نشریه علمی مهندسی فناوری اطلاعات مکانی

سال دهم، شماره دوم، تابستان ۱۴۰۱

Vol.10, No.2, Summer 2022

۸۹ -۱۰۴

مقاله پژوهشی

DOR: 20.1001.1.20089635.1401.10.2.3.9

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]

[DOR: 20.1001.1.20089635.1401.10.2.3.9]

[DOI: 10.52547/jgit.10.2.89]

نقشه‌برداری درختان در محیط‌های شهری با استفاده از شبکه عصبی عمیق کانولوشنی و فیلترهای مورفولوژیکی در پردازش تصاویر هوایی با قدرت تفکیک مکانی بالا

مسلم درویشی[۱]، رضا شاه حسینی[۲*]

۱- دانشجوی دکتری، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشکدگان فنی، دانشگاه تهران

۲- استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشکدگان فنی، دانشگاه تهران

تاریخ دریافت مقاله: ۱۴۰۰/۰۶/۳۱ تاریخ پذیرش مقاله: ۱۴۰۱/۰۶/۳۰

چکیده

فضای سبز و درختان در بهبود کیفیت زندگی انسان نقشی انکارناپذیر دارند. درختان شهری علاوه بر کمک به تعادل گرمایی و تصفیه هوا، از نظر زیبایی بصری نیز حائز اهمیت هستند. استفاده از سکوهای بدون سرنشین در تصویربرداری هوایی، جهت دستیابی به تصاویر با قدرت تفکیک مکانی بالا، روشی متداول محسوب می‌شود. در پردازش داده‌های حجیم حاصل از این نوع تصویربرداری و چالش‌های خاص آن، روش‌های یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی بسیار کارآمد هستند. در این تحقیق یک معماری برای شبکه عصبی کانولوشنی بر پایه شبکه‌های رمزگذار - رمزگشا پیشنهاد شده است که نتایج آن بر روی داده‌های مرجع ارائه شده توسط انجمن بین‌المللی فتوگرامتری و سنجش‌ازدور از شهر پوتسدام کشور آلمان جهت آشکارسازی درختان شهری، دقت کلی ۹۶٫۱۰٪ و امتیاز *F1* برابر با ۸۰٫۷۲٪ را ارائه می‌دهد. همچنین موضوع جایگزینی باند مادون‌قرمز به‌جای باند آبی در تصاویر مرئی مورد بررسی قرار گرفت که موجب بهبود در نتایج طبقه‌بندی شد؛ البته این بهبود دقت بسیار کمتر از حد انتظار بود که دلیل اصلی آن را می‌توان مرتبط با زمان تصویربرداری دانست؛ زیرا تصاویر مورد استفاده در این تحقیق مربوط به فصل زمستان است که درختان منطقه عموماً عاری از هرگونه سبزینگی هستند. بحث دیگر در این تحقیق ایجاد تصویر تفاضلی از نقشه واقعیت زمینی و نقشه تخمین حاصل از الگوریتم پیشنهادی است که با بررسی نقشه تفاضلی، اعمال فیلتر مورفولوژی بر نتایج الگوریتم پیشنهادی موجب افزایش دقت در طبقه‌بندی نهایی شده است. موضوع آخر استفاده از پارامترهای شبکه حاصل از آموزش توسط ترکیب باندی قرمز، سبز و آبی به‌عنوان پارامترهای اولیه برای شبکه با ورودی ترکیب باندی مادون‌قرمز نزدیک، قرمز و سبز است؛ در این حالت زمان دستیابی به بهترین عملکرد شبکه در حدود ۸۳٪ کاهش یافته و این پیش آموزش پارامترهای شبکه عصبی، موجب همگرایی سریع شبکه موردنظر شده است.

کلید واژه‌ها : یادگیری عمیق، سنجش‌ازدور، شبکه‌های عصبی کانولوشنی، تصاویر هوایی با قدرت تفکیک مکانی بالا.

* نویسنده مکاتبه کننده: گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، دانشکدگان فنی، دانشگاه تهران، تهران، ایران.

تلفن: ۰۲۱۶۱۱۱۴۵۲۷

Email: rshahosseini@ut.ac.ir

۱- مقدمه

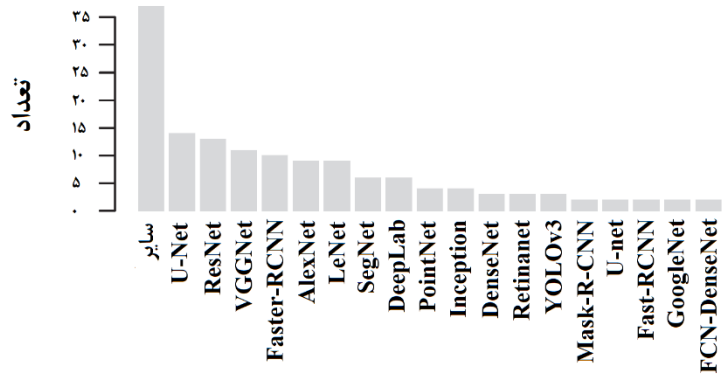
استفاده از سکوهای هوایی بدون سرنشین با توجه به انعطاف بالای آنها در رابطه با زمان تصویربرداری و همچنین قدرت تفکیک بسیار مناسب همواره موردتوجه محققان از جمله شیفر و همکاران (۲۰۲۰) بوده است [۱].

یادگیری عمیق، اخیراً در مطالعات سنجش‌ازدور گیاهان از جمله در کارهای برودریک و همکاران (۲۰۱۹) و ما و همکاران (۲۰۱۹) موردتوجه قرار گرفته است [۲ و ۳]. مؤثرترین الگوریتم در یادگیری عمیق برای پردازش تصاویر با قدرت تفکیک مکانی بسیار بالا، شبکه‌های عصبی کانولوشنی (*CNNs*)[1] هستند [۱]. تورس و همکاران (۲۰۲۰)، هارتلینگ و همکاران (۲۰۱۹) و سانتوز و همکاران (۲۰۱۹) نتایج مناسبی را در استفاده از یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی جهت پردازش تصاویر با قدرت تفکیک مکانی بالا برای آشکارسازی درختان در محیط‌های شهری ارائه داده‌اند [۴، ۵ و ۶]. یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی جهت پردازش تصاویر با قدرت تفکیک مکانی بالا برای شناسایی گیاهان در سال‌های اخیر به‌شدت موردتوجه محققان قرار گرفته است به‌گونه‌ای که حجم بسیار بالایی از مقالات در این زمینه به پردازش تصاویر حاصل از سکوهای پرنده بدون سرنشین پرداخته‌اند؛ کاتن برن و همکاران (۲۰۲۱) مروری بر کارهای متعدد انجام شده در این زمینه انجام داده‌اند [۷]. از نظر شکل کانولوشن اعمالی بر تصویر، شبکه‌های عصبی کانولوشنی را می‌توان تک، دو و سه‌بعدی دانست که یک تحقیق مروری که اخیراً انجام شده است نشان می‌دهد که شبکه‌های دوبعدی با سهم ۸۸ درصدی از کل روش‌های موجود، بیشترین استفاده را داشته‌اند، همچنین از نظر معماری، شبکه‌های عصبی کانولوشنی از تنوع بسیار بالایی برخوردارند که شکل(۱) فراوانی استفاده از معماری‌های مختلف طی یک تحقیق مروری در این زمینه را نشان می‌دهد [۷].

شناسایی و شمارش درختان با عمر کم یا اصطلاحاً نهال‌های درخت، موضوع تحقیق پیرز و همکاران (۲۰۲۰) بوده است که بر مبنای شبکه‌های عصبی کانولوشنی در پردازش تصاویر هوایی با قدرت تفکیک بسیار بالا به دقت‌های بالا ٪۹۰ دست‌یافته‌اند [۸]. البته شناسایی درختان در محیط‌های شهری با پیچیدگی زمینه بسیار بالا و درهم‌تنیدگی شاخوبرگ درختان نسبت به حالت نهال‌های باغ‌های مثمر کشاورزی متفاوت بوده و مقایسه دقت عملکرد نیز بدون در نظر گرفتن شرایط متفاوت درست نیست، لذا در این تحقیق، معماری موردنظر در مقایسه با معماری‌های متداول و بر روی یک داده مرجع پیاده‌سازی شده است. اسکو و همکاران (۲۰۲۰) در پردازش تصاویر با قدرت تفکیک مکانی بالا فراهم شده توسط سکوی پرنده بدون سرنشین جهت شمارش درختان توسط روش یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی استفاده از باند مادون‌قرمز نزدیک در کنار باندهای قرمز و سبز را حاوی نتایج بهتر نسبت به ترکیب رنگی قرمز، سبز و آبی دانسته‌اند [۹]. شمارش درختان در باغ‌ها با استفاده از روش‌های شبکه عصبی کانولوشنی موضوع تحقیقات لی و همکاران (۲۰۱۷) و جیانگ و همکاران (۲۰۱۷) نیز بوده است [۱۰ و ۱۱]. امپادزدیس و پارتل (۲۰۱۹) در یک پروژه تحقیقاتی دقت ٪۹۹٫۹ برای استخراج و شمارش درختان با استفاده از تصاویر سکوهای هوایی بدون سرنشین و یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی را ادعا کرده‌اند، البته در این تحقیق هدف شمارش نهال‌های درختان مثمر در یک منطقه بوده است که با نظم خاص و با فاصله استاندارد نسبت به یکدیگر کاشته شده‌اند و عملاً دستیابی به چنین دقتی با توجه به الگوی قرارگیری منظم درختان و وضعیت خاص آنها بدون درهم‌تنیدگی شاخوبرگ و یک پس‌زمینه ساده و بدون پیچیدگی، قابل پیش‌بینی است [۱۲].

[1] *Convolutional Neural Networks (CNNs)*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]



شکل ۱: معماری‌های مورداستفاده [۷]

استخراج اطلاعات سه‌بعدی از تصاویر هوایی موضوع تحقیق امینی امیرکلایی و عارفی (۲۰۱۹) بوده است که بر پایه شبکه‌های عصبی کانولوشنی به تخمین ارتفاع عوارض مختلف با استفاده از تک تصویر هوایی با دقت مکانی بسیار بالا پرداخته‌اند [۱۳].

برچسب‌گذاری معنایی تصاویر هوایی با قدرت تفکیک مکانی بسیار بالا موضوع داغ موردبحث لیو و همکاران (۲۰۱۸)، لیو و همکاران (۲۰۲۰) و چای و همکاران (۲۰۲۰) بوده است که استخراج درختان در فضای شهری در کنار سایر عوارض ساخته شده توسط انسان مورد نظر بوده است؛ ولی عموماً در روش‌های جامع ارائه شده، دقت استخراج عوارض طبیعی نسبت به عوارض مصنوعی و ساخته بشر نظیر ساختمان‌ها کمتر بوده است [۱۴، ۱۵ و ۱۶].

شبکه‌های عصبی کانولوشنی با معماری‌های مختلف ارائه شده‌اند که روش پیشنهادی ما در استخراج درختان شهری در داده‌های مرجع مورداستفاده با معماری‌های شبکه عصبی کانولوشنی کامل[1] [۱۷]، یونت[2] [۱۸] و سگنت[3] [۱۹] مورد مقایسه قرار می‌گیرد.

در روش پیشنهادی یک معماری متداول شبکه عصبی کانولوشنی رمزگذار - رمزگشا با تغییر در پارامترهای آن نظیر تعداد لایه‌ها و تعداد نرون‌های مورد استفاده در هر لایه و توابع فعال‌سازی و توابع هزینه مختلف مورد استفاده قرار گرفته است که با پیاده‌سازی بر روی یک مجموعه‌داده مرجع مورد بررسی قرار گرفته است. همچنین بحث تغییر در ویژگی‌های ورودی به شبکه عصبی با تغییر باندهای مختلف تصویر مورد بررسی قرار گرفته است. اعمال فیلترهای مختلف مورفولوژیکی جهت بهبود نتایج حاصله در بخش بعدی مورد ارزیابی قرار گرفته است، همچنین بحث کاهش زمان دستیابی به بهترین عملکرد شبکه با استفاده از پیش آموزش شبکه عصبی کانولوشنی مورد بحث، یکی دیگر از موارد مورد بررسی در این تحقیق است. به‌طورکلی در این تحقیق سعی شده است که با پیاده‌سازی راهکارهای پیشنهادی مختلف جهت بهبود عملکرد شبکه‌های عصبی کانولوشنی، نتایج حاصل به صورت یکجا بر روی یک داده مرجع، مورد ارزیابی قرار گیرد.

۲- منطقه موردمطالعه و داده‌ها

داده‌های مورداستفاده در این تحقیق، داده‌های مرجع ارائه شده توسط انجمن بین‌المللی فتوگرامتری و

[1] *Full Convolution Neural Network (FCN)*
[2] *U-Net*
[3] *SegNet*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

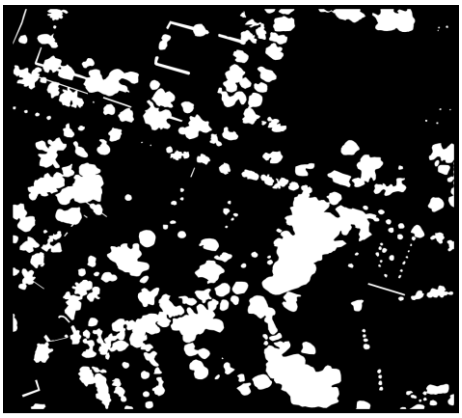
سنجش‌ازدور[1] برای شهر پوتسدام[2] است [۲۰].
این داده که حاصل از تصویربرداری هوایی است تصاویری با قدرت تفکیک مکانی ۵ سانتیمتر را ارائه می‌دهد، ابعاد هر تصویر قائم هوایی[3] آن ۶۰۰۰ در ۶۰۰۰ پیکسل[4] است. در کنار هر تصاویر هوایی قائم یک نقشه واقعیت زمینی برچسب‌گذاری شده در چند کلاس نظیر ساختمان، ماشین، گیاهان کوتاه، درختان و ... ارائه شده است که ما با یک پردازش، صرفاً تصویر حاوی برچسب کلاس درختان را مورداستفاده قرار داده‌ایم.
ضمناً در کنار تصاویر ارائه شده در چهار باند آبی، سبز، قرمز و مادون‌قرمز نزدیک، مدل رقومی سطح زمین[5] ارائه شده است و کلیه خروجی‌ها زمین مرجع سازی[6] شده‌اند.
در این تحقیق منطقه موردمطالعه با چندین نوار تصویربرداری و هر نوار شامل چندین تصویر با ابعاد ۶۰۰۰ در ۶۰۰۰ پیکسل پوشش داده شده است که پس از کنار هم قرار دادن تصاویر مربوط به هر نوار، تصویر کلی تشکیل شده از تصاویر اولیه جهت معرفی به شبکه عصبی کانولوشنی به اجزاء تصویر با ابعاد کوچک‌تر تقسیم‌بندی شده است که ابعاد هر جزء تصویر مورد پردازش ۵۱۲ در ۵۱۲ پیکسل در نظر گرفته شده است، به‌طورکلی ۲۹۲۰ تصویر در سه باند با ابعاد ۵۱۲ در ۵۱۲ ایجاد شده است که از این تعداد ۲۴۲۰ تصویر برای آموزش شبکه و ۴۸۴ تصویر جهت ارزیابی عملکرد شبکه مورداستفاده قرار گرفته‌اند.
به بیان دیگر منطقه با چهار نوار شامل ۲۴ عکس هوایی با ابعاد ۶۰۰۰ در ۶۰۰۰ پیکسل پوشش داده شده است که از عکس‌های موجود در هر نوار یکی به‌عنوان تصویر ارزیاب و باقی به‌عنوان تصاویر آموزشی مورد استفاده قرار گرفته‌اند. انتخاب تصاویر آموزشی کمتر از تعداد مورد بحث، دقت مورد نظر در مرحله آموزش شبکه عصبی کانولوشنی را تأمین نمی‌کند؛ لذا در نهایت ۲۰ تصویر کلی ۶۰۰۰ در ۶۰۰۰ پیکسل پس از تفکیک به جزء تصویر ۵۱۲ در ۵۱۲ پیکسل، به‌عنوان ۲۹۲۰ تصویر آموزشی و ۴ تصویر کلی ۶۰۰۰ در ۶۰۰۰ پیکسل پس از تفکیک به جزء تصویر ۵۱۲ در ۵۱۲ پیکسل، به‌عنوان ۴۸۴ تصویر ارزیاب مورد استفاده قرار گرفته‌اند.
جهت پردازش و پیاده‌سازی شبکه عصبی کانولوشنی از سامانه گوگل کولب[7] بر پایه زبان برنامه‌نویسی پایتون[8] استفاده شده است که پردازش اولیه بدون استفاده از پیش آموزش شبکه با داده‌های اولیه تخمینی در حدود ۳ روز زمان‌بر بوده است.
شکل (۲) بخشی از داده‌های مورداستفاده در این تحقیق را نمایش می‌دهد. شکل (۲-الف) ترکیب رنگی حقیقی سه باندی بخشی از منطقه موردمطالعه را نمایش می‌دهد که به‌وسیله تبدیل مرکاتور معکوس روی بیضوی جهانی[9] زمین مرجع سازی شده است و هر باند در ۲۵۶ درجه روشنایی ارائه می‌شود.
شکل (۲-ب) نقشه واقعیت زمینی تک کلاسه درختان، استخراج شده از نقشه چند کلاسه واقعیت زمینی ارائه شده توسط انجمن بین‌المللی فتوگرامتری و سنجش‌ازدور برای منطقه موردمطالعه است که به‌صورت تک باندی ارائه شده است و در آن پیکسل‌های با برچسب درخت با درجه روشنایی ۲۵۵ و پیکسل‌های زمینه با درجه روشنایی صفر نشان داده شده‌اند.

[1] *ISPRS*
[2] *Potsdam*
[3] *True Orthophoto (TOP)*
[4] *Pixel*
[5] *Digital Surface Model (DSM)*
[6] *Georeference*
[7] *Google Colab*
[8] *Python*
[9] *UTM WGS84*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]



(الف)



(ب)

شکل ۲: نمونه داده‌های مورداستفاده

۳- روش‌شناسی

در ابتدا فلوچارت روش پیشنهادی ارائه خواهد شد، سپس معماری شبکه عصبی کانولوشنی مورد استفاده تشریح خواهد شد و در نهایت موضوع استفاده از فیلترهای مورفولوژیکی جهت بهبود نتایج مورد بحث قرار خواهد گرفت.

۳-۱- فلوچارت روش پیشنهادی

در این بخش به تشریح فلوچارت روش پیشنهادی مطابق با شکل (۳) و مبانی تحلیلی آن پرداخته خواهد شد. مطابق با فلوچارت شکل(۳)، داده‌ها در دو ترکیب رنگی حقیقی و کاذب به شبکه عصبی کانولوشنی معرفی می‌شوند، ابتدا ترکیب سه باندی قرمز، سبز و آبی به شبکه معرفی شده و در ادامه ترکیب سه باندی مادون‌قرمز نزدیک، قرمز و سبز، هدف از این کار مقایسه عملکرد الگوریتم در جایگزینی باند مادون‌قرمز به‌جای باند آبی است، با توجه به رفتار طیفی گیاهان و تغییرات زیاد در رفتار انعکاسی آنها از باند قرمز به مادون‌قرمز، انتظار می‌رود این تغییر ترکیب باندهای ورودی موجب افزایش دقت شبکه شود که نتایج در بخش پنجم ارائه خواهد شد. معماری شبکه عصبی کانولوشنی مورداستفاده در زیربخش بعدی توضیح داده خواهد شد. درهرصورت خروجی این شبکه تصویر دودویی،[1] درختان در تصاویر است که پیکسل‌هایی که درخت تشخیص داده شده‌اند در یک تصویر ۸ بیتی[2] با درجه روشنایی ۲۵۵ و پیکسل‌های غیر درخت یا پس‌زمینه با درجه روشنایی صفر نمایش داده می‌شوند. بررسی نقشه‌های تخمینی[3] و نقشه‌های واقعیت زمینی[4] نشان‌دهنده این است که اعمال یک فیلتر مورفولوژیکی[5] می‌تواند موجب افزایش دقت آشکارسازی شود که روند اعمال این فیلتر در ادامه ارائه خواهد شد.

در این تحقیق حدود ۱۷ درصد از داده‌ها جهت ارزیابی نهایی بدون مشارکت در فرایند آموزش کنار گذاشته می‌شوند و پس از تعیین پارامترهای شبکه عصبی کانولوشنی و نهایی‌سازی آن این شبکه بر داده‌های ارزیاب[6] اعمال می‌شود، ضمن اینکه از داده‌های باقیمانده ۸۲ درصد جهت آموزش[7] و ۱۸ درصد برای اعتبارسنجی[8] مورداستفاده قرار گرفته‌اند.

در انتها نتایج حاصل از داده‌های ارزیاب با نتایج سه معماری متفاوت از معماری پیشنهادی برای شبکه عصبی کانولوشنی مورد مقایسه قرار می‌گیرند.

[1] *binary*

[2] *Bit*

[3] *Prediction Map*

[4] *Ground Truth Map*

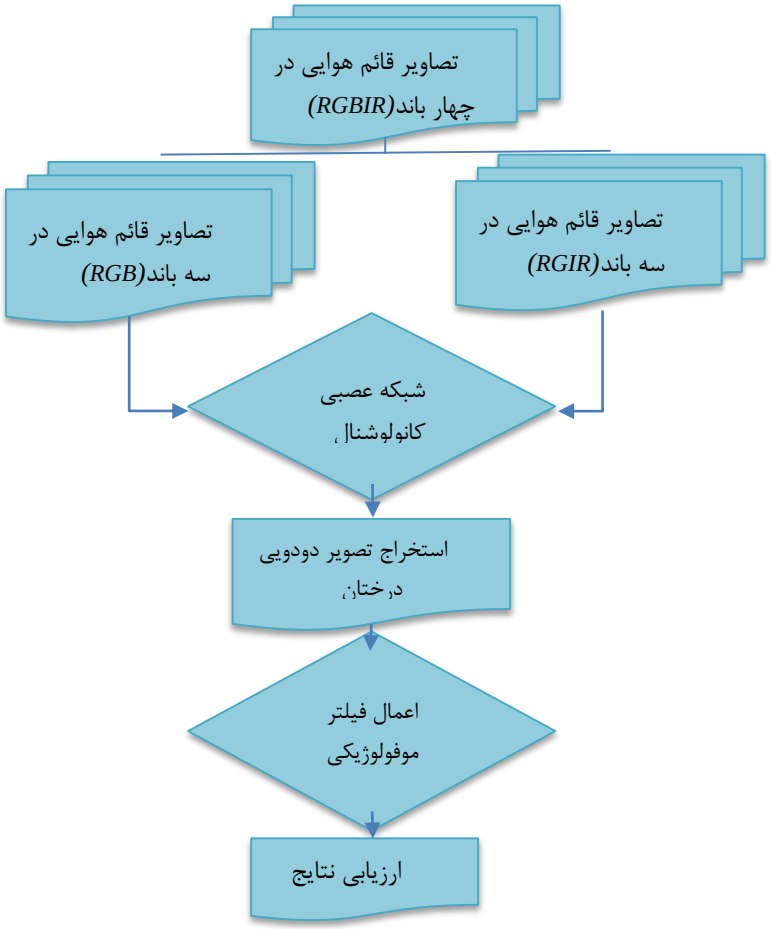
[5] *Morphological Filter*

[6] *Test*

[7] *Train*

[8] *Validation*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]



شکل ۳: فلوچارت روش پیشنهادی

۳-۲- معماری شبکه عصبی کانولوشنی

شبکه عصبی کانولوشنی طراحی شده از نوع رمزگذار - رمزگشا[1] بوده که مطابق با ماهیت این طراحی ابتدا ابعاد تصویر کاهش یافته و عمق آن افزایش می‌یابد و در بخش دوم مجدداً فرایند معکوس اعمال شده و در نهایت نقشه تخمین با ابعاد تصویر ورودی ساخته می‌شود، در این شبکه از کرنل کانولوشن دوبعدی سه در سه استفاده شده است و تابع فعال‌سازی[2] از نوع رلو[3] و تابع هزینه[4] از نوع کراس آنتروپی[5] است. انتخاب تابع فعال‌سازی و تابع هزینه بر مبنایی همگرایی نتایج عملکرد در مرحله آموزش است که عملاً توابع مورد بحث بهترین نتایج را از نظر افزایش دقت و کاهش خطا به همراه داشته‌اند. جداول (۱) و (۲) لایه‌های مختلف در معماری شبکه موردنظر در دو بخش رمزگذار و رمزگشا را نمایش می‌دهد. این جداول یکی از گزارش‌های خروجی شبکه عصبی کانولوشنی مورد استفاده است که معماری این شبکه را در دو مرحله رمزگذاری و رمزگشایی تشریح می‌کند.

[1] *Decoder*
[2] *Activation Function*
[3] *ReLU*
[4] *Loss Function*
[5] *Cross Entropy*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

جدول ۱: معماری شبکه در بخش رمزگذار

عنوان لایه در معماری شبکه عصبی	ابعاد ماتریس خروجی هر لایه	تعداد پارامترهای هر لایه
input_۱	(۵۱۲،۵۱۲،۳)	۰
Conv۲d	(۵۱۲،۵۱۲،۱۶)	۴۴۸
batch_normalization	(۵۱۲،۵۱۲،۱۶)	۶۴
activation	(۵۱۲،۵۱۲،۱۶)	۰
Conv۲d_۱	(۵۱۲،۵۱۲،۱۶)	۲۳۲۰
batch_normalization_۱	(۵۱۲،۵۱۲،۱۶)	۶۴
Concatenate	(۵۱۲،۵۱۲،۳۲)	۰
activation_۱	(۵۱۲،۵۱۲،۳۲)	۰
max_pooling۲d	(۲۵۶،۲۵۶،۳۲)	۰
Conv۲d_۲	(۲۵۶،۲۵۶،۳۲)	۹۲۴۸
batch_normalization_۲	(۲۵۶،۲۵۶،۳۲)	۱۲۸
activation_۲	(۲۵۶،۲۵۶،۳۲)	۰
Conv۲d_۳	(۲۵۶،۲۵۶،۳۲)	۹۲۴۸
batch_normalization_۳	(۲۵۶،۲۵۶،۳۲)	۱۲۸
concatenate_۱	(۲۵۶،۲۵۶،۶۴)	۰
activation_۳	(۲۵۶،۲۵۶،۶۴)	۰
max_pooling۲d_۱	(۱۲۸،۱۲۸،۶۴)	۰
Conv۲d_۴	(۱۲۸،۱۲۸،۶۴)	۳۶۹۲۸
batch_normalization_۴	(۱۲۸،۱۲۸،۶۴)	۲۵۶
activation_۴	(۱۲۸،۱۲۸،۶۴)	۰
Conv۲d_۵	(۱۲۸،۱۲۸،۶۴)	۳۶۹۲۸
batch_normalization_۵	(۱۲۸،۱۲۸،۶۴)	۲۵۶
concatenate_۲	(۱۲۸،۱۲۸،۱۲۸)	۰
activation_۵	(۱۲۸،۱۲۸،۱۲۸)	۰
max_pooling۲d_۲	(۶۴،۶۴،۱۲۸)	۰
Conv۲d_۶	(۶۴،۶۴،۱۲۸)	۱۴۷۵۸۴
batch_normalization_۶	(۶۴،۶۴،۱۲۸)	۵۱۲
activation_۶	(۶۴،۶۴،۱۲۸)	۰
Conv۲d_۷	(۶۴،۶۴،۱۲۸)	۱۴۷۵۸۴
batch_normalization_۷	(۶۴،۶۴،۱۲۸)	۵۱۲
concatenate_۳	(۶۴،۶۴،۲۵۶)	۰
activation_۷	(۶۴،۶۴،۲۵۶)	۰
max_pooling۲d_۳	(۳۲،۳۲،۲۵۶)	۰
Conv۲d_۸	(۳۲،۳۲،۲۵۶)	۵۹۰۰۸۰
batch_normalization_۸	(۳۲،۳۲،۲۵۶)	۱۰۲۴
activation_۸	(۳۲،۳۲،۲۵۶)	۰
Conv۲d_۹	(۳۲،۳۲،۲۵۶)	۵۹۰۰۸۰
batch_normalization_۹	(۳۲،۳۲،۲۵۶)	۱۰۲۴
concatenate_۴	(۳۲،۳۲،۵۱۲)	۰
activation_۹	(۳۲،۳۲،۵۱۲)	۰

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

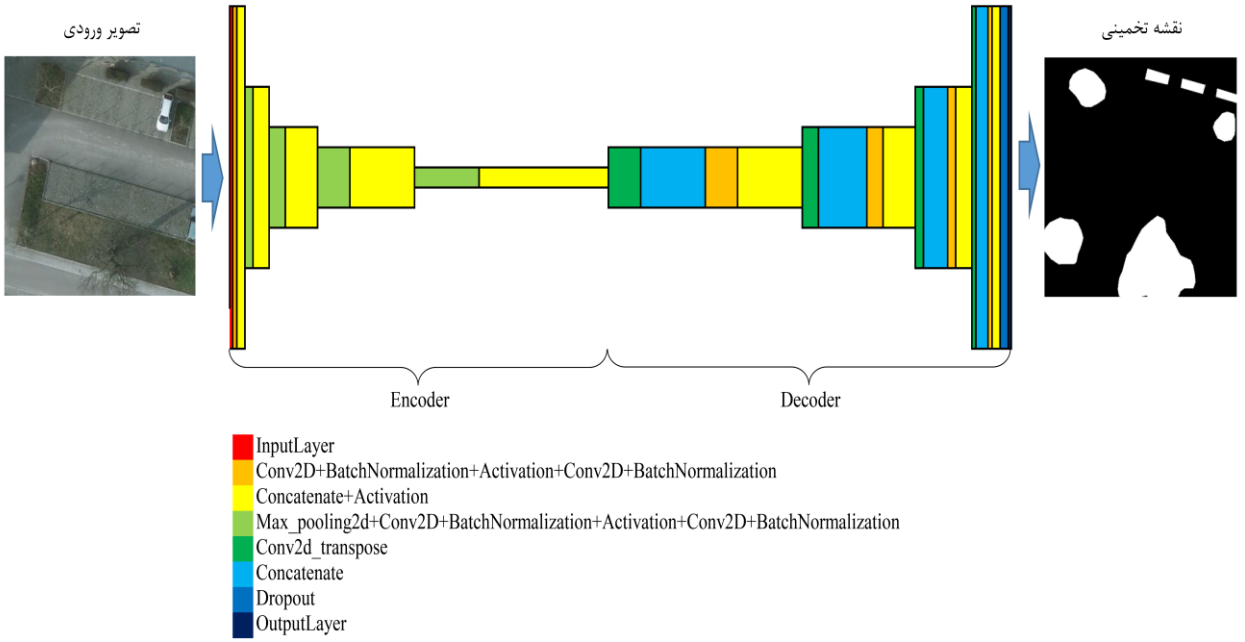
جدول ۲: معماری شبکه در بخش رمزگشا

عنوان لایه در معماری شبکه عصبی	ابعاد ماتریس خروجی هر لایه	تعداد پارامترهای هر لایه
Conv۲d_transpose	(۶۴،۶۴،۱۲۸)	۵۸۹۹۵۲
concatenate_۵	(۶۴،۶۴،۳۸۴)	۰
Conv۲d_۱۰	(۶۴،۶۴،۱۲۸)	۴۴۲۴۹۶
batch_normalization_۱۰	(۶۴،۶۴،۱۲۸)	۵۱۲
activation_۱۰	(۶۴،۶۴،۱۲۸)	۰
Conv۲d_۱۱	(۶۴،۶۴،۱۲۸)	۱۴۷۵۸۴
batch_normalization_۱۱	(۶۴،۶۴،۱۲۸)	۵۱۲
concatenate_۶	(۶۴،۶۴،۲۵۶)	۰
activation_۱۱	(۶۴،۶۴،۲۵۶)	۰
Conv۲d_transpose_۱	(۱۲۸،۱۲۸،۶۴)	۱۴۷۵۲۰
concatenate_۷	(۱۲۸،۱۲۸،۱۹۲)	۰
Conv۲d_۱۲	(۱۲۸،۱۲۸،۶۴)	۱۱۰۶۵۶
batch_normalization_۱۲	(۱۲۸،۱۲۸،۶۴)	۲۵۶
activation_۱۲	(۱۲۸،۱۲۸،۶۴)	۰
Conv۲d_۱۳	(۱۲۸،۱۲۸،۶۴)	۳۶۹۲۸
batch_normalization_۱۳	(۱۲۸،۱۲۸،۶۴)	۲۵۶
concatenate_۸	(۱۲۸،۱۲۸،۱۲۸)	۰
activation_۱۳	(۱۲۸،۱۲۸،۱۲۸)	۰
Conv۲d_transpose_۲	(۲۵۶،۲۵۶،۳۲)	۳۶۸۹۶
concatenate_۹	(۲۵۶،۲۵۶،۹۶)	۰
Conv۲d_۱۴	(۲۵۶،۲۵۶،۳۲)	۲۷۶۸۲
batch_normalization_۱۴	(۲۵۶،۲۵۶،۳۲)	۱۲۸
activation_۱۴	(۲۵۶،۲۵۶،۳۲)	۰
Conv۲d_۱۵	(۲۵۶،۲۵۶،۳۲)	۹۲۴۸
batch_normalization_۱۵	(۲۵۶،۲۵۶،۳۲)	۱۲۸
concatenate_۱۰	(۲۵۶،۲۵۶،۶۴)	۰
activation_۱۵	(۲۵۶،۲۵۶،۶۴)	۰
Conv۲d_transpose_۳	(۵۱۲،۵۱۲،۱۶)	۹۲۳۲
concatenate_۱۱	(۵۱۲،۵۱۲،۴۸)	۰
Conv۲d_۱۶	(۵۱۲،۵۱۲،۱۶)	۶۹۲۸
batch_normalization_۱۶	(۵۱۲،۵۱۲،۱۶)	۶۴
activation_۱۶	(۵۱۲،۵۱۲،۱۶)	۰
Conv۲d_۱۷	(۵۱۲،۵۱۲،۱۶)	۲۳۲۰
batch_normalization_۱۷	(۵۱۲،۵۱۲،۱۶)	۶۴
concatenate_۱۲	(۵۱۲،۵۱۲،۳۲)	۰
activation_۱۷	(۵۱۲،۵۱۲،۳۲)	۰
dropout	(۵۱۲،۵۱۲،۳۲)	۰
Conv۲d_۱۸	(۵۱۲،۵۱۲،۲)	۶۶
تعداد کل پارامترها: ۳۱۴۳۸۴۲		
تعداد پارامترهای آموزشی: ۳۱۴۰۸۹۸		
تعداد پارامترهای غیرآموزشی: ۲۹۴۴		

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

به‌طورکلی این شبکه دارای ۳۱۴۳۸۴۲ پارامتر جهت برآورد است.

شکل (۴) معماری شبکه عصبی کانولوشنی رمزگذار - رمزگشا پیشنهادی را نمایش می‌دهد که ابتدا با کاهش ابعاد و افزایش عمق در تصویر ورودی فرایند رمزگذاری انجام شده و در ادامه عکس این عمل در کاهش عمق و افزایش ابعاد تا رسیدن به ابعاد برابر با تصویر ورودی طی عملیات رمزگشایی صورت می‌پذیرد.



شکل ۴: معماری شبکه عصبی کانولوشنی پیشنهادی

۳-۳- فیلتر مورفولوژیکی

همان‌طور که پیش‌ازاین توضیح داده شد یکی از اهداف فرعی این تحقیق بررسی اثر اعمال فیلتر مورفولوژیکی در بهبود نتایج حاصل از عملکرد الگوریتم پیشنهادی است؛ لذا با تفاضل‌گیری از دو نقشه تخمینی و واقعیت زمینی یک نقشه تفاضل فراهم شده است که با تحلیل بصری این نقشه می‌توان دریافت که الگوریتم پیشنهادی در شناسایی هسته درختان خوب عمل کرده است؛ ولی در پوشش کامل آنها عملکرد ایده آلی ندارد؛ لذا ایده استفاده از یک فیلتر مورفولوژیکی که نتایج را در لبه‌ها گسترش دهد مطرح شد، انواع فیلترهای مورفولوژیکی با کرنل‌های متفاوت و شعاع‌های مختلف بر روی نتایج پیاده‌سازی شد که بهترین عملکرد مرتبط با یک فیلتر مورفولوژیکی با محدوده اثر یک دایره با شعاع ۲۰ پیکسل بود، بنابراین فیلتر مورفولوژیکی بستن[1] با شعاع ۲۰ پیکسل بر نتایج حاصل از شبکه عصبی اعمال شد، این فیلتر را می‌توان به‌صورت رابطه(۱) تعریف نمود:

$$A \bullet B = (A \oplus B) \ominus B \qquad \text{رابطه(۱)}$$

در رابطهٔ (۱)، A تصویر دودویی اولیه و B کرنل موردنظر است که تصویر نهایی حاصل از اعمال یک عملگر فرسایش[2] بر نتیجه اعمال یک عملگر گشایش[3]

[1] *Closing*
[2] *Erosion*
[3] *Dilation*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

بر تصویر اولیه است.

لازم به ذکر است که دلیل اصلی انتخاب فیلتر مورفولوژیکی بستن پرکردن حفره‌های موجود در تاج درختان و توسعه لبه‌های مناطق مربوط به پوشش درختان است به‌گونه‌ای که اعمال سایر عملگرها، در نتایج اثر معکوس گذاشته و دقت را کاهش می‌دهند و صرفاً فیلتر بستن موجب افزایش دقت تشخیص می‌شود. شکل کرنل مورد استفاده نیز یک دایره به شعاع ۲۰ پیکسل در نظر گرفته شده است که با استفاده از آزمون‌وخطا انتخاب شده و عملاً بهترین عملکرد را به همراه داشته است.

۴- پیاده‌سازی

با توجه به اهداف تعریف شده برای این پژوهش، پیاده‌سازی الگوریتم پیشنهادی را در دو بخش یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی و اعمال فیلتر مورفولوژیکی پیش می‌بریم.

۴-۱- شبکه عصبی کانولوشنی

روش پیشنهادی در سه حالت مختلف اعمال شده است، در حالت اول تصاویر سه باندی با ترکیب باندهای قرمز، سبز و آبی ورودی شبکه بوده است، در حالت دوم ترکیب باندی مادون‌قرمز نزدیک، قرمز و سبز ورودی شبکه را تشکیل داده است و در حالت سوم بر روی نتایج خروجی از مرحله دوم یک فیلتر مورفولوژیکی اعمال شده است و سپس نتایج وارد فرایند ارزیابی دقت شده‌اند.

در شکل (۵) نتایج بصری اعمال شبکه پیشنهادی در حالات سه‌گانه نمایش داده شده است، به‌طورکلی با حرکت از حالت اول تا سوم دقت به صورت بسیار کم افزایش یافته است این در حالی است که انتظار می‌رفت در حالت ترکیب باندی دوم، یعنی جایگزینی باند مادون‌قرمز به‌جای باند آبی نتایج تغییر بسیار بیشتری داشته باشند که این موضوع در بخش بعدی مورد بررسی قرار می‌گیرد. با توجه به نتایج ارائه شده در شکل (۵)، شبکه عصبی کانولوشنی پیشنهادی در تشخیص درختان دارای سبزینگی و همچنین درختان عاری از برگ و سبزینگی با قامت متوسط و بزرگ به‌خوبی عمل نموده است (ردیف‌های ۱، ۶، ۷، ۸ و ۹)، معایب و اختلاف نتایج به‌دست‌آمده با نقشهٔ واقعیت زمینی بیشتر در شناسایی درختچه‌ها و گیاهان مرتفع است که در نقشه واقعیت زمینی تحت عنوان درخت برچسب خورده‌اند (ردیف‌های ۳، ۵ و ۹)، مشکل دیگر عدم تطابق با حاشیه‌های کامل یک درخت با وجود شناسایی هسته مرکزی درخت است که سعی شده است با اعمال یک عملگر مورفولوژیکی تا جای ممکن این اثر تعدیل شود (ردیف‌های ۱ و ۲). همچنین استفاده از فیلتر مورفولوژیکی در برخی موارد موجب گسترش خطا شده است (ردیف ۴)، این در حالی است که عملکرد کلی این فیلتر مورد قبول بوده است.

۴-۲- اعمال فیلتر مورفولوژیکی

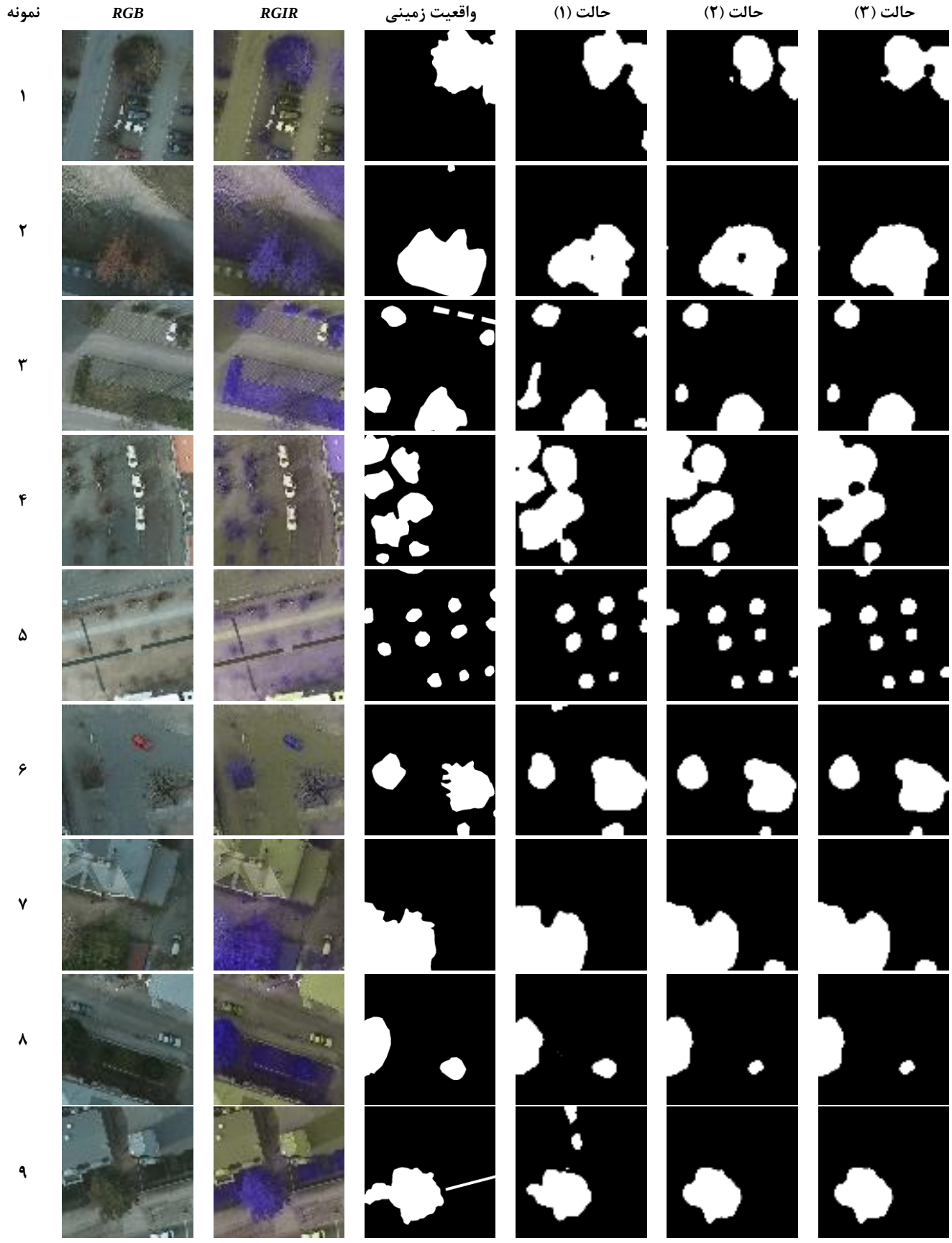
شکل(۶) نتیجه تفاضل دو تصویر خروجی از مرحله دوم و نقشه واقعیت زمینی را نمایش می‌دهد، مطابق با این شکل می‌توان انتظار داشت که در صورت گسترش تصویر می‌توان انتظار بهبود نتایج طبقه‌بندی را داشت، در این حالت با اعمال یک فیلتر مورفولوژیکی بستن[1] با شعاع ۲۰ پیکسل بر نتیجه طبقه‌بندی خروجی حالت دوم، نقشه تخمین حالت سوم پدید آمده است.

با دقت در نقشه تفاضلی این موضوع مشخص می‌شود که روش پیشنهادی در تشخیص هسته‌های آشکارسازی بسیار مناسب عمل نموده است و صرفاً در محدوده لبه‌های کناری و حاشیه‌ها دچار اشکال بوده و محدوده کمتری را آشکار کرده است، لذا با اعمال فیلتر مورفولوژی جهت گسترش تصویر می‌توان انتظار افزایش دقت طبقه‌بندی را داشت، در این تحقیق انواع فیلترهای مورفولوژی با کرنل‌های متفاوت و ابعاد متخلف بر روی تصویر اعمال شده و نتیجه مورد ارزیابی قرار گرفته است که کرنل بستن با شعاع ۲۰ پیکسل بهترین نتیجه را فراهم نموده است.

[1] *Closing*

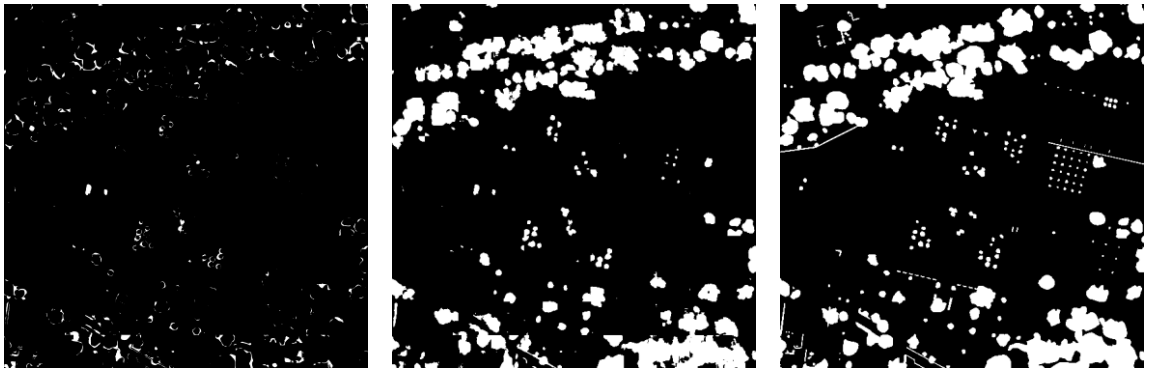
[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

شکل۵:تایج بصری اعمال شبکه پیشنهادی در حالات سه‌گانه



[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

نکته دیگر در رابطه با روش پیشنهادی مدت زمان لازم جهت اعمال الگوریتم یادگیری عمیق و برآورد حدود ۳ میلیون پارامتر شبکه است که با استفاده از حالت رایگان در سامانه گوگل کولب در حدود ۷۰ ساعت زمان لازم است تا در تکرارهای متعدد بهترین دقت عملکرد روی داده‌های ارزیابی اتفاق بیفتد؛ ولی درصورتی‌که از پارامترهای خروجی از حالت اول برای پیش آموزش پارامترهای شبکه در حالت دوم استفاده شود شاهد خواهیم بود که مدت زمان رسیدن به بهترین عملکرد و کمترین خطا در داده‌های ارزیابی در حدود ۸۳٪ کاهش یافته و تنها با ۱۱ ساعت آموزش، شبکه به حالت بهینه خود می‌رسد.



شکل ۶: نقشه واقعیت زمینی (راست)، تخمینی (وسط) و تفاضلی (چپ)

۵- ارزیابی نتایج

هرگونه مدل‌سازی و طبقه‌بندی همواره با مقادیر اجتناب‌ناپذیر از خطاها همراه است. جهت ارزیابی عملکرد روش پیشنهادی همان گونه که پیش‌ازاین توضیح داده شد در حدود ۱۷٪ از داده‌های اولیه که شامل ۴۸۴ تصویر با ابعاد ۵۱۲ در ۵۱۲ پیکسل می‌شود به‌عنوان داده‌های ارزیابی و بدون مشارکت در مرحله آموزش و اعتبارسنجی شبکه یادگیری عمیق پیشنهادی مورداستفاده قرار گرفته‌اند، در ادامه مهم‌ترین پارامترهای ارزیابی که در این تحقیق مورداستفاده قرار گرفته‌اند در روابط (۲) تا (۶) بررسی شده و نتایج ارائه خواهد شد.

رابطه(۲) $$OA = \frac{TP + TN}{P + N}$$

رابطه(۳) $$\mathrm{Pr}ecision = \frac{TP}{TP + FP}$$

رابطه(۴) $$\mathrm{Re}call = \frac{TP}{TP + FN}$$

رابطه(۵) $$IOU = \frac{TP}{TP + FP + FN}$$

رابطه(۶) $$F1 = 2 * \frac{\mathrm{Pr}ecision * \mathrm{Re}call}{\mathrm{Pr}ecision + \mathrm{Re}call}$$

در روابط (۲) تا (۶) ،P تعداد نمونه‌های کلاس هدف، N تعداد نمونه‌های کلاس زمینه، TP تعداد نمونه‌های کلاس هدف که به‌درستی برچسب هدف خورده‌اند، FP تعداد نمونه‌های کلاس زمینه که به‌غلط برچسب هدف خورده‌اند، TN تعداد نمونه‌های کلاس زمینه که به‌درستی برچسب زمینه خورده‌اند، FN تعداد نمونه‌های کلاس هدف که به‌غلط برچسب زمینه خورده‌اند و تعداد نمونه‌ها در هر تصویر برابر تعداد پیکسل‌های تصویر است. روابط (۲) تا (۶) دقت کلی[۱]، دقت آشکارسازی کلاس هدف[۲] یا دقت کاربر[۳]، میزان حساسیت[۴] یا دقت تولیدکننده[۵]، شاخص IOU[۶] و امتیاز $F1$[۷] را نشان می‌دهد.

---

۱ *Overall Accuracy*

2 *Precision*

۳ *User's Accuracy*

4 *Recall or Sensitivity*

۵ *Producer's Accuracy*

6 *Intersection over Union*

۷ *F1 Score*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

جدول (۳) نتایج ارزیابی بر روی داده‌ها را نشان می‌دهد، که حالت (۱) ترکیب باندی قرمز، سبز و آبی، حالت (۲) ترکیب باندی مادون‌قرمز نزدیک، قرمز و سبز و حالت (۳) اعمال عملگر مورفولوژیکی بر نتایج حاصل از حالت (۲) است.

جدول ۳: معیارهای ارزیابی در حالات پیشنهادی

معیار ارزیابی	حالت (۱)	حالت (۲)	حالت (۳)
OA	۹۵٫۷۶	۹۶٫۰۶	۹۶٫۱۰
Precision	۷۹٫۹۹	۸۳٫۱۶	۸۲٫۱۶
Recall	۷۸٫۴۲	۷۷٫۴۸	۷۹٫۳۴
IOU	۶۵٫۵۶	۶۶٫۹۷	۶۷٫۶۸
F۱	۷۹٫۲۰	۸۰٫۲۲	۸۰٫۷۲

جدول (۴) نتایج حاصل از روش پیشنهادی را در کنار نتایج حاصل از معماری‌های بسیار متداول دیگر شبکه‌های عصبی کانولوشنی ارائه شده در کار محققان دیگر بر روی همین داده‌ها را نمایش می‌دهد [۱۵].

مطابق با انتظار در حالت شماره (۳) یعنی در ترکیب باندی که باند مادون‌قرمز نزدیک جایگزین باند آبی شده است و یک فیلتر مورفولوژیک بر خروجی شبکه اعمال شده است، نتایج بهتری را نشان می‌دهد؛ ولی این تغییر بسیار جزئی است که می‌توان اصلی‌ترین دلیل آن تصویربرداری در فصل زمستان یا اواخر پاییز دانست، همان گونه که در تصاویر مشخص اکثر درختان عاری از پوشش سبز بوده و درختان عملاً فاقد برگ هستند و تنه درخت و شاخه‌های آن تشکیل‌دهنده کلاس درخت هستند در این حالت فرض جایگزینی باند مادون‌قرمز نزدیک به‌جای باند آبی با توجه به عدم وجود سبزینگی و کلروفیل در درختان عملاً تغییر محسوسی را در دقت طبقه‌بندی ایجاد نکرده است. ضمناً جدول شماره (۴) کارایی روش پیشنهادی را نسبت به معماری‌های متداول در طراحی شبکه‌های عصبی کانولوشنی نمایش می‌دهد.

جدول ۴: مقایسه روش پیشنهادی با سایر معماری‌ها

روش پیشنهادی	[۱۵]			معیار ارزیابی
	U-Net	***SegNet***	***FCN***	
۹۶٫۱۰	۸۳٫۲۰	۸۲٫۶۹	۸۴٫۲۵	***OA***
۶۷٫۶۸	۶۱٫۶۲	۶۰٫۲۵	۶۴٫۳۶	***IOU***
۸۰٫۷۲	۷۵٫۴۱	۷۳٫۷۴	۷۶٫۵۱	***F1***

۶- بحث و نتیجه‌گیری

درختان در زندگی انسان نقش بسیار مهم و تأثیرگذاری دارند به‌گونه‌ای که بسیاری جنگل‌ها را ریه‌های زمین می‌نامند. درختان شهری علاوه ایفای نقش تصفیه هوا از نظر زیبایی بصری نیز بسیار حائز اهمیت هستند. مدیریت شهری در حال حاضر برای درختان شهری هویت مستقلی تعریف می‌کند و ازبین‌بردن درختان شهری همراه با مجازات‌های سنگین خواهد بود، بنابراین نظارت و پایش درختان شهری از اهمیت بسیار بالایی برخوردار است. پایش درختان

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

شهری بر پایه یک سامانه سنجش‌ازدوری باتوجه‌به قابلیت پایش مناطق وسیع در زمان‌های دلخواه این گزینه را برای مدیران شهری غیر قابل چشم‌پوشی می‌کند. تعیین دقیق موقعیت و وضعیت درختان بر پایه علم سنجش‌ازدور می‌تواند در کوتاه‌ترین زمان هرگونه تغییر در وضعیت سلامت یا ازبین‌بردن درختان را گزارش نماید، طبیعتاً اطلاعات به هنگام در این زمینه می‌تواند مدیریت را ثمربخش‌تر نماید.

پرنده‌های بدون سرنشین با قابلیت تصویربرداری با قدرت تفکیک مکانی بسیار بالا و در زمان دلخواه مجموعه‌ای از داده را فراهم می‌کنند که پردازش آن نیازمند استفاده از روش‌های نوین پردازش تصویر نظیر یادگیری عمیق بر پایه شبکه‌های عصبی کانولوشنی است، این نوع از پردازش تصویر دررابطه‌با تصاویر با قدرت تفکیک مکانی بسیار بالا، مناسب عمل می‌کنند.

در این تحقیق با استفاده از روش یادگیری عمیق مبتنی بر شبکه‌های عصبی کانولوشنی یک معماری برای این شبکه پیشنهاد شد و مورد بررسی قرار گرفت که نتایج آن قدرت بالای این روش در آشکارسازی هدف را نشان می‌دهد که در مقایسه با معماری‌های متداول در این زمینه نتایج قابل قبول هستند.

یکی از مهم‌ترین مسائل در یادگیری عمیق نیازمندی به تعداد بسیار بالای داده آموزشی جهت برآورد پارامترهای میلیونی این شبکه‌ها است، بنابراین به نظر می‌رسد با افزایش تعداد داده‌های آموزشی دقت روش پیشنهادی بیشتر شود.

با مقایسه نقشه واقعیت زمین و نقشه‌های تخمینی مشاهده می‌شود که الگوریتم پیشنهادی در شناسایی گیاهان درخت گونه نظیر پرچین‌های شمشادی چندان خوب عمل نمی‌کند؛ ولی در شناسایی درختان بزرگ و به بیان دیگر استاندارد به‌خوبی عمل می‌کند ضمن اینکه درختانی که همچنان برگ‌های سبز خود را از دست نداده‌اند با دقت مناسبی استخراج شده‌اند، لذا پیش‌بینی می‌شود جایگزینی باند مادون‌قرمز نزدیک به‌جای باند آبی به‌ویژه در فصولی که درختان از سبزینگی برخوردارند موجب افزایش قابل‌توجه دقت آشکارسازی خواهد شد.

نکته دیگر بحث پیش آموزش شبکه و تعیین مقادیر اولیه برای پارامترهای شبکه عصبی کانولوشنی است. در این تحقیق وقتی که خروجی‌های پارامتری حالت یک به‌عنوان مقادیر اولیه پارامترها در حالت شماره دو مورداستفاده قرار گرفت، شاهد بودیم که زمان رسیدن شبکه به‌دقت بهینه در حدود ۸۳٪ کاهش می‌یابد و با تعداد تکرار کمتر، عملاً بهترین دقت در داده‌های ارزیاب با کمترین خطا مشاهده می‌شود.

به‌طورکلی دقت آشکارسازی در حدود ۸۰٪ در داده‌ها مورد قبول بوده هرچند محققان دیگر جهت افزایش دقت آشکارسازی استفاده از داده‌های کمکی، نظیر مدل رقومی سطح در کنار داده‌های چند باندی را پیشنهاد داده‌اند.

با دقت در نتایج ارائه شده در جدول (۴)، کارایی روش پیشنهادی نسبت به سایر روش‌های متداول موجود مشخص می‌شود، در این روش دقت کلی برابر ۹۶٫۱۰، شاخص *IOU* برابر ۶۷٫۶۸ و شاخص *F1* برابر با ۸۰٫۷۲ است. جهت بررسی کارایی شبکه عصبی کانولوشنی ارائه شده، پیشنهاد می‌شود این روش با مجموعه داده‌های مرجع دیگر مورد ارزیابی قرار گیرد، همچنین تأثیر افزونگی داده و تقویت داده ورودی در نتایج روش پیشنهادی نیز موضوع قابل بررسی بوده که می‌تواند در کارهای تحقیقاتی آتی مورد بررسی قرار گیرد.

انتظار می‌رود استفاده از روش پیشنهادی به‌ویژه ترکیب باندی مادون‌قرمز نزدیک، قرمز و سبز برای مناطق شهری در فصول سبز سال نظیر بهار و تابستان منجر به نتایج بهتری شود، عملاً تغییر باندهای ورودی از آبی به مادون‌قرمز نزدیک کارایی خود را باید در زمان سبزینگی گیاهان نمایش دهد هرچند کارایی روش پیشنهادی در شناسایی درختان خشک به دلیل شرایط فصلی موجود در تصاویر قابل قبول است.

در تشخیص درختان با استفاده از روش پیشنهادی در برخی موارد مشاهده می‌شود که با یک حفره شناسایی

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

در تاج درخت و یا عدم شناسایی کامل حاشیه‌ها مطابق با نقشه واقعیت زمینی مواجه هستیم، برای حل این مشکل در تحقیق پیش رو پیشنهاد شد از فیلترهای مورفولوژیکی استفاده شود که انواع فیلترهای مورفولوژیکی با شکل کرنل متفاوت و ابعاد مختلف بر روی تصویر اعمال گردید که بهترین تغییر حاصل از اعمال فیلتر مورفولوژیکی بستن با شکل کرنل دایره‌ای با شعاع ۲۰ پیکسل بود، برای تحقیقات آتی پیشنهاد می‌شود، اثر فیلترهای مورفولوژیکی در بهبود نتایج روش پیشنهادی بیش از این مورد بحث و بررسی قرار گرفته و توسعه داده شود.

مراجع

[1] *F. Schiefer, T. Kattenborn, A. Frick et al., "Mapping forest tree species in high resolution UAV-based RGB-imagery by means of convolutional neural networks," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 170, pp. 205-215, 2020.*

[2] *P. G. Brodrick, A. B. Davies, and G. P. Asner, "Uncovering ecological patterns with convolutional neural networks," Trends in ecology & evolution, vol. 34, no. 8, pp. 734-745, 2019.*

[3] *L. Ma, Y. Liu, X. Zhang et al., "Deep learning in remote sensing applications: A meta-analysis and review," ISPRS journal of photogrammetry and remote sensing, vol. 152, pp. 166-177, 2019.*

[4] *D. Lobo Torres, R. Queiroz Feitosa, P. Nigri Happ et al., "Applying fully convolutional architectures for semantic segmentation of a single tree species in urban environment on high resolution UAV optical imagery," Sensors, vol. 20, no. 2, pp. 563, 2020.*

[5] *S. Hartling, V. Sagan, P. Sidike et al., "Urban tree species classification using a WorldView-2/3 and LiDAR data fusion approach and deep learning," Sensors, vol. 19, no. 6, pp. 1284, 2019.*

[6] *A. A. d. Santos, J. Marcato Junior, M. S. Araújo et al., "Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs," Sensors, vol. 19, no. 16, pp. 3595, 2019.*

[7] *T. Kattenborn, J. Leitloff, F. Schiefer et al., "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 173, pp. 24-49, 2021.*

[8] *G. D. Pearse, A. Y. Tan, M. S. Watt et al., "Detecting and mapping tree seedlings in UAV imagery using convolutional neural networks and field-verified data," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 168, pp. 156-169, 2020.*

[9] *L. P. Osco, M. d. S. de Arruda, J. M. Junior et al., "A convolutional neural network approach for counting and geolocating citrus-trees in UAV multispectral imagery," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 160, pp. 97-106, 2020.*

[10] *W. Li, H. Fu, L. Yu et al., "Deep learning based oil palm tree detection and counting for high-resolution remote sensing images," Remote Sensing, vol. 9, no. 1, pp. 22, 2017.*

[11] *H. Jiang, S. Chen, D. Li et al., "Papaya tree detection with UAV images using a GPU-accelerated scale-space filtering method," Remote Sensing, vol. 9, no. 7, pp. 721, 2017.*

[12] *Y. Ampatzidis, and V. Partel, "UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence," Remote Sensing, vol. 11, no. 4, pp. 410, 2019.*

[13] *H. A. Amirkolaee, and H. Arefi, "Height estimation from single aerial images using a deep convolutional encoder-decoder network," ISPRS journal of photogrammetry and remote sensing, vol.*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

149, pp. 50-66, 2019.

[14] Y. Liu, B. Fan, L. Wang et al., "Semantic labeling in very high resolution images via a self-cascaded convolutional neural network," ISPRS journal of photogrammetry and remote sensing, vol. 145, pp. 78-95, 2018.

[15] W. Liu, Y. Zhang, H. Fan et al., "A New Multi-Channel Deep Convolutional Neural Network for Semantic Segmentation of Remote Sensing Image," IEEE Access, vol. 8, pp. 131814-131825, 2020.

[16] D. Chai, S. Newsam, and J. Huang, "Aerial image semantic segmentation using DCNN predicted distance maps," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 161, pp. 309-322, 2020.

[17] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431-3440, 2015.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", In International Conference on Medical image computing and computer-assisted intervention, pp. 234-241. 2015.

[19] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," IEEE transactions on pattern analysis and machine intelligence, vol. 39, no. 12, pp. 2481-2495, 2017.

[20] International Society for Photogrammetry and Remote Sensing," 2D semantic labeling contest." [Online]. Available: http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html.

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.10, No.2, Summer 2022

Research Paper

Mapping trees in urban environments using deep convolutional neural networks and morphological filters in the processing of aerial images with high spatial resolution

Moslem Darvishi [1], Reza Shahhosseini [2*]

1- Ph.D. candidate of remote sensing in Department of Surveying and Geospatial Engineering, College of Engineering, University of Tehran
2- Assistant Professor in Department of Surveying and Geospatial Engineering, College of Engineering, University of Tehran

Abstract

Green areas and trees play an undeniable role in improving the quality of human life. In addition to their natural role, urban trees are also very important in terms of visual beauty. Aerial imagery using unmanned platforms with very high spatial resolution is available today. Deep learning methods based on convolutional neural networks are very efficient in processing large amounts of data from this type of imaging and its special challenges. The current paper proposed a convolutional neural network architecture on the basis of the encoder_decoder networks, the results of which on the reference data provided by the International Society of Photogrammetry and Remote Sensing from Potsdam, Germany for urban tree detection, show 96.10% overall accuracy and their F1 score equals 80.72%. However, very little change in accuracy can be attributed to the time of imaging, at which the trees in the study region generally did not have any greenery. Another discussion in this research is to create a differential image of the terrestrial reality map and the output estimation map of the proposed algorithm. According to this differential map, the application of a morphological filter seems to be useful, which in practice has increased the accuracy of the final classification. The final issue is the use of training network parameters by red, green and blue band combination as the primary parameters for the network with near, red and green infrared band input, in case the time to achieve the best network performance is reduced by about 83% and this pre-training of neural network parameters has led to the rapid convergence of the target network.

Key words: *Deep Learning, Remote Sensing, Convolutional Neural Network, High Spatial Resolution Aerial images.*

Correspondence Address *: Remote Sensing Group, Department of Surveying and Geospatial Engineering, College of Engineering, University of Tehran.*
Tel*: +98 21 61114527.*
Email*: rshahosseini@ut.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-11-13]
[DOR: 20.1001.1.20089635.1401.10.2.3.9]
[DOI: 10.52547/jgit.10.2.89]