



تشخیص و استخراج درختان نخل از تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالای گوگل ارث بر مبنای شبکه‌های یادگیری عمیق باقیمانده

مصطفی کابلی‌زاده^{۱*}، کاظم رنگزن^۲، محمد عباسی^۳

دریافت: ۱۴۰۱/۰۶/۱۴ / پذیرش: ۱۴۰۱/۰۹/۰۶ / دسترسی اینترنتی: ۱۴۰۳/۱۲/۱۰

چکیده

شناسایی موقعیت درختان اولین گام جهت مدیریت فضای سبز، باغ‌ها و جنگل‌ها است. تهیه نقشه موقعیت درختان می‌تواند با عملیات زمینی نقشه‌برداری که نیاز به هزینه و زمان زیادی دارد یا با استفاده از تصاویر هوایی یا ماهواره‌ای انجام شود. در این پژوهش از تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالای گوگل ارث برای تشخیص و استخراج درختان نخل با توجه به نقش و اهمیت درخت نخل در مناطق جنوبی ایران استفاده شده است، اما تشخیص خودکار درخت از تصاویر ماهواره‌ای یک چالش است. در این راستا روش‌های یادگیری عمیق به عنوان یک راه‌حل مهم برای استخراج اشیا از تصاویر مطرح هستند.

در این تحقیق از روش‌های یادگیری عمیق باقیمانده با تعداد لایه‌های ۱۸، ۳۴ و ۵۰ استفاده شده است. ابتدا بیش از ۳۰۰۰ نمونه تصویر در دو کلاس حاوی درخت نخل و بدون درخت نخل با ابعاد ۶۴ در ۶۴ پیکسل بریده شده و سپس مدل‌ها با ۸۰ درصد نمونه‌ها برای یادگیری و ۲۰ درصد برای اعتبارسنجی با ۳۰ دوره تکرار و دقت بالای ۹۹ درصد برای هر سه مدل آموزش داده شده است. مدل آموزش دیده بر روی ۵۰۰ نمونه تست اجرا شده و نتایج ارزیابی هر سه مدل نشان داد که معیار دقت بیش از ۰.۹۶ و معیار بازیابی برابر ۱ و معیار F1Score بیش از ۰.۹۸ است. اجرای مدل‌ها بر روی تصاویر ماهواره‌ای گوگل ارث با حرکت پنجره ۶۴ در ۶۴ پیکسل با گام ۱۶ پیکسل و سپس اعمال روش سرکوب غیرحداکثری نشان می‌دهد که می‌توان از تصاویر ماهواره‌ای سامانه گوگل ارث برای تهیه نقشه موقعیت درختان نخل استفاده نمود. با توجه به زمان پردازش و امکان برآورد بهتر تعداد و استخراج موقعیت درختان نخل، مدل یادگیری عمیق باقیمانده با ۳۴ لایه پیشنهاد شد.

طرح مسئله: داده‌های مکانی و شمارش درختان نخل برای بررسی منطقه کاشت، پیش‌بینی عملکرد میوه، مدیریت و

مصطفی کابلی‌زاده^۱ (✉)، کاظم رنگزن^۲، محمد عباسی^۳

۱. دانشیار گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم زمین،

دانشگاه شهید چمران اهواز، اهواز، ایران

۲. استاد گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم زمین،

دانشگاه شهید چمران اهواز، اهواز، ایران

۳. دانشجوی دکتری سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم

زمین، دانشگاه شهید چمران اهواز، اهواز، ایران

DOI: 10.30495/girs.2022.697278

پست الکترونیکی مسئول مکاتبات: Amir.Ghayebi@shahroodut.ac.ir

ارث دانلود شده در باند مرئی (RGB) بوده و فاقد باندهای دیگر در تصاویر اصلی کوئیک برد می باشد که چالش تشخیص و استخراج درخت را بدلیل عدم امکان استفاده از سایر باندها بیشتر می کند. نمونه های آموزشی در دو کلاس شامل تصاویر حاوی درخت خرما و بدون درخت نخل تهیه گردید. نمونه های آموزشی طوری از تصاویر اصلی بریده شده اند که درخت نخل در وسط تصویر باشد. با توجه به برآورد اندازه درختان خرما، بیش از ۱۶۰۰ نمونه حاوی درخت نخل برای یادگیری با ابعاد ۶۴ در ۶۴ پیکسل استخراج گردید. همچنین بیش از ۱۶۰۰ نمونه تصویری بدون درخت نخل تهیه گردید. در نمونه های استخراج شده کیفیت تصویر، رنگ، کنتراست یا حذف سایه تغییری نکرده است. در نهایت نمونه ها در دو کلاس با درخت و بدون درخت برچسب گذاری شده اند. از نمونه های انتخاب شده ۲۰ درصد به صورت تصادفی برای اعتبارسنجی در مدل استفاده شده اند. در نهایت شبکه های یادگیری عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه آموزش داده شده اند و با حرکت پنجره جستجو در تصویر موقعیت مکانی درختان نخل با ضریب اطمینان برای هر قطعه تصویر محاسبه شده است. برای حذف تکراری قطعه های دارای همپوسانی از روش سرکوب غیر حداکثری استفاده شده است و نقشه موقعیت درختان نخل تهیه شده است.

نتایج و بحث: در این پژوهش برای آموزش مدل از روش های یادگیری عمیق باقیمانده با تعداد لایه های ۱۸ (ResNet18)، ۳۴ (ResNet34) و ۵۰ (ResNet50) استفاده شده است. پارامترهای مدل اجرا شده در این مطالعه شامل تعداد دوره های آموزشی برابر ۳۰ اپک می باشد. آموزش مدل های پیشنهادی بر روی داده های آموزشی با یک کامپیوتر شخصی لپ تاپ با پردازشگر مرکزی نسل هشت Core i7 و RAM برابر ۱۲ گیگابایت انجام شده است. زمان پردازش حدود ۱ ساعت و ۲۰ دقیقه برای مدل ResNet18 و حدود ۲ ساعت و ۱ دقیقه برای مدل ResNet34 و حدود ۳ ساعت و ۱۶ دقیقه برای مدل ResNet50 طول کشیده است. هرچه تعداد لایه ها بیشتر باشد، شبکه عمیق تر شده و زمان یادگیری بیشتر خواهد شد. میانگین

برنامه ریزی هوشمند باغ ها بسیار مهم هستند. این مهم به ابزار و فناوری نیاز دارد که اطلاعات قابل اعتماد را با دقت قابل قبول ارائه دهد. استفاده از روش های برداشت زمینی با استفاده از دوربین های نقشه برداری یا گیرنده های تعیین موقعیت ماهواره ای هر چند دارای دقت بالایی می باشند اما چالش های زیادی علاوه بر زمان و هزینه بالا دارند. در این راستا استفاده از تصاویر هوایی و ماهواره ای می تواند یک راه حل بهتر با در نظر گرفتن زمان و هزینه باشد. مطالعات در تشخیص و استخراج درخت برای جنگلداری و کشاورزی دقیق مانند درختان نخل در چند دهه اخیر به ویژه با استفاده از سنجش از دور در حال رشد بوده است. مدل های تشخیص اشیا در یادگیری عمیق که به طور گسترده در بینایی کامپیوتری استفاده می شوند، می توانند فرصتی برای تشخیص درختان نخل فراهم کنند که برای بدست آوردن سریع داده ها و کاهش خطاهای عملیات انسانی ضروری است. اما تشخیص خودکار درختان در تصاویر نوری همچنان یک چالش است. به ندرت مطالعه ای برای تشخیص درخت با استفاده از منابع اینترنتی مجانی و در دسترس مانند سامانه گوگل ارث انجام شده است. برای به دست آوردن اطلاعات پوشش درختی به صورت مقرون به صرفه و کارآمد، در این پژوهش، مدل های یادگیری عمیق بر مبنای شبکه های باقیمانده برای تشخیص درختان نخل از تصاویر گوگل ارث ارائه شده است.

هدف: هدف اصلی این تحقیق تشخیص درختان خرما از تصاویر ماهواره ای با قدرت تفکیک بالای گوگل ارث است که در نتیجه آن بتوان موقعیت مکانی، تعداد و توزیع مکانی درختان را برآورد نمود. با دستیابی به این اهداف، داده های زیادی برای بروزرسانی درختان نخل بدست خواهد آمد. همچنین بینش جدیدی ارائه می دهد که با استفاده از روش های مبتنی بر یادگیری عمیق، می توان اطلاعات در مقیاس بزرگ را از تصاویر ماهواره ای در دسترس به دست آورد.

روش تحقیق: در ابتدا تصاویر ماهواره ای با قدرت تفکیک مکانی بالا از نرم افزار گوگل ارث دانلود شد. تصویر گوگل

دقت داده‌های اعتبارسنجی مدل‌ها بالای ۹۹ درصد می‌باشد که نشان دهنده دقت بالای مدل‌ها در تشخیص درخت نخل در شرایط این پژوهش می‌باشد. برای ارزیابی مدل ۵۰۰ نمونه به عنوان داده‌های تست که در روند یادگیری مدل‌ها شرکت نکرده‌اند، تهیه گردید. ارزیابی نتایج بر روی داده‌های تست نشان می‌دهد با افزایش لایه‌های مدل لزوماً دقت افزایش نیافته است. ضریب اطمینان نمونه‌های تشخیصی اشتباه کمتر از ۰.۸ بوده است.

برای تشخیص و استخراج درخت از تصاویر ماهواره‌ای گوگل ارث، تصویر ماهواره‌ای ورودی در قطعات ۶۴ در ۶۴ پیکسل با گام یک چهارم اندازه هر قطعه یعنی ۱۶ پیکسل بریده شده است. بدیهی است کاهش گام می‌تواند دقت نهایی را افزایش دهد، اما زمان پردازش مقرون به صرفه نخواهد بود. سپس تصاویر به مدل‌های آموزش دیده جهت تشخیص کلاس آن‌ها وارد شده است، مدل‌ها برای هر قطعه نوع کلاس با و بدون درخت نخل را با ضریب اطمینان بین ۰ تا ۱ محاسبه می‌کنند. موقعیت هر قطعه با یک کادر مستطیلی بر روی تصویر ماهواره‌ای ژئورفرنس شده، نشان داده می‌شود. از بین قطعاتی که دارای همپوشانی زیادی هستند، قطعاتی که دارای ضریب اطمینان بالاتری هستند با روش سرکوب غیرحداکثری انتخاب و بقیه حذف شده‌اند. بررسی نتایج بر روی تصاویر مختلف نشان می‌دهد که هر سه مدل امکان تشخیص درخت نخل از تصاویر ماهواره‌ای گوگل ارث را با دقت مناسب بالای ۹۰ درصد دارند. اما مدل ۳۴ لایه با توجه به زمان پردازش و دقت بالاتر پیشنهاد می‌شود. مدل ۱۸ لایه نسبت به مدل ۳۴ لایه دقت کمتری دارد. بطور کلی مدل ۳۴ لایه برای شمارش درختان نتایج بهتری نشان می‌دهد.

نتیجه‌گیری: در این پژوهش زمان لازم برای یادگیری مدل برای مدل باقیمانده ۵۰ لایه تقریباً سه برابر مدل باقیمانده ۱۸ لایه و برای مدل ۳۴ لایه نیز تقریباً دو برابر ۱۸ لایه بوده است. تعداد دوره آموزشی برای هر سه مدل، ۳۰ اپک در نظر گرفته شده است. با توجه به رفتار تابع ضرر بر روی نمونه‌های آموزشی و اعتبارسنجی، هر سه مدل بخوبی آموزش دیده‌اند. مقدار

پارامترهای ارزیابی مدل شامل معیار دقت، معیار بازیابی و معیار F1Score برای هر سه مدل بیش از ۹۹ درصد می‌باشد. مدل شبکه عصبی باقیمانده ۵۰ لایه نتایج بهتری بر روی داده‌های اعتبارسنجی نشان داده است. برای ارزیابی مدل‌ها با داده‌های تست، بیش از ۵۰۰ نمونه تصویری که در فرآیند یادگیری شرکت نکرده‌اند، استخراج گردید. نتایج اجرای مدل‌های شبکه عصبی پیشنهادی بر روی داده‌های تست نشان می‌دهد که مدل شبکه عصبی باقیمانده ۱۸ لایه نسبت به مدل ۳۴ و ۵۰ لایه نتایج بهتری دارد. معیارهای ارزیابی دقت برای هر سه مدل بیش از ۹۶ درصد می‌باشد که نشان‌دهنده توانایی هر سه مدل در تشخیص درخت نخل از تصاویر ماهواره‌ای گوگل ارث است. برای تشخیص و استخراج موقعیت درختان نخل از تصاویر ماهواره‌ای یک پنجره با ابعاد ۶۴ در ۶۴ پیکسل بر روی تصویر ورودی حرکت داده شده است. برای کاهش زمان پردازش گام حرکت برابر یک چهارم ابعاد پنجره یعنی ۱۶ پیکسل در نظر گرفته شد. با ورود قطعات تصویری به مدل‌های پیشنهادی تعلق هر قطعه به یکی از دو کلاس ورودی با محاسبه و تعریف ضریب اطمینان آن بدست آمده است. برای تشخیص درخت نخل فقط قطعه‌های با ضریب اطمینان ۹۰ درصد انتخاب شد. چون قطعات با هم همپوشانی دارند از روش سرکوب غیرحداکثری برای حذف قطعات تکراری با امکان پوشش ۲۵ درصدی قطعات استفاده گردید. نتایج نشان می‌دهد که هر سه مدل با دقت بالای ۹۰ درصد امکان شمارش و تعیین موقعیت درختان نخل را دارند. در مدل شبکه عصبی باقیمانده ۱۸ لایه تعداد بیشتری از درختان نسبت به مدل‌های ۳۴ و ۵۰ لایه تشخیص داده نشده‌اند. مدل ۵۰ لایه چون حساسیت بیشتری دارد، قطعات نهایی دارای همپوشانی بیشتری هستند و در نتیجه تعداد درختان از حالت واقعی بیشتر خواهد بود. با توجه به زمان پردازش و بررسی دقت نتایج در تهیه نقشه موقعیت درختان نخل مدل شبکه یادگیری عمیق باقیمانده ۳۴ لایه پیشنهاد می‌شود. در صورت استفاده از تصاویر هوایی و همچنین افزایش تعداد نمونه‌های آموزشی

دقت مدل افزایش خواهد داشت. همچنین می‌توان با ادغام تصاویر چندطیفی کارکرد مدل را بهبود بخشید.

واژگان کلیدی: یادگیری ماشین، تشخیص خودکار، استخراج اهداف، شبکه عصبی کانولوشنال

لطفاً به این مقاله استناد کنید: کابلی‌زاده، م.، رنگزن، ک.، عباسی، م. تشخیص و استخراج درختان نخل از تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالای گوگل ارث بر مبنای شبکه‌های یادگیری عمیق باقیمانده. نشریه سنجش‌ازدور و سامانه اطلاعات جغرافیایی در منابع طبیعی، ۱۵(۴): ۶۷-۴۸.

مقدمه

تشخیص و استخراج موقعیت درختان می‌تواند به کشاورزی دقیق به‌ویژه در صنعت کاشت و مدیریت کارآمد باغ‌ها و جنگل‌ها کمک کند. این مهم به ابزار و فناوری نیاز دارد که اطلاعات قابل‌اعتماد را با دقت قابل‌قبول ارائه دهد. استفاده از روش‌های برداشت زمینی با استفاده از دوربین‌های نقشه‌برداری یا گیرنده‌های تعیین موقعیت ماهواره‌ای هرچند دارای دقت بالایی می‌باشند اما چالش‌های زیادی علاوه بر زمان و هزینه بالا دارند. در این راستا استفاده از تصاویر هوایی و ماهواره‌ای می‌تواند یک راه‌حل بهتر با در نظر گرفتن زمان و هزینه باشد. تحقیقات اخیر با استفاده از پیشرفت فناوری سنجش‌ازدور و ادغام الگوریتم‌های مختلف، این هدف را تأمین می‌کند. تشخیص درخت براساس تصاویر سنجش‌ازدور یک فناوری حیاتی برای ایجاد یک پایگاه داده درختان و پایش منابع کاشت درخت است که برای برنامه‌ریزی منظر شهری و نظارت بر محیط‌زیست محیطی اهمیت زیادی دارد (۱). پوشش درختی در خدمات اکوسیستمی نقش ویژه‌ای به‌عنوان جزء اصلی زیرساخت‌های فضای سبز خصوصاً در محیط‌های شهری ایفا می‌کنند (۲۲). کمیت‌های کلیدی ساختار و توزیع مکانی پوشش درختی می‌توانند درک عملکرد اکوسیستم و انعطاف‌پذیری آن را تسهیل کنند (۱۱). محققان در حال توسعه روش‌های دقیق و مؤثری برای شناسایی ویژگی‌های پوشش درختی و کمی کردن عملکرد آن‌ها هستند (۶). مطالعات در تشخیص و استخراج درخت برای جنگلداری و کشاورزی دقیق مانند درختان نخل در چند دهه اخیر به‌ویژه با استفاده از سنجش‌ازدور در حال رشد بوده است. در مطالعات مناطق دارای پوشش درختی، سنجش‌ازدور عموماً برای نقشه‌برداری، شناسایی و ترسیم تاج درختان، ارزیابی ترکیب گونه‌های درختی، تشخیص و شمارش درخت، تخمین سن، تشخیص آفات یا بیماری و همچنین برای پایش استفاده می‌شود (۲، ۱۲، ۲۰ و ۲۱). در طول دهه‌ها، روش‌های مختلفی برای تشخیص و استخراج درخت از پردازش تصاویر سنتی در سال‌های اولیه تا آخرین روش‌های مبتنی بر یادگیری عمیق در ۱۰ سال

گذشته استفاده شده است. توسعه سریع الگوریتم جدید اخیراً موانع با پیشرفت‌های تکنولوژیکی در زمینه سنجش از دور بوده است. با این حال، این فرض هنوز در نشان دادن رابطه بین الگوریتم‌های خاص با داده‌های سنجش از دور مشخص نیست، زیرا هر مطالعه ملاحظات خاص خود را دارد. انواع مختلفی از فن‌آوری‌های تصویربرداری به سرعت در حال رشد هستند، که می‌تواند شامل تصویربرداری هوایی معمولی (فتوگرامتری کلاسیک)، پرنده‌های هدایت‌پذیر از دور (پهپاد) و تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالا باشد. بطور کلی روش‌های موجود برای تشخیص درخت از تصاویر را می‌توان به سه گروه کلی طبقه‌بندی نمود: اول روش‌های مبتنی بر پردازش تصویر، دوم روش‌های مبتنی بر یادگیری ماشین و سوم روش‌های مبتنی بر یادگیری عمیق. از روش‌های مبتنی بر پردازش تصویر می‌توان روش‌های فیلتر حداکثر محلی، تشخیص و قطعه‌بندی لبه، عملگرهای مورفولوژیکی، تطبیق الگو و شاخص‌های پوشش گیاهی را نام برد (۱۳). روش‌های تشخیص درخت مبتنی بر یادگیری ماشین (ML) معمولاً به استخراج ویژگی، قطعه‌بندی تصویر، آموزش طبقه‌بندی‌کننده و پیش‌بینی نیاز دارند (۹).

طبقه‌بندی‌کننده‌ها که مهم‌ترین بخش در یادگیری ماشین هستند را می‌توان به دو گروه نظارت شده یا بدون نظارت طبقه‌بندی کرد. روش‌های جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM)، درخت تصمیم (DT) و شبکه عصبی برخی از نمونه‌های طبقه‌بندی‌کننده نظارت‌شده هستند. در حالی که نزدیک‌ترین همسایگی K (k-NN) و K-means دسته‌بندی‌کننده‌های بدون نظارت هستند. یادگیری ماشین نظارت شده نیاز به پایگاه داده تصاویر با برچسب مثبت و منفی با کلاس‌های مربوطه (به عنوان مثال درخت و پس زمینه) هستند که برای نمونه‌های آموزشی استفاده می‌شوند. پیشرفت‌های اخیر در یادگیری عمیق تأثیر زیادی بر سنجش از دور به طور کلی (۲۵) و به طور خاص، بر طبقه‌بندی پوشش زمین (۷) داشته است. یادگیری عمیق امکان شناسایی خودکار موقعیت تک تک درختان نخل در مناطق بزرگ را در یک زمان

داده‌های مکانی و شمارش درختان نخل برای بررسی منطقه کاشت، پیش‌بینی عملکرد میوه، مدیریت و برنامه‌ریزی هوشمند باغ‌ها بسیار مهم هستند. این داده‌ها معمولاً از طریق بررسی دستی و آماری و با تلاش زمان‌بر و پرهزینه به دست می‌آیند. مدل‌های تشخیص اشیا در یادگیری عمیق که به طور گسترده در بینایی کامپیوتری استفاده می‌شوند، می‌توانند فرصتی برای تشخیص دقیق درختان نخل فراهم کنند که برای بدست آوردن سریع داده‌ها و کاهش خطاهای عملیات انسانی ضروری است. تشخیص درختان در تصاویر نوری همچنان یک چالش است. به‌دورت مطالعه‌ای برای تشخیص درخت با استفاده از منابع اینترنتی مجانی و در دسترس مانند سامانه گوگل ارث انجام شده است. برای به‌دست آوردن اطلاعات پوشش درختی به صورت مقرون به صرفه و کارآمد، در این پژوهش، یک مجموعه مدل یادگیری عمیق بر مبنای شبکه‌های باقیمانده برای تشخیص درختان نخل از تصاویر گوگل ارث ارائه شده است.

هدف اصلی این تحقیق تشخیص درختان نخل از تصاویر ماهواره‌ای با قدرت تفکیک بالای گوگل ارث است که در نتیجه آن بتوان موقعیت مکانی، تعداد و توزیع مکانی درختان را برآورد نمود. با دستیابی به این اهداف، داده‌های زیادی برای بروزرسانی درختان نخل بدست خواهد آمد. همچنین بینش جدیدی ارائه می‌دهد که با استفاده از روش‌های مبتنی بر یادگیری عمیق، می‌توان اطلاعات در مقیاس بزرگ را از تصاویر ماهواره‌ای در دسترس به دست آورد.

روش تحقیق

منطقه مورد مطالعه

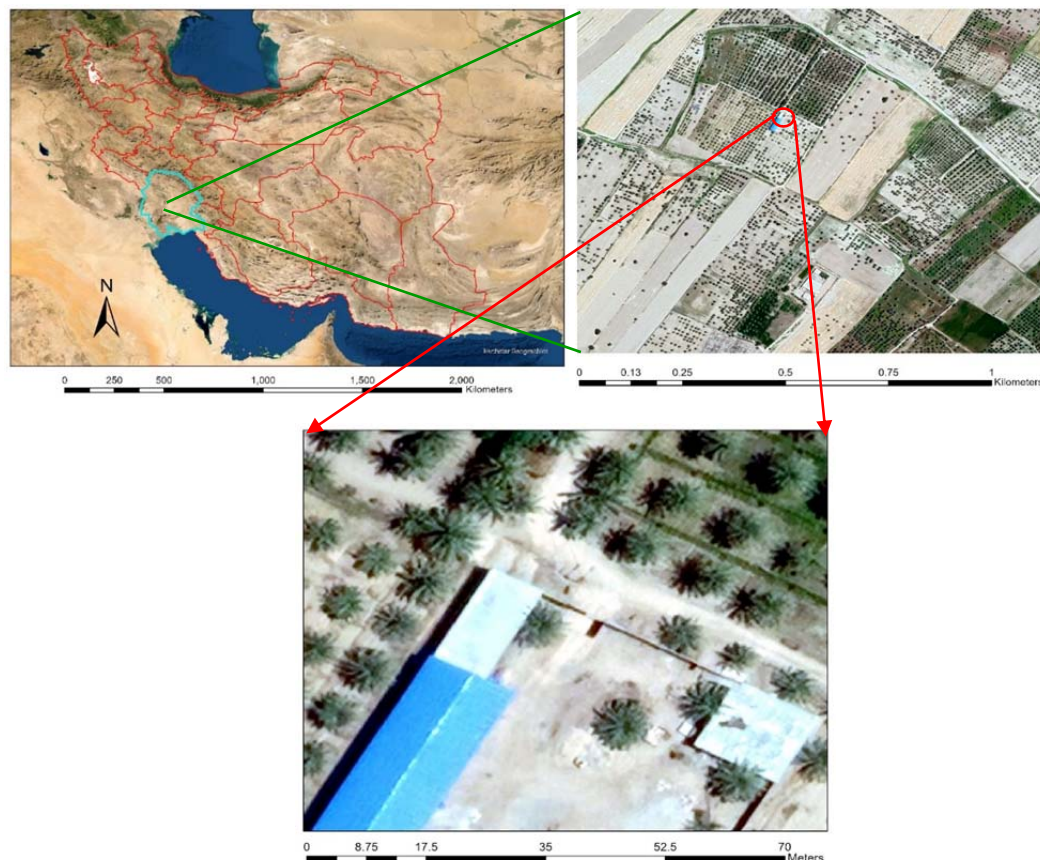
منطقه مورد مطالعه در این تحقیق درختان نخل در محدوده حریم شهری، شهر اهواز در استان خوزستان می‌باشد. در این تحقیق از تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالا استفاده شده است. چون در ایران تنها تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالای در دسترس تصاویر ماهواره‌ای در پایگاه داده گوگل ارث می‌باشد، از تصاویر ماهواره‌ای گوگل ارث استفاده شده است. تصویر ماهواره‌ای منطقه مورد مطالعه

معقول ارائه می‌دهد. چنین داده‌های برای ذینفعان مختلف مورد توجه است. کشاورزان بهتر می‌توانند توسعه مزارع خود را نظارت کرده و فرآیندهای مدیریتی خود را تنظیم کنند (۳).

اخیراً، چندین مدل تشخیص و کلاس‌بندی اشیاء یادگیری عمیق به تدریج در تشخیص و کلاس‌بندی درختان منفرد مانند درختان زیتون، نخل و نارگیل بر اساس تصاویر با قدرت تفکیک مکانی بالا و لیدار به دست آمده از ماهواره‌ها و پهپادها به کار گرفته شده است (۵، ۸ و ۲۴). واینستین و همکاران (۲۳) یک خط لوله یادگیری عمیق نیمه نظارت شده برای تشخیص تاج درختان بر اساس داده‌های مرئی و لیدار پیشنهاد کردند و به ترتیب ۶۹ و ۶۰ درصد دقت تولید کننده و دقت کاربر را به دست آوردند. سرستاساتیرن و راکواتین (۱۹) از تصاویر کوئیک برد (Quickbird) و وردویو (WorldView-2) با وضوح مکانی ۶۰ سانتی متر و چهار باند طیفی در تایلند استفاده کردند. آنها موقعیت نخل را از یک شاخص پوشش گیاهی، با استفاده از تبدیل داده، استخراج کردند. با استفاده از این رویکرد، آنها به امتیاز F1 بین ۸۹.۷ و ۹۹.۳ درصد رسیدند. با این حال، ذکر این نکته مهم است که آن‌ها الگوریتم خود را در مزارعی اعمال کردند که در آن نخل‌های جداگانه به خوبی بدون تاج‌های همپوشانی از هم جدا شده بودند و مرزهای کاشت قبلاً مشخص شده بود. لی و همکاران (۱۰) یک طبقه‌بندی شبکه عصبی کانولوشنال (CNN) را با استفاده از چندین هزار نمونه تصویر ۱۷ در ۱۷ پیکسلی با وضوح ۶۰ سانتیمتر آموزش دادند. شبکه قطعه‌های تصویر کوچک با (یا بدون) درخت در مرکز خود را به عنوان ورودی دریافت می‌کند و یک احتمال را برای هر قطعه تصویر حاوی یک درخت محاسبه می‌کند. پنجره ورودی کوچک روی کل تصویر جابجا می‌شود و در هر موقعیت احتمال مربوطه ثبت می‌شود. با این روش، یک نقشه احتمال درخت ایجاد شد. با استفاده از روش سرکوب غیرحداکثری، موقعیت درخت تعیین شد. این رویکرد نتایج بسیار خوبی به همراه داشت. آنها به امتیازات F1 بین ۹۲.۲ و ۹۷.۱ درصد رسیدند.

عرض جغرافیایی ۳۱.۲۵۰ درجه، حداقل طول جغرافیایی ۴۸.۶۱۰ درجه، حداکثر عرض جغرافیایی ۳۱.۲۶۰ درجه و حداکثر طول جغرافیایی ۴۸.۶۲۵ درجه می‌باشد. شکل ۱ موقعیت تصاویر ماهواره‌ای دانلود شده از گوگل ارث را نشان می‌دهد.

از گوگل ارث توسط نرم‌افزار Google Earth Images Downloader با بزرگنمایی ۲۱ دانلود شد. زمان اخذ تصاویر دانلود شده ۲۰۲۱/۲۶/۶ است. اندازه تصویر اصلی دانلود شده برابر ۶۳۷۷ × ۲۸۰۶۲ پیکسل است. موقعیت مکانی محدوده تصویر اصلی در سیستم مختصات جغرافیایی شامل حداقل



شکل ۱. منطقه مورد مطالعه و نمونه تصاویر ماهواره‌ای دانلود شده از گوگل ارث

Fig. 1- The study area and sample satellite images downloaded from Google Earth

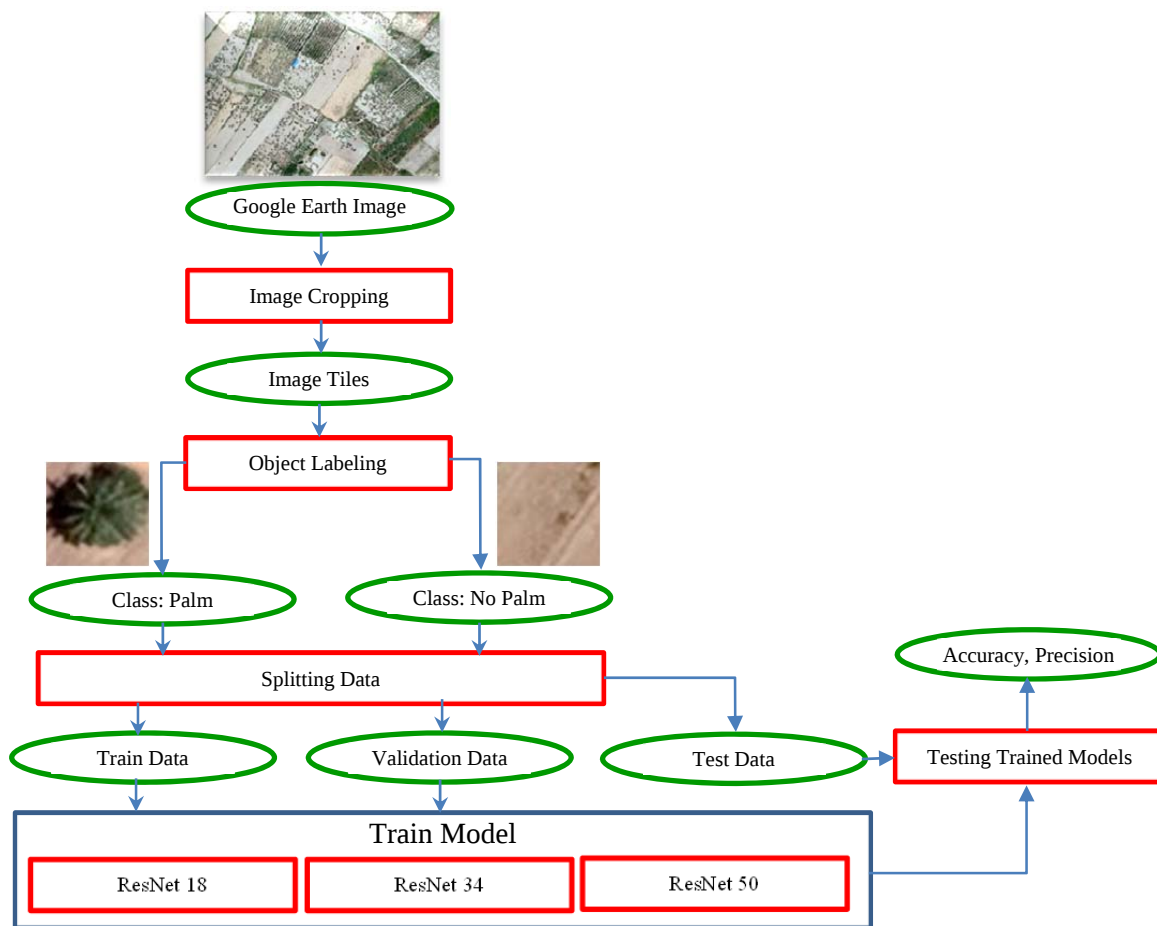
در این پژوهش با توجه به اینکه تصاویر ماهواره‌ای گوگل ارث بصورت مجانی در دسترس است، هدف تشخیص و استخراج درختان نخل از تصاویر ماهواره‌ای با قدرت تفکیک بالای مکانی گوگل ارث می‌باشد. این تصاویر در باند مرئی در دسترس هستند و فاقد باند مادون قرمز نزدیک هستند که چالش شناسایی درختان را بیشتر می‌کند. با توجه به اینکه روش‌های یادگیری عمیق ابزار مهمی در تشخیص اشیا در تصاویر هستند (۸)، در این پژوهش برای آموزش مدل از

روند کلی پژوهش

در این پژوهش با توجه به نقش مهم درختان نخل در مناطق جنوبی ایران، استخراج درختان نخل از تصاویر ماهواره‌ای مورد توجه قرار گرفته است. برای تعیین موقعیت درختان دو روش کلی عملیات میدانی با استفاده از ابزارهایی مانند گیرنده‌های تعیین موقعیت ماهواره‌ای و دوربین‌های توتال استیشن و عملیات هوایی بر مبنای سکوها ماهواره‌ای، هواپیماها و پرنده‌های هدایت‌پذیر از دور (پهپاد) وجود دارد.

موفقیت این مدل‌ها استفاده شده است (۱۴). شکل (۲) روند کلی این پژوهش را نشان می‌دهد.

روش‌های یادگیری عمیق باقیمانده با تعداد لایه‌های ۱۸ (ResNet18)، ۳۴ (ResNet34) و ۵۰ (ResNet50) با توجه به



شکل ۲. روند کلی پژوهش

Fig. 2- The general process of research

کشیده است. هرچه تعداد لایه‌ها بیشتر باشد، شبکه عمیق‌تر شده و زمان یادگیری بیشتر خواهد شد.

داده‌های مورد استفاده

در ابتدا تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالا از نرم‌افزار گوگل ارث دانلود شده است. تصویر گوگل ارث دانلود شده در باند مرئی (RGB) بوده و فاقد باندهای دیگر در تصاویر اصلی کوئیک برد می‌باشد که چالش تشخیص و استخراج درخت را بدلیل عدم امکان استفاده از سایر باندها بیشتر می‌کند. نمونه‌های آموزشی در دوکلاس شامل تصاویر حاوی درخت نخل و بدون درخت نخل (پس‌زمینه) تهیه

از آنجایی که درختان نخل وقتی از بالا دیده می‌شوند، شکل ساده‌ای مثل «ستاره» دارند، امکان تشخیص آن‌ها با استفاده از یک مدل یادگیری عمیق وجود دارد. پارامترهای مدل اجرا شده در این مطالعه شامل تعداد دوره‌های آموزشی برابر ۳۰ اپک می‌باشد. آموزش مدل‌های پیشنهادی بر روی داده‌های آموزشی با یک کامپیوتر شخصی لپ‌تاپ با پردازشگر مرکزی نسل هشت Core i7 و RAM برابر ۱۲ گیگابایت انجام شده است. زمان پردازش حدود ۱ ساعت و ۲۰ دقیقه برای مدل ResNet18 و حدود ۲ ساعت و ۱ دقیقه برای مدل ResNet34 و حدود ۳ ساعت و ۱۶ دقیقه برای مدل ResNet50 طول

کیفیت تصویر، رنگ، کنتراست یا حذف سایه تغییری نکرده است. در نهایت نمونه‌ها در دو کلاس با درخت و بدون درخت برچسب‌گذاری شده‌اند. از نمونه‌های انتخاب شده ۲۰ درصد به صورت تصادفی برای اعتبارسنجی در مدل استفاده شده‌اند. شکل (۳) تعدادی از نمونه‌های آموزشی در منطقه مورد مطالعه را نشان می‌دهد.

گردید. نمونه‌های آموزشی طوری از تصاویر اصلی بریده شده‌اند که درخت نخل در وسط تصویر باشد. با توجه به برآورد اندازه درختان نخل، بیش از ۱۶۰۰ نمونه حاوی درخت نخل برای یادگیری با ابعاد ۶۴ در ۶۴ پیکسل استخراج گردید. همچنین بیش از ۱۶۰۰ نمونه تصویری بدون درخت نخل به عنوان پس‌زمینه تهیه گردید. در نمونه‌های استخراج شده



نمونه تصاویر آموزشی بدون درخت (پس‌زمینه)

نمونه تصاویر آموزشی با درخت

شکل ۳. نمونه تصاویر آموزشی برای ورود به مدل یادگیری عمیق

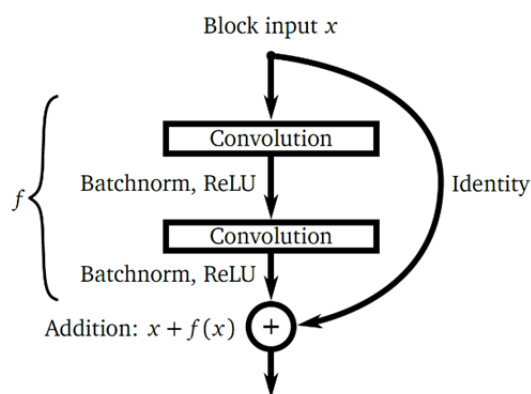
Fig. 3- Examples of educational images to enter the deep learning model

کانولوشن برای تشخیص اشیا در بینایی کامپیوتری و تجزیه و تحلیل تصویر ارائه شده‌اند (۱۶). شبکه‌های عصبی کانولوشنال یک کلاس از شبکه‌های عصبی مصنوعی عمیق را تشکیل می‌دهند که بر کانولوشن‌ها (عملیات خطی محلی) و به دنبال آن تبدیل‌های غیر خطی تکیه می‌کنند و نمایش داده‌های ورودی مختلف را ایجاد می‌کنند. لایه‌های کانولوشن به عنوان استخراج

شبکه‌های عصبی کانولوشن

در دهه اخیر، با پیشرفت تجهیزات سخت افزاری کامپیوتر و توسعه سریع فناوری هوش مصنوعی (AI)، شبکه‌های عصبی کانولوشن (CNN) در یادگیری عمیق، راه‌های جدیدی را برای تشخیص اشیا و استخراج ویژگی‌ها ارائه کرده‌اند. در تصاویر سنجش از دور بسیاری از معماری‌های شبکه‌های عصبی

تبدیل هویت آسان‌تر است، زیرا بهینه‌ساز کافی است که فقط وزن‌ها را در لایه‌های کانولوشن به صفر برساند.



شکل ۴: بلوک ساختمان مدل ResNet (۴)

Fig. 4- ResNet model building block (4)

نتایج مشاهداتی حاصل از مدل‌های شبکه‌های عصبی کانولوشن قبلی نشان می‌داد که افزودن لایه‌های بیشتر به یک شبکه موجود لزوماً عملکرد را افزایش نمی‌دهد. اما نقش کلیدی شبکه‌های باقیمانده ایجاد یک لایه یا بلوک بود که بتواند به راحتی تغییر هویت را یاد بگیرد (به عنوان مثال با تنظیم $f(x)=0$). چیدمان این بلوک‌ها در بالای شبکه موجود در حالت ایده‌آل نباید عملکرد آن را بدتر کند، زیرا بلوک‌های جدید می‌توانند تغییر هویت را یاد بگیرند و در نتیجه عملکرد مدل پایه را حفظ کنند. اما بلوک‌های جدید اضافه شده تعداد درجات آزادی و عمق شبکه را افزایش می‌دهند، بنابراین ظرفیت یادگیری ویژگی‌های پیچیده را افزایش می‌دهند. همچنین در صورتی که نمایش بهتری از داده‌ها وجود داشته باشد، شبکه جدید باید قادر به یادگیری آن باشد.

دو لایه کانولوشن در بلوک معمولاً مجموعه‌ای از پارامترهای مشابه مانند تعداد فیلترها، اندازه هسته و غیره را به اشتراک می‌گذارند. مجموعه‌ای از بلوک‌ها را که فرآیندهای یکسانی دارند، را می‌توان یک واحد نامید. برای بدست آوردن شبکه باقیمانده کامل، چندین واحد از این واحدها روی هم چیده می‌شوند و به دنبال آن یک لایه کاملاً متصل قرار می‌گیرد. بنابراین، شبکه باقیمانده را می‌توان با تعداد واحدها،

کننده ویژگی‌های تصاویر ورودی عمل می‌کنند. شبکه‌های عصبی کانولوشنال در ابتدا برای تشخیص حضور اشیا در تصاویر و تولید خروجی بدون هیچ بعد مکانی توسعه یافتند (۱۷). شبکه‌های عصبی کانولوشنال به طور گسترده در وظایف تشخیص اشیا به عنوان استخراج کننده ویژگی با هدف تغذیه آن اطلاعات در مدل‌های یادگیری ماشین برای طبقه بندی تصاویر استفاده می‌شوند (۱۵). در این پژوهش از شبکه‌های یادگیری عمیق باقیمانده استفاده شده است.

شبکه‌های یادگیری عمیق باقیمانده (ResNet)

معماری شبکه‌های یادگیری عمیق باقیمانده اولین بار در سال ۲۰۱۵ توسط هی و همکاران (۴) در تیم تحقیقاتی مایکروسافت طراحی گردید. در آن زمان، این معماری برنده چندین مسابقه شد و رکورد جدیدی را برای طبقه بندی مجموعه داده ImageNet به ثبت رساند. ایده اصلی معماری شبکه‌های یادگیری عمیق باقیمانده این است که بلوک‌های ساختمان آن برای «یادگیری توابع باقیمانده با ارجاع به ورودی‌های لایه، به جای یادگیری توابع غیر مرجع» طراحی شده‌اند. یک ورودی دلخواه x به یک لایه از یک شبکه عصبی و خروجی بهینه y را تصور کنید. در مورد یک تابع بدون مرجع، شبکه مستقیماً تابع f را که $y = f(x)$ را تولید می‌کند، تقریب می‌زند. منظور از «تابع باقیمانده» در این زمینه، این است که شبکه جدید اکنون $y = x + f(x)$ را تقریب می‌زند. بنابراین شبکه به جای اینکه یاد بگیرد که چگونه ورودی باید تبدیل شود تا خروجی مورد نظر را ایجاد کند، تفاوت بین ورودی و خروجی بهینه را یاد می‌گیرد. شکل ۴ نشان می‌دهد که چگونه این مفهوم بر روی لایه‌های یک شبکه کانولوشن اعمال می‌شود. بلوک اصلی یک ورودی دریافت می‌کند که در آن یک کانولوشن اعمال می‌شود، به دنبال آن نرمال سازی دسته‌ای و تابع فعال سازی ReLU اعمال می‌شود که مسیر اصلی نامیده می‌شود. همچنین این عمل برای بار دوم نیز تکرار می‌شود و ورودی مستقیم به نقشه ویژگی حاصل اضافه می‌شود که مسیر میانبر نامیده می‌شود. با این معماری یادگیری

کاهش شدید نمونه‌گیری ورودی می‌شود، اعمال می‌شود. سپس یک واحد با دو بلوک پایه اعمال می‌شود. واحد دوم نیز شامل دو بلوک است. اما در اینجا، کانولوشن اول مقدار پرش برابر دو گام است، بنابراین نقشه‌های ویژگی را کم‌نمونه می‌کند. شکل ۵ ساختار کلی شبکه‌های یادگیری عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه را نشان می‌دهد.

تعداد بلوک‌ها در هر واحد، فرآیندهای بلوک‌ها در واحدها و با تعداد گره‌ها در اتصال کامل مشخص نمود. لایه خروجی شبکه باقیمانده بر اساس تعداد کل لایه‌ها نامگذاری می‌شود. به عنوان مثال، ResNet18 دارای ۱۸ لایه و ResNet34 دارای ۳۴ لایه است. ابتدا، یک کانولوشن با پرش دو گام بر روی تصویر ورودی، و به دنبال آن تابع ادغام ماکزیمم، که منجر به



شکل ۵. ساختار مدل‌های یادگیری عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه

Fig. 5- Structure of residual deep learning models with 18, 34 and 50 layers

مجموعه داده آموزشی ارزیابی می‌کند. هدف از آموزش به حداقل رساندن خطای تابع ضرر است. آنتروپی متقاطع تابع softmax به صورت معادله ۱ تعریف می‌شود:

$$E = - \sum_{n=1}^N \sum_{k=1}^K t_{n,k} \log y_k \quad [1]$$

که t نشان دهنده بردار داده‌های آموزشی است، K نشان دهنده کلاس محتمل و N نشان دهنده تعداد کل نمونه‌ها است.

برآورد دقت یادگیری

مدل در هر دوره یادگیری با استفاده از توابع دقت و ضرر اعتبارسنجی می‌شود. تابع دقت نشان‌دهنده دقت مدل در طبقه-

آموزش و استنتاج شبکه‌های عصبی باقیمانده مانند هر طبقه‌بندی‌کننده دیگری آموزش داده می‌شود. شبکه تصاویر متعلق به کلاس‌های مشخصی را دریافت می‌کند. این داده‌ها به عنوان یک بردار بارز کدگذاری می‌شوند که در آن همه ورودی‌ها صفر هستند، به جز ورودی کلاس واقعی که روی یک تنظیم شده است. استنباط نیز به همین ترتیب انجام می‌شود. شبکه یک تصویر دریافت می‌کند و بردار طبقه‌بندی را با احتمالات برای هر کلاس خروجی می‌دهد.

برای ارزیابی عملکرد شبکه، باید یک تابع ضرر (loss) نیز تعریف شود. تابع ضرر، اثربخشی شبکه را در مدل‌سازی

هستند، اما شناسایی نشده‌اند. معیار F1- Score یک معادل بین این دو پارامتر ارزیابی مدل است و به عنوان میانگین هارمونیک آنها تعریف می‌شود (معادله ۴):

$$F1_{score} = 2 \times \frac{P \times R}{P + R} \quad [4]$$

نتایج

پس از اجرای مدل‌های پیشنهادی بر روی داده‌های آموزشی و اعتبارسنجی، روند یادگیری مدل بر مبنای تابع ضرر (شکل ۶) نشان می‌دهد که به دنبال افزایش تعداد دوره‌های آموزشی، مقادیر دقت افزایش و مقادیر تابع ضرر کاهش یافته است. این نتایج نشان می‌دهد که همه مدل‌ها بیش از حد به مجموعه داده‌ها برازش نشده‌اند و با موفقیت ویژگی‌های تصاویر قطعه شده را یاد گرفته‌اند.

جدول ۱ ماتریس ابهام را برای هر سه مدل شبکه‌های عمیق باقیمانده بر روی داده‌های اعتبارسنجی در فرآیند آموزش مدل را نشان می‌دهد.

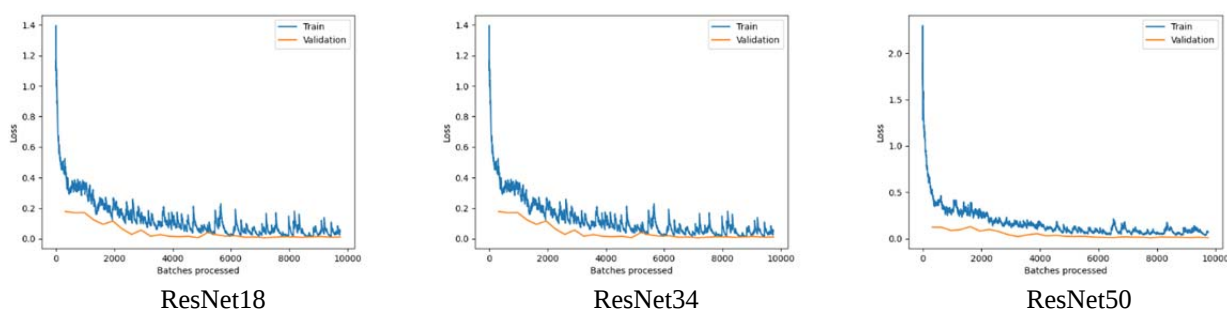
معیارهای ارزیابی برای داده‌های اعتبارسنجی در جدول ۲ نشان داده شده است.

بندی تصاویر اعتبارسنجی است، در حالی که تابع ضرر نشان‌دهنده عدم دقت پیش‌بینی توسط مدل است. اگر یادگیری مدل موفقیت‌آمیز باشد، مقدار تابع ضرر، کم و مقدار تابع دقت زیاد است. با این حال، اگر مقدار تابع ضرر در طول یادگیری زیاد شود، نشان‌دهنده بیش برازش است. معیار دقت (Precision) و معیار بازیابی یا حساسیت (Recall) رایج‌ترین شاخص‌های ارزیابی برای تشخیص اهداف در یادگیری عمیق هستند. هر چه مقدار این معیارها بالاتر باشد، توانایی پیش‌بینی قوی‌تر است. فرمول‌های محاسبه دقت تشخیص و نرخ بازیابی به ترتیب در معادلات ۲ و ۳ نشان داده شده است:

$$P = \frac{TP}{TP + FP} \quad [2]$$

$$R = \frac{TP}{TP + FN} \quad [3]$$

TP، TN و FP بر اساس ماتریس ابهام محاسبه می‌شوند که تعداد تشخیص‌های مختلف را نشان می‌دهند. TP (مثبت واقعی) تعداد اهدافی را نشان می‌دهد که به درستی شناسایی شده‌اند و FP (مثبت کاذب) تعداد اهدافی را نشان می‌دهد که هدف نبوده‌اند، اما به اشتباه به عنوان هدف شناسایی شده‌اند. FN (منفی کاذب) تعداد اهدافی را نشان می‌دهد که هدف



شکل ۶. تابع ضرر بر اساس دوره‌های یادگیری (دسته‌های پردازش شده)

Fig. 6- Loss function based on learning periods (batches processed)

جدول ۱. ماتریس ابهام برای مدل‌های شبکه عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه برای داده‌های اعتبارسنجی

Table 1. Confusion matrix for residual deep network models with 18, 34, and 50 layers for validation data

		ResNet18		ResNet34		ResNet50	
Actual	Palm	327	2	328	1	329	0
	No Palm	0	321	1	320	0	321
	Predicted						

جدول ۲. معیارهای ارزیابی دقت برای مدل‌های شبکه عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه برای داده‌های اعتبارسنجی

Table 2. Accuracy evaluation criteria for residual deep network models with 18, 34 and 50 layers for validation data

نام مدل	معیار دقت (Precision)	معیار بازیابی (Recall)	معیار F1 Score
ResNet18	1	0.994	0.997
ResNet34	0.997	0.997	0.997
ResNet50	1	1	1

میانگین دقت داده‌های اعتبارسنجی مدل‌ها بالای ۹۹ درصد می‌باشد که نشان دهنده دقت بالای مدل‌ها در تشخیص درخت نخل در شرایط این پژوهش می‌باشد. برای ارزیابی مدل‌ها، ۵۰۰ نمونه به عنوان داده‌های تست که در روند یادگیری مدل‌ها شرکت نکرده‌اند، تهیه گردید. مدل‌ها بر روی تصاویر داده‌های تست اجرا گردید. جدول ۳ ماتریس ابهام نتایج را نشان می‌دهد.

معیارهای ارزیابی برای داده‌های تست در جدول ۴ نشان داده شده است.

جدول ۳. ماتریس ابهام برای مدل‌های شبکه عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه برای داده‌های تست

Table 3. Confusion matrix for residual deep network models with 18, 34, and 50 layers for test data

		ResNet18		ResNet34		ResNet50	
Act ual	Palm	250	0	250	0	250	0
	No Palm	4	246	6	244	10	240
		Predicted					

جدول ۴. معیارهای ارزیابی دقت برای مدل‌های شبکه عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه برای داده‌های تست

Table 4. Accuracy evaluation criteria for residual deep network models with 18, 34 and 50 layers for test data

نام مدل	معیار دقت (Precision)	معیار بازیابی (Recall)	معیار F1 Score
ResNet18	0.984	1	0.992
ResNet34	0.976	1	0.988
ResNet50	0.961	1	0.980

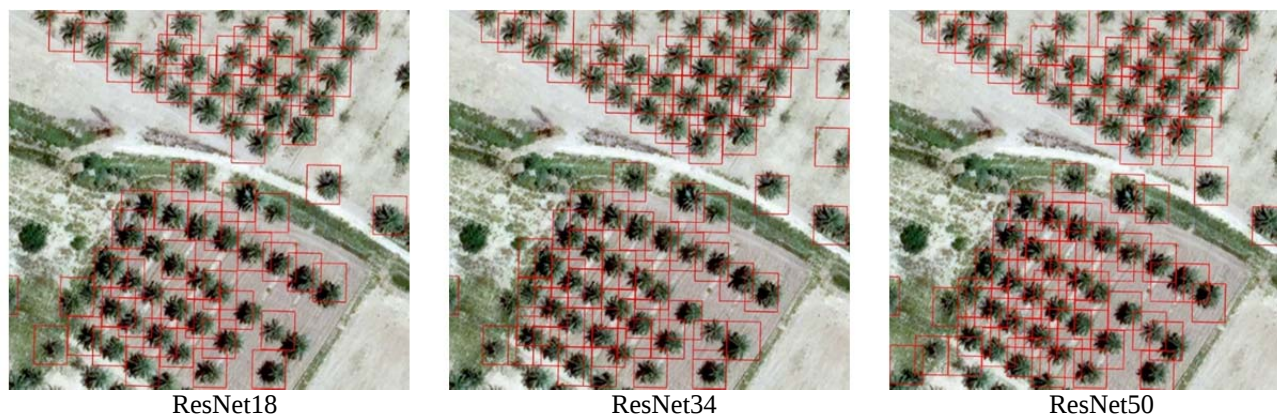
سپس تصاویر به مدل‌های آموزش دیده جهت تشخیص کلاس آن‌ها وارد شده است، مدل‌ها برای هر قطعه نوع کلاس با و بدون درخت نخل را با ضریب اطمینان بین ۰ تا ۱ محاسبه می‌کنند. موقعیت هر قطعه با یک کادر مستطیلی بر روی تصویر ماهواره‌ای ژئورفرنس شده، نشان داده شده است. از بین قطعاتی که دارای همپوشانی زیادی هستند، قطعاتی که دارای ضریب اطمینان بالاتری هستند انتخاب و بقیه حذف می‌شوند. برای اینکار از روش سرکوب غیرحداکثری (NMS) برای حذف قطعه‌های تکراری که روی همان شیء همپوشانی دارند، اعمال می‌شود (۱۸). پارامتر امکان همپوشانی روش سرکوب

ارزیابی نتایج بر روی داده‌های تست نشان می‌دهد که با افزایش لایه‌های مدل لزوماً دقت افزایش نیافته است. ضریب اطمینان نمونه‌های تشخیصی اشتباه کمتر از ۰.۸ بوده است. امتیاز اطمینان عددی بین ۰ تا ۱ است که درجه تعلق هر نمونه به کلاس را نشان می‌دهد.

برای تشخیص و استخراج درخت از تصاویر ماهواره‌ای گوگل ارث، تصویر ماهواره‌ای ورودی در قطعات ۶۴ در ۶۴ پیکسل با گام یک چهارم اندازه هر قطعه یعنی ۱۶ پیکسل بریده شده است. بدیهی است کاهش گام می‌تواند دقت نهایی را افزایش دهد، اما زمان پردازش مقرون به صرفه نخواهد بود.

اطمینان بالای ۰.۹ بوده‌اند، برای نمایش نهایی موقعیت درختان نخل استفاده شده است.

غیرحداکثری برابر ۲۵ درصد اعمال شده است. شکل ۷ نتیجه اعمال مدل‌های پیشنهادی بر روی یک نمونه تصویر ماهواره‌ای گوگل ارث را نشان می‌دهد. فقط نتایجی که دارای ضریب



شکل ۷. نمونه‌ای از نتیجه استخراج مکان درختان نخل بر مبنای مدل‌های یادگیری عمیق باقیمانده با ۱۸، ۳۴ و ۵۰ لایه

Fig. 7- Example of the result of extracting palm trees based on the residual deep learning models with 18, 34, and 50 layers

به‌طور بالقوه، چه مقدار برای احتراق در صورت آتش‌سوزی در دسترس است، می‌تواند مفید باشد. با توجه به رشد سریع شهرها، شناسایی درختان در محدوده شهری برای حفظ پوشش درختی شهری ضروری است. به این دلایل، تلاش قابل‌توجهی برای یافتن راه‌های سریع و ارزان برای به دست آوردن چنین داده‌هایی صورت گرفته است. مدیریت هوشمند پوشش درختی مانند باغ‌ها و پارک‌ها به داده‌های مربوط به مکان و ویژگی‌های تک‌درختان در منطقه پوشش درختی نیاز دارد، زیرا این داده‌ها نقش مهمی در بررسی دقیق منطقه کاشت، پیشگیری و کنترل بیماری‌ها و آفات، و پیش‌بینی عملکرد میوه دارند. به‌طور سنتی، از تحقیقات میدانی و عملیات زمینی برای جمع‌آوری این داده‌ها از جمله مکان، توزیع مکانی و غیره در باغ‌ها استفاده می‌شود. این تحقیقات زمان‌بر، نیاز به کار گسترده و پرهزینه هستند، اما نمی‌توانند نیازهای مدیریت هوشمند مناطق دارای پوشش درختی را برآورده کنند. توسعه روشی سریع، ارزان و دقیق برای به دست آوردن این داده‌ها ضروری است. تصاویر درختان در یک باغ نسبتاً بزرگ را می‌توان با تصویربرداری ماهواره‌ای یا هوایی اخذ نمود. اما برای مناطق

زمان پردازش برای مدل ۵۰ لایه تقریباً دو برابر مدل ۱۸ لایه بوده و زمان پردازش مدل ۳۴ لایه حدود یک و نیم برابر مدل ۱۸ لایه بوده است. بررسی نتایج بر روی تصاویر مختلف نشان می‌دهد که هر سه مدل امکان تشخیص درخت نخل از تصاویر ماهواره‌ای گوگل ارث را با دقت مناسب بالای ۹۰ درصد دارند. اما مدل ۳۴ لایه با توجه به زمان پردازش و دقت بالاتر پیشنهاد می‌شود. مدل ۱۸ لایه نسبت به مدل ۳۴ لایه دقت کمتری دارد. بطور کلی مدل ۳۴ لایه برای شمارش درختان نتایج بهتری نشان می‌دهد.

بحث و نتیجه‌گیری

به دست آوردن مکان، نوع و اندازه درختان در یک منطقه خاص براساس نوع کاربرد ضروری است. به‌عنوان مثال، یک باغدار ممکن است نیاز به شمارش درختان خود داشته باشند تا برای فعالیت‌های برداشت خود برنامه‌ریزی کند. برای مدیریت بهینه جنگل‌ها اغلب به اطلاعاتی در مورد سن، تراکم و انواع درختان در مناطق جنگلی نیاز است. همچنین دانستن اینکه چه مقدار چوب برای برداشت در یک منطقه خاص و

خیلی بزرگ نیاز به تصاویر ماهواره‌ای با قدرت تفکیک بالای مکانی وجود دارد. در مورد استفاده از تصاویر ماهواره‌ای، هوای ابری در ابتدا چالش بزرگی است و به دلیل کیفیت پایین تصاویر گرفته شده، تشخیص درختان را دشوارتر می‌کند. این مطالعه رویکردی را برای تشخیص درختان نخل و نقشه‌برداری توزیع مکانی آن‌ها با ادغام یادگیری عمیق با تصاویر ماهواره‌ای پیشنهاد می‌کند. یک مجموعه داده تصویری از نمونه‌های یادگیری عمیق درختان نخل از طریق تفسیر بصری بر اساس تصاویر درختان نخل از تصاویر ماهواره‌ای با قدرت تفکیک بالای مکانی داندلود شده از گوگل ارث ساخته شد. در این پژوهش برای تشخیص اشیا از روش‌های یادگیری عمیق باقیمانده، استفاده شده است. در این پژوهش بدلیل عدم دسترسی به تصاویر ماهواره‌ای با قدرت تفکیک بالای مکانی مانند کوئیک برد و ژئوآی از تصاویر ماهواره‌ای گوگل ارث استفاده شده است. چالش اصلی این تصاویر عدم دسترسی به همه باندهای تصاویر ماهواره‌ای است و فقط تصاویر باند مرئی قابل داندلود است. در این پژوهش تصویر مناسب از بین تصاویر در دسترس انتخاب و با نرم‌افزارهای داندلود تصاویر گوگل در تاریخ مورد نظر با بزرگنمایی بالا داندلود گردید. سپس نمونه تصاویر آموزشی در دو کلاس با درخت نخل و بدون درخت نخل با ابعاد ۶۴ در ۶۴ پیکسل با بیش از ۱۵۰۰ نمونه برای هر کلاس استخراج گردید. از نمونه‌های آموزشی ۸۰ درصد برای آموزش مدل‌های پیشنهادی و ۲۰ درصد برای اعتبارسنجی استفاده شده است. سپس مدل‌های شبکه یادگیری عمیق باقیمانده با تعداد لایه ۱۸، ۳۴ و ۵۰ با نمونه‌های آموزشی، آموزش داده شد. زمان لازم برای یادگیری مدل برای مدل باقیمانده ۵۰ لایه تقریباً سه برابر مدل باقیمانده ۱۸ لایه و برای مدل ۳۴ لایه نیز تقریباً دو برابر ۱۸ لایه باشد. تعداد دوره آموزشی برای هر سه مدل، ۳۰ اپک در نظر گرفته شده است. با توجه به رفتار تابع ضرر بر روی نمونه‌های آموزشی و اعتبارسنجی، هر سه مدل بخوبی آموزش دیده‌اند. مقدار پارامترهای ارزیابی مدل شامل معیار دقت، معیار بازیابی و معیار F1Score برای هر سه مدل بیش از ۹۹ درصد می‌باشد

که دارای دقت بهتری از روش لی و همکاران (۱۰) که به دقت بین ۹۲.۱ تا ۹۷.۱ برای معیار F1Score دست یافتند، است. در میان مدل‌های پیشنهادی، مدل شبکه عصبی باقیمانده ۵۰ لایه نتایج بهتری بر روی داده‌های اعتبارسنجی نشان داده است. برای ارزیابی مدل‌ها با داده‌های تست، بیش از ۵۰۰ نمونه تصویری که در فرآیند یادگیری شرکت نکرده‌اند، استخراج گردید. نتایج اجرای مدل‌های شبکه عصبی پیشنهادی بر روی داده‌های تست نشان می‌دهد که مدل شبکه عصبی باقیمانده ۱۸ لایه نسبت به مدل ۳۴ و ۵۰ لایه نتایج بهتری دارد. هر چند معیارهای ارزیابی دقت برای هر سه مدل بیش از ۰.۹۶ درصد می‌باشد که نشان‌دهنده توانایی هر سه مدل در تشخیص درخت نخل از تصاویر ماهواره‌ای گوگل ارث است. برای تشخیص و استخراج موقعیت درختان نخل از تصاویر ماهواره‌ای یک پنجره با ابعاد ۶۴ در ۶۴ پیکسل بر روی تصویر ورودی حرکت داده شده است. برای کاهش زمان پردازش گام حرکت برابر یک چهارم ابعاد پنجره یعنی ۱۶ پیکسل در نظر گرفته شد. با ورود قطعات تصویری به مدل‌های پیشنهادی تعلق هر قطعه به یکی از دو کلاس ورودی با محاسبه و تعریف ضریب اطمینان آن بدست آمده است. برای تشخیص درخت نخل فقط قطعه‌های تصویری با ضریب اطمینان بالای ۹۰ درصد انتخاب شد. چون قطعات با هم همپوشانی دارند از روش سرکوب غیرحداکثری برای حذف قطعات تکراری با امکان پوشش ۲۵ درصدی قطعات استفاده گردید. نتایج نشان می‌دهد که هر سه مدل با دقت بالای ۹۰ درصد امکان شمارش و تعیین موقعیت درختان نخل را دارند. در مدل شبکه عصبی باقیمانده ۱۸ لایه تعداد کمتری از درختان نسبت به مدل‌های ۳۴ و ۵۰ لایه تشخیص داده می‌شود. مدل ۵۰ لایه چون حساسیت بیشتری دارد، قطعات نهایی دارای همپوشانی بیشتری هستند و در نتیجه تعداد درختان از حالت واقعی بیشتر خواهد بود. با توجه به زمان پردازش و بررسی دقت نتایج در تهیه نقشه موقعیت درختان نخل مدل شبکه یادگیری عمیق باقیمانده ۳۴ لایه پیشنهاد می‌شود. در صورت استفاده از تصاویر هوایی و همچنین افزایش تعداد نمونه‌های آموزشی، دقت مدل افزایش

تقدیر و تشکر بدین وسیله از حمایت مالی معاونت پژوهش و فناوری دانشگاه شهید چمران اهواز در قالب پژوهانه شماره (SCU.EG1401.26151) در انجام این تحقیق تشکر و قدردانی می‌گردد.

خواهد داشت. همچنین می‌توان با ادغام تصاویر چندطیفی کارکرد مدل را بهبود بخشید.

منابع مورد استفاده

1. Ahl, R., Hogland, J. and Brown, S., 2019. A comparison of standard modeling techniques using digital aerial imagery with National Elevation Datasets and airborne LiDAR to predict size and density forest metrics in the Sapphire Mountains MT, USA. *ISPRS international journal of geo-information*, 8(1), p.24.
2. Bulatov, D., Wayand, I. and Schilling, H., 2016. Automatic tree-crown detection in challenging scenarios. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, pp.575-582.
3. Chong, K.L., Kanniah, K.D., Pohl, C. and Tan, K.P., 2017. A review of remote sensing applications for oil palm studies. *Geo-spatial Information Science*, 20(2), pp.184-200.
4. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
5. Hu, J., Zhang, Y., Zhao, D., Yang, G., Chen, F., Zhou, C. and Chen, W., 2022. A robust deep learning approach for the quantitative characterization and clustering of peach tree crowns based on UAV images. *IEEE Transactions on Geoscience and Remote Sensing*, 60, pp.1-13.
6. Hu, R., Bournez, E., Cheng, S., Jiang, H., Nerry, F., Landes, T., Saudreau, M., Kastendeuch, P., Najjar, G., Colin, J. and Yan, G., 2018. Estimating the leaf area of an individual tree in urban areas using terrestrial laser scanner and path length distribution model. *ISPRS journal of photogrammetry and remote sensing*, 144, pp.357-368.
7. Iglovikov, V., Mushinskiy, S. and Osin, V., 2017. Satellite imagery feature detection using deep convolutional neural network: A kaggle competition. *arXiv preprint arXiv:1706.06169*.
8. Jintasuttisak, T., Edirisinghe, E. and Elbattay, A., 2022. Deep neural network based date palm tree detection in drone imagery. *Computers and Electronics in Agriculture*, 192, p.106560.
9. Li, W., Dong, R., Fu, H. and Yu, L., 2018. Large-scale oil palm tree detection from high-resolution satellite images using two-stage convolutional neural networks. *Remote Sensing*, 11(1), p.11.
10. Li, W., Fu, H. and Yu, L., 2017, July. Deep convolutional neural network based large-scale oil palm tree detection for high-resolution remote sensing images. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 846-849). IEEE.
11. Lin, J., Kroll, C.N., Nowak, D.J. and Greenfield, E.J., 2019. A review of urban forest modeling: Implications for management and future research. *Urban forestry & urban greening*, 43, p.126366.
12. Maschler, J., Atzberger, C. and Immitzer, M., 2018. Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. *Remote Sensing*, 10(8), p.1218.
13. McKinnon T, Hoff P. 2017. Comparing RGB-Based Vegetation Indices With NDVI For Drone Based Agricultural Sensing. *AgrobotixLLC AGBX021-17*, 1-8.
14. Natesan, S., Armenakis, C. and Vepakomma, U., 2019. Resnet-based tree species classification using uav images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, pp.475-481.
15. Olivares R J L. 2019. PALM TREE IMAGE CLASSIFICATION A convolutional and machine learning approach.
16. Osco, L.P., Junior, J.M., Ramos, A.P.M., de Castro Jorge, L.A., Fathollahi, S.N., de Andrade Silva, J., Matsubara, E.T., Pistori, H., Gonçalves, W.N. and Li, J., 2021. A review on deep learning in UAV remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 102, p.102456.
17. Ponti, M.A., Ribeiro, L.S.F., Nazare, T.S., Bui, T. and Collomosse, J., 2017, October. Everything you wanted to know about deep learning for computer vision but were afraid to ask. In *2017 30th SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T)* (pp. 17-41). IEEE.
18. Ren, S., He, K., Girshick, R. and Sun, J., 2016. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on pattern analysis and machine intelligence*, 39(6), pp.1137-1149.
19. Srestasathiern, P. and Rakwatin, P., 2014. Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sensing*, 6(10), pp.9749-9774.

20. Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O. and Peña, J.M., 2015. High-throughput 3-D monitoring of agricultural-tree plantations with unmanned aerial vehicle (UAV) technology. *PloS one*, 10(6), p.e0130479.
21. Tuominen, S., Näsi, R., Honkavaara, E., Balazs, A., Hakala, T., Viljanen, N., Pölönen, I., Saari, H. and Ojanen, H., 2018. Assessment of classifiers and remote sensing features of hyperspectral imagery and stereo-photogrammetric point clouds for recognition of tree species in a forest area of high species diversity. *Remote Sensing*, 10(5), p.714.
22. Wang, X., Wang, Y., Zhou, C., Yin, L. and Feng, X., 2021. Urban forest monitoring based on multiple features at the single tree scale by UAV. *Urban forestry & urban greening*, 58, p.126958.
23. Weinstein, B.G., Marconi, S., Bohlman, S., Zare, A. and White, E., 2019. Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks. *Remote Sensing*, 11(11), p.1309.
24. Yao, L., Liu, T., Qin, J., Lu, N. and Zhou, C., 2021. Tree counting with high spatial-resolution satellite imagery based on deep neural networks. *Ecological Indicators*, 125, p.107591.
- Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F. and Fraundorfer, F., 2017. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*, 5(4), pp.8-36



Recognition and extraction of palm trees from satellite images with high spatial resolution of Google Earth based on residual deep learning networks

Mostafa Kabolizadeh^{1*}, Kazem Rangzan², Mohammad Abbasi³

Received: 2022-09-05 / Accepted: 2022-11-27 / Published: 2025-02-28

Abstract

Identifying the location of trees is the first step to managing green spaces, gardens, and forests. The preparation of the location map of the trees can be done by ground mapping operations, which require a lot of money and time, or by using aerial or satellite images. In this research, satellite images with the high spatial resolution of Google Earth have been used to detect and extract palm trees, considering the role and importance of palm trees in the southern regions of Iran, but automatic recognition of trees from satellite images is a challenge. In this regard, deep learning methods are considered an important solution for extracting objects from images. In this research, residual deep learning methods with the number of layers 18, 34, and 50 have been used. First, more than 3000 image samples were cut into two classes containing palm trees and without palm trees with dimensions of 64 x 64 pixels, then the models were trained with 80% samples for learning and 20% for validation with 30 epochs.

The training accuracy of the models has been above 99%. The trained model was implemented on 500 test samples and the evaluation results of all three models show that the precision is more than 0.96, the recall is equal to 1, and the F1Score is more than 0.98. Running the models on Google Earth satellite images by moving the 64 x 64-pixel window with a step of 16 pixels and applying the non-maximum suppression method shows that the satellite images of the Google Earth system can be used to prepare a map of palm trees. Considering the processing time and the possibility of better estimating the number and extracting the position of palm trees, the residual deep learning model with 34 layers is suggested.

Statement of the Problem: Spatial data and counting of palm trees are very important for surveying the planting area, predicting fruit yield, management, and intelligent planning of orchards. This requires tools and technology that provide reliable information with acceptable accuracy. Using land Surveying methods using mapping cameras or satellite positioning receivers, although they have high accuracy, they have many challenges in addition to time and high cost. In this regard, using aerial and satellite images can be a better solution considering time and cost. Studies in the detection and extraction of trees for forestry and precision agriculture, such as palm trees, have been growing in recent decades, especially using remote sensing. Object recognition models in deep learning, which are widely used in computer vision, can provide an opportunity to recognize palm trees, which is

Mostafa Kabolizadeh¹(✉), Kazem Rangzan², Mohammad Abbasi³

1. Associate Professor, Department of Remote Sensing and Geographical Information Systems, Faculty of Earth Sciences, Shahid Chamran University of Ahvaz, Ahvaz, Iran

2. Professor, Department of Remote Sensing and Geographical Information Systems, Faculty of Earth Sciences, Shahid Chamran University of Ahvaz, Ahvaz, Iran

3. Phd. Student of Remote Sensing and Geographical Information Systems, Faculty of Earth Sciences, University of Shahid Chamran Ahvaz, Ahvaz, Iran

DOI: 10.30495/girs.2022.697278

e-mail: Amir.Ghayebi@shahroodut.ac.ir

necessary for rapid data acquisition and reduction of human operation errors. Automatic detection of trees in optical images is still a challenge. Rarely, a study has been done to identify a tree using free and available internet resources such as Google Earth. To obtain tree cover information cost-effectively and efficiently, in this research, deep learning models based on residual networks are presented to recognize palm trees from Google Earth images.

Purpose: The main goal of this research is to identify palm trees from high-resolution satellite images of Google Earth so that the location, number, and distribution of trees can be estimated. By achieving these goals, a lot of data will be obtained to update palm trees. It also provides a new insight that by using methods based on deep learning, large-scale information can be obtained from available satellite images.

Methodology: At first, satellite images with high spatial resolution were downloaded from Google Earth software. The downloaded Google Earth image is in the visible band (RGB) and lacks other bands in the original Quickbird images. i.e., increases the challenge of tree detection and extraction due to the impossibility of using other bands. Learning samples were prepared in two classes, including pictures containing palm trees and without palm trees. The training examples are cut from the original images so that the palm tree is in the middle of the image. According to the estimation of the size of palm trees, more than 1600 samples containing palm trees were extracted for learning with dimensions of 64 x 64 pixels. Also, more than 1600 image samples without palm trees were prepared. In the extracted samples, image quality, color, contrast, or shadow removal have not changed. Finally, the samples are labeled in two classes with trees and without trees. 20% of the selected samples were randomly used for validation in the model. Finally, the remaining deep learning networks were trained with 18, 34, and 50 layers and by moving the search window in the image, the location of the palm trees was calculated with the confidence factor for each image piece. Non-maximum suppression method has been used to repeatedly remove overlapping parts, and a location map of palm trees has been prepared.

Results and discussion: In this research, residual deep learning methods with the number of (ResNet18) 18, (ResNet34) 34, and (ResNet50) 50 layers have been used to train the model. The model parameters implemented in this study include the number of training epochs equal to 30. The training of the proposed models has been done on the training data with a personal laptop computer with the eighth generation Core i7 central processor and 12 GB of

RAM. The processing time took about 1 hour and 20 minutes for the ResNet18 model, about 2 hours and 1 minute for the ResNet34 model, and about 3 hours and 16 minutes for the ResNet50 model. The higher the number of layers, the deeper the network and the longer the learning time. The average accuracy of the validation data of the models is above 99%, which shows the high accuracy of the models in detecting the palm tree in the conditions of this research. To evaluate the model, 500 samples were prepared as test data that did not participate in the model learning process. The evaluation of the results on the test data shows that the accuracy has not necessarily increased with the increase of model layers. The reliability coefficient of wrongly diagnosed samples was less than 0.8. To detect and extract the tree from Google Earth satellite images, the input satellite image is cut into 64 x 64-pixel pieces with a step of a quarter of the size of each piece, i.e. 16 pixels. Reducing the step can increase the final accuracy, but the processing time will not be economical. Then the images are entered into the trained models to recognize their class, the models calculate the type of class with and without palm trees for each plot with a confidence factor between 0 and 1. The position of each piece is shown with a rectangular box on the georeferenced satellite image. Among the parts that have a lot of overlap, the parts that have a higher confidence factor are selected by the non-maximum suppression method, and the rest are removed. Examining the results on different images shows that all three models can recognize palm trees from Google Earth satellite images with a suitable accuracy of over 90%. However, the 34-layer model is recommended due to the processing time and higher accuracy. The 18-layer model is less accurate than the 34-layer model. In general, the 34-layer model shows better results for counting trees.

Conclusion: In this research, the time required to learn the model for the 50-layer residual model was almost three times that of the 18-layer residual model, and for the 34-layer model, it was almost twice as much as 18 layers. The number of Learning periods for all three models is 30 epochs. According to the behavior of the loss function on the training and validation samples, all three models are well-trained. The value of model evaluation parameters including precision, recall, and F1Score for all three models is more than 99%. The residual neural network model of 50 layers has shown better results. To evaluate the models with test data, more than 500 image samples that did not participate in the learning process were extracted. The results of the implementation of the proposed neural network models on the test data show that the remaining 18-layer neural network model has better results than the 34 and 50-layer models. The accuracy evaluation criteria for all three models are more than 96%, which indicates the

ability of all three models to recognize palm trees from Google Earth satellite images. To detect and extract the position of palm trees from satellite images, a window with dimensions of 64 x 64 pixels has been moved to the input image. The movement step equal to a quarter of the window dimensions, i.e., 16 pixels, was considered to reduce the processing time. By entering image parts into the proposed models, the belonging of each part to one of the two input classes has been obtained by calculating and defining its reliability coefficient. To detect the palm tree, only pieces with a confidence factor of 90% were selected. Because the parts overlap, the non-maximum suppression method was used to remove duplicate parts with the possibility of covering 25% of the parts. The results show that all three models can count and determine the position of palm trees with an accuracy of over 90%. In the residual

neural network model of 18 layers, more trees are not detected than in the 34 and 50-layer models. Because the 50-layer model is more sensitive, the final parts have more overlap, and as a result, the number of trees will be more than in the 18-layer model. Considering the processing time and checking the accuracy of the results in preparing the position map of palm trees, the remaining 34-layer deep learning network model is suggested. If aerial images are used and the number of training samples is increased, the accuracy of the model will increase. It is also possible to improve the performance of the model by integrating multispectral images.

Keywords: Machine learning, Automatic recognition, Features extraction, Convolutional Neural Network.

Please cite this article as: Kabolizadeh M, Rangzan K, Abbasi M. Recognition and extraction of palm trees from satellite images with high spatial resolution of Google Earth based on residual deep learning networks. *Journal of RS and GIS for Natural Resources*, 15(4): 48-67