALS covered area. The base map displays forest cover (%) across Canada, as represented by the gray color derived from MODIS MOD44B imagery. Ecozones are also delineated using outlined polygons.

CNN models compared to a conventional metric-based random forest model. This improvement aligns with findings by Oehmcke et al. (2024), further validating the efficacy of deep learning in this domain. Beyond this, the broader adoption of deep learning for scaling individual-tree attributes in ALS forest studies remains absent. Current research predominantly focuses on plot-level inventory aggregates rather than exploiting the comprehensive potential of TLS tree extraction as a precise reference dataset. There is a clear pathway forward to advance from plot-level summaries to detailed, individual-tree attribute mapping, leveraging the fidelity of TLS data to guide ALS attribute scaling efforts effectively.

Overall, existing research on mapping individual tree crown attributes from ALS is limited by the lack of authentic ground truth data and the absence of a comprehensive 3D deep neural network tailored for vegetation classification, individual-tree segmentation, and attribute scaling. In this study, a transformer-based approach is introduced to enhance tree top area recognition and individual-tree extraction, along with the calibration of ALS individual-tree diameter-at-breast-height (DBH) for biomass mapping, using TLS data as the ground truth, while concurrently benchmarking against conventional techniques. This comprehensive suite of algorithms addresses the necessary processes for transforming raw-resolution ALS data into a classified and annotated tree point cloud dataset spanning two mountainous valley areas surveyed with ALS.

2. Data and preprocessing

2.1. TLS/ALS data collection

The integration of TLS and ALS data over the same plot areas requires close to synchronous acquisition. Ten plots from 2015 to 2021 were selected as joint ALS and TLS sites, including seven lodgepole pine plots (*Pinus contorta*, LPine#1–7), one trembling aspen (*Populus tremuloides*, TAspen), one red pine (*Pinus resinosa*, RPine), and one sugar maple plot (*Acer saccharum*, SMaple) (Table 2). Representative plots are illustrated in Fig. 1. LPine#1&2 were situated in semi-arid Canadian Prairies forest areas, characterized by conglomerate and sandstone beds. LPine#3–7 in Jasper National Park belonged to the Montane Cordillera ecozone,

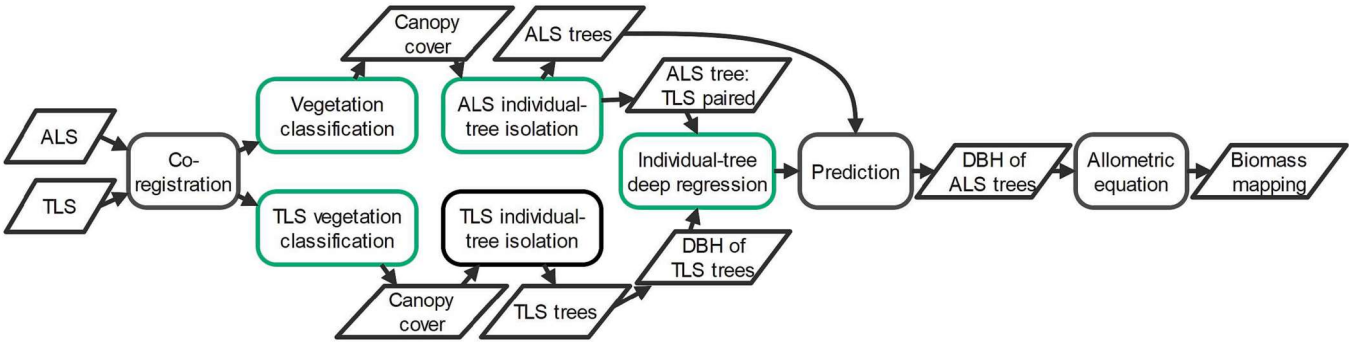

Fig. 2. Scheme of ALS/TLS integration for large-scale individual-tree biomass estimation. Green boxes indicate the key steps involving deep learning optimization. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where above-ground biomass tends to be denser than most other regions in Canada, except for the Pacific Maritime (Wulder et al., 2020). TAspen, a leaf-off plot with minor snow cover, was located in the Eastern Slopes of the Canadian Rockies near Castle Mountain Resort (McCaffrey and Hopkinson, 2020). RPine and SMaple plots were situated in the mixed wood plains transition ecozone. LPine#1&2 exhibited low understory height, flat terrain, and homogeneous tree heights, while the other five plots were more complex, featuring low tree density and dense understory (Table 2).

The ten plots were scanned using a Teledyne Optech TLS (Ilris HD 1535 nm, 1064 nm, or Polaris 1550 nm), and a multi-channel Teledyne Optech Titan ALS (532 nm, 1064 nm, and 1550 nm) (Fernandez-Diaz et al., 2016). The TLS plot setup included a 360° center scan and four corner scans with a narrow field of view (FOV) of 40-60° boresighting the plot center. For each plot, the five scans were co-registered into one plot scan manually. The co-registration was then adjusted using an iterative closest point (ICP) routine (Besl and McKay, 1992) through CloudCompare (v2.11) (Girardeau-Montaut, 2021), with a standard error of 3 cm on average. The resulting plot scan was cropped from the plot center with a fixed radius (20 or 25 m) and resampled regularly to a 2-cm spatial resolution. The variation of plot size and shape was intended to maximize the availability of less-occluded samples for the TLS-to-ALS georeferencing. The TLS plot scans were manually georeferenced to the corresponding ALS scans based on mutual link points using CloudCompare. The TLS-to-ALS alignment error from the three plots varied between 7 and 20 cm subject to the quality and availability of link points. Typically, these points include tree tops and distinct crown edges. This study did not explore the multi-spectral dimension; as a result, the ALS points from all three channels were consolidated into a single point cloud.

The objective of using ten joint TLS/ALS plots was to define a statistical relationship for tree attributes, using DBH as a representative example, across these two different point cloud scales. Once established, this relationship could then be extrapolated to an ALS-only dataset. The ALS sites where the model was applied consist of two montane valley polygons continuously scanned with ALS: the Jasper site, covering approximately 360 km^2, and the Castle site, spanning around 180 km^2, as shown in Fig. 1.

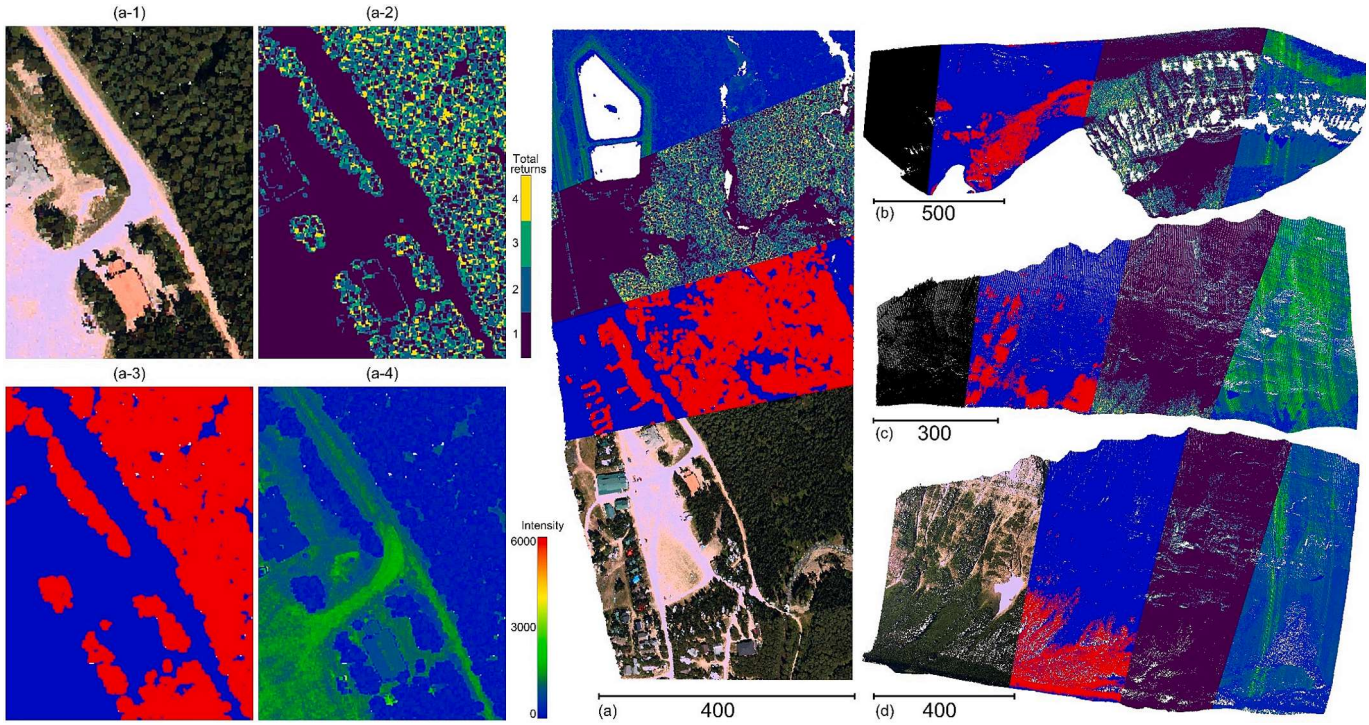

Fig. 3. An overview of four ALS sample areas within the Castle site (a-d) colored by RGB, tree classification (red: trees, blue: others), number of returns, and intensity, respectively, with example texture reference in (a-1–4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

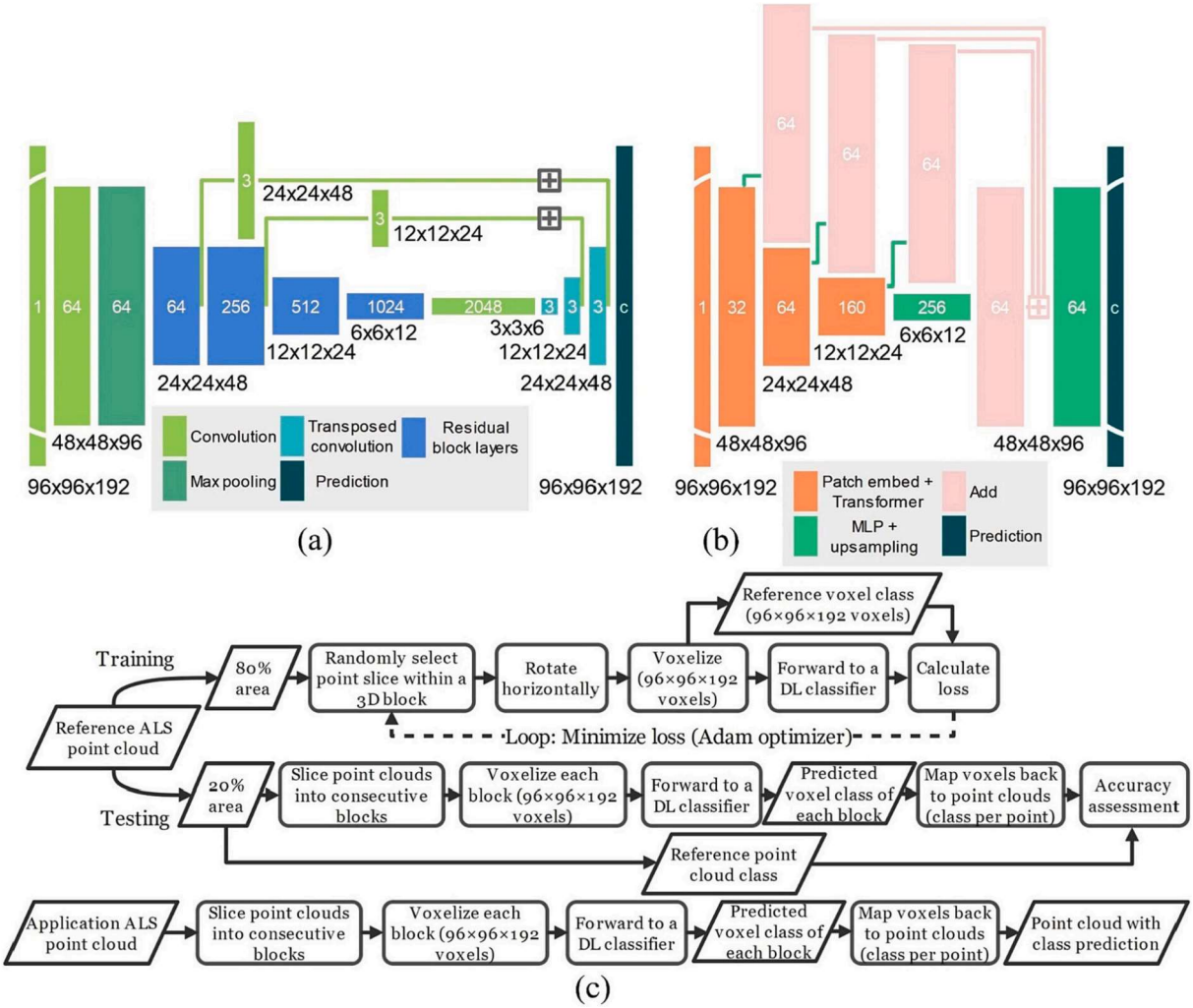

Fig. 4. The network layer configurations of ResNet (a), SegFormer (b), and the major practice steps (c). Different colors represent different layer operations, and their resulting block size is noted as the numbers below. The letter 'c' in the last layer denotes the number of predicted classes (three in our case).

2.2. Field inventory

The *in-situ* mensuration of individual tree DBH within LPine plots#3–7 and TAspen was conducted on August 2, 2016 (Xi et al., 2018) and July 21–24, 2021 (Gerrand et al., 2021). The inventory mensuration of LPine trees was matched with TLS trees through human interpretation, based on approximate within-quadrant position, tree height, DBH, and foliage percentage measurement. TAspen tree stem cross-sections were not perfectly circular, so both maximum and minimum DBHs were recorded for the same cross section at breast height, with an average difference of 1.6 cm per tree (6 % relative). The mean of maximum and minimum DBH values was used as the final DBH measurement for each TAspen tree.

3. ALS/TLS integration methods

The primary ALS/TLS integration workflow is presented in Fig. 2. Following plot scan georeferencing, ALS point clouds were classified into tree and non-tree components through deep learning methods, which are further explained in Section 3.1. The ALS tree point clouds were then segmented into individual trees through a hybrid approach that combines deep learning classification and graph-based clustering, discussed in Section 3.2. The same tree classification processes were applied to the TLS dataset. Individual tree segmentation from TLS point clouds was performed using a CloudCompare plugin *treeiso* (Xi and Hopkinson, 2022). Following the automatic segmentation, tree point clouds were manually clipped, merged, and edited to establish clear boundaries between trees. Finally, DBH was extracted from each manually-adjusted TLS tree, serving as the target variable. A deep learning regression model was then employed to train the 3D voxels corresponding to each ALS individual tree point cloud against this target variable. After training, the model was applied to the remaining individual trees in the ALS dataset. Leveraging tree height information extracted from ALS scans, a 3D individual-tree biomass map was generated across the entire area covered by the ALS flight survey. For brevity, detailed TLS processing will not be discussed further in this section.

3.1. Tree point classification from ALS

The term "tree points" primarily refers to trees in the dominant layer, excluding understory shrubs. The classification of some intermediate-height saplings (approximately 5 m) as belonging to the understory or tree class was challenging. Any saplings with distinct structure and with the centroid elevation above the average point elevation in the sub-canopy layer were considered tree points. To create a classification reference for ALS trees, four sample point clouds were clipped (Fig. 3a–d), with one spanning a recreational area and the other three covering hillslope areas within the Castle site. An ALS tree classification (tree: class 2, non-tree: class 1) was generated by manual (or "expert")

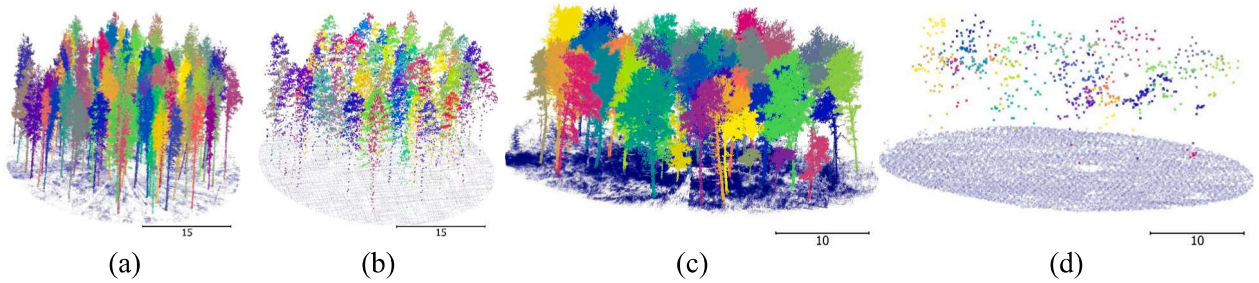

Fig. 5. Manual preparation of reference point cloud data: (a) individual-tree TLS point clouds manually isolated from LPine#1, (b) ALS point cloud with individual-tree points corresponding to TLS tree positions clipped to the TLS plot size, (c) manual segmentation of TLS individual-tree points from TAspen, and (d) the corresponding ALS individual-tree points from TAspen.

interpretation and filtering of tree and non-tree point clusters. LAZ formats (version 1.4) were used for point cloud data, and non-standard attributes, such as the reference class label, were stored as extra byte fields (ASPRS, 2013). The RGB point clouds, created by combining coincident air photos with the LiDAR point cloud, are visualized as a discrete section or 'slice' in each of Fig. 3(a-d). The other slices represent manually isolated tree point clouds and point clouds colorized based on the number of returns and return intensity, respectively. Each slice with different textures was labeled from 1 to 4, e.g., Fig. 3(a–1 to a–4).

Two deep learning classifiers, ResNet (Fig. 4a) (He et al., 2016) and SegFormer (Fig. 4b) (Xie et al., 2021), were adapted into 3D versions for tree point classification. The ResNet served as a baseline model for comparison with the SegFormer.

The 3D network adaptation was achieved by replacing 2D convolution and pooling layers with their 3D counterparts available from PyTorch (Paszke et al., 2019; PyTorch, 2019). ResNet was further adapted into a fully-connected-network (FCN) structure, as shown in Fig. 4a. The final output was voxel-to-voxel class prediction (2: tree, and 1: non-tree) with sufficient spatial resolution. In contrast, SegFormer replaced convolution layers with a vision Transformer (ViT) and incorporated several techniques to enhance positional encoding and computational efficiency (Xie et al., 2021). Compared to convolutional neural network (CNN) structures, the transformer-based model excels in the realm of natural language processing as a more powerful interpreter, based on contextual dependency (Chu et al., 2021) between words, image patches, or 3D block subsets. SegFormer utilized a network structure similar to FCN. The two classifiers were implemented with PyTorch using Intel® Core™ i7-9700 K (8 × 3.60 GHz), 64 GB RAM, and NVIDIA GeForce GTX 2070 (8 GB).

A deep learning classifier required a fixed-sized input into the network. Because inputting an entire point cloud would greatly amplify and complicate computation requirements, the point cloud was divided into blocks of voxels. An ALS block had 96 × 96 × 192 voxels with voxel resolution of 1 m. Each voxel was assigned a value of 1 if occupied by any point, otherwise with 0 if vacant. The setting of 192 voxels, or equivalently 192 m, along the z-dimension was intended to cover the huge variation of ALS point elevation in the mountainous area. No additional attribute such as intensity and height information was added to the ALS tree classifier, which might be species-, site-, or sensor-specific. Each of the four ALS reference point clouds was divided into training and testing areas (Fig. 4c). The training point clouds covered 80 % of the common area coverage, and 20 % was reserved for testing. The data augmentation approach followed the same method from Xi et al. (2023), utilizing random point bootstrapping and neighborhood rotation techniques, consistent with deep learning best practices (Choi et al., 2021). These rotated points were then converted to a block of voxels and passed on to the deep learning model for prediction. The disagreement between the prediction and reference was measured using cross-entropy loss (Bridle, 1990), which was iteratively minimized using an Adam optimizer (Kingma and Ba, 2014). In the testing phase, the voxel predictions of classes were mapped in reverse to the point cloud format to get per-point prediction of classes.

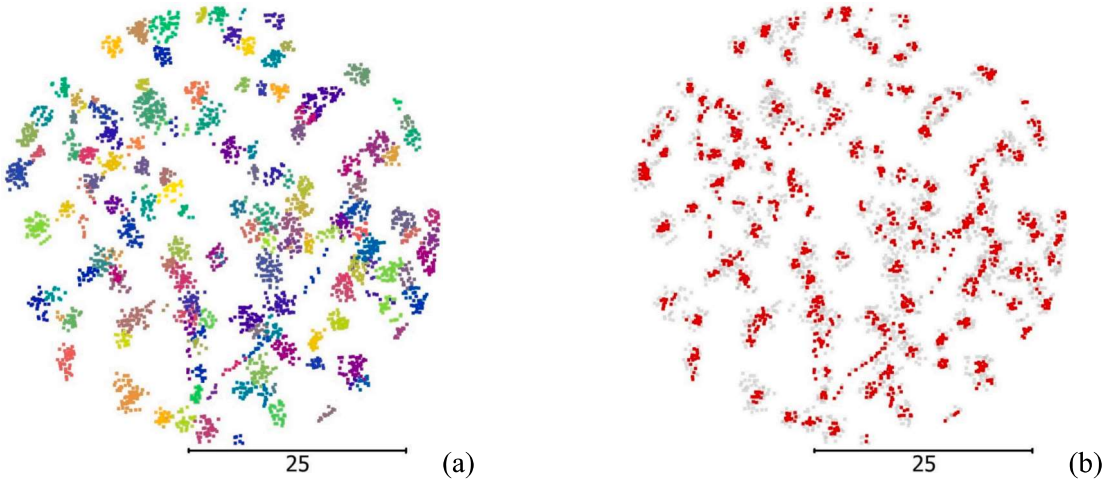

Fig. 6. Reference data preparation for semi-supervised ALS crown segmentation, LPine#6 as the example: (a) ALS reference crowns generated from the matched TLS crowns, each crown with a random color, and (b) near-center points assigned with class 1 (in red) and the remaining points with class 2 (in gray) for classification. The classification of near-center points facilitates subsequent segmentation steps. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

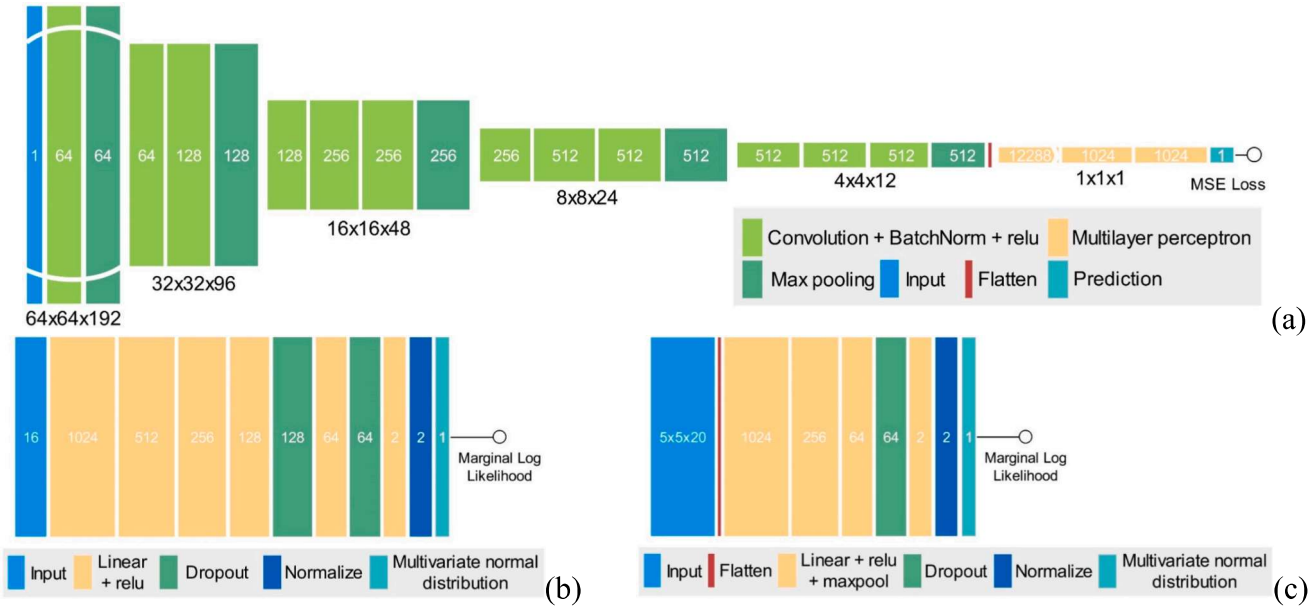

Fig. 7. The structural configuration of regression models for tree attribute estimation: (a) a CNN regression network (CNNR) and (b) a neural-network-based Gaussian Process Regression model (NN-GPR), and (c) a CNN-like Gaussian Process Regression model (CNN-GPR).

3.2. Individual-tree segmentation from ALS

Individual trees through automatic isolation and manual refinement (e.g., Fig. 5a,c) served as references for extracting ALS trees (e.g., Fig. 5b,d) based on point proximity. The majority of ALS points are located at the canopy top and ground level, with substantial occlusion effects on the main stems. On average, 3 % of trees in the ALS plots were highly occluded with fewer than four points, so were discarded to reduce the computational complexity involved in training the deep learning model. The corresponding TLS trees with the same tree IDs were also removed. The extracted ALS trees were considered reference ALS trees.

For ALS tree segmentation, a major challenge is the point density variation in the crown region, e.g., TAspen's 0.6 points/m^2 versus LPine#1's 40 points/m^2. Conventional segmentation approaches based on tree top search or point density would likely prove ineffective in such cases (Kaartinen et al., 2012). With TLS crown reference, supervised segmentation is more suitable for this complex scenario. Our strategy involved configuring a deep learning model (i.e., SegFormer) with a straightforward aim of classifying the fuzzy center region of each crown, serving as a more flexible alternative to R-CNN. Fig. 6 displays an example of fuzzy center classification for LPine#6, a fuel-rich plot with numerous dead, inclined, or clustered thin trees. In comparison with the original point clouds (Fig. 6a), the near-center points defined the critical part of a crown with less overlapped boundaries and clear gaps suitable for segmentation (Fig. 6b). After the near-center points were classified using SegFormer, the remaining points were isolated into meaningful groups using the graph-based cut-pursuit algorithm (nearest neighbor $K = 10$, and regularization $\lambda = 1.0$) (Landrieu and Obozinski, 2017), and then assigned the segment ID from their closest near-center points. This workflow is referred to as the fuzzy center segmentation model (FCS) with more details provided in Appendix A.

3.3. Calibrating ALS individual-tree attributes

Manual extraction of DBH from TLS, achieved by fitting a circle to the stem cross-section at 1.3 m breast height (Hopkinson et al., 2004) with CloudCompare, served as the reference for ALS tree attribute regression. It is important to note that several TLS trees with significant stem occlusion near breast height were manually removed from both training and testing datasets to ensure impartiality of DBH regression and accuracy evaluation. DBH values smaller than 5 cm were also discarded. Subsequently, the DBH extractions from TLS trees were assigned to the corresponding ALS trees as reference. In total, 544 and 223 trees were selected for training and testing, respectively.

Three deep learning methods were compared for directly inferring DBH from individual-tree point clouds (Fig. 7): a conventional 3D CNN-based regression network (CNNR) adapted from the Visual Geometry Group (VGG) (Simonyan and Zisserman, 2014), an attribute-based neural network combined with Gaussian Process Regression (NN-GPR) as the final layer (Lee et al., 2017), and a 3D voxel-based neural network combined with Gaussian Process Regression (CNN-GPR). An ALS tree point cloud was converted into a 0.2 m resolution voxel block (64 × 64 × 192) to serve as the CNNR input. Each block had a volume of 12.8 × 12.8 × 38.4 m^3, encompassing all tree sizes within our sampled forest environments. For different environments, voxel resolution would need to be customized to include the maximum tree size. The last two layers of CNNR were fully connected layers (also known as multilayer perceptron layers), with the output being a numerical prediction of DBH. The mean squared error (MSE) between the two DBHs was employed as the loss function to be minimized.

The input for NN-GPR consisted of a height-based profile, including mean, maximum, standard deviation, kurtosis, skewness, and percentiles (25 %, 50 %, 75 %, 90 %, 95 %), as well as point counts within specific height intervals (0–5 m, 5–10 m, 10–15 m, 15–20 m, 20–25 m, and > 25 m) (Chen, 2015; Sheridan et al., 2015) for each tree. The first six multilayer perceptron layers of NN-GPR were designed for high-dimensional feature encoding. The final layer utilized a multivariate normal distribution with a radial basis function kernel (RBF) to fit the extracted features. The marginal log likelihood between the output normal distribution and the reference DBH distribution was maximized iteratively by an Adam optimizer during the training process (Gardner et al., 2018). CNN-GPR simply altered the input to be a block of 5 × 5 × 20 voxels at a 2 m resolution. The voxel value was determined by the number of points within the voxel divided by the total number of points of the tree. The other layers mimicked the CNN network, with the final layer being the GPR regression. No other features were input to CNN-GPR, which aimed to verify whether the CNN layers could replicate manually extracted point cloud features.

Owing to the absence of ground truth biomass for regression, we employed allometric equations (Lambert et al., 2005) to estimate tree-

Table 3
Accuracy assessment of ALS tree point classification (proposed method in bold).

Plot	Slope mean (°)	Tree* (%)	TRI* mean (m)	mIoU of tree and non-tree point classification						
				Intensity-based	Return-based	LAStools	Terrascan	Random forest	ResNet	SegFormer
Sample area#1	14	61.0	0.50	0.77	0.66	0.87	0.91	0.78	0.93	0.96
Sample area#2	33	8.3	1.11	0.73	0.04	0.52	0.40	0.75	0.91	0.94
Sample area#3	35	4.9	0.75	0.66	0.59	0.43	0.38	0.67	0.82	0.88
Sample area#4	33	10.6	0.68	0.63	0.59	0.47	0.54	0.63	0.73	0.80
Average	29	21.2	0.76	0.70	0.47	0.57	0.56	0.71	0.84	0.89

* Tree (%) is the percentage of tree voxels with the resolution of 1 m to the total number of voxels. Slope and terrain ruggedness index (TRI) were based on a 3 m resolution raster digital elevation model from reference non-tree points.

wise biomass across the Castle and Jasper sites using predicted DBH and ALS tree height. The ALS tree height was extracted as the vertical difference between lowest and highest tree points. Species information was sampled from ground, but not yet sufficient to support deep learning requirement. Therefore, we did not implement tree-wise species classification in the biomass calculation. Instead, we derived five candidate biomass values for each tree based on the allometric coefficients from three dominant (majority) plot-level species. These calculations were made using Equation (1),

$$y_{pine} = 0.0202{*}D^{1.7179}H^{1.2078} + 0.0099{*}D^{1.6049}H^{0.7456} + 0.044{*}D^{3.719}H^{-2.0399} + 0.0785{*}D^{2.5377}H^{-1.1213}$$

$$y_{aspen} = 0.0142{*}D^{1.9389}H^{1.0572} + 0.0063{*}D^{2.0819}H^{0.6617} + 0.0137{*}D^{2.927}H^{-0.6221} + 0.027{*}D^{1.6183}$$

$$y_{spruce} = 0.0265{*}D^{1.7952}H^{0.9733} + 0.0124{*}D^{1.6962}H^{0.6489} + 0.0325{*}D^{2.8573}H^{-0.9127} + 0.202{*}D^{2.3802}H^{-1.1103}$$

$$y_{tree} = \frac{0.22{*}y_{pine} + 0.07{*}y_{aspen} + 0.03{*}y_{spruce}}{0.22 + 0.07 + 0.03} \quad (1)$$

which includes four biomass components for wood, bark, foliage, and branches. The final biomass estimate for each tree was obtained by averaging these candidate values, weighted by the respective species' population proportion in Alberta forests, as per Zhang et al. (2014). While there will be opportunities to enhance the accuracy of biomass calculations in the future, the primary goal at this developmental stage was to evaluate and demonstrate the feasibility of applying deep learning for broad-scale biomass mapping.

3.4. Accuracy assessment

The point-level predictions of the tree class were compared with our reference, and the accuracy was measured using the mean Intersection-over-Union (mIoU) for the tree and non-tree classes (Garcia-Garcia et al., 2017). In addition to the two deep learning classifiers, this study also compared five other ALS classifiers: random forest, LASTools, Terrascan, intensity threshold filter, and return threshold filter. Details are omitted here for brevity but can be found in the Appendix B.

The accuracy of ALS and TLS tree segmentation was assessed using the mIoU for all trees in each plot. To match predicted trees with reference trees, the most frequent class prediction among all points of a reference tree was considered the matched predicted tree ID. For each tree, intersection-over-union (IoU) was defined as the number of points with match classes between the reference and predicted tree point clouds divided by the total number of combined points from both reference and prediction. The mIoU among all trees within a plot was defined as the final tree segmentation accuracy (Equation (2),

$$\mathrm{mIoU} = \frac{1}{m}\sum_{R}\frac{Q \cap R}{Q \cup R} \quad (2)$$

It is important to note that the mIoU definition for tree segmentation differs from the tree classification situation. Detection rate is also used, defined as the number of matched trees between reference and prediction divided by the total number of reference trees. A match is considered any trees with an IoU greater than 0.5, a commonly accepted threshold for evaluating object detection (Padilla et al., 2020). Outlier ALS trees with fewer than four points were excluded from the evaluation.

Alongside FCS, this study also evaluated four additional tree segmentation methods: Segma (St-Onge et al., 2015), lidR-li2012 (Li et al., 2012), and lidR-itcSegment (Dalponte and Coomes, 2016), and Mask R-CNN (He et al., 2017). Detailed information is provided in Appendix B.

The accuracy of scaling DBH through deep learning regression was gauged using the RMSE and coefficient of determination (r^2). We also compared three widely-used machine learning regressors: Gaussian process regression (GPR), support vector regression (SVR), and random

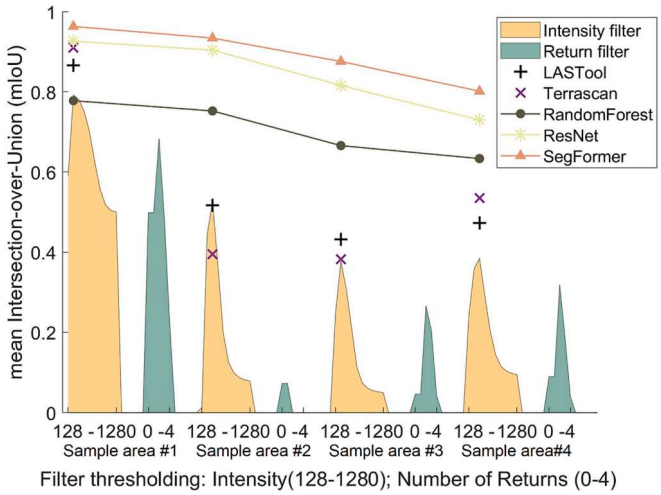

Fig. 8. Accuracy assessment of six ALS classifiers for tree point classification over four testing areas. Colored areas denote a range of accuracies when different thresholds were applied, where intensity thresholds range between 128 and 1280, and number of returns between 0 and 4. For example, a threshold of 256 on the x-axis means that the intensity between 0 and 256 is used. Intensity below 128 or beyond 1280 produced decreasing accuracy and thus not shown here.

Table 4
Accuracy assessment of TLS and ALS tree segmentation (proposed method in bold).

	ALS point density (/m^2)		TLS point density(/m^2)		ALS (Segma)		ALS (lidR-li2012)		ALS (lidR-itcSegment)		ALS (Mask R-CNN)		**ALS (FCS)**		TLS (*treeiso*)	
Plot	Tree points	All points	Tree points	All points	Rate*	mIoU	Rate	mIoU	Rate	mIoU	Rate	mIoU	Rate	mIoU	Rate	mIoU
LPine#1	48.9	71.0	1970	1992	12 %	0.24	43 %	0.48	58 %	0.53	77 %	0.66	97 %	0.80	97 %	0.95
LPine#2	41.4	60.1	2671	2907	5 %	0.08	24 %	0.31	32 %	0.35	58 %	0.48	75 %	0.66	89 %	0.89
LPine#3	1.0	5.3	5768	9135	62 %	0.54	77 %	0.72	74 %	0.73	81 %	0.77	81 %	0.76	85 %	0.85
LPine#4	0.8	5.4	4231	6875	46 %	0.43	54 %	0.60	71 %	0.70	78 %	0.73	88 %	0.81	79 %	0.79
LPine#5	5.1	9.9	10,982	13,718	17 %	0.23	23 %	0.31	29 %	0.36	40 %	0.38	59 %	0.57	69 %	0.69
LPine#6	2.1	5.7	6959	10,272	29 %	0.30	37 %	0.45	49 %	0.52	60 %	0.54	71 %	0.66	79 %	0.79
LPine#7	2.2	6.0	8230	11,997	28 %	0.29	40 %	0.49	56 %	0.58	67 %	0.63	79 %	0.71	84 %	0.84
TAspen	0.8	11.7	1443	3390	15 %	0.32	76 %	0.67	66 %	0.59	76 %	0.69	85 %	0.77	69 %	0.66
RPine	11.6	17.8	5799	7331	4 %	0.27	28 %	0.43	46 %	0.51	72 %	0.60	87 %	0.70	89 %	0.86
SMaple	9.6	12.4	5008	7306	30 %	0.38	44 %	0.42	30 %	0.41	81 %	0.66	67 %	0.54	75 %	0.70
Average	12.4	20.5	5306	7492	25 %	0.31	45 %	0.49	51 %	0.53	69 %	0.61	79 %	0.70	81 %	0.80

* 'Rate' is the detection rate. Point density is the ratio of point number to the area of the 2D convex hull surrounding the whole points.

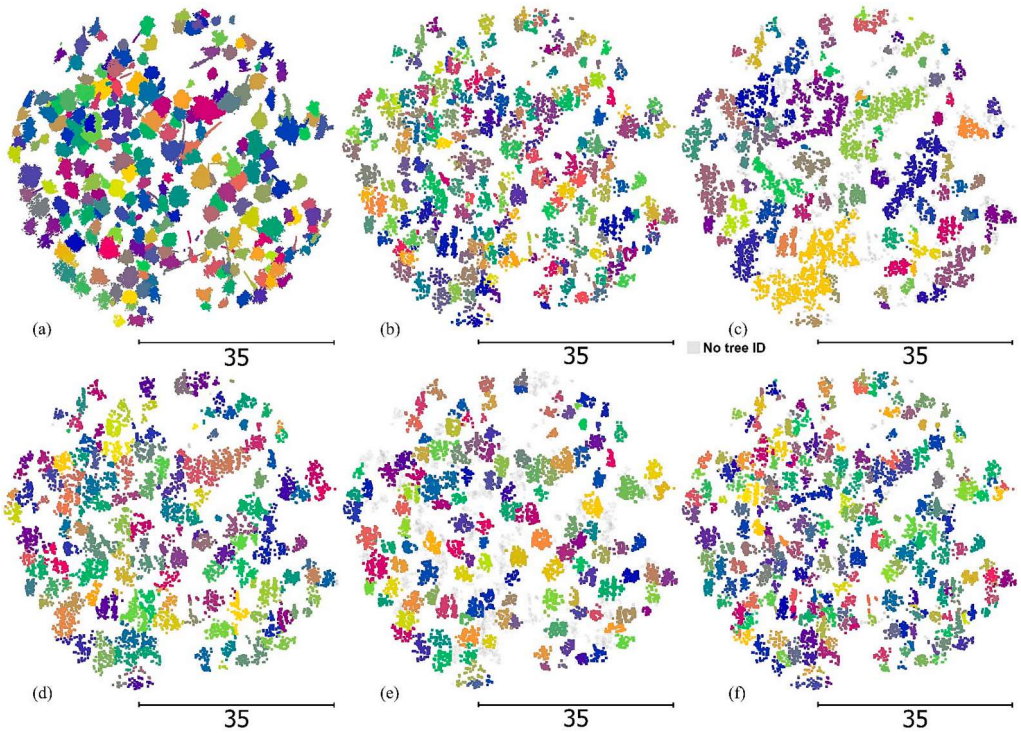

Fig. 9. ALS segmentation of LPine#5: (a) manual reference from TLS, (b) ALS reference trees matched from TLS, tree segmentation with (c) Segma, (d) lidR-itcSegment, (e) Mask R-CNN, and (f) FCS.

forest regression (RFR). These machine learning regressors utilized the same height-based metrics as NN-GPR on an individual-tree basis. All three machine learning regressions were automatically optimized for their hyperparameters by minimizing the five-fold cross-validation loss with MATLAB (MATLAB, 2020). This study does not aim to exhaustively benchmark these methods, but rather to compare them in a common setting. Feature selection and fine-tuning of height profiles are not considered here.

4. Results

4.1. Tree point classification

Comparative analysis deepens our insight into the efficacy of the proposed methods. The average mIoUs across the four ALS samples were 0.70 (intensity), 0.47 (number of returns), 0.57 (LAStools), 0.56 (Terrascan), 0.71 (random forest), 0.84 (ResNet), and 0.89 (SegFormer) (Table 3). The two deep learning classifiers outperformed all other classifiers for all four samples, with SegFormer proving more accurate than ResNet. The resulting tree points from SegFormer were cleaner with less non-tree noise than those of other classifiers, though it did mistakenly include building walls points in the tree class. Fig. 8 illustrates the ALS tree classification accuracies among the four testing samples. The random forest and deep learning classifiers were supervised, while the other four were unsupervised. The intensity- or return-based filters' best accuracies, which varied with threshold changes (colored areas in Fig. 8), did not surpass other classifiers in most cases. The number of returns was less indicative of trees across all samples relative to intensity, as evidenced by the comparatively lower peak mIoUs. The intensity-based filter tended to include trails and slopes in the tree class (see Appendix Fig. B1). In contrast, the return-based filter missed many tree points near top of canopy, misclassified building edges as trees, and failed to capture the sparse vegetation in sample #2's valley (Fig. B1). The random forest classifier achieved an mIoU of 0.71,

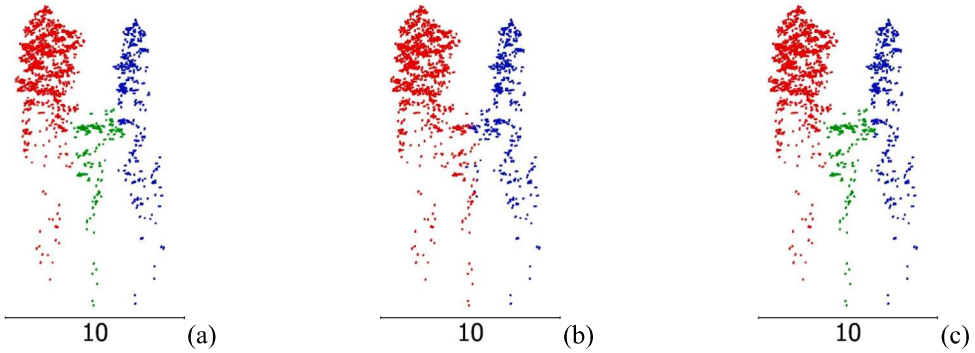

Fig. 10. Vertical view for example trees in LPine#1: (a) reference, (b) lidR-itcSegment, and (c) FCS.

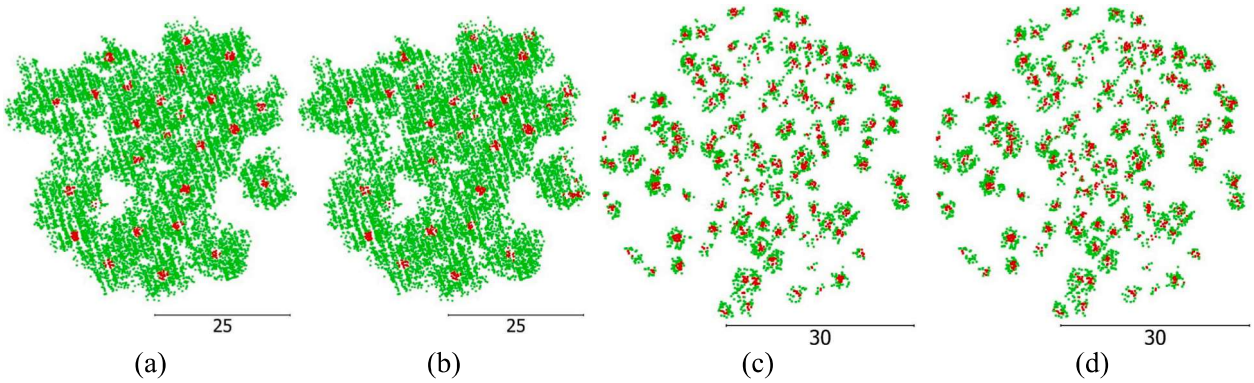

Fig. 11. Near center point detection with center points in red, and others in green: (a) reference based on Equation (1) from SMaple, (b) FCS detection from SMaple, (c) reference from LPine#7, and (d) FCS detection from LPine#7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

equivalent to the optimal accuracy from threshold-based filters (0.70), automating the search for optimal thresholds through decision trees. It exhibited more stable accuracies than all unsupervised classifiers while inheriting the intensity-based filter's limitations.

LAStools and Terrascan, both representative methods based on pure morpho-geometrical constraints, had classification accuracy susceptible to terrain geometry variation. They achieved high accuracy of ~ 0.9 on sample #1, an area with lodges, flat terrain, and homogenous stands. However, their accuracies dropped to ~ 0.4 when applied to samples with steep slopes (>33°) and sparse tree fractions (<11 %). This does not imply that the two methods fail in extreme situations, but rather that their settings require manual intervention, such as slope separation, to avoid abrupt accuracy drops. The need of location-specific intervention is justified by the evidence that local terrain roughness - even with the

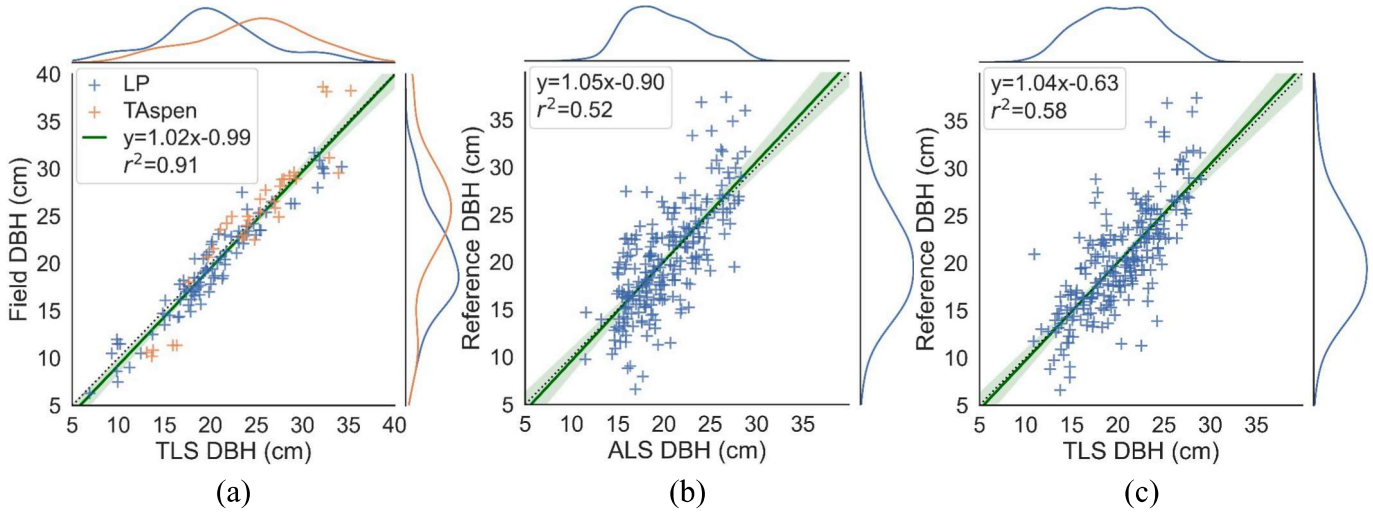

Fig. 12. Scatter plots with dashed 1:1 reference and fitted regression lines between reference TLS DBH and (a) field DBH, (b) ALS DBH estimate from NN-GPR, and (c) TLS DBH estimate from NN-GPR. The green buffer is the 95% confidence interval. The scatter points in (a) are from six validation plots, and in (b-c) from all ten plots. Curve histograms of dependent and independent variables are shown on top and right of each chart. The histograms in (a) are colored by the lodgepole pine plots and TAspen. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
Accuracy assessment of regression models for individual-tree DBH estimation using the independent testing data (proposed methods in bold). Mean RMSE, RMSE%, and r^2 are for all ten plots.

	Profile-based machine learning			Voxel-based deep learning		Profile-based deep learning	
Data(Method)	ALS(RFR)	ALS(SVR)	ALS(GPR)	**ALS (CNNR)**	**ALS (CNN-GPR)**	**ALS (NN-GPR)**	TLS(NN-GPR)
Mean RMSE (cm)	4.9	5.7	4.9	4.9	4.2	4.2	4.0
Mean RMSE%	22.0	24.9	21.7	22.8	19.1	18.9	18.0
Mean r^2	0.34	0.43	0.41	0.25	0.37	0.36	0.44

same slope - also influenced the performance of LASTools and Terrascan (Table 3).

4.2. Individual-tree segmentation

As shown in Table 4, TLS tree segmentation achieved a tree detection accuracy of 81 % and an average segmentation accuracy of 0.80. ALS tree segmentation displayed lower mIoUs compared to TLS and varied significantly among different algorithms. The highest accuracy was achieved by the FCS method, with a detection rate of 79 % and an mIoU of 0.70. The second highest accuracy, 69 % and 0.61, was obtained using Mask R-CNN, followed by lidR-itcSegment's 51 % and 0.53. Area-based point density was not strongly correlated with segmentation accuracy, as indicated in Table 4. The LPine#4 and TAspen, with the lowest point densities, did not yield the lowest segmentation accuracies. Tree geometry and distribution were more influential factors for explaining segmentation accuracy variation. Fig. 9 presents the classification results of all methods for LPine#5. With an appropriate setting of small resolution and moderate smoothing threshold, lidR-li2012 and lidR-itcSegment were capable of differentiating trees in closely clustered stands without introducing significant noise. However, using TLS as a reference, many ALS trees were overlapped in the subcanopy without well-defined 2D boundaries (Fig. 10a), a common scenario where raster-based methods would struggle to detect and delineate trees, such as the lidR-itcSegment in Fig. 10b. FCS partially overcame the overlap problem by first detecting tree center regions and then clustering the remaining points (Fig. 10c). Fig. 11 contrasts the results of detection and reference for SMaple and LPine#7, with FCS detection, which generally matches the reference while exhibiting a slight degree of over-detection among dense crowns or thin trees.

4.3. ALS tree DBH estimation and biomass mapping

The TLS-derived DBH, obtained through circle fitting, was compared with *in-situ* measurements from the TAspen and LPine#3–7 plots (Fig. 12a). Ground inventory for SMaple and RPine plots were omitted

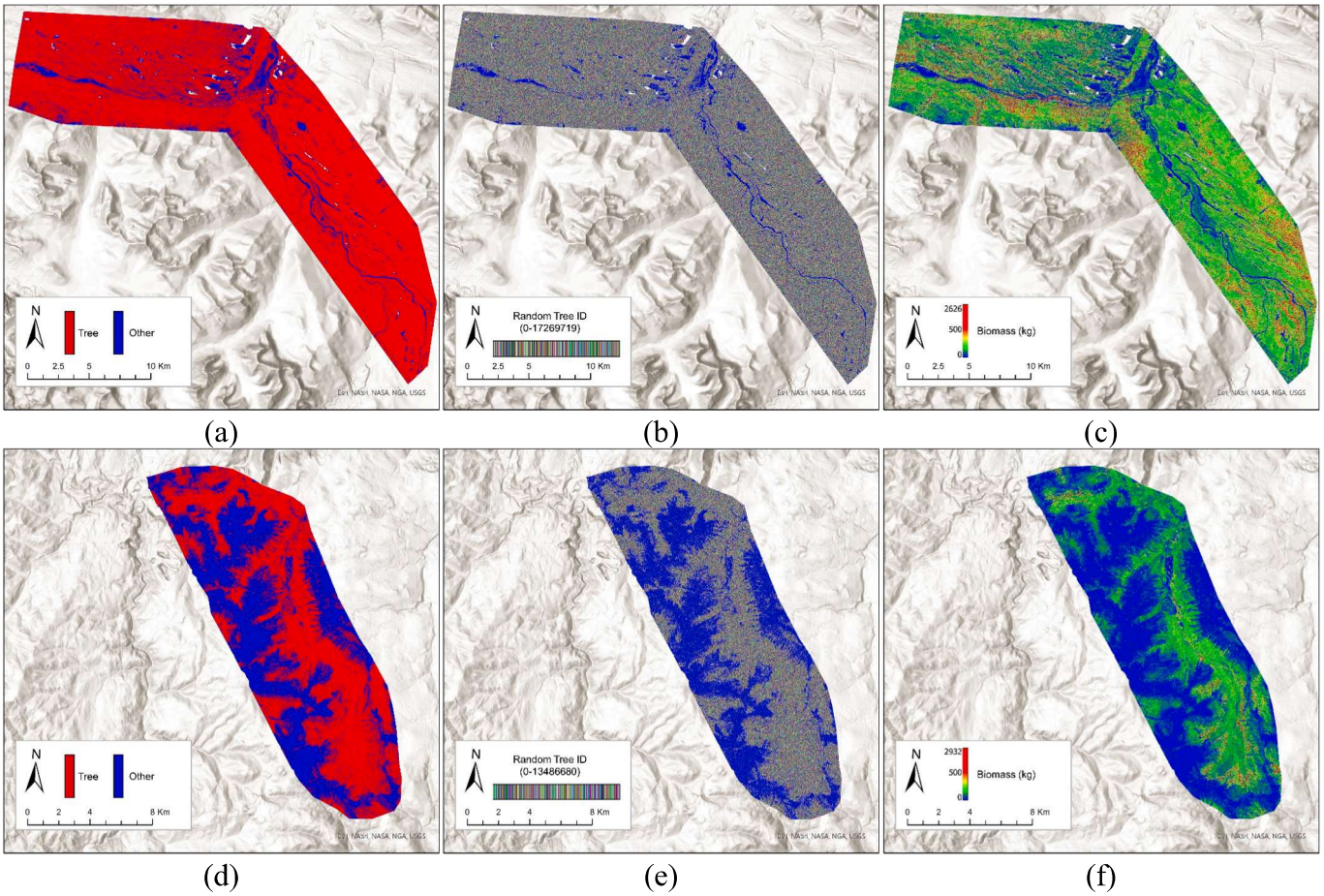

Fig. 13. ALS mapping of Jasper (a-c) and Castle sites (d-f): (a) and (d) tree point classification, (b) and (e) individual-tree segmentation with random colour to differentiate trees, and (c) and (f) individual-tree biomass estimation in (kg per tree).

Table 6
Summary statistics for trees with DBH > 5 cm across the Castle and Jasper sites.

Site	Tree count	Height mean (std)(m)	DBH mean (std)(cm)	Biomass mean (std)(kg)
Jasper	17,269,719	24.1(4.7)	15.9(5.2)	206.3(115.7)
Castle	13,486,680	20.4(3.7)	10.7(4.7)	128.6(71.0)

*std: standard deviation

because they were not part of the mapping case study area in western Canada. The TAspen plot contributed several large DBH values. A robust fit was observed among the 120 trees, with an r^2 value of 0.91. The RMSE was 2.0 cm, or 9.5 % relative to the measurement. The bias for TLS estimated DBH was close to zero for TAspen and + 3.6 % for LPine plots. Rigid *in-situ* biomass reference is desired for future work but was not available for this study.

Utilizing 223 testing trees, the relationship between the regression-based DBH estimates and TLS DBH reference is presented in Table 5, which shows the mean values of RMSE, RMSE%, and r^2 over all plots. Among the three machine-learning regressors, both RFR and GPR achieved the highest testing accuracy with an RMSE of 4.9 cm, or a relative RMSE of ~ 22 %. Only RFR overfitted the training data with a significantly lower RMSE of 2.9 cm, while the other two regressors exhibited similar RMSEs for both training and testing data. The CNNR did not surpass the top machine-learning regressors in accuracy; however, modifying the final layer to a Gaussian Process Regressor (CNN-GPR) significantly reduced the RMSE from 4.9 cm (22.8 %) to 4.2 cm (19.1 %). The attribute-based NN-GPR achieved a similarly low RMSE of 4.2 cm (18.9 %). These findings suggest that the CNN layers could effectively extract features in lieu of manual extraction, if not better, but in our study, the CNN layers alone did not play a more significant role than the GPR layer in addressing the regression problem.

The ALS DBH estimates of the NN-GPR, plotted against the TLS DBH reference in Fig. 12b, produced a relative RMSE for conifer plots of 27 %, which was higher than the 21 % for deciduous plots, suggesting that crown information is a crucial source for accurate inference. Substituting ALS with TLS (Fig. 12c and Table 5) reduced the RMSE to 4.0 cm (plot average) or 3.6 cm (overall) and increased the r^2 to 0.44 (plot average) or 0.58 (overall). However, the accuracy improvement achieved through this regression approach was minimal and not comparable to manual DBH extraction from TLS, indicating that the regressors in this study were not trained to effectively utilize points around the stem diameter for DBH inference. Future studies should develop a deep learning model designed to automatically focus on the appropriate height of stem points for feature extraction.

By integrating the suite of tree classification, segmentation, and regression methods, attributed individual-tree point clouds were generated for the two ALS flight polygons (Fig. 13). The point clouds were encoded in the LAZ 1.4 format, with DBH, height, and biomass attributes stored as extended variables for each point. Subsequently, tree attribute statistics were calculated for the entire valley forest coverages, encompassing over ten million trees in these two sites (Table 6).

5. Discussion

5.1. Comparing tree point classification methods

In this study, SegFormer achieved the highest mIoU of 0.89, compared to ResNet's mIoU of 0.84 and RF's 0.71. Morpho-geometric methods, LASTools and Terrascan with fixed parameters, exhibited a contrasting mIoU of ~ 0.9 in a relatively flat recreational area and ~ 0.4 in slopes with abrupt elevation variation. Intensity- and return-based filters were less affected by landform variation but were susceptible to local noise. The SegFormer classifier was also employed to classify TLS tree points, attaining a high mIoU of 0.92. It was discovered that morpho-geometrical methods, which employ hard geometric thresholds effectively for terrain extraction, falter when adapted for tree point classification due to the inherent variability in tree geometries. Maltezos and Ioannidis (2015) noted that lasclassify struggled to differentiate roof points clearly from surrounding vegetated areas. Zhang et al. (2020) underscored the accuracy and robustness of graph-based convolutional neural networks for ALS ground filtering. In contrast, Terrascan struggled to adapt to uneven terrain, and the Cloth Simulation Filtering (CSF) algorithm performed poorly in areas with low vegetation. Based on our evaluation, we recommend reconsidering the use of lasclassify and other methods relying on simple morpho-geometric constraints in complex mountainous environments where forests coexist with steep slopes and rock outcrops, i.e., areas with significant vertical variability. Deep learning tree classifiers offer parameter-free models without the need for explicit parameter tuning. However, this does not imply that deep learning requires less optimization time. In fact, the success of neural networks is contingent upon the quality of the data structure from training samples (Lines et al., 2022), necessitating not only a substantial sample size but also clear boundaries between classes. Editing and troubleshooting class annotations or labels from a large number of training samples for deep learning can be more laborious and unmanageable than simply optimizing geometric rules for higher classification

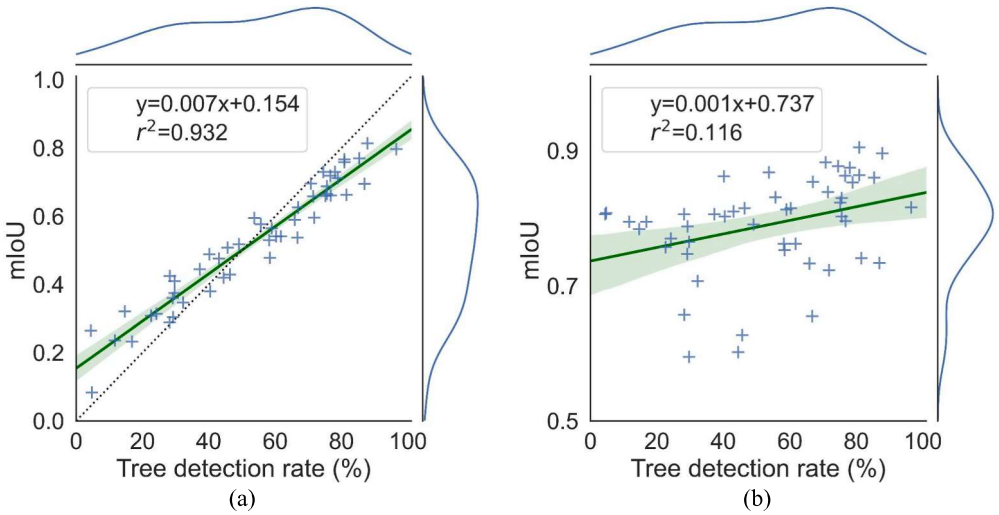

Fig. 14. The relationship between tree detection rate (%) and (a) tree segmentation accuracy mIoU, and (b) mIoU within detected trees. Data points are the results of five benchmarked algorithms over ten plots.

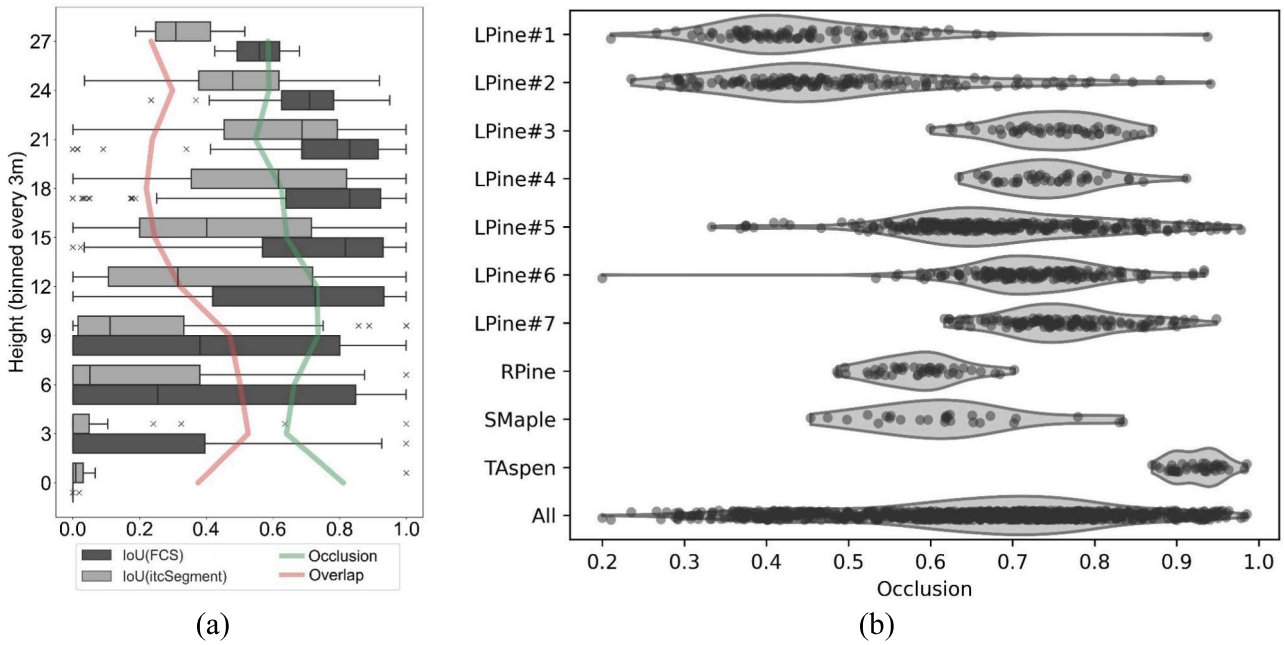

Fig. 15. Occlusion and overlapping factors for segmentation accuracy: (a) Tukey box plot of IoU for the lidR-itcSegment and FCS algorithms, along with the mean occlusion ratio (Occlusion) and mean crown overlap ratio (Overlap), categorized by 3-meter height bins. Vertical axis labels indicate height ranges: '0′ represents 0–3 m, '3′ represents 3–6 m, and so forth; (b) Plot-wise ALS occlusion distribution.

accuracy. An interactive or guided solution is needed to avoid inefficient sample manipulation.

5.2. Importance of tree detection to segmentation accuracy

Individual-tree segmentation from ALS has been extensively investigated in various studies but most methods have been applied to specific forest regions or types without considering the combination of solutions and sites comprehensively. A common problem with the four unsupervised solutions examined in this study is the under-segmentation of trees, consistent with findings from Aubry-Kientz et al. (2019) using Segma. When two trees are adjacent, the lower tree top is frequently missed by the algorithms (Vauhkonen et al., 2012), resulting in excessive expansion of a tree segment. One potential solution is to constrain tree size, as the four algorithms do, which mitigates but does not rectify tree top misdetection, leading to inaccurate shape segmentation. There are also two limitations of using a raster CHM. First, the minimum resolution of a CHM grid cell cannot be too small, as this would lead to tree top noise and crown gaps, increasing the risk of over-segmenting large trees. Smoothing methods are typically applied to remove crown gaps but also suppress the search for tree tops. These parameters become contradictory in plots with heterogeneous tree sizes. Second, 2D-raster-based methods overlook the multi-layered nature of forests and are incapable of delineating understory or leaning trees. Kaartinen et al. (2012) and Wang et al. (2016) suggested that smaller trees can be segmented more accurately using the original point cloud rather than the derived 2D raster. Indeed, as Fig. 14a demonstrates, individual-tree segmentation accuracy is strongly dependent on tree detection rate for nearly all benchmarked algorithms ($r^2 = 0.93$). However, for trees that were successfully detected (IoU $>$ 0.5), this dependency reduces (Fig. 14b), suggesting that the congruence of boundaries accounts for only 11.6 % of the mIoU variance in their linear relationship.

The FCS method ignores the assumptions of raster or CHM, focusing instead on detecting tree centers without limits on crown size heterogeneity, and subsequently applying 3D density-based clustering to address the problem of isolating subcanopy trees. This two-stage approach has gained increasing recognition in recent studies (Dersch et al., 2021; Krzystek et al., 2020; Reitberger et al., 2009; Wang et al., 2020). Windrim and Bryson (2020) are among the pioneers in adopting deep learning R-CNN for stem detection alone. However, imperfections were still evident from our experiment: (1) the R-CNN methods extract trees with boxes, which cannot clearly separate subcanopy trees, (2) labeling stems as training data is impractical when ALS point density is low, which is common for high-altitude flights over mountainous areas, and (3) the matching degree of bounding boxes between R-CNN and ground truth is weak, as also concluded by Xi and Hopkinson (2021) using high-resolution TLS trees, which affects segmented tree boundaries. This weak matching arises because R-CNN employs deep learning regression to obtain bounding box coordinates, but as revealed in this study, conventional deep learning networks demonstrate suboptimal accuracy when applied to regression-oriented problems. To leverage the higher precision of deep learning in classification, we propose an FCS that reinterprets the bounding box regression challenge as a center point classification problem. As a result, FCS was successfully adapted to

Table 7
Pearson's correlation between individual-Tree LiDAR attributes and segmentation/regression accuracies. Attributes from either TLS or ALS are selected based on their level of precision.

ALS processing accuracy	Height (ALS)	Area* (ALS)	Occlusion (ALS)	Overlap (TLS)	NND* (TLS)	TA* (TLS)	DBH (TLS)
Segmentation IoU (FCS)	0.41	0.10	−0.30	−0.43	0.36	−0.23	0.29
Regression error (NN-GPR)	−0.39	−0.16	0.14	0.08	−0.15	0.11	−0.54

* Area is the ALS crown area measured by convex hull. NND is the distance between the focal stem and the nearest neighbour stem from TLS. TA is the tilting angle of the focal tree from TLS in degree.

Table B1
Comparison of other ALS tree classification and segmentation algorithms.

Type	Method	Key steps and parameter settings
ALS tree classification	LAStools	lasground, lasheight, lasclassify: settings all default
	Terrascan	Classify ground, building, surface, and low-height points (<0.5 m) with the remained class to be trees
	random forest	200 bagged decision trees, all other settings default
	intensity	Optimal threshold between the minimum and maximum range
	number of returns	Optimal threshold between the minimum and maximum range
	ResNet	Input block size: 96 × 96 × 192 voxels; Voxel size: 1 m; Adam optimizer
	SegFormer	Input block size: 96 × 96 × 192 voxels; Voxel size: 1 m; Adam optimizer
ALS tree segmentation	Segma	Gaussian kernel sigma: 0.5Height slice (m)
		: 5,10,15,20, until max heightSlice search distance (m)
		:1.5, 2.0, 2.5, 3.0, 4.0
		Minimum tree height and crown area: 1 m and 1 m^2
	lidR-li2012	Threshold 1: 1.5; Threshold 2: 2; Search radius: 2 *m*
		Conditional point elevation: 15 *m*
		Minimum tree height: 2 *m*
		Maximum crown diameter = 10 *m*
	lidR-itcSegment	Raster resolution: 0.5 *m*
		Circle radius for the Digital Surface Model: 0.2 *m*
		Moving window diameter to detect local maxima: 3 *m*
		Growing threshold 1: 0.45; Growing threshold 2: 0.55
		Minimum tree height: 2 *m*
		Maximum crown diameter = 10 *m*
	Mask R-CNN	Backbone: ResNet50FPN
		Input CHM resolution: 0.5 m; Patch size: 128 × 128 pixels
		Batch size: 2, Channel count: 1
		Confidence level as successful detection: 0.8
		IoU threshold as successful detection: 0.5
		Number of classes: 2 (tree v.s. non-tree)
		Maximally allowed overlapping: 0.6
		Prior mean of a CHM image: 10Prior standard deviation of a CHM image: 10
	FCS (proposed)	Input block size: 128 × 128 × 128 × 2 voxels; Voxel size: 30 cm
		Nearest neighbor search K: 10
		Regularization parameter λ: 1.0
		Near-center distance threshold for creating training reference: 75 %

various forest scenes without significant loss of accuracy or additional parameter engineering. Donager et al. (2021) reported their TLS and ALS tree detection rates within a ponderosa pine forest to be 87 % (±14 %) and 68 % (±25 %), respectively. Considering this reference, our average ALS detection rate of 79 % using FCS was higher, based on the seven lodgepole pine plots in a comparable forest setting.

5.3. *Factors affecting tree segmentation and regression accuracy*

To better understand segmentation accuracy, we calculated the occlusion ratio for individual trees at a 1 m resolution using ALS and TLS. Occlusion was calculated with MATLAB as the ratio of the volume of differing voxels at a 1 m^3 resolution between ALS and TLS to the volume of TLS voxels for each individual tree. In Fig. 15, we grouped tree heights into 3 m bins and averaged the occlusion ratios for each. As tree height increases, occlusion decreases (r = -0.40). Trees between 0 and 3 m had the highest occlusion at 80 %, declining to 59 % for trees 27–30 m tall. The drop in occlusion among the large trees is explainable by increased visibility of upper crowns from the air (Kükenbrink et al., 2017; Schneider et al., 2019). Both lidR-itcSegment and FCS algorithms achieved peak accuracy in the 21–24 m height range. The decline in accuracy for taller or shorter trees is partly due to varying occlusion levels. Specifically, lidR-itcSegment showed a sharper IoU decline below 18 m, while FCS remained stable until below 12 m. This is attributed to FCS's better tree center detection across 3D space, including short trees often neglected by CHM-based models (Huo et al., 2022; Kaartinen et al., 2012). Interestingly, the decline in accuracy was more severe for short trees (3–9 m) than tall trees (24–30 m) though occlusion levels were similar. This is due to increased crown overlap in shorter trees, which complicates segmentation. The mean overlapping ratio is also depicted in Fig. 15 (a), where it is grouped by 3 m height bins similar to the occlusion ratio. This ratio is calculated as the area of overlap among neighboring TLS tree convex hulls relative to the area of the focal tree's convex hull, as outlined by Xi and Hopkinson (2021). Using TLS as reference, it can be found that ALS tree occlusion varies greatly among different species and sites indicated in Fig. 15 (b). The most frequent occlusion is 0.7 among all trees. Note that ALS trees with fewer than four points were excluded for segmentation and regression. Incorporating such sparse data would have decreased the mIoU by an average of 0.17. The actual occlusion is expected to be higher and the regression accuracy lower.

In addition to occlusion and overlapping ratios, nearest neighbour stem distance (NND) and tilting angle (TA) for each individual tree were also analyzed based on TLS metrics (Xi and Hopkinson, 2021). These four metrics reveal that shorter trees are more occluded and tend to have higher clustering, overlapping ratios, and inclination. Correlation analysis with tree height yielded correlation coefficients of r = -0.4 for occlusion, 0.33 for NND, −0.25 for overlapping ratio, and −0.29 for TA, underscoring the increased complexity of segmenting shorter trees. As concluded in Table 7, tree geometric factors of height, overlapping, inclination, neighbouring, all significantly impact individual-tree segmentation accuracy. Height emerges as a more significant factor than area for segmentation accuracy, corroborating findings from TLS tree segmentation (Xi and Hopkinson, 2021). Overlapping and neighbouring are the other key influencers, aligning with results from the TLS segmenter, *treeiso* (Xi and Hopkinson, 2022). Trees with larger and overlapping crowns, such as sugar maple, demonstrated lower mIoU compared to narrower conifers. Conifers typically exhibited higher delineation accuracy than their deciduous counterparts (Hastings et al., 2020). However, densely clustered stems, like LPine#5, also resulted in low detection rate and segmentation accuracy.

The DBH extraction accuracy from TLS trees falls within the typical RMSE range of 1–3 cm or 5–15 %, as reported in the benchmarking of fourteen DBH extraction algorithms (Liang et al., 2018). The RMSE% range of DBH scaling using the deep learning regressor (Table 5) was consistent at approximately 20–25 %, lower than the 40–70 % range

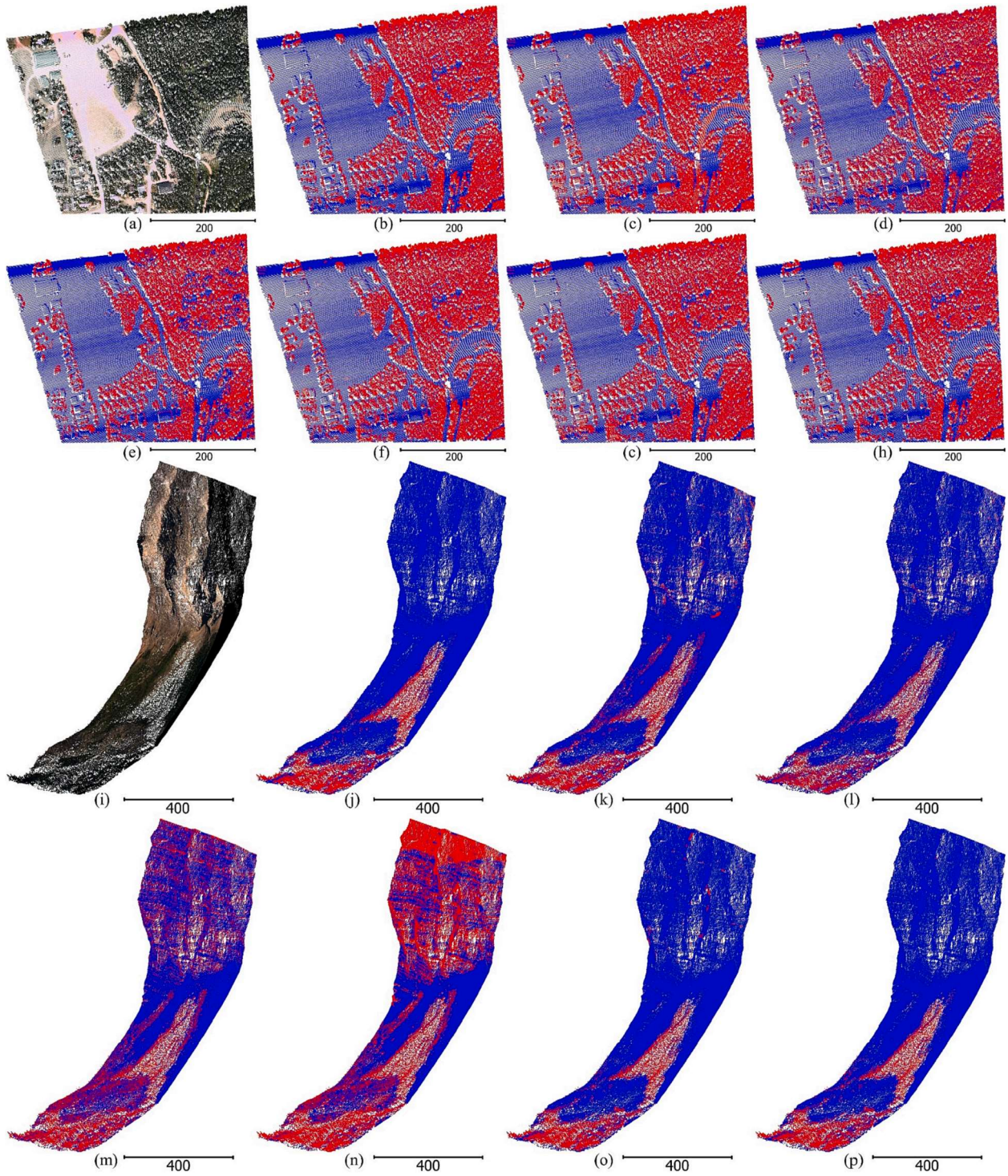

Fig. B1. Example ALS testing areas for tree point cloud classification: (a-h) Sample area #1, and (i-p) Sample area #2, including (a,i) RGB point clouds, and classification from (b,j) manual reference with tree points in red, non-trees in blue, (c,k) intensity-based, (d) return-based, (l) random forest, (e,m) LASTools, (f,n) Terrascan, (g,o) ResNet, and (h,p) SegFormer. Without showing are the random forest classification of Sample area #1 identical to that from (c), and the return-based classification of Sample area #2 with only few tree points.

achieved by the DBH estimator using allometric regression that solely relied on height and crown diameter as indicators (Aubry-Kientz et al., 2019), but higher than the 13 % attained by the RFR DBH estimator for Scot pines (Pyörälä et al., 2019). The modest increase in regression accuracy achieved by deep learning methods over the RFR was also observed in Seely et al. (2023), reporting a 5 % enhancement in r^2. Table 5 suggests that incorporating CNN layers might not necessarily enhance DBH regression accuracy. Instead, integrating machine learning layers appears more effective, resonating with findings from Liu et al. (2020). Their comparison of deep neural networks with multiple linear regression revealed even lower DBH regression accuracy from the latter. The DBH scaling error demonstrated a tendency to overestimate smaller DBHs and underestimate larger DBHs, a pattern also observed by Lindberg et al. (2012) when employing TLS variables as the training

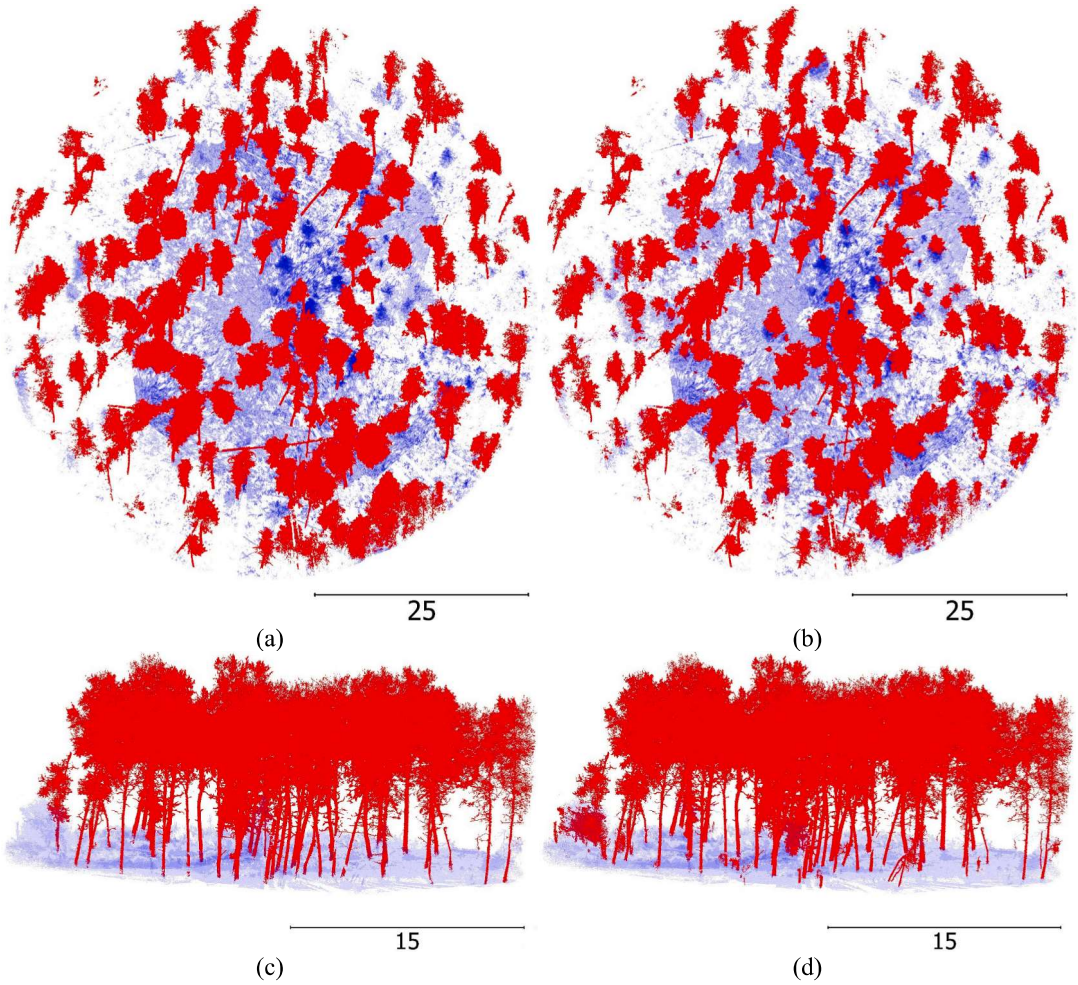

Fig. C1. Visualization of tree classification from two testing samples: (a) LPine#6 reference tree class, (b) LPine#6 SegFormer classification, (c) TAspen reference tree class, and (d) TAspen SegFormer classification. Tree points are in red, and other points in blue with a transparency of 40 % for clear visualization of the ground layer.

dataset for their regression model. Based on Table 7, regression accuracy is less impacted by tree geometric complexity than segmentation accuracy, but is mostly a function of DBH and tree height.

6. Conclusions

ALS and TLS point clouds represent distinct view geometries and 3D density characteristics within forest environments. This study proposed a sequence of processes for ALS and TLS integration, including tree point classification, individual-tree segmentation, and ALS tree biomass estimation. Two deep learning classifiers, SegFormer and ResNet, were evaluated for classifying ALS tree points from four test areas. The optimal accuracy was obtained using the SegFormer classifier in a benchmark comparison with other six classifiers. Using manually isolated TLS trees from ten plots as reference, the proposed FCS method facilitated individual-tree segmentation by precise tree center classification via the deep learning network, with a high success rate among crowns where tree tops or stem features were hardly detectable. When comparing multiple ALS segmentation methods, it was evident that the accuracy of tree detection serves as a critical factor in determining the overall accuracy of tree segmentation ($r^2 = 0.93$). A key limitation of existing segmentation methods is their reduced accuracy when dealing with shorter trees. Three distinct neural network structures, CNNR, CNN-GPR, and NN-GPR, were also compared for optimal estimation of individual-tree DBH values from ten plots. Reference individual-tree DBH values were extracted from joint TLS plots. Both CNN-GPR and NN-GPR achieved similar high accuracies: RMSE of 4.2 cm and relative RMSE of 19 %, outperforming the three machine-learning regressors. The CNN structure effectively extracted implicit tree features as a proxy for manually designated features but did not exceed traditional machine learning in addressing regression problems.

Overall, deep learning algorithms have improved upon traditional methods of tree classification, segmentation, and scaling, achieving acceptable DBH estimates, which are essential for individual-tree biomass modeling and mapping. Transferring crown, stem, or understory feature information from TLS to the broader scale of ALS not only calibrates tree-level attribute retrieval but also enables the imputation of attributes typically unattainable from ALS. While incorporating large generative models marks a pivotal future direction, the methodologies introduced for transferring forest attributes from TLS to ALS provide a solid automated framework for comprehensive forest inventory, and biomass, fuel, or carbon monitoring. This innovative approach seamlessly scales detailed tree stem-level attributes to broader regional scales, meeting the needs of management decision-making and fulfilling reporting obligations efficiently.

Funding.

The authors acknowledge research funding from an Alberta Environmental Protection [grant #GRAEM26], NSERC Discovery [grant numbers 2017–04362, 2017–04492], Mitacs Accelerate [grant number IT27605], Foothills Research Institute (fRI) - Federal-Provincial MPB Research Partnership [grant number 247.15], and NSERC Canada Wildfire [grant number SPG-N - NETGP 548629–19]. Laser scanner and

GNSS equipment was funded via Canada Foundation for Innovation [grant number 32436], and Western Economic Diversification Canada [grant number 000015316]. The authors extend their appreciation to Tristan Skretting, Emily Jones, Saied Parsian, Linda Flade, Viviana Lartiga for their assistance in the field, Dr. Chris Watson and Dr. Landon Shepherd from Jasper National Park for providing field support and access to sites, and Dr. Jonathan Boucher of the Canadian Forest Service for providing insights for fuel sampling.

CRediT authorship contribution statement

Zhouxin Xi: Data curation, Writing – original draft, Writing – review and editing, Visualization, Investigation, Formal analysis, Validation, Methodology. **Chris Hopkinson:** Conceptualization, Funding Aquisition, Data curation, Investigation, Methodology, Supervision, Resources, Project administration, Software. **Laura Chasmer:** Funding Aquisition, Data curation, Investigation, Supervision, Resources, Project administration, Software.

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.

Appendix A:. Fuzzy center segmentation settings

The input of SegFormer consisted of a 128 × 128 × 128 × 2 voxel block with a 30-cm voxel resolution and normalized intensity (scaled by the maximum intensity of 4096) as an additional attribute. The output maintained the same block size of voxels with class labels (class 0: blank, 1: near-center, 2: peripheral). Training and testing the SegFormer model followed the steps illustrated in Fig. 4c, with reference data comprising point clouds with class labels. The center point of a crown was defined by maximizing an index that combined normalized point intensity ρ_i, normalized height ρ_h, and distance to the crown centroid ρ_d (Equation A). Points within 25 % horizontal distance to the center point were assigned to class 1 (i.e., near-center), while all other points were assigned to class 2. The chosen threshold of 25 % could be arbitrary, but it should be contingent upon ensuring that the tree centers from the reference dataset do not overlap. A lower threshold would result in a smaller center region, thereby increasing the difficulty for the deep learning model in detecting center points. Conversely, a higher threshold might obscure the boundaries between proximate trees.

$$\text{center} = \text{argmax}(\exp[-(\rho_i - 1)^2] + \exp[-(\rho_h - 1)^2] + \exp[-{\rho_d}^2]) \quad \text{(A)}$$

The key steps and parameters of six classifiers and five segmenters are presented in Table B1. While most methods have tunable parameters depending on the landscape type, a classifier with fixed parameters applicable to a wide range of areas without manual intervention is more desirable in practice. Therefore, our benchmarking standard compared the mIoU accuracy across the four testing samples with fixed classifier parameters. The LAStools classification employed lasground, lasheight, and lasclassify modules to detect ground points, extract point heights, and identify clustered points over 2 m above ground, respectively (Varlik and Uray, 2023). Terrascan, another often-used ALS classification tool, primarily focused on extracting reliable tree points, leaving many residual crown points. An alternative approach involved classifying ground, building, surface, and low points (height below 0.3 m above ground), with the remaining points designated as tree points (Brovelli and Lucca, 2012). Default values from LASTools and Terrascan modules were considered their most generalizable parameter settings and were used for benchmarking. The RF classifier utilized intensity and returns, normalized by their maximum values, as independent variables. Following initial optimization using four training samples in Fig. 3, the RF classifier was then used in benchmarking. Intensity, return number, and their derivatives provided valuable information for identifying vegetation areas, species, or leaf area index (Ørka et al., 2009). Tree points typically exhibited lower intensity values and multiple return numbers compared to non-tree classes. Due to their classification capabilities, intensity and return number were added to the benchmarking as baseline classifiers. The intensity-based classifier simply filtered tree points with intensities below a threshold, while the return-based classifier identified points with total returns exceeding a threshold. It's noteworthy that the ALS data we utilized was a composite of point clouds from three spectral bands, whereas the intensity distributions for trees differed across bands. Although it would be possible to set band-specific intensity thresholds, tree points uniformly fell within a similar low-intensity range across all bands. For the sake of generalizability, this study chose not to incorporate spectral information into classification, even though multi-channel data was available. Instead, we opted for a single intensity threshold applicable across all bands.

Our ALS tree segmentation tool was compared with Segma (v 0.3.2) (St-Onge et al., 2015),lidR (Roussel et al., 2020) and MaskRCNN (He et al., 2017). Essentially, Segma detected tree apex locations from the canopy height model (CHM) and applied a watershed algorithm for crown segmentation. The input CHM rasters were created using LAStools (lasground, lasheight, and lasgrid), within-crown cavity with fewer than three pixels were filled, and subsequently smoothed using a Gaussian filter using a 5 × 5 grid cell window. The other settings for Segma were based on trial and error to optimize overall accuracy using our testing samples. The lidR tool provided the most updated and automated solutions for crown segmentation. The tested algorithms included the region-growing method lidR-li2012 from Li et al. (2012) and the tree-top searching with the region-growing method lidR-itcSegment from Dalponte and Coomes (2016). The MaskRCNN used the default settings from PyTorch (v2.1) with a ResNet50-FPN backbone (He et al., 2016; Lin et al., 2017). The training sample was 2D CHM image patches converted from a point cloud randomly clipped and rotated. The output crown masks were first converted to polygons whose enclosed points from the plot scans were assigned unique crown IDs. The patch size was 64 × 64 m^2 covering the our plot size; so no need to worry about the over-segmentation problem for crowns near the plot edge in our benchmarking. All listed segmentation tools exhibited a certain degree of tree omission in our testing plots; therefore, the resolution and scale-related parameters were adjusted to maximize tree detectability without introducing noise detection. Table B1 presents the final settings fixed for the benchmarking tests (Fig. B1).

Appendix C:. TLS tree classification

The same two classifiers were employed to train and test tree point classification for the TLS plots. The block size was adjusted to 128 × 128 × 128 voxels with a 5-cm voxel resolution. Any point with a height greater than 6.4 m, uncovered by the block size, was considered a tree point. Out of the ten TLS plots, four plots (LPine#1, LPine#3, LPine#6, and TAspen) were designated as testing plots, while the remaining six plots served as training plots for the tree classification models.

Fig. C1 visualizes the tree classification results for LPine#6 and TAspen by SegFormer, benchmarked with manual reference. The classification outcome closely matched the reference, with only the upper parts of a few understory saplings misclassified as trees. The classifiers did not rely on rigid elevation point thresholds or tree geometric constraints for tree detection. Consequently, some tilted stems and stem bases within thick ground layers were successfully identified as tree points.

The testing mIoU of TLS tree classification using SegFormer was 0.99 (LPine#1), 0.98 (LPine#3), 0.78 (LPine#6), and 0.94 (TAspen), yielding an average of 0.92. The corresponding mIoU using ResNet was 0.97, 0.98, 0.78, and 0.94, respectively, resulting in an identical average of 0.92.

References

Alonso-Rego, C., Arellano-Pérez, S., Guerra-Hernández, J., Molina-Valero, J.A., Martínez-Calvo, A., Pérez-Cruzado, C., Castedo-Dorado, F., González-Ferreiro, E., Álvarez-González, J.G., Ruiz-González, A.D., 2021. Estimating stand and fire-related surface and canopy fuel variables in pine stands using low-density airborne and single-scan terrestrial laser scanning data. Remote Sensing 13, 5170.

Asprs LAS Specification Version 1.4-R13. The American Society for Photogrammetry & Remote Sensing 2013 Bethesda, MD, USA.

Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J., Coomes, D., Piboule, A., Vincent, G., 2019. A comparative assessment of the performance of individual tree crowns delineation algorithms from ALS data in tropical forests. Remote Sensing 11, 1086.

Bang, K.I., Habib, A.F., Kusevic, K., Mrstik, P., 2008. Integration of terrestrial and airborne LiDAR data for system calibration. Proc. Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 391–398.

Besl, P.J., McKay, N.D., 1992. Method for registration of 3-D shapes, Robotics-DL tentative. International Society for Optics and Photonics 586–606.

Bridle, J.S., 1990. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. Neurocomputing. Springer 227–236.

Brovelli, M.A., Lucca, S., 2012. Comparison of GRASS-LiDAR modules–TerraScan with respect to vegetation filtering. Applied Geomatics 4, 123–134.

Chasmer, L., Hopkinson, C., Treitz, P., 2004. Assessing the three-dimensional frequency distribution of airborne and ground-based lidar data for red pine and mixed deciduous forest plots. Int Arch Photogramm Remote Sens Spat Inf Sci 36, W2.

Chasmer, L., Hopkinson, C., Treitz, P., 2006. Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial lidar. Canadian Journal of Remote Sensing 32, 116–125.

Chen, Q., 2015. Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing 106, 95–106.

Chen, Y., Xing, Y., Li, X., Gao, W., 2023. DGCN-ED: dynamic graph convolutional networks with encoder–decoder structure and its application for airborne LiDAR point classification. International Journal of Remote Sensing 44, 3489–3506.

Choi, J., Song, Y., Kwak, N., 2021. In: Part-Aware Data Augmentation for 3d Object Detection in Point Cloud, pp. 3391–3397.

Chu, X., Tian, Z., Wang, Y., Zhang, B., Ren, H., Wei, X., Xia, H., Shen, C., 2021. Twins: Revisiting the design of spatial attention in vision transformers. arXiv preprint arXiv: 2104.13840 1, 3.

Coomes, D.A., Dalponte, M., Jucker, T., Asner, G.P., Banin, L.F., Burslem, D.F., Lewis, S. L., Nilus, R., Phillips, O.L., Phua, M.-H., 2017. Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. Remote Sensing of Environment 194, 77–88.

Corte, A.P.D., Souza, D.V., Rex, F.E., Sanquetta, C.R., Mohan, M., Silva, C.A., Zambrano, A.M.A., Prata, G., Alves de Almeida, D.R., Trautenmüller, J.W., Klauberg, C., de Moraes, A., Sanquetta, M.N., Wilkinson, B., Broadbent, E.N., 2020. Forest inventory with high-density UAV-Lidar: Machine learning approaches for predicting individual tree attributes. Computers and Electronics in Agriculture 179, 105815.

Dalponte, M., Coomes, D.A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods in Ecology and Evolution 7, 1236–1245.

Dersch, S., Heurich, M., Krueger, N., Krzystek, P., 2021. Combining graph-cut clustering with object-based stem detection for tree segmentation in highly dense airborne lidar point clouds. ISPRS Journal of Photogrammetry and Remote Sensing 172, 207–222.

Diab, A., Kashef, R., Shaker, A., 2022. Deep learning for LiDAR point cloud classification in remote sensing. Sensors 22, 7868.

Donager, J.J., Sánchez Meador, A.J., Blackburn, R.C., 2021. Adjudicating perspectives on forest structure: how do airborne, terrestrial, and mobile lidar-derived estimates compare? Remote Sensing 13, 2297.

Fernandez-Diaz, J.C., Carter, W.E., Glennie, C., Shrestha, R.L., Pan, Z., Ekhtari, N., Singhania, A., Hauser, D., Sartori, M., 2016. Capability assessment and performance metrics for the titan multispectral mapping lidar. Remote Sensing 8, 936.

Ferraz, A., Bretar, F., Jacquemoud, S., Gonçalves, G., Pereira, L., Tomé, M., Soares, P., 2012. 3-D mapping of a multi-layered Mediterranean forest using ALS data. Remote Sensing of Environment 121, 210–223.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Garcia-Rodriguez, J., 2017. A Review on Deep Learning Techniques Applied to Semantic Segmentation. arXiv preprint arXiv:1704.06857.

Gardner, J., Pleiss, G., Weinberger, K.Q., Bindel, D., Wilson, A.G., 2018. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. Advances in neural information processing systems 31.

Gerrand, S., Aspinall, J., Jensen, T., Hopkinson, C., Collingwood, A., Chasmer, L., 2021. Partitioning carbon losses from fire combustion in a montane Valley. Alberta Canada. Forest Ecology and Management 496, 119435.

Girardeau-Montaut, D., 2021. CloudCompare, 2.12, Beta ed. GPL software, p. p..

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587.

Gleason, C.J., Im, J., 2012. Forest biomass estimation from airborne LiDAR data using machine learning approaches. Remote Sensing of Environment 125, 80–91.

Greaves, H.E., Vierling, L.A., Eitel, J.U., Boelman, N.T., Magney, T.S., Prager, C.M., Griffin, K.L., 2017. Applying terrestrial lidar for evaluation and calibration of airborne lidar-derived shrub biomass estimates in Arctic tundra. Remote Sensing Letters 8, 175–184.

Gupta, S., Weinacker, H., Koch, B., 2010. Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave lidar data. Remote Sensing 2, 968–989.

Hamedianfar, A., Mohamedou, C., Kangas, A., Vauhkonen, J., 2022. Deep learning for forest inventory and planning: a critical review on the remote sensing approaches so far and prospects for further applications. Forestry: an International Journal of Forest Research 95, 451–465.

Hastings, J.H., Ollinger, S.V., Ouimette, A.P., Sanders-DeMott, R., Palace, M.W., Ducey, M.J., Sullivan, F.B., Basler, D., Orwig, D.A., 2020. Tree species traits determine the success of LiDAR-Based crown mapping in a mixed temperate forest. Remote Sensing 12, 309.

Hauglin, M., Gobakken, T., Astrup, R., Ene, L., Næsset, E., 2014. Estimating single-tree crown biomass of Norway spruce by airborne laser scanning: a comparison of methods with and without the use of terrestrial laser scanning to obtain the ground reference data. Forests 5, 384–403.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn, Proceedings of the IEEE international conference on computer vision, pp. 2961-2969.

Hilker, T., van Leeuwen, M., Coops, N.C., Wulder, M.A., Newnham, G.J., Jupp, D.L.B., Culvenor, D.S., 2010. Comparing canopy metrics derived from terrestrial and airborne laser scanning in a Douglas-fir dominated forest stand. Trees 24, 819–832.

Hilker, T., Coops, N.C., Newnham, G.J., van Leeuwen, M., Wulder, M.A., Stewart, J., Culvenor, D.S., 2012. Comparison of terrestrial and airborne LiDAR in describing stand structure of a thinned lodgepole pine forest. Journal of Forestry 110, 97–104.

Holmgren, J., Lindberg, E., 2019. Tree crown segmentation based on a tree crown density model derived from Airborne Laser Scanning. Remote Sensing Letters 10, 1143–1152.

Hopkinson, C., Chasmer, L., Young-Pow, C., Treitz, P., 2004. Assessing forest metrics with a ground-based scanning lidar. Canadian Journal of Forest Research 34, 573–583.

Hopkinson, C., Chasmer, L., Colville, D., Fournier, R.A., Hall, R.J., Luther, J.E., Milne, T., Petrone, R.M., St-Onge, B., 2013a. Moving toward consistent ALS monitoring of forest attributes across canada. Photogrammetric Engineering & Remote Sensing 79, 159–173.

Hopkinson, C., Lovell, J., Chasmer, L., Jupp, D., Kljun, N., van Gorsel, E., 2013b. Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index. Remote Sensing of Environment 136, 301–314.

Hu, X., Hu, C., Han, J., Sun, H., Wang, R., 2023. Point cloud segmentation for an individual tree combining improved point transformer and hierarchical clustering. Journal of Applied Remote Sensing 17, 034505.

Huo, L., Lindberg, E., Holmgren, J., 2022. Towards low vegetation identification: a new method for tree crown segmentation from LiDAR data based on a symmetrical structure detection algorithm (SSD). Remote Sensing of Environment 270, 112857.

Isenburg, M., 2014. LAStools - efficient LiDAR processing software.

Jin, S., Su, Y., Zhao, X., Hu, T., Guo, Q., 2020. A point-based fully convolutional neural network for airborne lidar ground point filtering in forested environments. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, 3958–3974.

Jung, S.-E., Kwak, D.-A., Park, T., Lee, W.-K., Yoo, S., 2011. Estimating crown variables of individual trees using airborne and terrestrial laser scanners. Remote Sensing 3, 2346–2363.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sensing 4, 950–974.

Kelley, J., Trofymow, J.A., Bone, C., 2022. Combining area-based and individual tree metrics for improving merchantable and non-merchantable wood volume estimates in coastal douglas-fir forests. Remote Sensing 14, 2204.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Krzystek, P., Serebryanyk, A., Schnörr, C., Červenka, J., Heurich, M., 2020. Large-scale mapping of tree species and dead trees in šumava national park and bavarian forest national park using lidar and multispectral imagery. Remote Sensing 12, 661.
Kükenbrink, D., Schneider, F.D., Leiterer, R., Schaepman, M.E., Morsdorf, F., 2017. Quantification of hidden canopy volume of airborne laser scanning data using a voxel traversal algorithm. Remote Sensing of Environment 194, 424–436.
Lambert, M., Ung, C., Raulier, F., 2005. Canadian national tree aboveground biomass equations. Canadian Journal of Forest Research 35, 1996–2018.
Landrieu, L., Obozinski, G., 2017. Cut pursuit: fast algorithms to learn piecewise constant functions on general weighted graphs. SIAM Journal on Imaging Sciences 10, 1724–1766.
Lee, J., Bahri, Y., Novak, R., Schoenholz, S.S., Pennington, J., Sohl-Dickstein, J., 2017. Deep neural networks as gaussian processes. arXiv preprint arXiv:1711.00165.
Li, Y., Bu, R., Sun, M., Wu, W., Di, X., Chen, B., 2018. PointCNN: convolution On X-transformed points. Advances in Neural Information Processing Systems 828–838.
Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. Photogrammetric Engineering and Remote Sensing 78, 75–84.
Li, N., Kähler, O., Pfeifer, N., 2021. A comparison of deep learning methods for airborne lidar point clouds classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 6467–6486.
Liang, X., Hyyppä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., Holopainen, M., Brolly, G., Francesco, P., Hackenberg, J., Huang, H., Jo, H.-W., Katoh, M., Liu, L., Mokroš, M., Morel, J., Olofsson, K., Poveda-Lopez, J., Trochta, J., Wang, D., Wang, J., Xi, Z., Yang, B., Zheng, G., Kankare, V., Luoma, V., Yu, X., Chen, L., Vastaranta, M., Saarinen, N., Wang, Y., 2018. International benchmarking of terrestrial laser scanning approaches for forest inventories. ISPRS Journal of Photogrammetry and Remote Sensing 144, 137–179.
Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature pyramid networks for object detection, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2117-2125.
Lindberg, E., Holmgren, J., Olofsson, K., Olsson, H., 2012. Estimation of stem attributes using a combination of terrestrial and airborne laser scanning. European Journal of Forest Research 131, 1917–1931.
Lines, E.R., Allen, M., Cabo, C., Calders, K., Debus, A., Grieve, S.W., Miltiadou, M., Noach, A., Owen, H.J., Puliti, S., 2022. AI applications in forest monitoring need remote sensing benchmark datasets. arXiv preprint arXiv:2212.09937 2022.
Liu, L., Pang, Y., Li, Z., Si, L., Liao, S., 2017. Combining airborne and terrestrial laser scanning technologies to measure forest understorey volume. Forests 8, 111.
Liu, H., Shen, X., Cao, L., Yun, T., Zhang, Z., Fu, X., Chen, X., Liu, F., 2020. Deep learning in forest structural parameter estimation using airborne lidar data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 1603–1618.
Lovell, J., Jupp, D.L., Culvenor, D., Coops, N., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. Canadian Journal of Remote Sensing 29, 607–622.
Maltezos, E., Ioannidis, C., 2015. Automatic Detection of Building Points from Lidar and Dense Image Matching Point Clouds. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 2.
Matlab MATLAB 2020b 2020 The MathWorks Inc. Natick, Massachusetts, United State.
McCaffrey, D., Hopkinson, C., 2020. Repeat oblique photography shows terrain and fire-exposure controls on century-scale canopy cover change in the alpine treeline ecotone. Remote Sensing 12, 1569.
Oehmcke, S., Li, L., Trepekli, K., Revenga, J.C., Nord-Larsen, T., Gieseke, F., Igel, C., 2024. Deep point cloud regression for above-ground forest biomass estimation from airborne LiDAR. Remote Sensing of Environment 302, 113968.
Okyay, U., Telling, J., Glennie, C.L., Dietrich, W.E., 2019. Airborne lidar change detection: an overview of Earth sciences applications. Earth-Science Reviews 198, 102929.
Ørka, H.O., Næsset, E., Bollandsås, O.M., 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. Remote Sensing of Environment 113, 1163–1174.
Osco, L.P., Marcato Junior, J., Marques Ramos, A.P., de Castro Jorge, L.A., Fatholahi, S. N., de Andrade Silva, J., Matsubara, E.T., Pistori, H., Gonçalves, W.N., Li, J., 2021. A review on deep learning in UAV remote sensing. International Journal of Applied Earth Observation and Geoinformation 102, 102456.
Padilla, R., Netto, S.L., Da Silva, E.A., 2020. A survey on performance metrics for object-detection algorithms. 2020 international conference on systems, signals and image processing (IWSSIP) IEEE 237–242.
Paris, C., Kelbe, D., Aardt, J.v., Bruzzone, L., 2017. A novel automatic method for the fusion of ALS and TLS LiDAR data for robust assessment of tree crown structure. IEEE Transactions on Geoscience and Remote Sensing 55, 3679–3693.
Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., 2019. Pytorch: an imperative style, high-performance deep learning library. Advances in neural information processing systems 32.
Popescu, S.C., 2007. Estimating biomass of individual pine trees using airborne lidar. Biomass and Bioenergy 31, 646–655.
Pyörälä, J., Saarinen, N., Kankare, V., Coops, N.C., Liang, X., Wang, Y., Holopainen, M., Hyyppä, J., Vastaranta, M., 2019. Variability of wood properties using airborne and terrestrial laser scanning. Remote Sensing of Environment 235, 111474.
PyTorch, PyTorch 1.12.1 ed. 2019 Curran Associates Inc.
Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. arXiv preprint arXiv:1706.02413.
Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788.
Reitberger, J., Schnörr, C., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing 64, 561–574.
Rizaldy, A., Persello, C., Gevaert, C., Oude Elberink, S., Vosselman, G., 2018. Ground and multi-class classification of airborne laser scanner point clouds using fully convolutional networks. Remote Sensing 10, 1723.
Rocha, K.D., Silva, C.A., Cosenza, D.N., Mohan, M., Klauberg, C., Schlickmann, M.B., Xia, J., Leite, R.V., Almeida, D.R.A.d., Atkins, J.W., Cardil, A., Rowell, E., Parsons, R., Sánchez-López, N., Prichard, S.J., Hudak, A.T., 2023. Crown-Level structure and fuel load characterization from airborne and terrestrial laser scanning in a longleaf pine (pinus palustris mill.) forest ecosystem. Remote Sensing 15, 1002.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.
Roussel, J.-R., Auty, D., Coops, N.C., Tompalski, P., Goodbody, T.R., Meador, A.S., Bourdon, J.-F., de Boissieu, F., Achim, A., 2020. lidR: an R package for analysis of Airborne Laser Scanning (ALS) data. Remote Sensing of Environment 251, 112061.
Sackov, I., Hlasny, T., Bucha, T., Juris, M., 2017. Integration of tree allometry rules to treetops detection and tree crowns delineation using airborne lidar data. iForest - Biogeosciences and Forestry 10, 459–467.
Schmohl, S., Narváez Vallejo, A., Soergel, U., 2022. Individual tree detection in urban ALS point clouds with 3D convolutional networks. Remote Sensing 14, 1317.
Schneider, F.D., Kükenbrink, D., Schaepman, M.E., Schimel, D.S., Morsdorf, F., 2019. Quantifying 3D structure and occlusion in dense tropical and temperate forests using close-range LiDAR. Agricultural and Forest Meteorology 268, 249–257.
Seely, H., Coops, N.C., White, J.C., Montwé, D., Winiwarter, L., Ragab, A., 2023. Modelling tree biomass using direct and additive methods with point cloud deep learning in a temperate mixed forest. Science of Remote Sensing 8, 100110.
Sheridan, R.D., Popescu, S.C., Gatziolis, D., Morgan, C.L.S., Ku, N.-W., 2015. Modeling forest aboveground biomass and volume using airborne LiDAR metrics and forest inventory and analysis data in the pacific northwest. Remote Sensing 7, 229–255.
Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
Soininen, A., 2020. Terrascan user's guide. Terrasolid, Helsinki, Finland.
St-Onge, B., Audet, F.-A., Bégin, J., 2015. Characterizing the height structure and composition of a boreal forest using an individual tree crown approach applied to photogrammetric point clouds. Forests 6, 3899–3922.
Straker, A., Puliti, S., Breidenbach, J., Kleinn, C., Pearse, G., Astrup, R., Magdon, P., 2023. Instance segmentation of individual tree crowns with YOLOv5: a comparison of approaches using the ForInstance benchmark LiDAR dataset. ISPRS Open Journal of Photogrammetry and Remote Sensing 9, 100045.
Strîmbu, V.F., Strîmbu, B.M., 2015. A graph-based segmentation algorithm for tree crown extraction using airborne LiDAR data. ISPRS Journal of Photogrammetry and Remote Sensing 104, 30–43.
Sun, C., Huang, C., Zhang, H., Chen, B., An, F., Wang, L., Yun, T., 2022. Individual tree crown segmentation and crown width extraction from a heightmap derived from aerial laser scanning data using a deep learning framework. Frontiers in Plant Science 13, 914974.
Tao, S., Guo, Q., Li, L., Xue, B., Kelly, M., Li, W., Xu, G., Su, Y., 2014. Airborne Lidar-derived volume metrics for aboveground biomass estimation: a comparative assessment for conifer stands. Agricultural and Forest Meteorology 198, 24–32.
van Leeuwen, M., Hilker, T., Coops, N.C., Frazer, G., Wulder, M.A., Newnham, G.J., Culvenor, D.S., 2011. Assessment of standing wood and fiber quality using ground and airborne laser scanning: a review. Forest Ecology and Management 261, 1467–1478.
Varlik, A., Uray, F., 2023. Filtering airborne LIDAR data by using fully convolutional networks. Survey Review 55, 21–31.
Vastaranta, M., Kankare, V., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H., 2012. Combination of individual tree detection and area-based approach in imputation of forest variables using airborne laser data. ISPRS Journal of Photogrammetry and Remote Sensing 67, 73–79.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Advances in Neural Information Processing Systems 5998–6008.
Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry 85, 27–40.
Wang, Y., Hyyppä, J., Liang, X., Kaartinen, H., Yu, X., Lindberg, E., Holmgren, J., Qin, Y., Mallet, C., Ferraz, A., 2016. International benchmarking of the individual tree detection methods for modeling 3-D canopy structure for silviculture and forest ecology using airborne laser scanning. IEEE Transactions on Geoscience and Remote Sensing 54, 5011–5027.
Wang, Z., Li, P., Cui, Y., Lei, S., Kang, Z., 2023. Automatic detection of individual trees in forests based on airborne LiDAR data with a tree region-based convolutional neural network (RCNN). Remote Sensing 15, 1024.
Wang, X., Zhang, Y., Luo, Z., 2020. Combining trunk detection with canopy segmentation to delineate single deciduous trees using airborne LiDAR data. IEEE Access 8, 99783–99796.
Weinstein, B.G., Marconi, S., Aubry-Kientz, M., Vincent, G., Senyondo, H., White, E.P., 2020. DeepForest: a Python package for RGB deep learning tree crown delineation. Methods in Ecology and Evolution 11, 1743–1751.
Windrim, L., Bryson, M., 2019. Forest tree detection and segmentation using high resolution airborne LiDAR. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 3898–3904.

Windrim, L., Bryson, M., 2020. Detection, segmentation, and model fitting of individual tree stems from airborne laser scanning of forests using deep learning. Remote Sensing 12, 1469.
Wulder, M.A., Hermosilla, T., White, J.C., Coops, N.C., 2020. Biomass status and dynamics over Canada's forests: disentangling disturbed area from associated aboveground biomass consequences. Environmental Research Letters 15, 094093.
Xi, Z., Chasmer, L., Hopkinson, C., 2023. Delineating and reconstructing 3D forest fuel components and volumes with terrestrial laser scanning. Remote Sensing 15, 4778.
Xi, Z., Hopkinson, C., 2021. Detecting individual-tree crown regions from terrestrial laser scans with an anchor-free deep learning model. Canadian Journal of Remote Sensing 47, 228–242.
Xi, Z., Hopkinson, C., 2022. 3D graph-based individual-tree isolation (treeiso) from terrestrial laser scanning point clouds. Remote Sensing 14, 6116.
Xi, Z., Hopkinson, C., Chasmer, L., 2018. Filtering stems and branches from terrestrial laser scanning point clouds using deep 3-D fully convolutional networks. Remote Sensing 10, 1215.
Xiang, B., Wielgosz, M., Kontogianni, T., Peters, T., Puliti, S., Astrup, R., Schindler, K., 2023. Automated forest inventory: analysis of high-density airborne LiDAR point clouds with 3D deep learning. arXiv preprint arXiv:2312 15084.
Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M., Luo, P., 2021. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. arXiv preprint arXiv:2105.15203.
Zhang, J., Huang, S., Hogg, E.H., Lieffers, V., Qin, Y., He, F., 2014. Estimating spatial variation in Alberta forest biomass from a combination of forest inventory and remote sensing data. Biogeosciences 11, 2793–2808.
Zhang, J., Hu, X., Dai, H., Qu, S., 2020. DEM extraction from ALS point clouds in forest areas via graph convolution network. Remote Sensing 12, 178.
Zhang, Y., Liu, H., Liu, X., Yu, H., 2023. Towards intricate stand structure: a novel individual tree segmentation method for ALS point cloud based on extreme offset deep learning. Applied Sciences 13, 6853.