

Estimation of height and diameter at breast height of forest trees with multi-scale individual tree detection method and machine learning algorithms using airborne LiDAR data

M. Fallah¹, A.A. Matkan² and H. Aghighi^{3*}

1- Ph.D. Student, Remote Sensing and GIS Research Center, Faculty of Earth Science, Shahid Beheshti University, Tehran, Iran

2- Prof., Remote Sensing and GIS Research Center, Faculty of Earth Science, Shahid Beheshti University, Tehran, Iran

3*- Corresponding author, Assistant Prof., Remote Sensing and GIS Research Center, Faculty of Earth Science, Shahid Beheshti University, Tehran, Iran. E-mail: h_aghghi@sbu.ac.ir

Received: 24.10.2023

Accepted: 8.2.2024

Abstract

Background and Objectives: Estimating forest tree structural attributes such as height and diameter at breast height (DBH) is crucial for understanding the structure and management of forest resources. One important method for estimating these parameters is the individual tree detection (ITD) method using appropriate remote sensing data, such as airborne LiDAR data. However, it should be noted that different ITD methods have various limitations and capabilities and react differently to changes in forest tree species and the vertical structure of the canopy.

Methodology: This study presents a hybrid individual tree detection method that combines raster-based and point-based methods in a multi-scale framework to identify single trees from LiDAR data. In this method, tree crown scale levels are obtained from morphological filters in the canopy height model (CHM). Segmentation is then performed using a multi-scale method, and the results are merged. To better separate adjacent and understory trees, the point cloud inside the segments is analyzed using the probability density function, and tree crown segments are modified. After detecting single trees, DBH and height parameters were estimated using ground control data and extracted features from LiDAR data with machine learning algorithms, including random forest (RF), support vector machine (SVM), and cubist (CB), in the form of 10-fold nested cross-validation (10-fold NCV). The Boruta feature selection algorithm was used to identify the most important metrics based on the LiDAR point cloud, which played an effective role in improving the performance of machine learning algorithms. Due to limited access to LiDAR and ground data from Iran's forests, this study uses the NEWFOR single tree detection benchmark dataset, collected from forests of the Alpine region with a combination of different tree species and vertical canopy structures.

Results: Although understory trees cannot be extracted with the same accuracy as overstory trees, the results of this study showed that, on average, the developed multi-scale individual tree detection (MSITD) method detected 89% of the tree crowns in the highest height layer and the highest number of small overstory trees with a detection rate of 48% in the lowest height layer (2-5 meters). The analysis of the machine learning algorithms' results in estimating forest structural attributes showed that, despite slight differences in performance, the SVM algorithm performed better than the RF and CB algorithms in estimating both height and DBH attributes. For the height attribute, the mean values of RMSE, rRMSE, and R² in the SVM algorithm were 1.75 m, 9%, and 0.85, respectively. For the DBH attribute, the values obtained for RMSE, rRMSE, and R² were 4.74 cm, 19%, and 0.78, respectively.

Conclusion: The evaluation of the results showed that the methods presented in this study for identifying single trees and estimating forest tree structural attributes have high potential for practical applications.

Keywords: Canopy height model, forest, individual tree detection, LiDAR, machine learning, point cloud.



برآورد ارتفاع و قطر برابر سینه درختان جنگل با روش تشخیص تک درخت چندمقیاسی و الگوریتم‌های یادگیری ماشین با استفاده از داده‌های لایدار هوایی

محمد فلاح^۱، علی اکبر متکان^۲ و حسین عقیقی^{۳*}

۱- دانشجوی دکتری، مرکز پژوهش‌های سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران

۲- استاد، مرکز پژوهش‌های سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران

۳- نویسنده مسئول، استادیار، مرکز پژوهش‌های سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران. پست الکترونیک: h_aghghi@sbu.ac.ir

تاریخ پذیرش: ۱۴۰۲/۱۱/۱۹

تاریخ دریافت: ۱۴۰۲/۰۸/۰۲

چکیده

سابقه و هدف: برآورد مشخصه‌های ساختاری درختان جنگل همچون ارتفاع و قطر برابر سینه (DBH)، اهمیتی کلیدی در مدیریت منابع جنگلی دارد. یکی از روش‌های مطرح در برآورد این مشخصه‌ها، روش تشخیص تک درخت با استفاده از داده‌های سنجش از دوری مناسب مانند لایدار هوایی (Airborne LiDAR) است. با این حال، باید توجه داشت که روش‌های متفاوت تشخیص تک درخت توسعه داده شده با محدودیت‌ها و قابلیت‌های مختلف، عملکردهای متفاوتی را نسبت به تغییر در گونه‌های درختان جنگل و ساختار عمودی تاج از خود نشان می‌دهند.

مواد و روش‌ها: در این پژوهش، یک روش ترکیبی تشخیص تک درخت ارائه شده است که روش‌های رستریپایه و ابرنقطه مینا را در یک چهارچوب چندمقیاسی برای شناسایی تک درختان از داده‌های لایدار ترکیب می‌کند. در این روش، سطوح مقیاس تاج درختان بر اساس اعمال فیلترهای ریخت‌شناسی بر مدل ارتفاعی تاج محاسبه می‌شود. سپس قطعه‌بندی به صورت چندمقیاسی صورت می‌گیرد و نتایج با یکدیگر ادغام می‌شوند. به منظور جداسازی بهتر درختان مجاور و اشکوب زیرین، با استفاده از تابع چگالی احتمال، ابرنقاط داخل قطعه‌ها آنالیز می‌شوند و قطعه‌های تاج درختان اصلاح می‌شود. پس از تشخیص تک درختان، برخی مشخصه‌های ساختاری درختان جنگل شامل ارتفاع و DBH با استفاده از داده‌های مرجع زمینی و ویژگی‌های استخراج شده از داده لایدار با استفاده از الگوریتم‌های یادگیری ماشین شامل جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM) و کویبست (CB) در قالب اعتبارسنجی متقابل تودرتو ده‌قسمتی برآورد شدند و نتایج حاصل مقایسه شدند. در این پژوهش، الگوریتم انتخاب ویژگی بوروتا برای شناسایی مهم‌ترین ویژگی‌های مستخرج از داده لایدار در برآورد مشخصه‌های ارتفاع و DBH استفاده شد. این الگوریتم، نقش مؤثری در بهبود عملکرد الگوریتم‌های رگرسیون ایفا کرد. با توجه به دسترسی محدود به داده‌های لایدار و زمینی از جنگل‌های ایران، به منظور ارزیابی روش‌های فوق‌الذکر، در این پژوهش از مجموعه داده معیار تشخیص تک درخت از پروژه تحقیقاتی NEWFOR استفاده شد که از جنگل‌های منطقه آلپ با گونه‌های مختلف درختان و ساختار عمودی تاج متفاوت جمع‌آوری شده است.

نتایج: اگرچه درختان اشکوب زیرین را نمی‌توان با دقت معادل درختان اشکوب غالب استخراج کرد، اما نتایج این پژوهش نشان داد که روش توسعه یافته در مناطق جنگلی مختلف، علاوه بر تشخیص ۸۹ درصد درختان در بالاترین لایه ارتفاعی، بیشترین تعداد درختان اشکوب زیرین را با نرخ تشخیص ۴۸ درصد در پایین‌ترین لایه ارتفاعی (دو تا پنج متر) تشخیص داده است. تجزیه و تحلیل نتایج حاصل از الگوریتم‌های یادگیری ماشین در برآورد ویژگی‌های ساختاری جنگل نشان داد که با وجود تفاوت ناچیز در عملکرد الگوریتم‌های مورد استفاده، الگوریتم ماشین بردار پشتیبان، عملکرد بهتری را نسبت به الگوریتم‌های جنگل تصادفی و کویبست در برآورد ارتفاع و قطر برابر سینه ارائه کرد. برای مشخصه ارتفاع، میانگین RMSE، r^2 و در الگوریتم SVM به ترتیب برابر با ۱/۷۵ متر، ۹ درصد و ۰/۸۵ بودند. از طرف دیگر، برای مشخصه DBH، مقدار معیارهای مذکور به ترتیب برابر با ۴/۷۴ سانتی‌متر، ۱۹ درصد و ۰/۷۸ به دست آمد.

نتیجه گیری کلی: ارزیابی نتایج به دست آمده نشان داد که به طور کلی، روش های ارائه شده در این پژوهش در خصوص تشخیص تک درختان و برآورد برخی مشخصه های ساختاری درختان جنگل، بهبود قابل توجهی را در مقایسه با الگوریتم های معیار ارائه می دهند و پتانسیل زیادی برای کاربردهای عملی دارند.

واژه های کلیدی: ابرنقطه، تشخیص تک درخت، جنگل، لایدار، مدل ارتفاعی تاج، یادگیری ماشین.

مقدمه

برآورد دقیق مشخصه های ساختاری جنگل، اهمیتی کلیدی در درک ساختار، نظارت بر جنگل ها و مدیریت منابع جنگلی دارد (Slik *et al.*, 2010; Poorazimy *et al.*, 2023; Sefidi & Jahdi, 2023). مهم ترین مشخصه های ساختاری جنگل شامل ارتفاع درخت، قطر برابر سینه (Diameters at Breast Height (DBH))، رویه زمینی (Basal area)، مساحت تاج (Crown area)، حجم سرپا (Stand volume) و زی توده روی زمین (Aboveground Biomass (AGB)) هستند. از میان این مشخصه ها، ارتفاع و قطر برابر سینه درخت در برآورد مشخصه های دیگر ساختاری جنگل، اهمیت بسیاری دارند (Ghaderi *et al.*, 2023). ارتفاع درخت، تابعی از قطر برابر سینه است و مشخصاتی مانند حجم و زی توده با معادله های آلومتریک براساس ارتفاع و قطر برابر سینه درختان برآورد می شود (Chave *et al.*, 2005). امروزه فناوری لایدار هوایی (Airborne LiDAR)، به عنوان یک سنجنده فعال سنجنش ازدور، ابزار مفیدی را در پژوهش ها و کاربردهای مربوط به جنگل از جمله برآورد مشخصه های ساختاری جنگل فراهم کرده است. داده های ابر نقاط لایدار جمع آوری شده را می توان برای درک ساختار سه بعدی تاج درختان در لایه های ارتفاعی مختلف به کار برد و نتایجی با دقت زیاد را تولید کرد (Hyypä *et al.*, 2012).

روش های برآورد مشخصه های ساختاری جنگل که مبتنی بر استفاده از داده های لایدار هوایی هستند را می توان به دو گروه عمده تقسیم کرد: روش های ناحیه مبنا (Area-based approach (ABA)) و روش های تشخیص تک درخت (Individual tree detection (ITD)). در روش های ناحیه مبنا، برآورد مشخصه های جنگل براساس روابط آماری و رگرسیون بین داده های مرجع زمینی و ویژگی های به دست آمده از داده های لایدار در سطح قطعه نمونه است (Hyypä *et al.*, 2012). در روش های تشخیص تک درخت، تک درختان شناسایی شده و سپس، مشخصه های ساختاری جنگل در سطح تک درخت استخراج می شوند، بنابراین روش های ITD صحت و قابلیت اطمینان بیشتری نسبت به روش های ناحیه مبنا دارند (Yu *et al.*, 2011). روش های ITD که مبتنی بر استفاده از داده های لایدار هوایی هستند به سه دسته کلی تقسیم می شوند: روش های رستر مبنا (Raster-based method)، روش های ابرنقطه مبنا (Point-based method) و روش های ترکیبی. روش های رستر مبنا از داده رستری حاصل از درون یابی ابر نقاط لایدار استفاده می کنند. رایج ترین داده رستری مورد استفاده، مدل ارتفاعی تاج (Canopy height model (CHM)) است (Wang *et al.*, 2016). روش های تشخیص تک درخت ابرنقطه مبنا به طور مستقیم شکل سه بعدی تاج تک درخت را از ابر نقاط شناسایی و استخراج می کنند. با توجه به استفاده از همه ابر نقاط سه بعدی لایدار، این

پیشین، الگوریتم‌های ML نسبت به الگوریتم‌های رگرسیون سنتی، مانند رگرسیون خطی، مزایای زیادی دارند. الگوریتم‌های ML می‌توانند به‌طور مؤثری خود را با مجموعه داده‌های چندبعدی و ویژگی‌های خاص داده‌های لایدار از جمله توزیع نرمال، غیرخطی بودن و وجود نقاط پرت تطبیق دهند (Dalla Corte et al., 2020). در سال‌های اخیر، الگوریتم‌های یادگیری ماشین متفاوتی همچون جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM)، شبکه‌های عصبی و k-نزدیک‌ترین همسایگی به‌منظور برآورد مشخصه‌های ساختاری جنگل با استفاده از داده‌های لایدار به‌کار گرفته شده‌اند (Dalla Corte et al., 2020). علی‌رغم عملکرد مناسب این الگوریتم‌ها در برآورد مشخصه‌های ساختاری درختان جنگل با استفاده از داده‌های لایدار، ارتقای دقت در فرایند برآورد مشخصه‌های ساختاری جنگل همواره مورد توجه پژوهشگران بوده است.

از سوی دیگر، روش‌های انتخاب ویژگی که سعی در شناسایی مهم‌ترین ویژگی‌های (متغیرهای مستقل توصیف‌گر) مؤثر بر برآورد مشخصه‌های ساختاری جنگل دارند، می‌توانند نقش مؤثری در بهبود عملکرد مدل‌های رگرسیونی داشته باشند (Frénay et al., 2013). انتخاب ویژگی می‌تواند به کاهش پیچیدگی محاسباتی و نیز افزایش دقت برآورد منجر شود. در واقع، اجرای فرایند انتخاب ویژگی، پیش از اجرای الگوریتم رگرسیونی سبب آموزش سریع‌تر مدل، کاهش پیچیدگی آن و در نهایت، افزایش دقت مدل رگرسیونی می‌شود. با توجه به اهمیت و گستردگی استفاده از فرایند انتخاب ویژگی در پژوهش‌های مختلف، تاکنون رویکردهای متفاوتی در این زمینه توسعه داده شده‌اند. در میان روش‌های مورد استفاده در این زمینه می‌توان به الگوریتم بوروتا (Boruta) اشاره کرد که به دلیل عملکرد

روش‌ها قابلیت مدل‌سازی اشکوب درختان در لایه‌های ارتفاعی مختلف را دارند (Wang et al., 2008). روش‌های خوشه‌بندی K-means و قطعه‌بندی بر مبنای وکسل (Voxel) را می‌توان به‌عنوان دو روش رایج ابرنقطه مینا نام برد.

با پیشرفت سیستم‌های لایدار و امکان برداشت ابرنقاط با تراکم زیاد، بسیاری از پژوهشگران، فواید ترکیب روش‌های رستر مینا و ابرنقطه مینا را بررسی کرده‌اند (Yang et al., 2021; Hui et al., 2020). در روش‌های ترکیبی تشخیص تک‌درخت، روش‌های ابرنقطه مینا و رستر مینا (مبتنی بر CHM) ترکیب می‌شوند تا از مزایای هر دو روش به‌طور هم‌زمان استفاده شود و دقت شناسایی تک‌درختان بهبود یابد. روش‌های ابرنقطه مینا به دلیل استفاده مستقیم از ابرنقاط زمان‌بر هستند. در مقابل، روش‌های رستر مینا به نسبت سریع هستند، از طرفی، روش‌های رستر مینا در تشخیص اشکوب زیرین، محدودیت دارند. در روش‌های ترکیبی ابتدا تک‌درختان در داده رستری شناسایی می‌شوند. سپس، به‌منظور جداسازی بهتر درختان مجاور و اشکوب زیرین، ابرنقاط داخل قطعه‌های تاج درختان آنالیز می‌شود و قطعه‌های تاج درختان اصلاح می‌شوند (Hui et al., 2021). پس از تشخیص تک‌درختان، مشخصه‌های ساختاری جنگل بر اساس روابط آماری و رگرسیون بین داده‌های مرجع زمینی و ویژگی‌های به‌دست آمده از داده‌های لایدار در سطح تک‌درخت برآورد می‌شوند. برای برآورد مقدار مرجع مشخصه‌هایی مانند زی‌توده بالای سطح زمین نیز از معادله‌های آلومتریک استفاده می‌شود (Hyypä et al., 2012; Yu et al., 2011; Yang et al., 2020). تاکنون الگوریتم‌های مختلف آماری و یادگیری ماشین (ML) برای برآورد ویژگی‌های ساختاری جنگل‌ها با استفاده از مجموعه داده‌های لایدار به‌کار گرفته شده‌اند. با توجه به پژوهش‌های

(Sparks *et al.*, 2022). Malek و همکاران (۲۰۱۹) برای برآورد DBH و AGB در سطح تک‌درخت از روش‌های SVM و RF استفاده کردند. آن‌ها ابتدا با روش تشخیص تک‌درخت رستر مینا (مبتنی بر CHM) تاج تک‌درختان را استخراج کردند. تاج تک‌درختان استخراج‌شده از لایدار به‌طور خودکار به درختان اندازه‌گیری‌شده میدانی منطبق شد و برای هر درخت شناسایی‌شده، ویژگی‌های ابر نقاط لایدار استخراج شد. نتایج پژوهش مذکور نشان داد که روش پیشنهادی در مقایسه با مدل رگرسیون خطی کمترین مربعات، دقت بهتری دارد. Yu و همکاران (۲۰۱۱) مشخصه‌های ساختاری جنگل مانند ارتفاع، DBH و حجم تنه را از داده‌های لایدار هوایی با روش جنگل تصادفی استخراج کردند. آن‌ها برای تشخیص تک‌درخت از روش ITD بر مبنای CHM استفاده کردند.

تعدادی از پژوهش‌های داخل کشور مربوط به داده‌های لایدار جنگل‌های هیرکانی شصت کلاته گرگان است که با تراکم چهار نقطه در مترمربع و با استفاده از سنجنده Riegl LMS Q560 برداشت شده است. به‌منظور ارزیابی قابلیت داده‌های لایدار هوایی در برآورد ارتفاع پایه‌های درختان جنگلی پلت (*Acer velutinum* Boiss.) و ممرز (*Carpinus betulus* L.) در جنگل‌های مذکور، ۱۰۰ پایه درختی از این دو گونه که در اشکوب چیره واقع شده بودند یا تداخل تاجی با پایه‌های مجاور نداشتند، انتخاب شدند (Khorami *et al.*, 2014). درختان برداشت‌شده در زمین با استفاده از مختصات مرکز و قطر تاج در ابرنقاط لایدار بازیابی شدند و استخراج ارتفاع به‌طور مستقیم از ابر نقاط لایدار صورت گرفت. نتایج حاصل از روابط رگرسیونی بین ارتفاع درختان اندازه‌گیری‌شده در عرصه و استخراج‌شده از ابرنقاط لایدار، دقت مطلوبی ارائه کرد. به‌منظور بررسی

بسیار مناسب آن به‌شکل ویژه‌ای موردتوجه پژوهشگران مختلف قرار گرفته است (Kursa & Rudnicki, 2010). از مسائل مهم دیگر در برآورد مشخصه‌های ساختاری جنگل با استفاده از داده‌های لایدار و الگوریتم‌های یادگیری ماشین، فرایند ارزیابی کارایی الگوریتم‌های مورد استفاده است (Hyypä *et al.*, 2012). در صورت کم بودن تعداد نمونه‌های زمینی در برآورد مشخصه‌های ساختاری جنگل، اعتبارسنجی متقابل تودرتو (Nested cross-validation) مورد توجه پژوهشگران قرار گرفته است. همان‌طور که در پژوهش‌های مختلف اشاره شده است، اعتبارسنجی متقابل تودرتو می‌تواند برآوردی منصفانه‌تر و دقیق‌تر از عملکرد یک مدل ارائه دهد (Krstajic *et al.*, 2014). در این رویکرد به‌صورت تصادفی و تکراری، بخشی از داده‌ها برای اعتبارسنجی مدل جدا می‌شوند و بقیه داده‌ها برای آموزش یک مدل استفاده می‌شوند. این فرایند چندین‌بار تکرار می‌شود تا همه داده‌ها برای اعتبارسنجی و آموزش به‌کار گرفته شوند. سپس، میانگین عملکردهای مدل در هر یک از این تکرارها به‌عنوان عملکرد مدل محاسبه می‌شود. گفتنی است که استفاده از همه داده‌های موجود برای اعتبارسنجی سبب کاهش تأثیر نویز و خطای نمونه‌ها در ارزیابی دقت و عملکرد مدل می‌شود (Vabalas *et al.*, 2019).

تاکنون، پژوهش‌های بسیاری درمورد قابلیت داده‌های لایدار هوایی برای برآورد مشخصه‌های ساختاری جنگل انجام شده است. ارزیابی عملکرد روش‌های مختلف تشخیص تک‌درخت (روش‌های رستر مینا، ابرنقطه مینا و ترکیبی) و برآورد ارتفاع درختان با استفاده از مدل رگرسیون خطی در توده‌های جنگلی مخروطی با ساختار و ترکیب متنوع نشان داد که همه روش‌های ITD، ارتفاع درختان شناسایی‌شده را با $RMSE$ کمتر از یک متر برآورد می‌کنند

از الگوریتم جنگل تصادفی (RF) در جنگل‌های شصت کلاته گرگان نشان داد که ترکیب داده‌های لایدار هوایی و تصاویر UltraCam-D سبب افزایش جزئی در دقت پیش‌بینی در مقایسه با مدل‌سازی جداگانه می‌شود (Mohammadi *et al.*, 2017).

باتوجه به پیشینه ارائه‌شده، پژوهش پیش‌رو بر چهار مرحله اصلی در برآورد ویژگی‌های درختان جنگل با استفاده از داده‌های لایدار تمرکز دارد. در مرحله اول، با معرفی یک رویکرد جدید، روش‌های رستریایه و ابرنقطه مبنا در یک چهارچوب چندمقیاسی برای شناسایی و جداسازی تک‌درختان در داده‌های لایدار ترکیب شدند که باعث افزایش معنی‌داری در جداسازی اشکوب‌های زیرین جنگل نسبت به روش‌های مرسوم شده است. سپس، از الگوریتم بوروتا برای انتخاب ویژگی‌های مناسب از داده‌های لایدار استفاده شد. در نهایت، الگوریتم‌های یادگیری ماشین به منظور برآورد مشخصه‌های ساختاری درختان جنگل شامل ارتفاع و DBH در قالب اعتبارسنجی متقابل تودرتو ده‌قسمتی (-10 fold NCV) برای آموزش و ارزیابی دقت مدل‌های رگرسیون به‌کار گرفته شدند. در بخش دوم به معرفی مناطق پژوهش، مجموعه داده‌های لایدار و زمینی مرجع مورد استفاده و پیش‌پردازش‌های لازم پرداخته می‌شود. همچنین، مراحل پیاده‌سازی الگوریتم‌های مورد استفاده در تشخیص تک‌درخت و برآورد مشخصه‌های ساختاری درختان جنگل با استفاده از داده‌های لایدار تشریح خواهند شد. در بخش سوم، نتایج حاصل از پیاده‌سازی روش‌ها ارائه می‌شوند. در بخش چهارم، نتایج حاصل از پژوهش موردبحث و بررسی قرار می‌گیرند.

قابلیت داده‌های لایدار هوایی در برآورد مساحت تاج و ارتفاع درختان جنگلی در بخشی از جنگل‌های طبیعی شصت کلاته گرگان، ۱۱۷ پایه درختی از درختان که دارای تاج مستقل بودند و تداخل تاجی با پایه‌های مجاور نداشتند به روش تصادفی انتخاب شدند (Seyed Mousavi *et al.*, 2017). برای جداسازی مرز دقیق تاج تک‌درختان و تهیه پلیگون مرز آن‌ها از تصاویر دوربین رقومی هوایی UltraCam-D منطقه مورد مطالعه با دقت ۰/۵ سانتیمتر استفاده شد. داده‌های لایدار هوایی مربوط به هر درخت نمونه با استفاده از مختصات مرکز درختان و پلیگون تهیه‌شده از تاج درختان نمونه نیز استخراج شد. به منظور تعیین ارتفاع و قطر بزرگ و کوچک تاج درختان نمونه از نرم‌افزار Fusion استفاده شد. نتایج حاصل از روابط رگرسیونی بین ارتفاع و مساحت تاج درختان نشان داد که داده‌های لایدار هوایی، قابلیت زیادی در برآورد مساحت تاج و ارتفاع درختان دارای تاج مستقل دارند. در پژوهش دیگری، به منظور بررسی قابلیت داده‌های لایدار هوایی در برآورد ارتفاع، حجم، رویه زمینی، قطر برابرسینه و مساحت تاج درختان جنگلی در همین داده، ۱۲۵ پایه درختی از درختان که دارای تاج مستقل بودند و تداخل تاجی با پایه‌های مجاور نداشتند، به روش تصادفی انتخاب شدند (Seyed Mousavi *et al.*, 2019). سپس، شاخص‌های ارتفاعی و تراکمی داده‌های لایدار هوایی استخراج شد. در نهایت، مشخصه‌های ساختاری جنگل با استفاده از الگوریتم‌های رگرسیون ناپارامتریک (جنگل تصادفی، نزدیک‌ترین همسایه، شبکه عصبی مصنوعی و ماشین بردار پشتیبان) و رگرسیون خطی چندمتغیره برآورد شد. بررسی قابلیت استفاده ترکیبی از داده‌های لایدار هوایی و تصاویر رقومی UltraCam-D برای مدل‌سازی حجم سرپا، تراکم درخت و رویه زمینی با استفاده

مواد و روش‌ها

منطقه مورد مطالعه و داده‌های پژوهش

باتوجه به دسترسی محدود به داده‌های ابرنقاط لایدار و داده‌های زمینی جامع از جنگل‌های ایران، از مجموعه داده معیار تشخیص تکدرخت لایدار از پروژه تحقیقاتی NEWFOR استفاده شد که ابرنقاط لایدار با تراکم متفاوت از جنگل‌های مختلف منطقه آلپ با تنوع گونه‌های درختان را پوشش می‌دهد (NEWFOR, 2015). این مجموعه داده به منظور ارتقا و ترویج استفاده از فناوری جدید لایدار برای استفاده پژوهشگران به طور عمومی در وبگاه پروژه مذکور منتشر شده است و تاکنون در پژوهش‌های مختلف مرتبط با روش‌های تشخیص تکدرخت استفاده شده است (Eysn et al., 2015; Dong et al., 2018; Hui et al., 2021; Wang et al., 2023). جنگل‌های منطقه آلپ در پنج کشور فرانسه، ایتالیا، سوئیس، اتریش و اسلوونی قرار دارند. از میان مناطق مطالعاتی موجود در این مجموعه داده، هشت منطقه مطالعاتی شامل منطقه یک واقع در Saint-Agnan از کشور فرانسه، مناطق دو و سه در Cotelivier از کشور ایتالیا، مناطق چهار و پنج واقع در Pellizzano از کشور ایتالیا و مناطق شش، هفت و هشت در Asiago از کشور ایتالیا به منظور ارزیابی عملکرد روش‌های تشخیص تکدرخت استفاده شدند (شکل ۱). داده‌های پروژه NEWFOR، متشکل از ابرنقاط لایدار هوایی با فرمت LAS است. تراکم ابرنقاط برداشتی در این پروژه از ۱۰ نقطه در مترمربع (pts/m^2) در منطقه Asiago تا ۱۲۱ نقطه در مترمربع در منطقه Pellizzano واقع در ایتالیا متفاوت است. علاوه بر داده‌های لایدار، داده‌های مرجع زمینی با فرمت شیپ‌فایل (Shape file) شامل مشخصه‌های ساختاری درختان مانند موقعیت مکانی درختان، قطر برابرسینه، ارتفاع درختان و نیز اطلاعاتی در مورد ترکیب

گونه‌ها و تعداد اشکوب درختان در لایه‌های ارتفاعی مختلف در همه مناطق با اندازه‌گیری‌های دقیق زمینی نیز به اشتراک گذاشته شده است. برای تعیین موقعیت مکانی درختان هر منطقه از سیستم تعیین موقعیت GPS و برداشت زمینی توتال استیشن و برای اندازه‌گیری ارتفاع درختان از سیستم‌های Vertex و فاصله‌یاب لیزری TruPulse 360 با دقت ± 1 متر استفاده شد. اطلاعات مرتبط با مناطق مورد مطالعه و داده‌های مورد استفاده در جدول ۱ ارائه شده است.

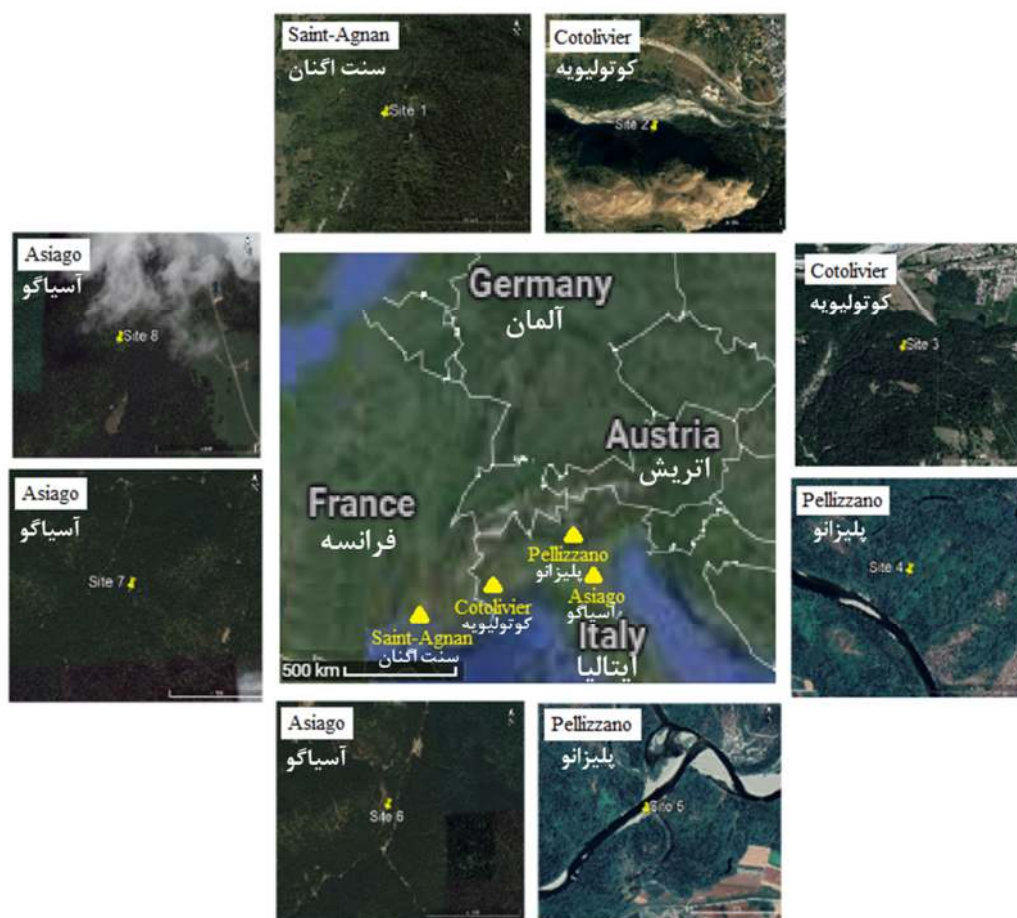
جنگل‌های منطقه پژوهش با تفسیر توزیع ارتفاعی درختان و نوع گونه‌های موجود در داده مرجع زمینی طبقه‌بندی شدند. این جنگل‌ها از نظر تعداد اشکوب درختان در لایه‌های ارتفاعی مختلف، به صورت جنگل‌های یک‌اشکوبه و چنداشکوبه و از نظر نوع گونه درختان به صورت سوزنی‌برگ و آمیخته نام‌گذاری شدند. با طبقه‌بندی انجام‌شده، چهار نوع جنگل در مناطق پژوهش وجود دارند: جنگل سوزنی‌برگ یک‌اشکوبه، جنگل آمیخته یک‌اشکوبه، جنگل سوزنی‌برگ چنداشکوبه و جنگل آمیخته چنداشکوبه. گونه‌های درختان موجود در جنگل‌های مورد پژوهش عبارت‌اند از: درختان سوزنی‌برگ مانند کاج نوئل (Spruce, Fir, *Abies alba* (L.) H. Karst. کاج نراد (Scots pine, *Pinus sylvestris* L.) (Mill.) کاج جنگلی (Larch, *Larix decidua* Mill.) و درختان پهن‌برگ (خزان‌کننده) مانند افرا (Sycamore, *Acer pseudoplatanus* L. نارون (Elm, *Ulmus glabra* (Huds.) صنوبر (Poplar, *Populus nigra* L.) و راش اروپایی (Beech, *Fagus sylvatica* L.).

جدول ۱- مشخصات کلی مناطق مورد مطالعه و داده لایدار مورد استفاده

Table 1. General characteristics of the study areas and LiDAR data

Area ID	Area name	Airborne LiDAR data		Mean tree height (m)	Mean diameter (cm)	Main tree species	Forest type
		Sensor	Point density				
1	Saint-Agnan	Riegl LMS-Q560	13	17.1	30.1	<i>Fir, beech</i>	ML/M
2	Cotolivier	Optech ALTM 3100	12	18.1	25.8	<i>Scots pine, larch, spruce</i>	ML/C
3	Cotolivier	Optech ALTM 3100	11	16.5	29.7	<i>Scots pine, larch</i>	SL/C
4	Pellizzano	Riegl LMS-Q680i	95	13.7	16.7	<i>Larch, spruce, fir, sycamore, poplar</i>	ML/M
5	Pellizzano	Riegl LMS-Q680i	121	25.6	40.9	<i>Spruce, larch, fir</i>	SL/C
6	Asiago	Optech ALTM 3100	11	16.9	23.7	<i>Spruce, fir, beech</i>	ML/M
7	Asiago	Optech ALTM 3100	11	13.9	16	<i>Spruce, fir, beech</i>	ML/M
8	Asiago	Optech ALTM 3100	10	23.6	29.5	<i>Spruce, fir, beech</i>	SL/M

SL/C: Single-Layered Coniferous forest; SL/M: Single-Layered Mixed forest; ML/C: Multi-Layered Coniferous forest; ML/M: Multi-Layered Mixed forest.

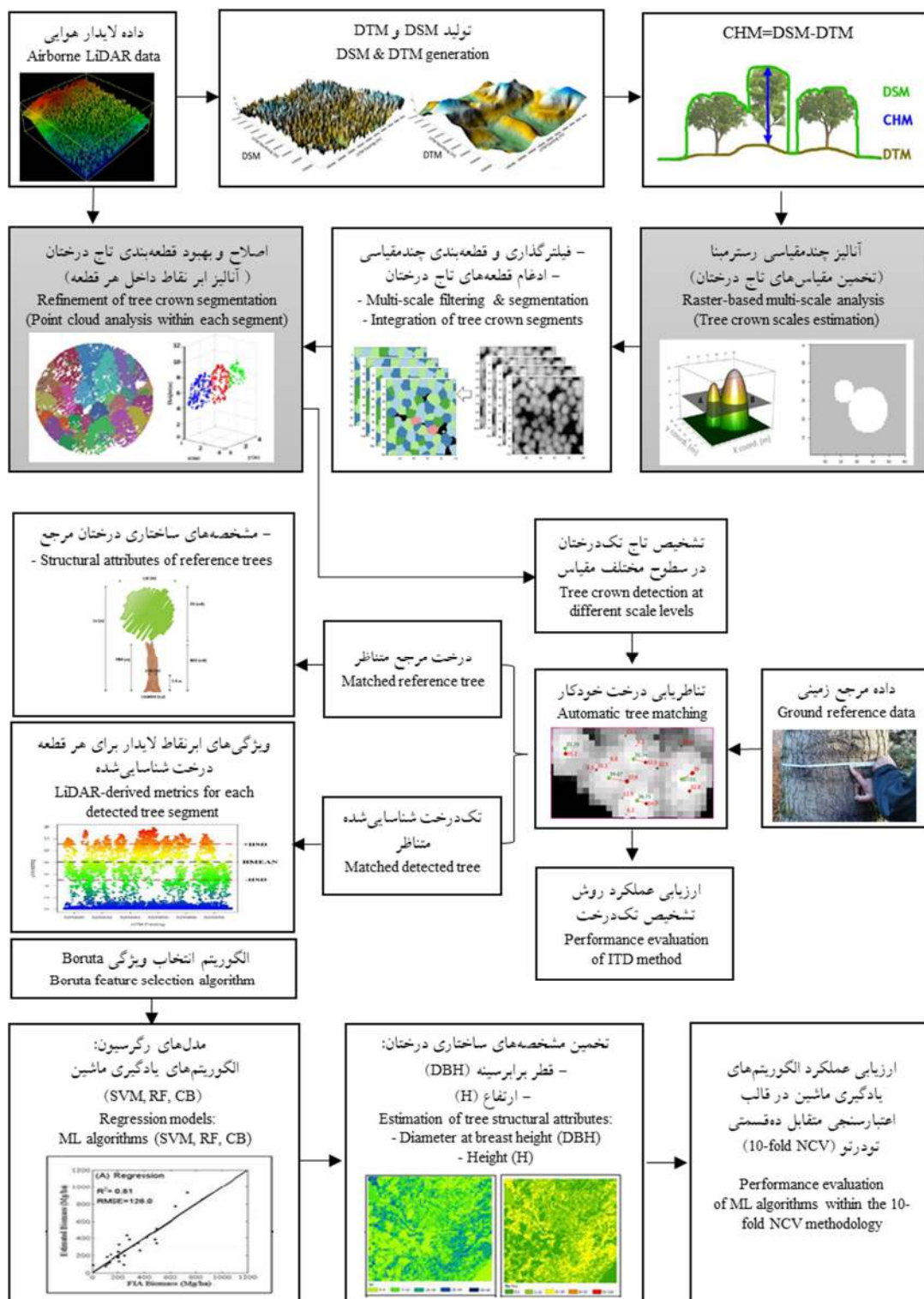


شکل ۱- موقعیت هشت منطقه مورد مطالعه

Figure 1. Location of the eight study areas

روش پژوهش
 مراحل اجرای پژوهش در شکل ۲ نشان داده شده است
 که در ادامه تشریح خواهند شد. گفتنی است که کدنویسی در
 این پژوهش در محیط نرم‌افزار RStudio انجام گرفت. از
 بسته‌های منبع باز نرم‌افزار R نیز مانند lidR (Roussel et al., 2018; Silva et al., 2017; Silva et al., 2018) استفاده شد که به منظور پردازش داده‌های لایدار برای کاربردهای مرتبط با جنگل در اختیار عموم قرار گرفته است.

روش پژوهش
 مراحل اجرای پژوهش در شکل ۲ نشان داده شده است
 که در ادامه تشریح خواهند شد. گفتنی است که کدنویسی در
 این پژوهش در محیط نرم‌افزار RStudio انجام گرفت. از



شکل ۲- مراحل اجرای پژوهش

Figure 2. Workflow of the methodology

لایدار انجام گرفت. این مرحله بسیار مهم است. زیرا بازگشت‌های دارای نویز می‌توانند بر دقت CHM تولیدشده و نتایج تشخیص درخت بعدی تأثیر بگذارند. برای رسیدن

پیش‌پردازش داده لایدار برای هر مجموعه از داده لایدار مورد استفاده (جدول ۱)، ابتدا پیش‌پردازش‌هایی مانند حذف نویز روی داده‌های

که در قسمت پیش‌پردازش داده لایدار توضیح داده شد. در مرحله دوم، تجزیه و تحلیل مقیاس برای برآورد اندازه تاج درختان در مدل ارتفاعی تاج با استفاده از فیلترهای ریخت‌شناسی صورت گرفت. در این مرحله برای آنالیز سطوح مقیاس تاج درختان از اپراتورهای ریخت‌شناسی opening با المان‌های ساختاری با شکل و اندازه مناسب (قطر سه تا ۵۱ پیکسل با طول گام دو پیکسل) استفاده شد تا درختان با اندازه مختلف را در تصویر CHM شناسایی کنند. با اینکه تغییرات اندازه تاج درختان یا همان سطوح مقیاس موجود در CHM زیاد است، اما چندین گروه اندازه غالب تاج درختان وجود دارد. برطبق نتایج حاصل از اعمال اپراتورهای ریخت‌شناسی با شکل و اندازه المان ساختاری مختلف روی CHM با قدرت تفکیک ۰/۵ متر، اندازه تاج درختان در مناطق مختلف مورد مطالعه در سه گروه هفت تا ۱۱ (۳/۵ تا ۵/۵ متر)، ۱۵ تا ۱۹ (۷/۵ تا ۹/۵ متر) و ۲۱ تا ۲۷ (۱۰/۵ تا ۱۳/۵ متر) پیکسل تقسیم‌بندی شدند که می‌توان به ترتیب به اندازه تاج درختان کوچک، متوسط و بزرگ تعبیر کرد. با در نظر گرفتن کمینه مقدار در هر گروه اندازه تاج درخت، سه سطح مقیاس هفت، ۱۵ و ۲۱ پیکسل به‌عنوان مناسب‌ترین مقدار به ترتیب برای تاج درختان کوچک، متوسط و بزرگ تعیین شدند.

مرحله سوم، فیلترگذاری چندمقیاسی است. در این مرحله، براساس برآورد اندازه تاج درختان که در مرحله آنالیز مقیاس تعیین شد، مطابق رابطه (۱) مجموعه‌ای از فیلترهای گوسین طراحی و روی CHM اعمال شد. در این پژوهش، اندازه پنجره فیلتر گوسین در سطح مقیاس k پیکسل (7, 15, 21) برابر $k \times k$ پیکسل و مقدار σ برای فیلتر، کسری از سطح مقیاس و به صورت تجربی برابر $0.2 \times k$ پیکسل انتخاب شدند (Jing et al., 2012).

به این هدف، از یک فیلتر ریخت‌شناسی پیش‌رونده (progressive) همان‌طور که در پژوهش Zhang و همکاران (۲۰۰۳) آمده است، استفاده شد. پس از حذف بازگشت‌های دارای نویز، نقاط زمینی با استفاده از الگوریتم طبقه‌بندی انحنای چندمقیاسی پیشنهادشده توسط Evans و Hudak (۲۰۰۷) شناسایی شدند. سپس، از همه نقاط لایدار (زمین و غیرزمین) برای تولید مدل رقومی سطح (DSM) و نقاط طبقه‌بندی‌شده زمینی برای تولید مدل رقومی زمین (DTM) با قدرت تفکیک مکانی ۰/۵ متر با استفاده روش درون‌یابی شبکه نامنظم مثلثی (TIN) استفاده شد. از تفاضل DSM و DTM، مدل ارتفاعی تاج (CHM) با قدرت تفکیک مکانی ۰/۵ متر تولید شد. در مرحله بعد با اعمال فیلتر گوسین پایین‌گذر با اندازه پنجره 5×5 ، نویزها و بیشینه‌های محلی اشتباه ناشی از شاخه‌های درختان در CHM حذف شد. نویزها نقاطی هستند که ارتفاع آن‌ها با نقاط همسایه همخوانی ندارد. از آن جمله می‌توان به برخورد پرتولیزر با عوارض مصنوعی بالای سطح زمین و یا خطاهای ارتفاعی مرتبط با الگوی اسکن سنجنده لایدار اشاره کرد (Jakubowski et al., 2013).

روش تشخیص تک‌درخت چندمقیاسی

روش شناسایی تک‌درخت چندمقیاسی (MSITD)، یک الگوریتم ترکیبی مبتنی بر به‌کارگیری CHM و ابرنقاط لایدار برای بهبود دقت شناسایی تک‌درخت است. در واقع الگوریتم ترکیبی توسعه‌یافته، الگوریتم‌های تشخیص تک‌درخت رستریایه و ابرنقطه مبنا را ترکیب می‌کند تا از قابلیت‌های این دو روش به‌طور هم‌زمان در یک چهارچوب چندمقیاسی که در ادامه توضیح داده می‌شود، استفاده کند. فرایند MSITD شامل پنج مرحله اصلی است. مرحله اول، تولید CHM است

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \text{رابطه (۱)}$$

است، شباهت کمتری به دایره دارد. با توجه به این مشاهده‌ها، قطعه بزرگ مقیاس (coarse-scale) را می‌توان با قطعه‌های متناظر کوچک مقیاس (fine-scale) از نظر دایره‌ای بودن (circularity) مقایسه کرد. با استفاده از روش شرح داده شده در Jing و همکاران (۲۰۱۲)، قطعه‌های مرتبط با سطح تاج درخت بزرگ در ابتدا با قطعه‌های مربوط به اندازه تاج متوسط ادغام شدند. سپس، قطعه‌های حاصل با آن‌هایی که مربوط به سطح تاج درخت کوچک بودند، ادغام شدند. نتیجه این فرایند، تولید نقشه کامل شامل قطعه‌های تاج درختان با اندازه‌های متفاوت است (Jing et al., 2012).

مرحله پنجم، اصلاح و بهبود قطعه‌های تاج درختان استخراج شده با روش ابرنقطه مینا است. از آنجایی که این احتمال وجود دارد که تعدادی از قطعه‌های درختان نزدیک باهم ادغام شوند، ابر نقاط لایدار مربوط به هر قطعه تاج درخت برش داده شد و قطعه‌بندی دقیق مبتنی بر ابرنقطه سه‌بعدی روی هر خوشه نقاط انجام شد. در این مرحله از تابع چگالی احتمال (pdf) مطابق رابطه (۲) برای آنالیز ابرنقاط داخل قطعه‌ها استفاده شد:

$$f(p) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{p-p_i}{h}\right) \quad \text{رابطه (۲)}$$

که در آن، n بیانگر تعداد نقاط در هر ناحیه قطعه‌بندی شده است. همچنین، h نشان‌دهنده پهناى باند، K تابع هسته گوسین و p بیانگر مختصات پس از تبدیل متعامد است. در این فرایند، درخت قطعه‌بندی شده که توزیع تابع چگالی احتمال آن یک‌وجهی (Unimodal) است، به‌عنوان یک تک‌درخت استخراج می‌شود. در صورتی که تابع چگالی احتمال چندوجهی (Multimodal) باشد، بیانگر وجود چند تک‌درخت در یک قطعه است که با تقسیم تابع چگالی

که در آن، x و y فاصله تا مرکز هسته هموارسازی و σ برای تنظیم مشخصه‌های تابع گوسین استفاده می‌شود. سه سطح تاج درخت که در مرحله آنالیز مقیاس با اعمال اپراتورهای ریخت‌شناسی تعیین شد، سبب سه فیلتر گوسین و در نتیجه، سه CHM فیلترشده با قدرت تفکیک‌های مختلف شد که هر سطح شامل درختانی با اندازه و شکل‌های مشابه بودند. استفاده از قدرت تفکیک‌های مختلف CHM برای تشخیص درختان در مقیاس‌ها و سطوح جزئیات مختلف مناسب است. مدل‌های ارتفاعی تاج با وضوح بیشتر، اطلاعات دقیق‌تر و کوچک مقیاس در مورد تک‌درختان و مدل‌های ارتفاعی تاج با وضوح کمتر، اطلاعات بزرگ مقیاس و کلی را از ساختار جنگل ارائه می‌دهند. در مرحله چهارم با استفاده از روش پیشینه محلی، نوک تاج درختان (treetop) به‌عنوان نشانگر (marker) در مقیاس‌های مختلف شناسایی می‌شوند. سپس، قطعه‌بندی حوضه آبخیز کنترل شده با نشانگر (marker-controlled watershed segmentation) برای قطعه‌بندی مدل‌های ارتفاعی تاج در مقیاس‌های مختلف به‌طور جداگانه استفاده شد. در نهایت، قطعه‌های تاج درختان در سطوح مختلف مقیاس باهم ادغام شدند و نقشه تاج درختان شامل قطعه‌های تاج درختان با اندازه‌های متفاوت تولید شد (Jing et al., 2012). فرایند ادغام با روش بالا به پایین (top-to-bottom) از بزرگ‌ترین تا کوچک‌ترین مقیاس انجام شد. به‌منظور ادغام قطعه‌ها در سطوح مختلف، باید بدانیم که آیا یک قطعه در مقیاس بزرگ، یک درخت است یا خوشه‌ای از درختان؟ شکل تاج یک درخت موجود در CHM به دایره‌ای بودن تمایل دارد، در حالی که خوشه‌ای از درختان که ترکیبی از چندین تاج

$$r = \frac{TP}{TP + FN} \quad \text{رابطه (۳)}$$

$$p = \frac{TP}{TP + FP} \quad \text{رابطه (۴)}$$

$$F = 2 \times \frac{r \times p}{r + p} \quad \text{رابطه (۵)}$$

احتمال در نقطه کمینه محلی، درختان مجاور از یکدیگر جداسازی می‌شوند. به‌طور کلی، تراکم نقطه مرکز یک تک‌درخت، بیشتر از تراکم منطقه حاشیه‌ای است. در نتیجه، توزیع چگالی نقطه‌ای یک تک‌درخت همیشه یک‌وجهی است.

ارزیابی عملکرد روش تشخیص تک‌درخت

عملکرد نتایج روش‌های تشخیص تک‌درخت (ITD) با مقایسه کمی هر درخت شناسایی شده از داده لایدار با نزدیک‌ترین درخت مرجع داده زمینی در هشت منطقه مطالعاتی به دست می‌آید. برای اتصال مکانی تک‌درختان شناسایی شده با روش ITD و درختان مرجع موجود در داده زمینی از الگوریتم تناظریابی درخت خودکار استفاده شده است. این الگوریتم از داده‌های اندازه‌گیری شده میدانی درختان از جمله مختصات GPS تنه، ارتفاع و قطر تاج درختان استفاده می‌کند (Eysn et al., 2015). اگر یک درخت به‌طور صحیح شناسایی شده باشد، مثبت - صحیح (TP) نامیده می‌شود، اما اگر یک درخت شناسایی نشده باشد و به نزدیک‌ترین درخت اختصاص داده شده باشد، منفی - غلط (FN) یا خطای Omission نامیده می‌شود. اگر یک درخت وجود نداشته باشد، اما توسط ابر نقاط لایدار شناسایی شده باشد، مثبت - غلط (FP) یا خطای Commission نامیده می‌شود. برای ارزیابی عملکرد روش‌های ITD، معیارهای recall (r) که بیانگر نرخ تشخیص درخت (tree detection rate)، precision (p) که نشان‌دهنده دقت درختان شناسایی شده و F-score (F) که بیانگر صحت کلی (Overall accuracy) است، با استفاده از معادله‌های (۳) تا (۵) محاسبه شدند (Sokolova et al., 2006):

آ. ید ویژگی‌های ساختاری جنگل با استفاده از
۱۲۴
الگوریتم‌های یادگیری ماشین

پس از تشخیص درختان با الگوریتم MSITD، برای برآورد ویژگی‌های ساختاری درختان جنگل شامل ارتفاع و DBH، ابتدا برای هر تک‌درخت شناسایی شده، چندین ویژگی از جمله بیشینه ارتفاع، کشیدگی توزیع ارتفاع و کرویت از داده لایدار استخراج شد. سپس الگوریتم انتخاب ویژگی بوروتا (Boruta) برای شناسایی مهم‌ترین ویژگی‌های مبتنی بر ابرنقطه لایدار استفاده شد. الگوریتم بوروتا در میان روش‌های پوشاننده (Wrapper) برای انتخاب بهترین ویژگی قرار می‌گیرد. در روش‌های پوشاننده سعی بر این است که از یک زیرمجموعه از ویژگی‌های اولیه استفاده شده تا یک مدل با استفاده از آن زیرمجموعه آموزش داده شود. سپس، براساس نتایج به دست آمده از آموزش مدل با مجموعه ویژگی‌های مختلف، مناسب‌ترین ویژگی‌ها معرفی می‌شوند (Kursa & Rudnicki, 2010). انتخاب ویژگی‌های بهینه، نقش مؤثری در بهبود عملکرد الگوریتم‌های یادگیری ماشین دارد. در نهایت، مشخصه‌های DBH و ارتفاع با استفاده از الگوریتم‌های یادگیری ماشین شامل RF، SVM و کویست (CB) در قالب اعتبارسنجی متقابل تودرتو ده‌قسمتی برآورد شدند و نتایج حاصل مقایسه شدند.

فرایند اجرایی اعتبارسنجی متقابل تودرتو ده‌قسمتی برای دستیابی به آموزش و اعتبارسنجی در دولا به انجام گرفت.

اندازه‌گیری شده زمینی و n بیانگر تعداد درختان مرجع هستند.

نتایج

نتایج حاصل از روش تشخیص تک‌درخت چندمقیاسی

در این بخش، نتایج روش MSITD با داده‌های لایدار هوایی در جنگل‌هایی با ساختار و ترکیب گونه‌ای متفاوت ارائه شده است. در این پژوهش، با ترکیب روش آنالیز چندمقیاسی بر مبنای اپراتورهای ریخت‌شناسی با تابع چگالی احتمال، روش ترکیبی جدیدی ارائه شده است. مطابق شکل ۳، پس از حذف نویز و نرم کردن CHM، با اعمال سه فیلتر گوسین با اندازه پنجره‌های 7×7 ، 15×15 و 21×21 پیکسل که در مرحله آنالیز مقیاس طراحی شدند، مدل ارتفاعی تاج در سه سطح مقیاس کوچک، متوسط و بزرگ تولید شد. سپس بیشینه‌های محلی در هر CHM استخراج شدند. الگوریتم مورد استفاده برای فیلتر بیشینه محلی (تابع "lmf" در پکیج lidR)، با توجه به مقدار ارتفاع هر پیکسل CHM، اندازه پنجره را برای جستجوی بیشینه محلی محاسبه می‌کند (Roussel et al., 2018). برای جلوگیری از شناسایی بیشینه‌های محلی اشتباه ناشی از پوشش گیاهی اطراف درختان، حدآستانه کمینه ارتفاع برای شناسایی نوک تاج درخت استفاده می‌شود. مقدار این حدآستانه روی کمینه ارتفاع درخت مرجع (۳/۸۵ متر) تنظیم شد. همان‌طور که در شکل ۳ مشاهده می‌شود، با افزایش مقیاس از سطح کوچک به متوسط و سپس به سطح بزرگ، تعداد بیشینه‌های محلی شناسایی شده کاهش یافت.

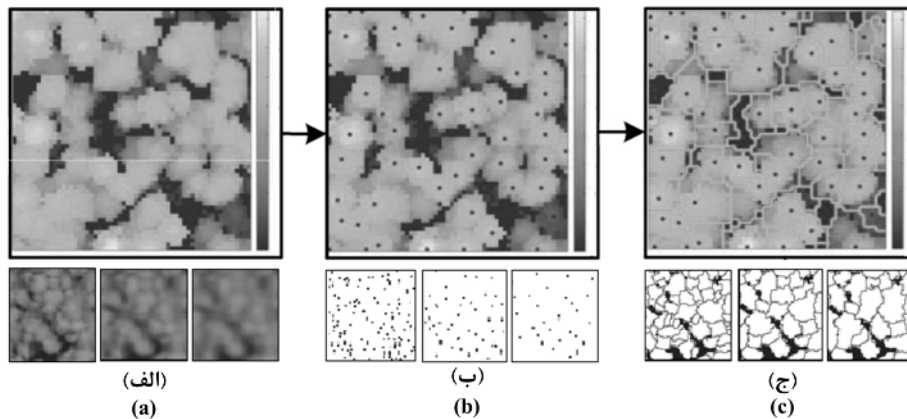
در لایه بیرونی، ۱۰ درصد از داده‌ها به صورت تصادفی برای اعتبارسنجی جدا شدند و بقیه داده‌ها برای ساخت الگوریتم یادگیری ماشین استفاده شدند. در لایه داخلی، ۹۰ درصد باقی‌مانده از داده‌ها برای انتخاب ویژگی با استفاده از الگوریتم بوروتا و نیز بهینه‌سازی مشخصه‌های الگوریتم یادگیری ماشین استفاده شدند. سپس، الگوریتم یادگیری ماشین ساخته شده با استفاده از ۱۰ درصد از داده‌هایی که در ابتدا تقسیم شده بودند (۱۰ درصد موجود در لایه بیرونی)، ارزیابی شدند. این فرایند ۱۰ مرتبه تکرار شد. در هر تکرار از ابتدا ۱۰ درصد مختلف از داده‌ها برای اعتبارسنجی و ۹۰ درصد مختلف از داده‌ها برای ایجاد مدل جدید انتخاب شدند. سپس، عملکرد کلی الگوریتم با توجه به میانگین عملکردهای ۱۰ مدل توسعه یافته در این ۱۰ تکرار مشخص شد. در این پژوهش، سه معیار مجذور میانگین مربعات خطا ($RMSE$)، مجذور میانگین مربعات خطای نسبی ($rRMSE$) و ضریب تبیین (R^2) مطابق رابطه‌های (۶) تا (۸) برای ارزیابی عملکرد استفاده شدند (Chicco et al., 2021).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad \text{رابطه (۶)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad \text{رابطه (۷)}$$

$$rRMSE = \frac{RMSE}{\bar{y}_i} \times 100 \quad \text{رابطه (۸)}$$

که در آن‌ها، y_i مقدار مرجع (اندازه‌گیری شده زمینی)، \hat{y}_i مقدار برآورد شده، \bar{y}_i نشان‌دهنده میانگین مقادیر



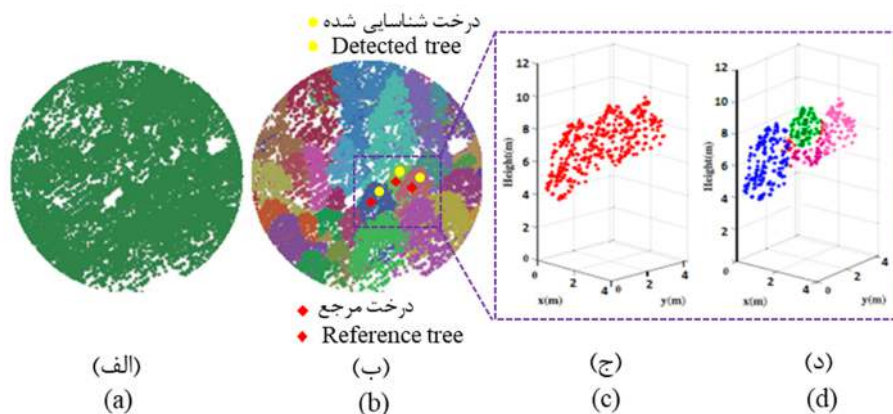
شکل ۳- (الف) فیلترگذاری چندمقیاسی روی CHM در سه سطح مقیاس کوچک، متوسط و بزرگ در بخشی از منطقه ۱، (ب) شناسایی نوک تاج درختان با روش بیشینه محلی در هر CHM فیلترشده و (ج) تولید نقشه قطعه‌بندی تاج تک‌درختان با ادغام قطعه‌های تاج درختان در سه سطح مقیاس

Figure 3. (a) Multi-scale filtering on CHM at the small, medium and large scale levels in part of study area #1, (b) Detection of treetops of crown by local maxima method in each filtered CHM, and (c) Production of the segmentation map of the single trees crown by merging the tree crown segments at three scale levels

و پس از بهبود قطعه‌بندی نشان می‌دهد که استفاده از تابع چگالی احتمال سبب بهبود نتایج حاصل از آنالیز چندمقیاسی بر مبنای اپراتورهای ریخت‌شناسی شده است. همان‌طور که در شکل ۴ نشان داده شده است، درختان مجاوری که در یک قطعه ادغام شده‌اند، پس از آنالیز ابر نقاط با تابع چگالی احتمال از یکدیگر جداسازی شده‌اند.

مطابق جدول ۲ در منطقه یک از ۳۵۲ درخت مرجع زمینی توسط الگوریتم توسعه‌یافته MSITD، ۲۸۱ درخت شناسایی شد. از این تعداد ۲۵۱ درخت به‌طور صحیح، ۱۰۱ درخت شناسایی نشد (خطای Omission) و ۳۰ درخت نیز اشتباه شناسایی شده‌اند (خطای Commission). به‌طور کلی از تعداد ۱۰۶۰ درخت مرجع موجود در داده زمینی در هشت منطقه مختلف، ۹۳۵ پایه توسط الگوریتم MSITD شناسایی شده است. از این تعداد ۷۹۳ درخت به‌طور صحیح شناسایی شدند، ۲۶۷ درخت شناسایی نشدند (۲۵/۱۹ درصد) و ۱۴۲ درخت نیز اشتباه شناسایی شده‌اند (۱۵/۱۹ درصد).

با استفاده از بیشینه‌های محلی شناسایی‌شده به‌عنوان نشانگر، روش قطعه‌بندی حوضه آبخیز کنترل‌شده با نشانگر که در پکیج lidR در نرم‌افزار R پیاده‌سازی شده است، روی CHM اعمال شد و نقشه قطعه‌بندی در سه سطح مقیاس تولید شد. هر سطح شامل درختان با اندازه و شکل شبیه به هم است. در نهایت، قطعه‌های تاج‌پوشش درختان در سه سطح مقیاس کوچک، متوسط و بزرگ با هم ادغام شدند و نقشه کامل تاج درختان تولید می‌شود. در مرحله بعد برای اصلاح و بهبود قطعه‌های تاج درختان استخراج‌شده، ابر نقاط لایدار موجود در هر قطعه توسط تابع چگالی احتمال آنالیز شد و تاج درختان در اشکوب زیرین شناسایی شد. نتایج ارزیابی دقت تشخیص تک‌درخت با روش MSITD در هشت منطقه مطالعاتی در جدول ۲ ارائه شده است. در شکل ۴، نتایج تشخیص تک‌درختان توسط روش MSITD در بخشی از منطقه ۵ نشان داده شده است. درختان استخراج‌شده با رنگ‌های مختلف نمایش داده شده است. مقایسه نتایج پیش



شکل ۴- (الف) ابر نقاط اولیه لایدار، (ب) نتایج قطعه‌بندی درختان با روش MSITD در بخشی از منطقه پنج. تناظریابی سه درخت مرجع و شناسایی شده، (ج) سه درخت مجاور ادغام شده پیش از بهبود قطعه‌بندی و (د) جداسازی سه درخت مجاور با تابع چگالی احتمال

Figure 4. (a) LiDAR point cloud, (b) Tree segmentation results of the MSITD method levels in part of study area #5. Matching of three reference and detected trees, (c) Three merged neighboring trees before segmentation refinement, and (d) separating the three neighboring trees with probability density function

نظر می‌گیرد از ۰/۶۴ تا ۰/۹۲ متغیر است و مقدار کلی آن ۰/۷۹ به دست آمد. همان‌طور که در جدول ۲ مشاهده می‌شود، الگوریتم MSITD در جنگل‌های سوزنی‌برگ تک‌اشکوبه (SL/C) (مناطق سه و پنج) در مقایسه با جنگل‌های آمیخته چنداشکوبه (ML/M) (مناطق یک، چهار، شش و هفت)، عملکرد بهتری دارد. در این پژوهش، مناطق چهار و پنج، بیشترین تراکم نقطه را داشتند (۹۵ تا ۱۲۱ نقطه در مترمربع).

دقت موقعیت مکانی درختان شناسایی شده (RMS_H) برابر ۱/۷۶ متر و دقت ارتفاعی (RMS_V) برابر ۱/۲۲ متر است. نرخ تشخیص درخت (F) نیز از ۰/۵۹ تا ۰/۸۷ متغیر است و مقدار کلی آن ۰/۷۵ به دست آمد. این بدان معنی است که از نظر آماری، ۷۵ درصد از درختان مرجع موجود به درستی تشخیص داده شده‌اند. مقدار p از ۰/۶۸ تا ۰/۹۸ متغیر است و مقدار کلی آن ۰/۸۵ به دست آمد. ضریب F-score که هر دو خطای Commission و Omission را در

جدول ۲- ارزیابی عملکرد روش MSITD در مناطق پژوهش

Table 2. Performance evaluation of the MSITD method at the study areas

Area ID (Forest type)	Nr. Ref. trees	Nr. Det. trees	TP	FP	FN	Com%	Om%	r	p	F	RMS_H (m)	RMS_V (m)
1 (ML/M)	352	281	251	30	101	10.68	28.69	0.71	0.89	0.79	1.91	1.63
2 (ML/C)	22	28	19	9	3	32.14	13.64	0.86	0.68	0.76	1.79	1.35
3 (SL/C)	49	45	41	4	8	8.89	16.33	0.84	0.91	0.87	1.61	1.01
4 (ML/M)	47	43	37	6	10	13.95	21.28	0.79	0.86	0.82	1.65	1.03
5 (SL/C)	235	210	205	5	30	2.38	12.77	0.87	0.98	0.92	1.52	0.95
6 (ML/M)	169	139	99	40	70	28.78	41.42	0.59	0.71	0.64	1.75	1.36
7 (ML/M)	107	92	73	19	34	20.65	31.78	0.68	0.79	0.73	1.91	1.08
8 (SL/M)	79	97	68	29	11	29.9	13.92	0.86	0.7	0.77	1.83	1.12
Overall	1060	935	793	142	267	15.19	25.19	0.75	0.85	0.79	1.76	1.22

SL/C: Single-Layered Coniferous forest; SL/M: Single-Layered Mixed forest; ML/C: Multi-Layered Coniferous forest; ML/M: Multi-Layered Mixed forest.

برای هر تک‌درخت شناسایی شده، ۴۰ متریک متفاوت ویژگی از جمله بیشینه ارتفاع، کشیدگی توزیع ارتفاع و کرویت از داده‌های لایدار استخراج شدند (جدول ۳). در این پژوهش با بهره‌گیری از الگوریتم انتخاب ویژگی بوروتا، مهم‌ترین ویژگی‌های حاصل از داده‌های لایدار برای برآورد ارتفاع و DBH مشخص شدند. مطابق با نتایج، مشخصه‌های مرتبط با توزیع ارتفاع، اطلاعات بیشتری را برای برآورد ارتفاع و DBH با استفاده از الگوریتم‌های یادگیری ماشین فراهم می‌کنند، درحالی‌که مشخصه‌های مرتبط با درصد بازگشت و نیز ویژگی‌های آماری ارتفاع، به مراتب اطلاعات کمتری را در این زمینه ارائه می‌کنند.

جدول ۴ نتایج مدل‌سازی و ارزیابی عملکرد سه الگوریتم یادگیری ماشین (RF، SVM و CB) برای برآورد مشخصه‌های ارتفاع و DBH را در مناطق مطالعاتی مختلف در قالب اعتبارسنجی متقابل تودرتو ده‌قسمتی نشان می‌دهد. گفتنی است که برای آموزش الگوریتم‌های رگرسیون فقط ویژگی‌های بهینه استخراج شده توسط الگوریتم بوروتا استفاده شدند.

از طرفی، بیشترین نرخ تشخیص به منطقه پنج تعلق داشت که بیانگر اثر زیاد تراکم نقطه در شناسایی درختان با روش MSITD است. برای ارزیابی دقت تشخیص در اشکوب زیرین، درختان مرجع زمینی در مناطق مورد مطالعه در لایه‌های ارتفاعی شامل دو تا پنج، پنج تا ۱۰، ۱۰ تا ۱۵، ۱۵ تا ۲۰ متر و بیشتر از ۲۰ متر جداسازی شدند. سپس، عملکرد روش تشخیص تک‌درخت چندمقیاسی در هرکدام از این لایه‌ها ارزیابی شد. روش MSITD در لایه دو تا پنج متر که شامل درختان اشکوب زیرین است، ۴۸ درصد درختان، در سه لایه میانی (پنج تا ۲۰ متر) به ترتیب ۵۹، ۶۹ و ۸۵ درصد و در بالاترین لایه ارتفاعی (بیشتر از ۲۰ متر) که دارای تاج مستقل بودند و تداخل تاجی با پایه‌های مجاور نداشتند، ۸۹ درصد از درختان را شناسایی کرد.

نتایج حاصل از برآورد ویژگی‌های ساختاری جنگل

پس از تشخیص درختان با روش تشخیص تک‌درخت چندمقیاسی، اقدام به الحاق مکانی درختان مرجع و تک‌درختان شناسایی شده از داده‌های لایدار با استفاده از الگوریتم تناظریابی درخت خودکار (Eysn *et al.*, 2015) براساس اختلاف ارتفاع و فاصله افقی بین درختان شد. سپس

جدول ۳- ویژگی‌های استخراج شده از داده‌های لایدار برای هر تک‌درخت

Table 3. LiDAR-derived features for each single tree

Metric	Definition	Metric	Definition
zpcum2	Cumulative percentage of return in the 2th layer	zmax	Maximum height
zpcum3	Cumulative percentage of return in the 3th layer	zmean	Mean height
zpcum4	Cumulative percentage of return in the 4th layer	zsd	Standard deviation of height distribution
zpcum5	Cumulative percentage of return in the 5th layer	zskew	Skewness of height distribution
zpcum6	Cumulative percentage of return in the 6th layer	zkurt	Kurtosis of height distribution
zpcum7	Cumulative percentage of return in the 7th layer	zentropy	Entropy of height distribution
zpcum8	Cumulative percentage of return in the 8th layer	pzabomean	Percentage of returns above zmean
zpcum9	Cumulative percentage of return in the 9th layer	pzabove2	Percentage of returns above 2
P1th	percentage 1th returns	zq5	5th percentile of height distribution
P2th	percentage 2th returns	zq10	10th percentile of height distribution
P3th	percentage 3th returns	zq20	20th percentile of height distribution
eigen_largest	maximum eigen value	zq30	30th percentile of height distribution
eigen_medium	mean eigen value	zq40	40th percentile of height distribution
eigen_smallest	minimum eigen value	zq50	50th percentile of height distribution
curvature	Surface curvature	zq60	60th percentile of height distribution
linearity	Eigenvalues linearity	zq70	70th percentile of height distribution
planarity	Eigenvalues planarity	zq80	80th percentile of height distribution
sphericity	Eigenvalues sphericity	zq90	90th percentile of height distribution
anisotropy	Eigenvalues anisotropy	zq95	95th percentile of height distribution
horizontality	Eigenvalues horizontality	zpcum1	Cumulative percentage of return in the 1th layer

جدول ۴- ارزیابی عملکرد الگوریتم‌های یادگیری ماشین در برآورد مشخصه‌های ارتفاع و DBH برای همه مناطق مورد مطالعه

Table 4. Performance evaluation of the ML algorithms for height and DBH estimation for all study areas

Evaluation metric	Height (m)			DBH (cm)		
	RF	SVM	CB	RF	SVM	CB
RMSE	2.12	1.75	1.95	4.94	4.74	5.04
rRMSE (%)	12	9	11	20	19	20
R ²	0.79	0.85	0.81	0.73	0.78	0.75

چندمقیاسی سبب افزایش دقت تشخیص تک درختان در لایه‌های ارتفاعی مختلف و به خصوص اشکوب زیرین درختان در انواع مختلف جنگل شد. نتایج این پژوهش نشان داد که روش توسعه یافته MSITD در مناطق جنگلی مختلف، علاوه بر تشخیص ۸۹ درصد تاج درختان در بالاترین لایه ارتفاعی، بیشترین تعداد درختان کوچک زیرین را با نرخ تشخیص ۴۸ درصد در پایین‌ترین لایه ارتفاعی (دو تا پنج متر) تشخیص داده است. نتایج پژوهش‌های دیگر نیز نشان دادند که شناسایی درختان اشکوب زیرین در تاج‌های مترام و چندلایه، چالش اساسی در کاربرد روش‌های ITD برای برآورد مشخصه‌های کمی درختان جنگل است (Eysn et al., 2016; Wang et al., 2015). نرخ تشخیص کم روش‌های ITD برای درختان اشکوب زیرین تاحدی به این دلیل است که پالس‌های لیزر محدودی از اشکوب زیرین به دلیل انسداد ناشی از درختان اشکوب غالب منعکس می‌شوند و در نتیجه، تراکم نقاط برداشتی لایدار کم است (Kaartinen et al., 2012; Sparks et al., 2022). برای پرداختن به این چالش به صورت عملی، الگوریتم‌هایی که مبتنی بر ساختار چندمقیاسی هستند، مورد توجه ویژه پژوهشگران قرار گرفته‌اند (Koch et al., 2006; Jing et al., 2012).

برطبق نتایج، الگوریتم MSITD در جنگل‌های سوزنی‌برگ تک‌اشکوبه (SL/C) (مناطق شماره سه و پنج) در مقایسه با جنگل آمیخته چنداشکوبه (ML/M) (مناطق

به‌طورکلی، الگوریتم‌های یادگیری ماشین مورد استفاده، نتایج قابل قبولی را در برآورد مقادیر ارتفاع و DBH با استفاده از داده‌های لایدار ارائه کردند. نتایج به‌دست‌آمده نشان‌دهنده اختلافات بسیار جزئی این سه روش در برآورد مشخصه‌های ارتفاع و DBH است. البته الگوریتم SVM بهترین عملکرد را نشان داد و پس از آن، الگوریتم‌های CB و RF قرار گرفتند. برطبق جدول ۴ در برآورد مشخصه ارتفاع، میانگین RMSE، rRMSE و R² در همه مناطق مورد مطالعه برای الگوریتم SVM به ترتیب برابر با ۱/۷۵ متر، ۹ درصد و ۰/۸۵ بودند. اگرچه تفاوت موجود بین الگوریتم‌های موردنظر در برآورد DBH قابل توجه نبودند، اما الگوریتم SVM بهترین عملکرد را در برآورد DBH داشت. میانگین RMSE، rRMSE و R² در الگوریتم SVM برای برآورد DBH در همه مناطق مطالعاتی به ترتیب برابر با ۴/۷۴ سانتی‌متر، ۱۹ درصد و ۰/۷۸ به‌دست آمد.

بحث

علی‌رغم عملکرد مناسب الگوریتم‌های تشخیص تک درخت ارائه‌شده، بهبود عملکرد و ارتقای دقت در فرایند برآورد مشخصه‌های ساختاری درختان جنگل همواره مورد توجه پژوهشگران بوده است. در پژوهش پیش‌رو، روش MSITD به‌عنوان یک روش ترکیبی برای تشخیص تک درختان از داده‌های لایدار ارائه شد. این الگوریتم با ترکیب روش‌های رستریپایه و ابرنقطه مینا در یک چهارچوب

شماره یک، چهار، شش و هفت)، عملکرد بهتری داشت. نتایج پژوهش‌های پیشین نیز نشان می‌دهد که استخراج درختان در جنگل‌های با ساختار تاج ساده (جنگل‌های سوزنی‌برگ تک‌اشکوبه) دقت زیادی دارند، درحالی‌که این کار برای جنگل‌های آمیخته چنداشکوبه با گونه‌های مختلف درختان که تاج‌های با شکل‌های مختلف، نامنظم و چندلایه دارند، سخت و پرچالش است (Eysn *et al.*, 2015; Dong *et al.*, 2018; Wang *et al.*, 2023). دلیل این موضوع به شکل پیچیده‌تر تاج در درختان پهن‌برگ نسبت به سوزنی‌برگ برمی‌گردد که سبب ادغام قطعه‌های تاج درختان مجاور و اشکوب زیرین می‌شود (Wang *et al.*, 2016; Sparks *et al.*, 2022). در این پژوهش، مناطق چهار و پنج، بیشترین تراکم نقطه را داشته‌اند (۹۵ تا ۱۲۱ نقطه در مترمربع). از طرفی، بیشترین نرخ تشخیص مربوط به منطقه پنج است که بیانگر اثر زیاد تراکم نقطه در شناسایی درختان با روش MSITD است. در برخی از پژوهش‌های پیشین نیز گزارش شده است که تراکم نقاط، تأثیر زیادی بر کارایی روش‌های مبتنی بر ابرنقاط دارد (Wang *et al.*, 2016).

اگرچه تفاوت موجود بین الگوریتم‌های یادگیری ماشین در برآورد ارتفاع و DBH قابل توجه نبودند، اما الگوریتم SVM عملکرد بهتری را در برآورد این مشخصه‌ها نشان داد. میانگین $RMSE$ ، $rRMSE$ و R^2 در همه مناطق مورد مطالعه برای الگوریتم SVM در برآورد مشخصه ارتفاع به ترتیب برابر با ۱/۷۵ متر، ۹ درصد و ۰/۸۵ و برای مشخصه DBH به ترتیب برابر با ۴/۷۴ سانتی‌متر، ۱۹ درصد و ۰/۷۸ بودند. در پژوهش Seyed Mousavi و همکاران (۲۰۱۹)، بهترین نتیجه مربوط به الگوریتم شبکه عصبی مصنوعی با درصد جذر میانگین مربعات خطا و انحراف معیار تفاوت‌ها به ترتیب برابر ۱۳/۳۹ درصد و ۱/۷۱ متر در برآورد ارتفاع و ۲۲/۲۴

درصد و ۹/۲ سانتی‌متر برای DBH بود. Khorami و همکاران (۲۰۱۴) نیز با برآورد ارتفاع دو گونه درخت پهن‌برگ در اشکوب غالب، مقادیر $RMSE$ و $rRMSE$ را به ترتیب برابر ۱/۶۷ متر و شش درصد گزارش کردند. بیشتر این معیارها در پژوهش پیش‌رو به اندازه‌گیری درختان با گونه‌های مختلف و نیز در لایه‌های مختلف ارتفاعی برمی‌گردد. هم‌راستا با نتایج پژوهش پیش‌رو، Malek و همکاران (۲۰۱۹) برای برآورد قطر برابر سینه با روش‌های SVM و RF، برتری بسیار جزئی روش SVM ($RMSE$) برابر ۴/۸۷ سانتی‌متر) را گزارش کردند.

تاکنون، روش‌های یادگیری ماشین متفاوتی در برآورد ویژگی‌های ساختاری درختان جنگل استفاده شده‌اند. در این میان، روش‌های یادگیری عمیق از طرف پژوهشگران مورد توجه ویژه قرار گرفته‌اند (Garcia-Gutierrez *et al.*, 2015; Zhang *et al.*, 2019). از این رو، پیشنهاد می‌شود که در پژوهش‌های آینده، روش‌های یادگیری عمیق برای برآورد مشخصه‌های ساختاری جنگل از جمله زی‌توده بالای سطح زمین، روبه زمینی و حجم سرپا در سطح تک‌درخت ارزیابی و مقایسه شوند. همچنین، پیشنهاد می‌شود که نسبت به جمع‌آوری داده‌های جامع از جنگل‌های ایران و پیاده‌سازی الگوریتم‌های مورد استفاده در این پژوهش برای کاربردهای عملی داخل کشور اقدام شود.

سپاسگزاری

نویسندگان مقاله از شرکای پروژه تحقیقاتی NEWFOR که مجموعه داده تشخیص تک‌درخت معیار را به منظور ارتقا و ترویج استفاده از فناوری جدید لایدار در مدیریت منابع جنگل در اختیار پژوهشگران قرار داده‌اند، تشکر و قدردانی می‌کنند.

منابع مورد استفاده

- Kursa, M.B. and Rudnicki, W.R., 2010. Feature Selection with the Boruta Package. *Journal of Statistical Software*, 36(11): 1-13.
- Malek, S., Miglietta, F., Gobakken, T., Næsset, E., Gianelle, D. and Dalponte, M., 2019. Prediction of stem diameter and biomass at individual tree crown level with advanced machine learning techniques. *iForest-Biogeosciences and Forestry*, 12: 323-329.
- Mohammadi, J., Shataee, S., Namiranian, M. and Nasset, E., 2017. Modeling biophysical properties of board-leaved stands in the Hyrcanian forests of Iran using fused airborne laser scanner data and UltraCam-D images. *International Journal of Applied Earth Observation and Geoinformation*, 61: 32-45.
- NEWFOR, 2015. NEWFOR Lidar Single Tree Detection Benchmark Dataset. Available at: <http://www.newfor.net/download-newfor-single-tree-detection-benchmark-dataset/>
- Poorazimy, M., Shataee Jouibary, Sh., Mohammadi, J. and Aghababaei, H., 2023. Feasibility of single-polarized TanDEM-X data for Hyrcanian forest height estimation (Case study: Shast-Kalateh forest). *Iranian Journal of Forest*, 15(3): 329-343 (In Persian with English Summary).
- Roussel, J.R., Auty, D., Boissieu, F.D., Meador, A.S., Jean-François, B., Demetrios, G., ... and St-Onge, B., 2018. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. version 1.4.1. Available at: <https://cran.r-project.org/web/packages/lidR/index.html>
- Sefidi, K. and Jahdi, R., 2023. Impact of Anthropogenic disturbance on the size diversity of trees in Arasbaran forests (Case study: Hatam-Meshasi Forest Reserve in Meshgin-Shahr county, Iran). *Iranian Journal of Forest and Poplar Research*, 31(3): 241-256 (In Persian with English Summary).
- Seyed Mousavi, S.Z., Mohammadi, J. and Shataee, Sh., 2017. The evaluation of potential airborne laser scanner data in estimating of individual canopy area and tree heights in part of Educational and Research Shast-Kalate Forests - Gorgan. *Ecology of Iranian Forest*, 5(9): 47-55 (In Persian with English Summary).
- Seyed Mousavi, S.Z., Mohammadi, J. and Shataee, Sh., 2019. Estimation of the some quantitative characteristics of individual tree using airborne laser scanning data in part of Shast-Kalate forests of Gorgan. *Journal of Wood and Forest Science and Technology*, 26(1): 1-19 (In Persian with English Summary).
- Silva, C.A., Crookston, N.L., Hudak, A.T. and Vierling, L.A., 2017. rLiDAR: An R package for reading, processing and visualizing lidar (Light Detection and Ranging) data. version 0.1. Available at: <http://cran.r-project.org/web/packages/rLiDAR/index.html>
- Silva, C.A., Klauber, C., Mohan, M.M. and Bright, B.C., 2018. LiDAR Analysis in R and RLiDAR for Forestry Applications. *Lidar Remote Sensing Environment Monitoring*, 404/504, 90p.
- Slik, J.W.F., Aiba, S.I., Brearley, F.Q., Cannon, C.H., Forshed, O., Kitayama, K., ... and van Valkenburg, J.L.C.H., 2010. Environmental correlates of tree biomass, basal area, wood specific gravity and stem density gradients in Borneo's tropical forests. *Global Ecology and Biogeography*, 19: 50-60.
- Sokolova, M., Japkowicz, N. and Szpakowicz, S., 2006. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation: 1015-1021. In: Sattar, A. and Kang, B.H. (Eds.). *AI 2006: Advances in Artificial Intelligence. Lecture Notes in Computer Science*, Vol. 4304, Berlin, Springer, 1303p.
- Sparks, A.M., Corrao, M.V. and Smith, A.M.S., 2022. Cross-comparison of individual tree detection methods using low and
- Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., ... and Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, 145: 87-99.
- Chicco, D., Matthijs, J.W. and Giuseppe, J., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7: e623.
- Corte, A.P.D., Souza, D.V., Rex, F.E., Sanquetta, C.R., Mohan, M., Silva, C.A., ... and 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.
- Evans, J.S. and Hudak, A.T., 2007. A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Transactions on Geoscience and Remote Sensing*, 45: 1029-1038.
- Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.M., Dalponte, M., ... and Pfeifer, N., 2015. A benchmark of Lidar-based single tree detection methods using heterogeneous forest data from the Alpine Space. *Forests*, 6: 1721-1747.
- Frénay, B., Doquire, G. and Verleysen, M., 2013. Is mutual information adequate for feature selection in regression? *Neural Networks*, 48: 1-7.
- Garcia-Gutierrez, J., Martinez-Alvarez, F., Troncoso, A. and Riquelme, J.C., 2015. A comparison of machine learning regression techniques for LiDAR-derived estimation of forest variables. *Neurocomputing*, 167: 24-31.
- Ghaderi, P., Mohammadi, J., Shataee, S., Rahmani, R. and Kariminejad, N., 2023. Efficiency of nonlinear mixed-effects model in determining height-diameter equations of velvet maple and ironwood trees. *Iranian Journal of Forest*, 14(4): 473-485 (In Persian with English Summary).
- Hui, Z., Jin, S., Xia, Y., Nie, Y., Xie, X. and Li, N., 2021. A mean shift segmentation morphological filter for airborne LiDAR DTM extraction under forest canopy. *Optics & Laser Technology*, 136: 106728.
- Hyypä, J., Holopainen, M. and Olsson, H., 2012. Laser scanning in forests. *Remote Sensing*, 4: 2919-2922.
- Jakubowski, M.K., Li, W., Guo, Q. and Kelly, M., 2013. Delineating individual trees from lidar data: a comparison of vector- and raster-based segmentation approaches. *Remote Sensing*, 5: 4163-4186.
- Jing, L., Hu, B., Li, J. and Noland, T., 2012. Automated delineation of individual tree crowns from lidar data by multi-scale analysis and segmentation. *Photogrammetric Engineering & Remote Sensing*, 78: 1275-1284.
- Kaartinen, H., Hyypä, J., Yu, X., Vastaranta, M., Hyypä, H., Kukko, A., ... and Wu, J.C., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4: 950-974.
- Khorami, R.A., Darvishsefat, A.A., Tabari Kochaksaraei, M. and Shataee Jouybari, Sh., 2014. Potential of LIDAR data for estimation of individual tree height of *Acer velutinum* and *Carpinus betulus*. *Iranian Journal of Forest*, 6(2): 127-140 (In Persian with English Summary).
- Koch, B., Heyder, U. and Weinacker, H., 2006. Detection of individual tree crowns in airborne LiDAR data. *Photogrammetric Engineering & Remote Sensing*, 72: 357-363.
- Krstajic, D., Buturovic, L.J., Leahy, D.E. and Thomas, S., 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, 6: 10.

- LiDAR data with a tree Region-Based Convolutional Neural Network (RCNN). *Remote Sensing*, 15: 1024.
- Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M. and Viitala, R., 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(1): 28-37.
 - Zhang, K., Chen, S.C., Whitman, D., Shyu, M.L., Yan, J. and Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41: 872-882.
 - Zhang, L., Shao, Z., Liu, J. and Cheng, Q., 2019. Deep learning based retrieval of forest aboveground biomass from combined LiDAR and Landsat 8 data. *Remote Sensing*, 11: 1459.
 - high pulse density airborne laser scanning data. *Remote Sensing*, 14: 3480.
 - Vabalas, A., Gowen, E., Poliakoff, E. and Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. *PLoS ONE*, 14: e0224365.
 - Wang, Y., Hyyppä, J., Liang, X., Kaartinen, H., Yu, X., Lindberg, E., ... and Alho, P., 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(9): 5011-5027.
 - Wang, Y., Weinacker, H. and Koch, B., 2008. A LiDAR point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. *Sensors*, 8: 3938-3951.
 - Wang, Z., Li, P., Cui, Y., Lei, S. and Kang, Z., 2023. Automatic detection of individual trees in forests based on airborne