

# مقایسه الگوریتم‌های یادگیری ماشین در طبقه‌بندی مکانی گونه‌های درختی پهن‌برگ

مژده میرکی<sup>۱</sup>، هرمز سهرابی<sup>۲\*</sup>، پرویز فاتحی<sup>۳</sup>، Mathias Kneubuehler<sup>۴</sup>

<sup>۱</sup> دانشجوی دکتری مدیریت جنگل - گروه جنگلداری - دانشکده منابع طبیعی و علوم دریایی - دانشگاه تربیت

مدرس

m.miraki@modares.ac.ir

<sup>۲</sup> دانشیار گروه جنگلداری - دانشکده منابع طبیعی و علوم دریایی - دانشگاه تربیت مدرس

hsohrabi@modares.ac.ir

<sup>۳</sup> استادیار گروه جنگلداری و اقتصاد جنگل - دانشکده کشاورزی و منابع طبیعی - دانشگاه تهران

parviz.fatehi@ut.ac.ir

<sup>۴</sup> استادیار گروه سنجش‌ازدور - دانشگاه زوریخ - زوریخ - سوئیس

mathias.kneubuehler@geo.uzh.ch

(تاریخ دریافت اسفند ۱۳۹۸، تاریخ تصویب خرداد ۱۳۹۹)

## چکیده

شناسایی گونه‌های درختی برای انواع فعالیت‌های مدیریت و نظارت بر منابع طبیعی از جمله آماربرداری، ارزیابی خطر آتش‌سوزی، نظارت بر تنوع زیستی و بهبود زیستگاه حیات وحش از اهمیت ویژه‌ای برخوردار است. در طی سالیان اخیر، پیشرفت در فن‌آوری سنجش از دور امکان طبقه‌بندی گونه‌های درختی در سطح تک‌درخت را با استفاده از سنجنده‌های مختلف فراهم کرده است. در این بین استفاده از تصاویر رنگی واقعی مبتنی بر پهپاد به دلیل سهولت در استفاده و مقرون به صرفه بودن به یک انتخاب محبوب تبدیل شده است. اما با توجه به نتایج متفاوت الگوریتم‌های شناسایی درختان در توده‌های مختلف، لازم است که تأثیر استفاده از الگوریتم‌ها بر دقت شناسایی درختان در گونه‌های مورد مطالعه داخل کشور مورد بررسی قرار گیرد. بنابراین هدف ما در این تحقیق شناسایی گونه‌های پهن‌برگ جنگلی با استفاده از تصاویر پهپاد با وضوح مکانی بالا و مقایسه عملکرد الگوریتم‌های ناپارامتری ماشین بردار پشتیبان، جنگل تصادفی و شبکه عصبی مصنوعی و الگوریتم پارامتری تحلیل تشخیص خطی در طبقه‌بندی درختان بود. به همین منظور تصویربرداری و برداشت زمینی درختان در خرداد ۱۳۹۸ در قسمت جلگه‌ای جنگل‌های شهرستان نور در مجاورت پارک جنگلی نور انجام شد. پس از پردازش تصاویر، اورتوموزاییک جنگل مورد مطالعه با اندازه پیکسل ۲/۵ سانتی‌متر با استفاده از الگوریتم ساختار حرکت مینا تولید شد. سپس باندهای اصلی، شاخص‌ها و نسبت‌های باندهای به دست آمده از اورتوموزاییک به عنوان داده ورودی در الگوریتم‌های طبقه‌بندی مورد استفاده قرار گرفت. نتایج نشان داد که الگوریتم پارامتری تحلیل تشخیص خطی با صحت کلی ۰/۸۷ بهترین نتیجه را برای طبقه‌بندی درختان ارائه داد. همچنین گونه سفیدپلت (صحت کاربر ۰/۹۷) و گونه انجیلی (صحت کاربر ۰/۷۲) به ترتیب بیشترین و کمترین صحت طبقه‌بندی را نشان دادند. در کل نتایج این تحقیق نشان داد که تصاویر پهپاد در جنگل‌های جلگه‌ای خزری قابلیت طبقه‌بندی و شناسایی درختان با الگوریتم‌های مختلف را دارد.

**واژگان کلیدی:** تحلیل تشخیص خطی، ماشین بردار پشتیبان، جنگل تصادفی، شبکه عصبی مصنوعی، پهپاد، شاخص‌های طیفی

## ۱- مقدمه

می‌گرفتند، نتوانسته‌اند اطلاعات مناسب در مقیاس زمانی و مکانی ریز ارائه دهند (۱۰).

امروزه به لطف پیشرفت‌های سریع و افزایش روند صعودی توسعه سامانه‌های هوایی بدون سرنشین نقطه عطفی در کاربرد سنجش از دور اتفاق افتاده است. این سامانه‌ها که به آن‌ها هواپیماهای بدون سرنشین<sup>۴</sup> یا پهپاد نیز گفته می‌شود، سکوه‌های جدید برای سنجنده‌های متفاوت در پیکربندی‌های مختلف هستند. پهپادهای کوچک که به اهداف نقشه‌برداری متنوع اختصاص داده شده‌اند، مقرون به صرفه و انعطاف‌پذیر هستند و بر حسب نیاز به سنجنده مورد نیاز مجهز شده و تصاویر با وضوح مکانی بسیار بالا را ارائه می‌دهند (۱۱). در طی سالیان اخیر، پیشرفت در فن‌آوری سنجش از دور امکان طبقه‌بندی گونه‌های درختی در سطح تک‌درخت را با استفاده از سنجنده‌های مختلف فراهم کرده است (۱۲). به طور مشخص طبقه‌بندی داده‌های ناشی از داده‌های فراطیفی و لایدار دارای دقت بالاتری نسبت به سایر داده‌ها هستند. Maschler و همکاران (۱۳) طبقه‌بندی ۱۶ گونه درختی متشکل از گونه‌های غالب بلوط، زبان‌گنجشگ، توسکا و راش در جنگل‌های معتدله اتریش را با استفاده از داده‌های فراطیفی انجام دادند. در این تحقیق از الگوریتم جنگل تصادفی شیء مبنا برای طبقه‌بندی درختان استفاده شده است. صحت کلی طبقه‌بندی برابر با ۹۱/۷ درصد بود که نشان‌دهنده قابلیت بالای داده‌های فراطیفی برای طبقه‌بندی درختان است. با این حال، استفاده از سنجنده‌های فراطیفی بسیار پرهزینه هستند. به همین علت، در اغلب پهپادها از دوربین‌های رنگی واقعی کم هزینه استفاده می‌شود (۱۴). همچنین ساختار حرکت مبنا به عنوان یک الگوریتم در تبدیل تصاویر دو بعدی به سه بعدی نشان داده است که ابر نقطه سه بعدی حاصل از این روش با استفاده از داده‌های پهپاد با آنچه که از طریق لایدار به دست می‌آید قابل مقایسه است و می‌تواند اطلاعات مشابه با آنچه که توسط داده‌های لایدار به دست آمده است، را ارائه دهد (۱۵). Nevalainen و همکاران (۲) عملکرد تصاویر فراطیفی و فتوگرامتری پهپاد را در طبقه‌بندی گونه‌های کاج، نوئل، لاریکس و توس در جنگل‌های بورال فنلاند مورد مطالعه قرار دادند. نتایج نشان داد که صحت کلی شناسایی درختان

طبقه‌بندی تک‌پایه‌های درختی با استفاده از داده‌های سنجش از دور، رؤیای بسیاری از دانشمندان و مدیران جنگل بوده است (۱). چرا که اطلاع از ترکیب گونه‌های درختی یک جنگل، می‌تواند ارزش اقتصادی جنگل را برآورد کرده و اطلاعات ارزشمندی را برای مطالعه اکوسیستم‌های جنگلی و بسیاری از کارهای محیط زیست، حفاظت و نظارت بر بخش‌های مختلف منابع طبیعی از جمله ارزیابی خطر آتش‌سوزی، نظارت بر تنوع‌زیستی و بهبود زیستگاه حیات‌وحش فراهم کند (۲) - (۴). در حالی که آماربرداری میدانی برای جمع‌آوری این اطلاعات، کاری پرهزینه و وقت‌گیر است و بویژه برای مطالعاتی که نیاز به جمع‌آوری داده‌های دوره‌ای دارند نیز قابل توصیه نیست (۵)، (۶). اما امروزه با ظهور فناوری‌های پیشرفته سنجش از دور به دلیل توانایی آن در استخراج اطلاعات دقیق از ویژگی‌های سطح زمین، می‌تواند جایگزین معقولی برای تعیین پارامترهای توده از جمله طبقه‌بندی درختان در نظر گرفته شود (۷). این فن‌آوری به طور فزاینده‌ای برای کمک به روش‌های سنتی میدانی در برآورد ویژگی‌های ساختاری جنگل در مناطق وسیع مورد استفاده قرار گرفته و منبع ارزشمندی از اطلاعات در نقشه‌برداری و نظارت بر ویژگی‌های جنگل فراهم می‌آورد.

نوع تصاویر و سکوی مورد استفاده فاکتورهای مهمی در طبقه‌بندی محسوب می‌شوند، زیرا وضوح مکانی و طیفی می‌تواند بر دقت طبقه‌بندی تأثیر بگذارد (۸). طبقه‌بندی گونه‌های درختی با استفاده از تصویربرداری چندطیفی<sup>۱</sup>، فراطیفی<sup>۲</sup> یا لیزر اسکنرها<sup>۳</sup> به طور گسترده‌ای مورد مطالعه قرار گرفته است. با این حال، داده‌ها عمدتاً توسط ماهواره‌ها یا هواپیماهای دارای سرنشین برداشت شده‌اند، بنابراین تمرکز این مطالعات بیشتر بر سطوح نسبتاً وسیع نظیر کل جنگل یا نهایتاً در سطح توده یا پلات بوده است (۲)، (۹). اما فنون سنجش از دور در زمینه علوم محیطی اگرچه برای مدت طولانی برای مطالعه پدیده‌های زیست‌محیطی مورد استفاده قرار می‌گرفتند، تغییر یافتند چرا که تصاویر ماهواره‌ای و عکس‌های هوایی که تاکنون مورد استفاده قرار

<sup>۱</sup> Multispectral

<sup>۲</sup> Hyperspectral

<sup>۳</sup> Laser scanning

<sup>۴</sup> Unmanned Aerial Vehicles

بهبود دقت برآورد همراه است؛ استفاده از آن در داخل کشور بسیار محدود بوده و در خارج از کشور نیز بر اساس مرور منابع بیشتر تحقیقات صورت گرفته بر روی گونه‌های سوزنی‌برگ انجام شده‌است، در حالی‌که با توجه به شکل تاج، نوع گونه و ناهمگونی توده جنگلی میزان کارایی و دقت می‌تواند متفاوت باشد. همچنین در تحقیقات صورت گرفته از الگوریتم‌های شناسایی متفاوتی استفاده شده‌است که با توجه به نتایج متفاوت آن‌ها در توده‌های مختلف، لازم است که تأثیر استفاده از آن‌ها بر دقت شناسایی درختان در گونه‌های مورد مطالعه داخل کشور مورد بررسی قرار گیرد.

بنابراین هدف این تحقیق مقایسه و بررسی عملکرد روش‌های طبقه‌بندی جنگل تصادفی<sup>۱</sup>، تحلیل تشخیص خطی<sup>۲</sup>، شبکه عصبی مصنوعی<sup>۳</sup> و ماشین بردار پشتیبان<sup>۴</sup> برای طبقه‌بندی گونه‌های درختی با استفاده از داده‌های به دست آمده از تصویربرداری پهپاد در یک جنگل با توزیع گونه‌های درختی ناهمگن است تا تأثیر استفاده از الگوریتم‌های مختلف بر صحت طبقه‌بندی درختان مورد سنجش قرار گیرد.

## ۲- مواد و روش‌ها

### ۲-۱- منطقه مورد مطالعه

تحقیق حاضر در محدوده‌ای از جنگل‌های جلگه‌ای هیرکانی در ۵ کیلومتری شرق شهرستان نور، استان مازندران انجام شد. مساحت محدوده مورد مطالعه حدود ۲ هکتار با دامنه شیب ۰ تا ۴ درصد است که در ارتفاع ۲۷- متر از سطح دریا قرار دارد. طبق طبقه‌بندی اقلیمی دومارتون، اقلیم منطقه مرطوب و متوسط بارندگی و دمای سالانه به ترتیب ۹۹۷ میلی‌متر و ۱۶/۴ درجه سانتی‌گراد است. پوشش غالب منطقه درختان پهن‌برگ خزان‌کننده شامل بلوط بلندمازو (*Quercus castaneifolia*)، انجیلی (*Parrotia persica*)، سفیدپلت (*Populus capsica*) و اوجا (*Ulmus minor*) می‌باشد.

از تصاویر فراطیفی بالای ۹۰ درصد و از ابر نقطه فتوگرامتری از ۴۰ تا ۹۵ درصد، بسته به ویژگی‌های منطقه، متغیر بود. Kuzmin و همکاران (۴) برای شناسایی گونه‌های موجود در جنگل‌های بورال از تصاویر پهپاد در ارتفاع پایین و الگوریتم تحلیل تشخیص خطی استفاده کردند که هدف آن جدا کردن سه گونه غالب در منطقه بود، طبقه‌بندی صحت کلی ۸۲٪ و ضریب کاپا ۰/۷۳ را نشان داد. در داخل کشور نیز پوراحمد و همکاران (۱۶) با استفاده از روش طبقه‌بندی شیء مینا و در ارتفاع پروازهای مختلف شناسایی گونه‌های درختی در جنگل‌های پهن‌برگ خزری را مورد مطالعه قرار دادند. نتایج نشان داد که نقشه حاصل از طبقه‌بندی شیء پایه در ارتفاع پرواز ۵۵ متر بهترین توانایی تشخیص گونه‌های درختی، با ضریب کاپای ۰/۸۱ و صحت کلی ۸۷ درصد را دارد.

علاوه بر داده، نوع الگوریتم طبقه‌بندی نیز در نرخ شناسایی گونه‌های درختی نقش بسیار مهمی دارد. این در حالی است که در مطالعات اندکی تأثیر رویکردهای مختلف طبقه‌بندی با هم مورد مقایسه قرار گرفته‌اند. Ballanti و همکاران (۸) طبقه‌بندی گونه‌های درختی با استفاده از داده‌های فراطیفی در جنگل‌های آمیخته در حوالی سانفرانسیسکو را با استفاده از دو الگوریتم غیر پارامتری جنگل تصادفی و ماشین بردار پشتیبان مورد مقایسه قرار دادند که تفاوتی بین دو الگوریتم مشاهده نشد اما زمانی که تعداد نمونه‌های آموزشی افزوده شد عملکرد الگوریتم ماشین بردار پشتیبان نیز بهبود یافت. Deng و همکاران (۱۷) برای طبقه‌بندی گونه‌های درختی با استفاده از داده‌های لیزر اسکنر هوایی و داده‌های رنگی واقعی از الگوریتم‌های ماشین بردار پشتیبان، جنگل تصادفی و شبکه عصبی مصنوعی استفاده کردند. آن‌ها به این نتیجه رسیدند که ترکیب داده‌ها به همراه استفاده از الگوریتم ماشین بردار پشتیبان بیشترین صحت طبقه‌بندی را ارائه می‌دهد. صادقی و سهرابی (۱۸) از روش طبقه‌بندی جنگل تصادفی و تحلیل تشخیص خطی با استفاده از تصاویر رنگی واقعی اخذ شده از پهپاد طبقه‌بندی درختان جنگلی را انجام دادند. نتایج این تحقیق نشان داد که استفاده از الگوریتم جنگل تصادفی نتایج بهتری را از تحلیل تشخیص خطی ارائه می‌دهد.

با وجود این‌که ارزیابی متغیرهای تک‌پایه‌های درختان و بهبود شناسایی آن‌ها، با کاهش هزینه‌ها و در بعضی موارد با

<sup>۱</sup> Random Forest (RF)

<sup>۲</sup> Linear Discriminant Analysis (LDA)

<sup>۳</sup> Artificial Neural Net (ANN)

<sup>۴</sup> Support Vector Machin (SVM)

از مشخص کردن نقاط کنترل زمینی<sup>۱</sup> و تنظیم شبکه پرواز، ارتفاع پرواز و همپوشانی در خرداد ماه ۱۳۹۸ در ارتفاع پرواز ۱۰۰ متر در یک ظهر آفتابی و بدون وزش باد انجام شد. خلاصه پارامترهای پرواز در جدول ۱ نشان داده شده است.

جدول ۱- خلاصه پارامترهای پرواز

پارامتر	مقدار
ارتفاع پرواز	۱۰۰ متر
اندازه پیکسل	۲/۵ سانتیمتر
همپوشانی طولی	۹۰
همپوشانی عرضی	۹۰
سرعت پرواز	۸ متر/ثانیه

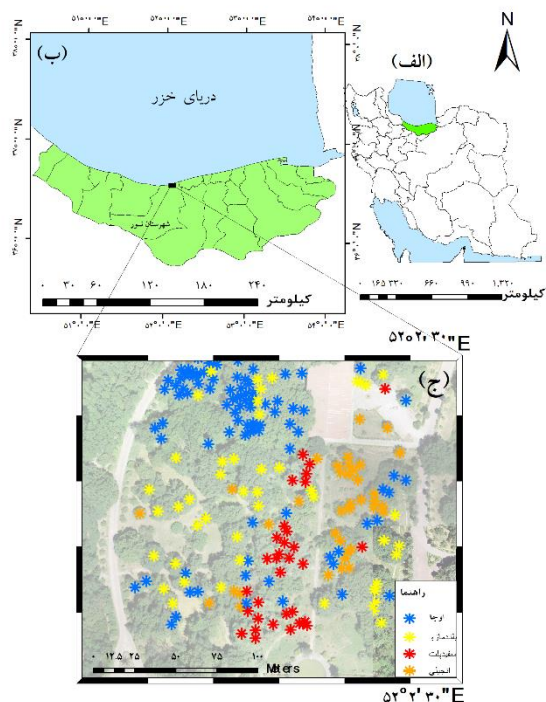
برای پردازش در گام اول تمامی تصاویر به صورت دستی یک بار مورد ارزیابی قرار گرفتند و تصاویری که واضح یا از کیفیت مناسب برخوردار نبودند از مراحل پردازش حذف شدند (۱۹). تصاویر باقیمانده در نرم‌افزار Agisoft PhotoScan v1.27 بازخوانی شدند. پس از مرتب کردن تصاویر در کنار یکدیگر این نقاط متراکم با استفاده از الگوریتم ساختار حرکت مینا تولید شد که برای تهیه موزاییک تصاویر مورد استفاده قرار گرفت. مراحل پردازش تصاویر به منظور تولید ابرنقطه متراکم و موزاییک تصویر مطابق با الگوهای معرفی شده در منابع (۱۱)، (۲۰) انجام شد.

#### ۲-۴- زمین مرجع کردن تصاویر

برای زمین مرجع کردن تصاویر از ۶ نقطه کنترل زمینی در اطراف و ۳ نقطه کنترل زمینی در داخل منطقه مورد مطالعه استفاده شد. برای نقاط کنترل قرار گرفته بر روی جاده از اسپری و نقاط داخل منطقه جنگلی از صفحه‌های پارچه‌ای سفید و قرمز مستقر در نقاط قابل دید از بالای تاج پوشش استفاده گردید. مختصات هر کدام از نقاط کنترل با استفاده از GPS-RTK مدل Gintec G10 برداشت شد.

#### ۲-۵- طبقه‌بندی

طبقه‌بندی و تعیین ترکیب جنگل در سنجش از دور بر این اصل استوار است که هر گونه درختی دارای الگوی طیفی<sup>۲</sup> خاص خود است و دوربین‌هایی مانند دوربین‌های رنگی واقعی، پاسخ طیفی یک درخت در طیف مرئی از طول



شکل ۱- موقعیت منطقه مورد مطالعه در ایران (الف)، استان مازندران (ب) و درختان اندازه‌گیری شده در جنگل نور (ج). نقاط با رنگ مشابه نشان دهنده گونه‌های درختی یکسان است

#### ۲-۲- آماربرداری و تهیه واقعیت زمینی

جمع‌آوری داده‌های زمینی در این تحقیق برای درختان اشکوب فوقانی که در موزاییک تصویر قابل مشاهده بودند انجام شد که در طی آن نوع گونه و موقعیت هر درخت ثبت گردید. به دلیل محدودیت دریافت سیگنال در زیر تاج درختان توسط گیرنده‌های سامانه موقعیت‌یاب جهانی، موقعیت هر درخت با روش آزیموت-فاصله ثبت گردید. در مجموع ۲۰۹ پایه درخت در منطقه مورد مطالعه به تفکیک ۳۷ پایه انجیلی، ۳۳ پایه سفیدپلت، ۴۸ پایه بلوط بلندمازو و ۹۱ پایه اوجا جهت تهیه واقعیت زمینی، مورد آماربرداری قرار گرفتند (درختان ازگیل، زالزالک و پرتقال که در موزاییک تصویر قابل مشاهده بودند، به دلیل تعداد کم مورد آماربرداری قرار نگرفتند).

#### ۲-۳- پردازش و برداشت تصاویر پهپاد

در این مطالعه برای برداشت تصاویر هوایی از پهپاد مدل Phantom 4 Pro با قابلیت ثبت تصاویر رنگی واقعی با فرمت JPG و کیفیت ۲۰ مگاپیکسل استفاده شد. تصویربرداری بعد

<sup>۱</sup> Ground Control Point

<sup>۲</sup> Spectral signature

این مدل با تبدیل داده‌های ورودی به یک فضای ابر صفحه چند بعدی توسط تابع کرنل، سبب گروه‌بندی داده‌های ورودی می‌شود (۲۳). مدلسازی در این روش نیازمند به انتخاب درست برخی پارامترها است که مقدار آن‌ها وابسته به ابعاد داده‌ها بوده و از یک طبقه‌بندی به طبقه‌بندی دیگر متفاوت است. برای تیونینگ پارامترها در الگوریتم‌ها از نمونه‌های آموزشی مورد استفاده در طبقه‌بندی استفاده می‌شود (۸). در مطالعه حاضر از آزمون و خطا برای تعیین پارامترها در الگوریتم‌های یادگیری ماشین غیرپارامتری استفاده شد. در این پژوهش الگوریتم ماشین بردار پشتیبان به صورت کرنل خطی و از ۹۰ بردار پشتیبان استفاده شد. مقدار پارامتر هزینه برای این الگوریتم ۱ در نظر گرفته شد. طبقه‌بندی ماشین بردار پشتیبان در این تحقیق در پکیج e1071 اجرا شد.

## ۲-۵-۳- طبقه‌بندی جنگل تصادفی (RF)

جنگل تصادفی یک الگوریتم طبقه‌بندی و مدل‌سازی ناپارامتری است که از یک مجموعه درختان تصمیم تشکیل شده است. در این الگوریتم، درختان جنگل توسط مجموعه‌ای از نمونه‌های تعلیمی تصادفی ایجاد می‌شوند. الگوریتم طبقه‌بندی جنگل تصادفی برای اجرا به چند پارامتر از جمله تعداد درختان و تعداد ویژگی‌های پیش-بینی کننده نیاز دارد و با حداقل تنظیمات برای پارامترها، به محاسبات و زمان اجرای کمتر نسبت به سایر طبقه‌بندی‌ها نیاز دارد، همچنین دارای دقت طبقه‌بندی بالایی در مدل‌های پیچیده است (۸). در فرآیند طبقه‌بندی تعداد بسیاری درخت تصمیم در طبقه‌بندی رشد می‌یابد و پدیده طبقه‌بندی نشده با توجه به ویژگی‌های همراه خود در یک طبقه قرار خواهند گرفت. این امر با عبور دادن پدیده از میان درختان تصمیم انجام خواهد شد به گونه‌ای که تعلق آن به یکی از طبقه‌ها در میان درختان تصمیم به رای گذاشته شود و جنگل، پدیده به رای گذاشته شده را به طبقه‌ای که دارای بیشترین رای درختان تصمیم است اختصاص می‌دهد. پیش فرض تعداد درخت تصمیم ۵۰۰ است که با آزمون و خطا بالاترین عملکرد با ۱۰۰ درخت تصمیم به دست آمد. همچنین کلیه درختان برداشت شده به عنوان داده آموزشی در نظر گرفته شدند. الگوریتم جنگل تصادفی در پکیج randomForest انجام شد.

موج خاص از تابش الکترومغناطیسی را ثبت نموده و تمایز گونه‌ها از اختلاف در پاسخ طیفی بین هر درخت را ممکن می‌سازند. در این تحقیق بعد از پردازش و ایجاد موزاییک تصویر نسبت‌های باندی<sup>۱</sup> و شاخص‌های پوشش گیاهی<sup>۲</sup> در نرم افزار Arcmap10.3 محاسبه شدند. سپس محدوده میانی تاج هر پایه درختی با استفاده از یک بافر یک متری ترسیم شد و آماره‌هایی مانند میانگین، مد، میانه، دامنه و انحراف معیار برای هر درخت بر روی باندهای اصلی<sup>۳</sup> (R, G, B)، نسبت‌های باندی (G/R, R/B) و شاخص‌های گیاهی (NGB, NRB) محاسبه شد. در نهایت طبقه‌بندی گونه‌های درختی با استفاده از الگوریتم‌های مذکور در محیط نرم‌افزار R(3.6.0) (۲۱) انجام گردید. برای ارزیابی صحت نتایج هر الگوریتم از روش اعتبارسنجی K-fold استفاده شد. به این صورت که مقدار K برابر ۱۰ در نظر گرفته شد و مجموعه داده‌های آموزشی به طور تصادفی به ۱۰ زیرنمونه با حجم یکسان تفکیک شد. در هر مرحله از فرایند، تعداد K-۱ از این لایه‌ها به عنوان داده آموزشی و یک لایه به عنوان داده اعتبارسنجی در نظر گرفته شد. برای دستیابی به مدل مناسب، این فرایند ۱۰ بار تکرار خواهد شد (۲). سپس از معیارهای صحت تولید کننده<sup>۴</sup>، صحت کاربر<sup>۵</sup> و صحت کلی<sup>۶</sup> برای ارزیابی مدل استفاده شد.

## ۲-۵-۱- طبقه‌بندی تحلیل تشخیص خطی (LDA)

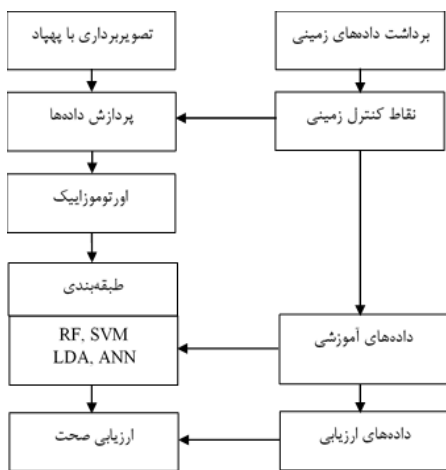
تحلیل تشخیص خطی یک الگوریتم پارامتری است که تفسیر آن ساده و سریع انجام می‌شود و برای اجرا نیاز به داده‌های آموزشی زیادی ندارد. این تحلیل یکی از فنون مؤثر در طبقه‌بندی داده‌های سنجش از دور است (۲۲). برای انتخاب مهمترین متغیرها در تحلیل تشخیص خطی، از روش گام به گام استفاده شد (۱۸)، (۴). طبقه‌بندی درختان با استفاده از این تحلیل در پکیج Mass انجام شد.

## ۲-۵-۲- طبقه‌بندی ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان یک الگوریتم یادگیری ماشین غیرپارامتری به منظور طبقه‌بندی و برآورد داده‌ها است.

۱ Band ratio  
 ۲ Vegetation index  
 ۳ Raw band  
 ۴ Producer's accuracy  
 ۵ User's accuracy  
 ۶ Overall accuracy

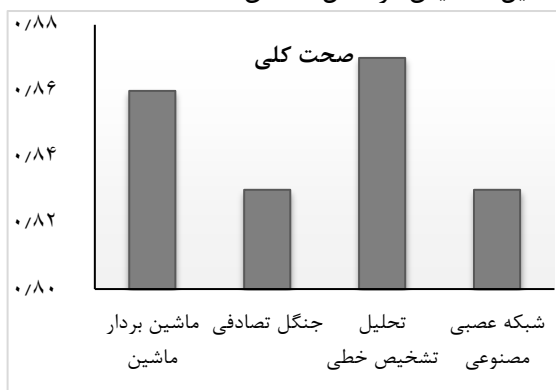
مراحل اجرای فرایند تحقیق در شکل ۴ نشان داده شده‌است.



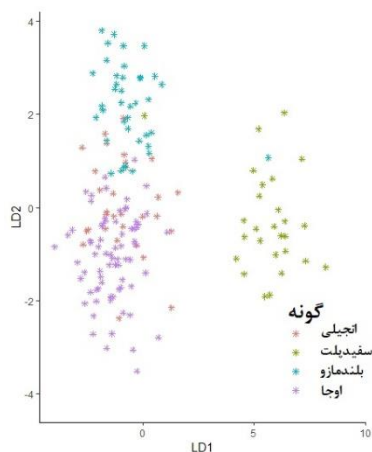
شکل ۴- نمودار مراحل انجام کار

### ۳- نتایج

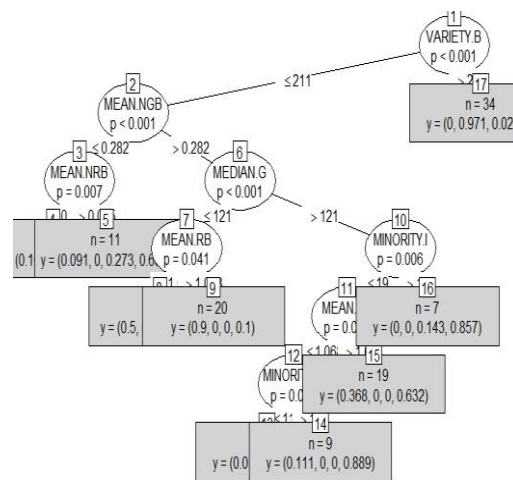
در این تحقیق عملکرد هر چهار الگوریتم به صورت جداگانه مورد بررسی قرار گرفت ( شکل ۵) و نتایج نشان داد که تحلیل تشخیص خطی با صحت کلی ۰/۸۷ بیشترین مقدار صحت کلی را دارد. نمودار طبقه‌بندی تحلیل تشخیص در شکل ۶ نشان داده شده است.



شکل ۵- نمودار عملکرد طبقه‌بندی درختان برای هر الگوریتم



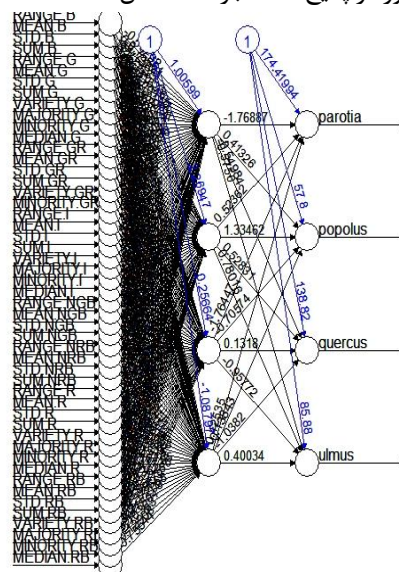
شکل ۶- نمودار طبقه‌بندی درختان با استفاده از تحلیل تشخیص خطی



شکل ۲- دیاگرام طبقه‌بندی جنگل تصادفی

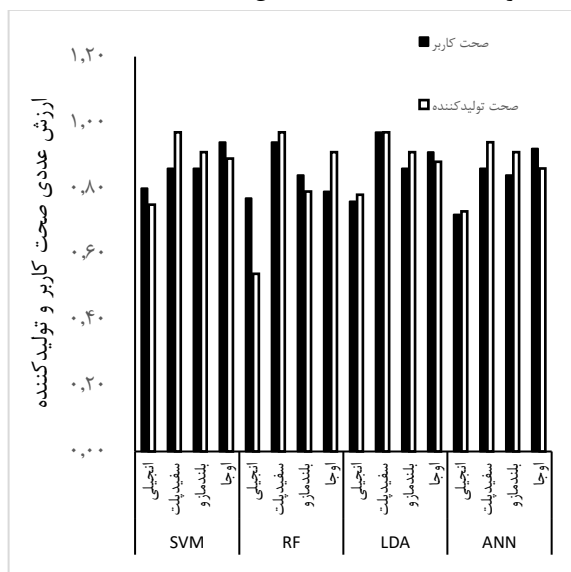
### ۲-۵-۴- شبکه عصبی مصنوعی (ANN)

شبکه عصبی مصنوعی یک روش خودآموز برای توصیف روابط پیچیده برای دستیابی به برآوردهای غیرخطی قوی است. از مزایای استفاده از این الگوریتم ساختار ساده، توانایی یادگیری قوی، دقت بالا و تحمل خطای بالا است. یکی از رایج‌ترین انواع شبکه عصبی شبکه چند لایه پرسپترون است که مبتنی بر یادگیری نظارت شده است و فرایند آموزش در آن نیاز به مجموعه‌ای از داده‌های آموزشی با ورودی‌ها و خروجی‌های مناسب دارد (۲۴). به منظور طبقه‌بندی درختان با استفاده از شبکه عصبی مصنوعی در این تحقیق پارامترهای شبکه (تعداد واحدها در لایه‌های میانی، تنزل و تعداد تکرار) بر اساس آزمون و خطا تغییر داده شد و در نهایت تعداد لایه پنهان ۱، تنزل ۰/۰۰۱ و تکرار ۱۰۰ در نظر گرفته شد. طبقه‌بندی درختان با روش مذکور در پکیج nnet اجرا شد (شکل ۳).



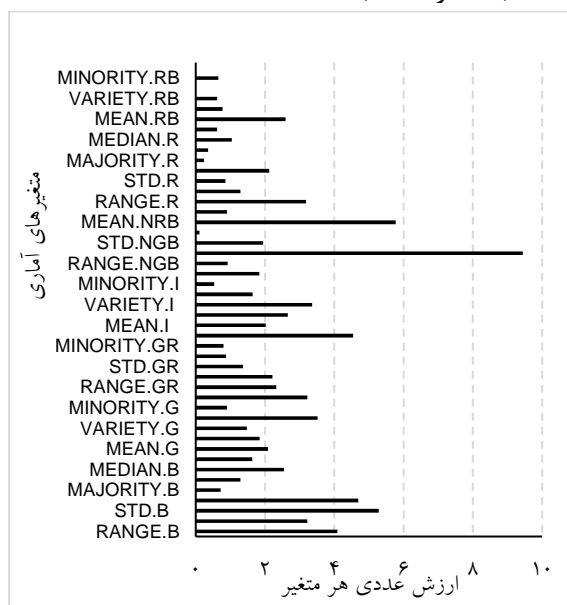
شکل ۳- دیاگرام طبقه‌بندی درختان در الگوریتم شبکه عصبی مصنوعی

بیشترین و کمترین مقدار صحت کاربر و صحت تولید کننده برای گونه‌های سفیدپلت و انجیلی به ترتیب با مقادیر صحت کاربر ۰/۹۷ و ۰/۷۲ و صحت تولیدکننده ۰/۹۷ و ۰/۵۴ به دست آمد (شکل ۷).



شکل ۷- مقدار صحت کاربر و صحت تولیدکننده برای هر الگوریتم

ضریب اهمیت متغیرهای استفاده شده در هر الگوریتم مقادیر عددی متفاوت را نشان داد. چنان‌که در شکل ۸ نشان داده شده است بالاترین مقدار برای میانگین شاخص NGB با مقدار ۹/۴۴ به دست آمد.



شکل ۸- ارزش عددی متغیرهای مورد استفاده در الگوریتم جنگل تصادفی

همچنین مهمترین متغیر در شناسایی درختان در استفاده از تحلیل تشخیص خطی میانگین باند آبی با مقدار

در الگوریتم تحلیل تشخیص خطی از ۳۳ پایه درخت سفیدپلت فقط یک پایه به عنوان بلوط شناسایی شد. در مورد شناسایی درختان بلوط بلندمازو با الگوریتم تحلیل تشخیص نیز ۴۴ پایه به عنوان درخت بلوط، یک پایه به عنوان سفیدپلت و چهار پایه بعنوان انجیلی و دو پایه به عنوان اوجا شناسایی شدند. بیشترین اشتباه طبقه‌بندی مربوط به گونه انجیلی بود که به عنوان گونه اوجا شناسایی شده بود که عکس آن نیز صادق بود. یعنی درخت اوجا به صورت اشتباه به عنوان درخت انجیلی شناسایی شد.

ماشین بردار پشتیبان نیز صحت بالای ۰/۸۶ را نشان داد. الگوریتم‌های شبکه عصبی مصنوعی و جنگل تصادفی با صحت کلی ۰/۸۳ نسبت به دو الگوریتم دیگر صحت کمتری را نشان دادند. ماتریس اختلاف برای الگوریتم‌های مورد استفاده در جدول شماره ۲ نشان داده شده است.

جدول شماره ۲- ماتریس اختلاف طبقه‌بندی SVM: ماشین بردار پشتیبان. RF: جنگل تصادفی. LDA: تحلیل تشخیص خطی. ANN: شبکه عصبی مصنوعی. محور عمودی واقعیت زمینی و محور افقی

پیش‌بینی را نشان می‌دهد

الگوریتم	گونه	انجیلی	سفید پلت	بلوط	اوجا	مجموع
SVM	انجیلی	۲۸	۰	۲	۵	۳۵
	سفیدپلت	۲	۳۲	۱	۲	۳۷
	بلوط	۳	۱	۴۴	۳	۵۱
	اوجا	۴	۰	۱	۸۱	۸۶
	مجموع	۳۷	۳۳	۴۸	۹۱	
RF	انجیلی	۲۰	۰	۰	۶	۲۶
	سفیدپلت	۱	۳۲	۱	۰	۳۴
	بلوط	۴	۱	۳۸	۲	۴۵
	اوجا	۱۲	۰	۹	۸۳	۱۰۴
	مجموع	۳۷	۳۳	۴۸	۹۱	
LDA	انجیلی	۲۹	۰	۰	۹	۳۸
	سفیدپلت	۰	۳۲	۱	۰	۳۳
	بلوط	۴	۱	۴۴	۲	۵۱
	اوجا	۴	۰	۳	۸۰	۸۷
	مجموع	۳۷	۳۳	۴۸	۹۱	
ANN	انجیلی	۲۷	۰	۱	۹	۳۷
	سفیدپلت	۲	۳۱	۱	۲	۳۶
	بلوط	۴	۲	۴۴	۲	۵۲
	اوجا	۴	۰	۲	۷۸	۸۴
	مجموع	۳۷	۳۳	۴۸	۹۱	

فیشر ۱۳/۴ به دست آمد. و در شبکه عصبی مصنوعی نیز مهمترین متغیر شامل میانگین و میانه باند آبی بود.

#### ۴- بحث

در مطالعه حاضر نتایج عملکرد چهار روش طبقه‌بندی ماشین بردار پشتیبان، جنگل تصادفی، تحلیل تشخیص خطی و شبکه عصبی مصنوعی برای شناسایی درختان جنگلی با استفاده از شاخص‌های طیفی مورد بررسی قرار گرفت و مشخص شد که استفاده از تصاویر پهپاد با وضوح مکانی بسیار بالا - تا وضوح سانتی‌متر - برای تمایز گونه‌های جنگلی بسیار امیدوار کننده است.

در این مطالعه تحلیل تشخیص خطی با صحت کلی ۸۷ درصد نتیجه بهتری را در مقایسه با سه روش دیگر برای شناسایی درختان ارائه داد. زیرا در روش تحلیل تشخیص به عنوان یک روش پارامتری نیازی به تنظیم پارامترهای آزاد نیست و این قابلیت خوب منجر به استفاده گسترده و عملی آن در برنامه‌های طبقه‌بندی سنجش از دور شده است (۴)، (۲۵)، (۲۲)، (۳). اما یافتن مناسب‌ترین ورودی و تنظیم پارامترهای مختلف برای روش‌های غیرپارامتری مانند جنگل تصادفی و شبکه عصبی مصنوعی به صورت آزمون و خطا مشخص می‌شود تا این‌که در نهایت مدل‌های مختلفی تولید و بررسی شوند (۲۴). بنابراین، استفاده از این روش‌ها باید با نظارت کارشناسانی انجام شود که می‌توانند روند یادگیری انتخاب مدل را بطور مناسب انجام دهند. علیرغم این‌که معیارهای مختلفی برای انتخاب پارامترهای آزاد در روش‌های غیر پارامتری در دسترس است، این روش‌ها نیازمند کارهای محاسباتی هستند، به ویژه هنگامی که باید چندین پارامتر آزاد (حساس) تنظیم شوند (۲۲). روش ماشین بردار پشتیبان نیز به عنوان یک روش غیرپارامتری دیگر قادر به شناسایی درختان با صحت ۸۶ درصد بود، که دلیل آن را می‌توان در عملکرد مناسب این الگوریتم در حضور داده‌های آموزشی کم دانست (۲۶). Kuzmin و همکاران (۴)،

Shang و Chisholm (۲۵)، Bandos و همکاران (۲۲)، Kim (۳) و Sadeghi و Sohrabi (۱۸) با استفاده از تحلیل تشخیص خطی نتایج قابل توجه‌ای به دست آوردند. در حالی که نتایج قابل قبول در مطالعات Ballanti و همکاران (۸)، Deng و همکاران (۱۷)، Korpela و همکاران (۲۷) با

استفاده از الگوریتم‌های مختلف غیرپارامتری به دست آمد. با توجه به سابقه تحقیق، بیشینه صحت برای طبقه‌بندی در تحقیقات متفاوت با استفاده از الگوریتم‌های مختلف به دست آمده است. در واقع روش‌های طبقه‌بندی مذکور برای دستیابی به دقت مناسب باید آموزش داده شوند این بدان معنا است که یکسری داده‌های آموزشی باید به صورت دستی تولید شوند تا مدل بر اساس آن‌ها پیش‌بینی را انجام دهد. به همین دلیل نحوه انتخاب داده‌های آموزشی بر دقت و صحت طبقه‌بندی بسیار مؤثر می‌باشد (۲۴).

به‌علاوه تعداد داده‌های آموزشی نیز در نحوه عملکرد هر مدل بسیار مؤثر است. در کل ما از ۲۰۹ مشاهده و ۴۶ هر مدل برای شناسایی درختان استفاده کردیم در حالی که با افزایش در تعداد داده‌های آموزشی عملکرد مدل تغییر خواهد کرد (۸). از سوی دیگر داده‌های مرجع نیز به اندازه کافی متعادل نبودند زیرا بیشتر داده‌های ورودی را گونه اوجا تشکیل می‌داد که این نیز می‌تواند بر روی داده‌های تعلیمی و ارزیابی عملکرد تأثیر بگذارد.

در مطالعه حاضر عملکرد بسیار مطلوبی از نظر دقت طبقه‌بندی به خصوص در مورد شناسایی گونه سفیدپلت مشاهده شد که می‌تواند با رنگ تاج متفاوت این گونه از گونه‌های دیگر توجیه شود (۴)، (۱۸)، (۲۸) به نحوی که از ۳۳ پایه، ۳۲ پایه به درستی به عنوان درخت سفیدپلت شناسایی شدند. بیشترین خطا در طبقه‌بندی تک‌درختان مربوط به گونه‌های انجیلی و اوجا بود. در شناسایی گونه انجیلی در طبقه‌بندی تحلیل تشخیص خطی ۹ پایه به عنوان درخت اوجا تشخیص داده شد. برای درختان اوجا نیز ۴ پایه به عنوان انجیلی و ۳ پایه به عنوان بلوط شناسایی شدند. این موضوع می‌تواند به دلایل مختلفی صورت گرفته باشد. دو گونه اوجا و انجیلی دارای شباهت ظاهری از نظر رنگ و بافت هستند که می‌تواند باعث همپوشانی الگوی طیفی آن‌ها شود. موارد دیگر مانند زاویه تابش و تاریخ و روز تصویربرداری نیز می‌تواند بر طبقه‌بندی اشتباه مؤثر باشد (۸).

مهمترین متغیر در شناسایی گونه‌های درختی بر پایه الگوریتم‌های مختلف میانگین بود که می‌توان علت را در انتخاب نمونه‌های آموزشی با حداقل آمیختگی طیفی دانست. چرا که انتخاب نمونه‌ها به گونه‌ای صورت گرفت که حد الامکان شامل اجزای خود گونه درختی و فاقد سایه، خاک و سایر پوشش گیاهی باشد.



## ۵- نتیجه گیری

به صورت کلی می توان بیان داشت که عملکرد هر چهار الگوریتم قابل توجه بود زیرا در این تحقیق فقط از تصاویر رنگی واقعی استفاده شد که به صورت گسترده در دسترس عموم قرار دارد در حالیکه در بسیاری از مطالعات از داده های گران قیمت فراطیفی و چند طیفی برای بهبود عملکرد استفاده شده است. در مورد مقیاس مکانی نیز، استفاده از پهپاد بیشتر از سایر سکوهای سنجش از دوری محدودیت دارد. اما کم هزینه بودن و کاربرد آسان استفاده از آن می تواند مدیران و کاربران را قادر به نظارت دوره ای کند.

## ۶- پیشنهادها

در این مطالعه از یک بافر دایره ای به شعاع یک متر برای جداسازی پیکسل های تاج درختان استفاده شد. بنابراین در تحقیقات آینده این موضوع باید در نظر گرفته شود که جداسازی تاج درختان به صورت اتوماتیک چه تغییری در نتایج حاصل از شناسایی درختان ایجاد خواهد کرد. همچنین برداشت اطلاعات مربوط به داده های آموزشی و تست در یک منطقه و در یک زمان انجام شد اما با توجه به این که رنگ و نورپردازی برگ و تاج درختان در فصول مختلف متفاوت است بهتر برای افزایش دقت شناسایی گونه ها، در تحقیقات آتی برداشت تصاویر از مکان ها و زمان های مختلف صورت گیرد.

## مراجع

- [1] M. Onishi and T. Ise, "Automatic classification of trees using a UAV onboard camera and deep learning," CoRR, arXiv preprint arXiv:1804.10390, 2018.
- [2] A. Nevalainen, N. Nilton, M. G. Antonio, "Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging," *Remote Sens.*, vol. 9, no. 3, p. 185, 2017.
- [3] S. Kim, "Individual tree species identification using LIDAR- derived crown structures and intensity data", Doctor of Philosophy Diss, University of Washington, 2007.
- [4] A. Kuzmin, L. Korhonen, T. Manninen, and M. Maltamo, "Automatic Segment-Level Tree Species Recognition Using High Resolution Aerial Winter Imagery," *Eur. J. Remote Sens.*, vol. 7254, pp. 238–259, 2017.
- [5] A. C. Birdal, U. Avdan, and T. Türk, "Estimating tree heights with images from an unmanned aerial vehicle," *Geomatics, Nat. Hazards Risk*, vol. 8, no. 2, pp. 1144–1156, 2017.
- [6] M. Mohan et al., "Individual Tree Detection from Unmanned Aerial Vehicle ( UAV ) Derived Canopy Height Model in an Open Canopy Mixed Conifer Forest," *Forests*, vol. 8, no. 9, pp. 1–17, Sep. 2017.
- [7] V. Yilmaz and O. Gungor, "Estimating crown diameters in urban forests with Unmanned Aerial System-based photogrammetric point clouds," *Int. J. Remote Sens.*, vol. 40, no. 2, pp. 468–505, 2019.
- [8] L. Ballanti, L. Blesius, E. Hines, and B. Kruse, "Tree Species Classification Using Hyperspectral Imagery: A Comparison of Two Classifiers," *Remote Sens.*, vol. 8, no. 6, p. 445, May. 2016.
- [9] C. Torresan, P. Corona, G. Scrinzi, and J. Valls Marsal, "Using classification trees to predict forest structure types from LIDAR data," *Ann. For. Res.*, vol. 59, no. 1, pp. 281–298, May. 2016.
- [10] K. Anderson and K. J. Gaston, "Lightweight unmanned aerial vehicles will revolutionize spatial ecology," *Front. Ecol. Environ.*, vol. 11, no. 3, pp. 138–146, Apr. 2013.
- [11] J. Lisein, A. Michez, H. Claessens, and P. Lejeune, "Discrimination of deciduous tree species from time series of unmanned aerial system imagery," *PLoS One*, vol. 10, no. 11, pp. 1–20, Nov. 2015.
- [12] A. Wallner, A. Elatawneh, T. Schneider, M. Kindu, B. Ossig, and T. Knoke, "Remotely sensed data controlled forest inventory concept," *Eur. J. Remote Sens.*, vol. 51, no. 1, pp. 75–87, Jan. 2018.
- [13] J. Maschler, C. Atzberger, M. Immitzer, J. Maschler, C. Atzberger, and M. Immitzer, "Individual Tree Crown Segmentation and Classification of 13 Tree Species Using Airborne Hyperspectral Data," *Remote Sens.*, vol. 10, no. 8, p. 1218, Aug. 2018.
- [14] A. E. Effiom, "UAV-RGB and Multispectral Pleiades images, for tree species identification and forest carbon estimation in Amtsvenn, Germany.," Master of Science, University of Twente, 2018.
- [15] L. Hernandez-Santin et al., "Identifying Species and Monitoring Understorey from UAS-Derived Data: A Literature Review and Future Directions," *Drones*, vol. 3, no. 1, p. 9, Jan. 2019.

- [16] M. Pourahmad, J. Oladi, and A. Fallah, "Detection of Tree Species in Mixed Broad-Leaved Stands of Caspian Forests Using UAV Images (Case study: Darabkola Forest)," *Ecol. Iran. For.*, vol. 6, no. 11, pp. 61–75, 2018., (in Persian).
- [17] S. Deng, M. Katoh, X. Yu, J. Hyypä, and T. Gao, "Comparison of tree species classifications at the individual tree level by combining ALS data and RGB images using different algorithms," *Remote Sens.*, vol. 8, no. 12, 2016.
- [18] S. Sadeghi and H. Sohrabi, "Tree species discrimination using RGB vegetation indices derived from UAV images," in *UAV Small Unmanned Aerial system for environmental research-6th Edition*, no. 1, p. 5, 2018.
- [19] O. Brovkina, E. Cienciala, P. Surový, and P. Janata, "Unmanned aerial vehicles (UAV) for assessment of qualitative classification of Norway spruce in temperate forest stands," *Geo-spatial Inf. Sci.*, vol. 21, no. 1, pp. 12–20, Jan. 2018.
- [20] Å. M. Klein Hentz and M. P. Strager, "Cicada Tree Damage Detection Based on UAV Spectral and 3D Data," *Nat. Sci.*, vol. 10, no. 01, pp. 31–44, Jan. 2018.
- [21] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2019.
- [22] T. Bandos, L. Bruzzone, G. Camps-valls, S. Member, G. Camps-valls, and S. Member, "Classification of Hyperspectral Images With Regularized Linear Discriminant Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 862–873, 2009.
- [23] A. Safari, H. Sohrabi, S. Powell, and S. Shataee, "A comparative assessment of multi-temporal Landsat 8 and machine learning algorithms for estimating aboveground carbon stock in coppice oak forests," *Int. J. Remote Sens.*, vol. 38, no. 22, pp. 6407–6432, Nov. 2017.
- [24] C. Poblete-Echeverría, G. Olmedo, B. Ingram, and M. Bardeen, "Detection and Segmentation of Vine Canopy in Ultra-High Spatial Resolution RGB Imagery Obtained from Unmanned Aerial Vehicle (UAV): A Case Study in a Commercial Vineyard," *Remote Sens.*, vol. 9, no. 3, p. 268, Mar. 2017.
- [25] X. Shang and L. A. Chisholm, "Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2481–2489, 2014.
- [26] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3. Elsevier, pp. 247–259, 01-May-2011.
- [27] I. Korpela, H. Ole Ørka, M. Maltamo, T. Tokola, and J. Hyypä, "Tree species classification using airborne LiDAR - effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type," *Silva Fenn.*, vol. 44, no. 2, pp. 319–339, 2010.
- [28] K. Rominger and S. Meyer, "Application of UAV-Based Methodology for Census of an Endangered Plant Species in a Fragile Habitat," *Remote Sens.*, vol. 11, no. 6, p. 719, Mar. 2019.