

یک سیستم بینایی هوشمند برای نظارت خودکار بر آتش‌سوزی جنگل‌ها

محمدصادق کیهان‌پناه و بهروز کوهستانی

زیرساخت‌های انسانی گسترش پیدا نکند [۳]. کشف زودهنگام و فرونشانی آتش‌سوزی‌های جنگل برای به حداقل رساندن تخریب‌هایی که ممکن است به علت انتشار سریع و همچنین احتراق طولانی‌مدت آنها ایجاد شود بسیار مهم است [۴].

رویکردهای سنجش از راه دور برای پایش و کشف آتش‌سوزی جنگل به سه دسته سیستم‌های مبتنی بر وسایل نقلیه با سرنشین، سیستم‌های زمینی و سیستم‌های ماهواره‌ای تقسیم می‌شوند [۵]. با این حال، هر یک از این سیستم‌ها مشکلات خاص خود را دارند. وسایل نقلیه با سرنشین معمولاً بزرگ و گران هستند. علاوه بر این، زندگی خلبان می‌تواند توسط محیط‌های خطرناک و خستگی اپراتور تهدید شود [۶]. جایگزین کردن حسگر به جای انسان به طوری که بتوان این کار را به صورت خودکار و با نظارت انسان انجام داد، از روش‌های بعدی مورد استفاده است که از معایب آن می‌توان به وجود محیط باز جنگل که مناسب برای تعداد محدود حسگر نیست، اشاره کرد و تعداد زیاد حسگر هم نیاز به هزینه بالایی دارد. حسگر زمانی آتش را شناسایی خواهد کرد که ویژگی مرتبط با آتش در محدوده تشخیص باشد. با افزایش میزان حساسیت برای استفاده با تعداد کمتر نیز حسگرها به کوچک‌ترین تغییرات حساس شده و نرخ تشخیص اشتباه در آنها افزایش پیدا می‌کند. همچنین این روش، توانایی ارائه اطلاعات درباره اندازه آتش و یا سرعت پیشرفت آن را ندارد. سیستم‌های ماهواره‌ای نیز در برنامه‌ریزی مسیر و به روز رسانی‌ها ضعف دارند و دقت مکانی و زمانی آنها ممکن است برای جمع‌آوری اطلاعات دقیق و مبارزه مؤثر با آتش‌سوزی جنگل کم باشد [۷].

پهپاد یا وسیله نقلیه بدون سرنشین (UAV) با سیستم‌های سنجش از راه دور مبتنی بر بینایی کامپیوتر، گزینه واقع‌بینانه‌ای است که امکان نمایش سریع، سیار و کم‌هزینه را برای پایش، کشف و حتی مبارزه با آتش‌سوزی جنگل فراهم می‌کند. همچنین استفاده از چند پهپاد به صورت هم‌زمان و یا استفاده از انواع حسگر نیز باعث برآورده کردن دقت‌های بینایی، مکانی و زمانی است و این امکان را به عنوان مکمل قدرتمند برای روش‌های موجود ارائه می‌دهد [۸] و [۹]. علاوه بر این، پهپادها اجازه اجرای وظایف طولانی‌مدت، یکنواخت و مکرر بیش از توانایی‌های انسان را می‌دهند و امروزه برای مصارف گوناگون زیست‌محیطی بر فراز جنگل‌ها در حال پرواز هستند. این برتری‌ها در سال‌های اخیر باعث افزایش توجه جهانی به پهپادها برای مقابله با آتش‌سوزی جنگل‌ها شده است [۱]. در فرایند نظارت خودکار بر آتش‌سوزی جنگل‌ها با استفاده از سیستم بینایی می‌توان از راه دور به صحت نتایج گرفته‌شده پی برد و اطلاعات بیشتری نظیر اندازه، حجم، مکان، سرعت و جهت گسترش را نیز برای پیش‌بینی

چکیده: مقابله با آتش‌سوزی جنگل‌ها برای جلوگیری از خطرات بالقوه آنها و همچنین حفاظت از منابع طبیعی به عنوان یک چالش در میان محققان مطرح است. هدف از این تحقیق، تشخیص ویژگی‌های آتش و دود از تصاویر بصری پهپاد برای دسته‌بندی، تشخیص شیء و قطعه‌بندی تصاویر است. از آنجا که جنگل‌ها محیط‌های بسیار پیچیده و غیر ساختاری هستند، استفاده از سیستم بینایی همچنان با مشکلاتی نظیر شباهت ویژگی‌های شعله با نور خورشید، گیاهان و حیوانات و یا پوشش شعله با دود مواجه است که باعث هشدارهای اشتباه می‌شوند. روش پیشنهادی در این تحقیق، استفاده از شبکه‌های عصبی کانولوشنی از روش‌های یادگیری عمیق است که به صورت خودکار، توانایی استخراج یا تولید ویژگی در لایه‌های مختلف خود را دارند. ابتدا به جمع‌آوری داده و افزایش آنها با توجه به روش‌های داده‌افزایی پرداخته شده و در ادامه، استفاده از یک شبکه ۱۲ لایه برای دسته‌بندی و همچنین روش یادگیری انتقالی برای قطعه‌بندی تصاویر پیشنهاد می‌شود. نتایج به دست آمده نشان می‌دهند که روش داده‌افزایی به کار برده شده با توجه به تغییر اندازه و آماده‌سازی تصاویر ورودی به شبکه از کاهش شدید ویژگی‌های موجود در تصاویر اولیه جلوگیری کرده و همچنین شبکه‌های عصبی کانولوشنی مورد استفاده می‌توانند به خوبی ویژگی‌های آتش و دود موجود در تصاویر را استخراج کنند و نهایتاً به تشخیص و محلی‌سازی آنها بپردازند.

کلیدواژه: آتش‌سوزی جنگل، پردازش تصویر، شبکه عصبی کانولوشنی، یادگیری عمیق.

۱- مقدمه

جنگل‌ها کارکردهای مهمی در طبیعت از جمله تصفیه آب، تثبیت خاک، چرخه عناصر غذایی، تعادل آب و هوایی و ذخیره کربن دارند و همچنین زیستگاه‌های حیات وحش را فراهم و محیط‌های طبیعی را در تنوع زیستی غنی می‌کنند. از لحاظ اقتصادی، محصولات صنعتی جنگل‌ها از صدها هزار شغل پشتیبانی و میلیاردها دلار به ثروت اقتصادی کشورها کمک می‌کنند [۱]. اما متأسفانه هر ساله میلیون‌ها هکتار از آنها با آتش‌سوزی نابود و صدها میلیون دلار برای خاموش کردن آنها صرف می‌شود [۲]. همچنین اگرچه آتش‌سوزی‌های کنترل‌شده می‌تواند به تشکیل جنگل‌های جدید کمک کند، مهم است اطمینان حاصل شود که آتش‌سوزی‌ها کنترل‌شده باشد و به سیستم‌های حساس محیط زیست و

این مقاله در تاریخ ۸ آذر ماه ۱۳۹۹ دریافت و در تاریخ ۲۹ آبان ماه ۱۴۰۰ بازنگری شد.

محمدصادق کیهان‌پناه، دانشکده مهندسی مکانیک، دانشگاه تبریز، تبریز، ایران، (email: kayhanpanah95@ms.tabrizu.ac.ir)

بهروز کوهستانی (نویسنده مسئول)، دانشکده مهندسی برق و کامپیوتر، دانشگاه تبریز، تبریز، ایران، (email: b.koohestani@tabrizu.ac.ir)

احساسی که با لمس یا تصویری که با دیدن یک سطح حاصل می‌شود است و معمولاً از طریق زبری، نرمی، سختی، ماتی، براقی و ... یا در یک تصویر با تغییرات محلی مانند رنگ، چرخش و شدت روشنایی مشخص می‌شود. تحقیقات فراوانی برای استخراج بافت از تصاویر صورت گرفته و توصیفگرهای گوناگونی مانند توصیفگر دودویی که به بررسی بافت پیرامون یک پیکسل در تصویر بر اساس اختلاف بین آنها می‌پردازد، معرفی شده‌اند [۱۲]. ویژگی دیگر، هندسه یا همان شکل نامنظم آتش است. با توجه به مرزهای ناهموار آتش در مقایسه با نواحی غیر آتش می‌توان اطلاعاتی درباره رفتار نواحی آتش و غیر آتش به دست آورد [۱۳]. منظور از مقدار ویژگی‌ها در گذر زمان یا ویژگی‌های زمانی، استفاده از دنباله‌های ویدئویی یا فریم‌های متوالی در گذر زمان است. بررسی تغییرات مساحت، تشخیص حرکت و بررسی رفتار فرکانسی نواحی آتش طی فریم‌های متوالی می‌تواند اطلاعات مفیدی از ویژگی‌های منحصر به آتش را بدهد که البته تغییر ناگهانی صحنه و حرکت دوربین مانع عملکرد مناسب این روش‌ها می‌شود. استفاده از این روش‌ها بیشتر برای موارد دوربین ثابت یا دیدبان رایج است. با استخراج مقادیر هر کدام از ویژگی‌ها و با به کار بردن آستانه‌گذاری‌های مختلف می‌توان از آنها به تنهایی یا در الگوریتم‌های مختلف هوش مصنوعی استفاده کرد. شبکه‌های عصبی مصنوعی از الگوریتم‌های هوش مصنوعی هستند که مدل‌های ریاضی و نرم‌افزاری متعددی را با الهام گرفتن از مغز انسان پیشنهاد می‌دهند و برای حل گستره وسیعی از مسایل علمی، مهندسی و کاربردی، در حوزه‌های مختلف کاربرد دارند. کارکرد اصلی این گونه شبکه‌ها، پردازش داده‌ها و اطلاعات به منظور یادگیری و ایجاد دانش است. شبکه‌های عصبی با دریافت بردارهای ویژگی مرتبط با نمونه‌های آموزشی در مرحله آموزش، توانایی یادگیری و دسته‌بندی نمونه‌های جدید را پیدا و در مرحله بعدی یا مرحله آزمایشی از مقادیر یا وزن‌های به دست آمده برای تعیین دسته نمونه جدید استفاده می‌کنند. همان گونه که ذکر شد این نوع از شبکه‌های عصبی مصنوعی ابتدا نیاز به استخراج ویژگی از تصاویر اولیه دارند اما نوع دیگر یادگیری که یادگیری عمیق نامیده می‌شود و از تکنیک‌های آن می‌توان به الگوریتم‌های شبکه‌های عصبی کانولوشنی^۱ (CNN) اشاره کرد نیز وجود دارد که به صورت خودکار، توانایی استخراج یا تولید ویژگی در لایه‌های مختلف خود را دارند و در واقع به کارگیری شبکه‌های عصبی بر روی حجم بسیار زیادی از داده‌ها می‌باشند.

شبکه عصبی کانولوشنی (CNN) از روش‌های یادگیری عمیق است که از ساختار بینایی موجودات زنده الهام گرفته شده و بیشتر از آن برای تحلیل تصاویر بصری استفاده می‌شود. CNN با یادگیری فیلترها، از طریق شیوه‌های مبتنی بر داده‌ها به صورت محلی به مانند غشای بینایی^۲ (پردازشگر اطلاعات بصری در مغز) نورون‌های درون خود را با استفاده از محرک‌های تولیدشده در نواحی محلی فعال می‌کند [۱۴]. CNN‌ها نسبت به سایر الگوریتم‌های طبقه‌بندی تصاویر از پیش‌پردازش نسبتاً کمتری استفاده می‌کنند، به این معنی که دیگر نیاز به ساختن جدای فیلترها یا ویژگی‌ها برای یادگیری نیست و این استقلال از تلاش انسانی برای طراحی ویژگی‌ها خود یک مزیت برای این گونه از شبکه‌ها حساب می‌شود. CNN‌ها به مانند شبکه‌های عصبی معمولی از سلول‌های عصبی ساخته شده‌اند و قابلیت یادگیری وزن‌ها و بایاس‌ها^۳ را که در اینجا همان

آتش فراهم کرد. سیستم بینایی به انواع استفاده از دوربین بصری و دوربین مادون قرمز، استفاده بر روی دکل دیدبانی به صورت ثابت یا پهناد به صورت متحرک و همچنین استفاده از تصاویر در روش‌های پردازش تصویر، استخراج ویژگی و استفاده از این ویژگی‌ها در شبکه‌های عصبی و یا بدون استخراج ویژگی تقسیم می‌شود. اساس کار دوربین‌های مادون قرمز برای استخراج ویژگی دما و در ادامه، پردازش تصویر یا استفاده از الگوریتم‌های هوش مصنوعی، اندازه‌گیری انرژی منتشرشده از ماده مشتعل طول موج آن است که انرژی منتشرشده از آتش بسته به ماده مشتعل دارای طیف طول موج خاصی است [۱۰]. در حالی که چشم انسان انرژی را در محدوده طول موج کوتاه $0.3 \mu\text{m}$ تا $0.7 \mu\text{m}$ تشخیص می‌دهد، انرژی آتش‌سوزی در طول موج بالاتر، به طور اسمی $2 \mu\text{m}$ تا $14 \mu\text{m}$ رخ می‌دهد [۱۰]. از رایج‌ترین طول موج‌های استفاده‌شده برای تشخیص انواع آتش، طیف طول موج $3 \mu\text{m}$ تا $5 \mu\text{m}$ و $8 \mu\text{m}$ تا $14 \mu\text{m}$ است [۱۰]. تشخیص‌دهنده‌ها یا حسگرهایی که در طول موج‌های کمتر کار می‌کنند، بیشتر انرژی منعکس شده یا بازتاب‌شده از جسم را نمایش می‌دهند که محدودیت استفاده از این سیستم‌ها را بیان می‌کند [۱۰]. از معایب استفاده از دوربین‌های مادون قرمز، تشخیص حرارت از روی انرژی موج بازتاب‌شده از جسم به جای انرژی موج منتشرشده درون آنها است و در نتیجه سطوح بسیار بازتابنده مثل صخره‌ها و سنگ‌ها، آب، برف، حیوانات، زمین بدون پوشش گیاهی، سطح براق برگ‌ها و ... می‌توانند با تابش خورشید در طول روز باعث ایجاد هشدارهای آتش‌سوزی اشتباه شوند. همچنین دوربین‌های کوچک مادون قرمز مورد استفاده هنوز حساسیت کمی دارند که برای تولید تصاویر باکیفیت، باعث ایجاد دوره نظارت طولانی‌تری برای پهنادها هستند و لرزش پهنادها هم باعث ایجاد تصاویر محو می‌شود [۱]. برای کاهش هشدارهای نادرست در دوربین‌های مادون قرمز می‌توان از طیف‌های طول موج مختلف از جمله طیف بالایی طول موج برای شفاف کردن دود و دیدن پس‌زمینه استفاده کرد، چرا که دود به مانند دوربین‌های بصری مانع دیدن آتش در طیف‌های طول موج بالای دوربین‌های مادون قرمز نیست. همچنین از تلفیق دوربین‌های مادون قرمز با دوربین‌های بصری برای استفاده از دیگر ویژگی‌های آتش (صرف نظر از هزینه بیشتر) نیز می‌توان استفاده کرد و نتایج دقیق‌تر و مطمئن‌تری به دست آورد.

با استفاده از دوربین‌های بصری یا همان دوربین‌های ارزان قیمت دیجیتال و پردازش تصویر می‌توان به استخراج ویژگی‌های رنگ، بافت و هندسه یا به اصطلاح مکانی منحصر به آتش (یا دود) و همچنین مقدار آنها در گذر زمان پرداخت و در ادامه از آنها به تنهایی یا در الگوریتم‌های مختلف هوش مصنوعی استفاده کرد. رنگ آتش از ویژگی‌های کلیدی تشخیص آن است که بیشتر از آن برای جداسازی آتش از پس‌زمینه استفاده می‌شود و وابسته به خصوصیات ماده در حال احتراق شامل زرد و قرمز، آبی، سبز و ... است که در مورد آتش‌سوزی جنگل‌ها می‌توان آن را در محدوده زرد و قرمز در نظر گرفت. برای استخراج ویژگی رنگ از فضاهای رنگی مختلف استفاده می‌شود. فضای رنگ RGB از فضاهای رنگ اولیه و رایج در وسایل نمایش بصری است که از ترکیب مقادیر اصلی قرمز، سبز و آبی که برای هر پیکسل تصویر هر کدام از آنها در بازه ۰ تا ۲۵۵ است به دست می‌آید. از معایب این فضا می‌توان به حساس نبودن به تغییرات شدت روشنایی و عدم تفکیک ویژگی رنگ و روشنایی، عدم تولید کلیه رنگ‌ها از ترکیب این سه رنگ اصلی و منفی شدن مقادیر رنگ و وابستگی این فضا به وسیله نمایش اشاره کرد [۱۱]. از ویژگی‌های دیگر آتش، بافت آن است. بافت در واقع همان

1. Convolutional Neural Network
2. Visual Cortex
3. Bias

مشخصه اصلی تصاویر انتخاب‌شده، زاویه دید پهنای آنها در پرواز بر فراز جنگل‌ها بود. این تصاویر در ساعت‌های مختلف شبانه‌روز و زاویه تابش خورشید، فصل‌های مختلف سال یا در انواع جنگل‌ها و پوشش گیاهی و جانوری آنها می‌تواند مختلف باشد. متأسفانه مجموعه داده استاندارد برای مقایسه شبکه‌ها با معماری‌های مختلف برای تشخیص آتش‌سوزی به کمک پهپادها وجود ندارد و به نظر می‌رسد که چالش اصلی برای حل این مسئله، استفاده از کمیت و کیفیت مناسب داده‌های آموزشی و آزمایشی است. با توجه به زاویه متفاوت تصاویر آتش گرفته شده بر روی زمین و تصاویر پهپادها، استفاده از داده‌های آموزشی استاندارد موجود تشخیص آتش، چندان گره‌گشا به نظر نمی‌رسد. برای حل مشکل کمبود داده‌ها می‌توان از روش‌های داده‌افزایی (قرینه، چرخش، برش تصادفی، تغییر رنگ، اضافه کردن نویز، هدررفت اطلاعات، تغییر درخشندگی و ...) استفاده کرد. در واقع روش‌های داده‌افزایی، روش‌هایی هستند که با به کار بردن آنها ویژگی‌های موجود در تصاویر حذف نمی‌شوند. علاوه بر مهم‌بودن زاویه دید تصاویر آتش، مسئله دیگر گم‌نبودن آتش و قابل تشخیص بودن آن در بین تل دود است. در واقع دود، همان طور که یک نشانه برای وجود آتش است، با احاطه آتش می‌تواند باعث حذف مشخصه‌های آن نیز شود. با توجه به این که معماری شبکه استفاده‌شده، اندازه تصاویر ورودی را مشخص می‌کند (که همه باید یا به یک اندازه مشخص باشند یا در بعضی از شبکه‌های خاص از اندازه‌های متفاوت تصاویر نیز می‌توان استفاده کرد)، تغییر اندازه تصاویر نیز می‌تواند باعث حذف ویژگی آتش شود. شکل ۱ نمونه‌ای از این حذف ویژگی را نشان داده است.

با توجه به هدف دسته‌بندی، تشخیص شیء و قطعه‌بندی تصاویر، مرحله بعد جمع‌آوری تصاویر و جداسازی داده‌های آموزشی، برچسب‌گذاری روی تصاویر است. در دسته‌بندی به شکل کلی باید تعیین کنیم که تصویر مربوط به کدام دسته است، حال آن که در تشخیص شیء و قطعه‌بندی تصاویر باید حتماً پیکسل مربوط به هر دسته در هر تصویر مشخص شود. می‌توان برای برچسب‌زنی از روش‌های مختلف پردازش تصویر استفاده کرد، اما با توجه به مشخصه‌های نه چندان مشخص دود و استفاده الگوریتم‌های مطرح تشخیص دود از تشخیص حرکت و دسترسی ما به تصاویر و نه فیلم‌های هوایی از آتش‌سوزی جنگل‌ها، برچسب‌زنی برای هر تصویر به صورت جدا و دستی انجام شد. پیکسل‌های تصاویر به دست آمده به سه بخش آتش، دود و پس‌زمینه (یا بدون آتش و دود) جدا شدند. برای جلوگیری از حذف ویژگی‌ها و همچنین افزایش داده‌های موجود از تصاویر با اندازه‌های بزرگ و تعداد بیشتر و گسسته ناحیه آتش، هر ناحیه از تصاویر برچسب زده شده را یک بار با در نظر گرفتن آتش و بار دیگر با در نظر گرفتن دود، با طول و عرض برابر جدا کردیم. در شکل ۲ یک نمونه از تصاویر ساخته‌شده آمده است.

تقریباً به اندازه تصاویر دسته‌بندی شده در دسته آتش، تصاویر با دسته عدم وجود آتش نیز به پایگاه داده اضافه گردید و از هر دسته به میزان برابر و به صورت تصادفی از پایگاه داده برای مرحله آزمایش نهایی آموزش انجام‌شده جدا کردیم. همچنین میزان استفاده از داده‌های اعتبارسنجی در مرحله آموزش ۲۰ درصد از داده‌های آموزشی می‌باشد. در جدول ۱ مقادیر داده‌های پایگاه داده آورده شده است.

۳- شبکه عصبی کانولوشنی پیشنهادی

مرحله بعد از آماده‌سازی تصاویر ورودی به شبکه، آموزش شبکه مورد نظر می‌باشد. در شکل ۳ ساختار کلی شبکه ۱۲ لایه پیشنهادی با ورودی یکی از تصاویر به دست آمده در شکل ۲ نشان داده شده است.

ویژگی‌ها می‌باشند، دارا هستند. هر نورون تعدادی ورودی دریافت و سپس حاصل ضرب وزن‌ها در ورودی‌ها را محاسبه کرده و در انتها با استفاده از یک تابع تبدیل (فعال‌سازی) مشتق‌پذیر غیر خطی نتیجه‌ای را ارائه می‌دهد. ورودی‌های اولیه، تصاویر سه‌بعدی تشکیل‌شده از طول، عرض و عمق هستند که در طول شبکه با نرخ مشخص یادگیری و اندازه مومنتوم^۱ (یا میزان تبعیت به روز رسانی‌ها از الگوهای قبلی) با حرکت رو به جلو و کاهش پارامترهای طول و عرض و افزایش پارامتر عمق، به تدریج باعث کاهش پارامترهای مربوط و افزایش (یا استخراج) ویژگی‌ها می‌شوند. در واقع تصاویر از ویژگی‌های سطح پایین مانند لبه به ویژگی‌های سطح بالاتر نظیر بافت و در آخر به دسته‌بندی مورد نظر می‌رسند. در ادامه، فرایند آموزش بعد از به دست آمدن امتیاز ورودی‌ها با محاسبه مقدار تابع هزینه برای بهینه‌سازی با روش پس‌انتشار و محاسبه گرادیان کاهشی^۲ به عقب برمی‌گردد تا وزن‌ها و بایاس‌ها به روز شوند.

در سال ۲۰۱۶، فریزی و همکاران [۱۵] یک شبکه عصبی کانولوشنی ۷ لایه با ورودی تصاویر RGB را برای دسته‌بندی تصاویر به سه دسته آتش، دود و غیر آتش و دود پیشنهاد دادند. نتایج به دست آمده با توجه به مشکلات استخراج ویژگی در دیگر روش‌ها امیدوارکننده و قابل قبول بود. همچنین در سال ۲۰۱۸، ژاو و همکاران [۱۶] برای جنگل‌های با پوشش‌دهی ضعیف شبکه‌های اینترنتی و مشکل ارسال اطلاعات به ایستگاه زمینی، برنامه فایرنِت^۳ شامل یک شبکه عصبی کانولوشنی عمیق ۱۱ لایه را برای استفاده بر روی پهپادهای سنتی نظارت بر آتش‌سوزی جنگل‌ها پیشنهاد دادند. در سال ۲۰۲۰ نیز پن و همکاران [۱۷] یک روش تشخیص آتش‌سوزی مبتنی بر یادگیری انتقالی را پیشنهاد دادند که با استفاده از تصاویر با وضوح بسیار بالای دوربین‌های دیدبان نظارتی موجود و تشکیل پنجره‌های کوچک‌تر، مکان دود را به دست می‌آورد. روش پیشنهادی با اعمال بر روی دنباله‌های ویدئویی، وجود دود را در همه ویدئوها تشخیص داده است اما به دلیل اضافه‌کردن ویژگی‌های ابر به شبکه برای کاهش هشدارهای اشتباه، در مواردی از فریم‌ها که دود بزرگ است به اشتباه ابر تشخیص می‌دهد.

در ادامه این تحقیق، از روی تصاویر هوایی گرفته‌شده از پهپاد به تشخیص وجود آتش و در صورت وجود محلی‌سازی آن پرداخته شده است. در بخش ۲ به نحوه تشکیل پایگاه داده، اعمال روش داده‌افزایی به کار برده شده و برچسب‌گذاری تصاویر ورودی به شبکه پرداخته شده و در ادامه در بخش ۳، معماری شبکه عصبی کانولوشنی پیشنهادی آمده است. در بخش ۴، نتایج عددی به دست آمده و توانایی شبکه در تشخیص وجود آتش و محلی‌سازی آن مورد بررسی قرار گرفته و نهایتاً در بخش ۵، نتیجه‌گیری ارائه شده است.

۲- ایجاد پایگاه داده

برای تشکیل پایگاه داده، از تصاویر خبری موجود در اینترنت با جستجو در گوگل استفاده کردیم. تعداد بیش از ۱۰۰۰ تصویر با اندازه‌های متفاوت از 300×200 تا 4000×3000 شامل آتش، دود، هر دو یا هیچ کدام به دست آمد. داده‌های آموزشی باید به تعداد مناسب و مشابه همان چیزی که بعد از یادگیری از شبکه انتظار داریم تا نتیجه حاصل از داده جدید یا مجموعه آزمایشی را به دست آورد، باشد. با توجه به هدف یادگیری،

1. Momentum
2. Gradient Decent
3. FireNet



شکل ۱: حذف ویژگی‌های آتش با تغییر اندازه تصویر.

در (۱) تا (۵)، β_1 مومتوم، β_2 گشتاور دوم، η نرخ یادگیری و ε برای جلوگیری از صفرشدن مخرج به کار رفته است. در واقع با محاسبه گرادینان تابع فعال‌ساز هر نورون یا در CNNها هر یک از ویژگی‌ها، مقادیر m و v به دست می‌آید و در ادامه به دلیل صفربودن m و v در ابتدا و نیاز به اصلاح داشتن در مراحل ابتدایی، مقادیر تصحیح \hat{m} و \hat{v} محاسبه شده تا نهایتاً مقادیر جدید وزن‌ها به روز شوند.

نوع دیگر استفاده از شبکه‌های کانولوشنی، تشخیص شیء و قطعه‌بندی روی هر تصویر است که برای تشخیص آتش، روشی برای محلی‌سازی آتش بر روی تصاویر می‌باشد. در این روش علاوه بر ویژگی‌های شیء در هر تصویر، خطوط جداکننده شیء برای روش تشخیص شیء و یا دسته هر پیکسل برای روش قطعه‌بندی، به شبکه اعمال و آموزش روی آنها نیز انجام می‌گیرد. با توجه به کمبودن داده‌های آزمایشی برای این مرحله، می‌توان از شبکه‌های کانولوشنی مشخص و آموزش دیده شده برای تشخیص دسته‌های مختلف یا یادگیری انتقالی و همچنین رابط برنامه‌نویسی کاربردی (API) ^۲ (TensorFlow) ^۳ [۱۸] استفاده کرد. استفاده در این روش به این صورت است که لایه آخر از شبکه آموزش دیده شده با داده‌های متنوع و فراوان مربوط به دسته‌های مجزا از هم، جدا شده و دسته‌بندی دلخواه که در اینجا سه دسته آتش، دود و غیر آتش و دود است به شبکه اضافه و با داده‌های موجود، شبکه دوباره آموزش داده می‌شود. در

جدول ۱: مقادیر داده‌های پایگاه داده.

دود	آتش	جداسازی بر پایه تصاویر اولیه آتش
۵۰۴		
۹۱۶	۲۱۱۰	تصاویر به دست آمده
۷۹۱	۱۹۸۵	تصاویر جدا شده برای آموزش
۱۲۵	۱۲۵	تصاویر جدا شده برای آزمایش نهایی

یکی از کاربردی‌ترین بهینه‌سازهای استفاده شده در شبکه‌های کانولوشنی و مورد استفاده در این تحقیق، بهینه‌ساز تخمین لحظه تطبیقی (Adam) ^۱ است.

$$m_w^{(t+1)} = \beta_1 m_w^{(t)} + (1 - \beta_1) \nabla_w L^{(t)} \quad (۱)$$

$$v_w^{(t+1)} = \beta_2 v_w^{(t)} + (1 - \beta_2) (\nabla_w L^{(t)})^2 \quad (۲)$$

$$\hat{m}_w = \frac{m_w^{(t+1)}}{1 - \beta_1^{t+1}} \quad (۳)$$

$$\hat{v}_w = \frac{v_w^{(t+1)}}{1 - \beta_2^{t+1}} \quad (۴)$$

$$w^{(t+1)} = w^{(t)} + \eta \frac{\hat{m}_w}{\sqrt{\hat{v}_w} + \varepsilon} \quad (۵)$$



شکل ۲: جداسازی تصویر بر پایه آتش و دود.

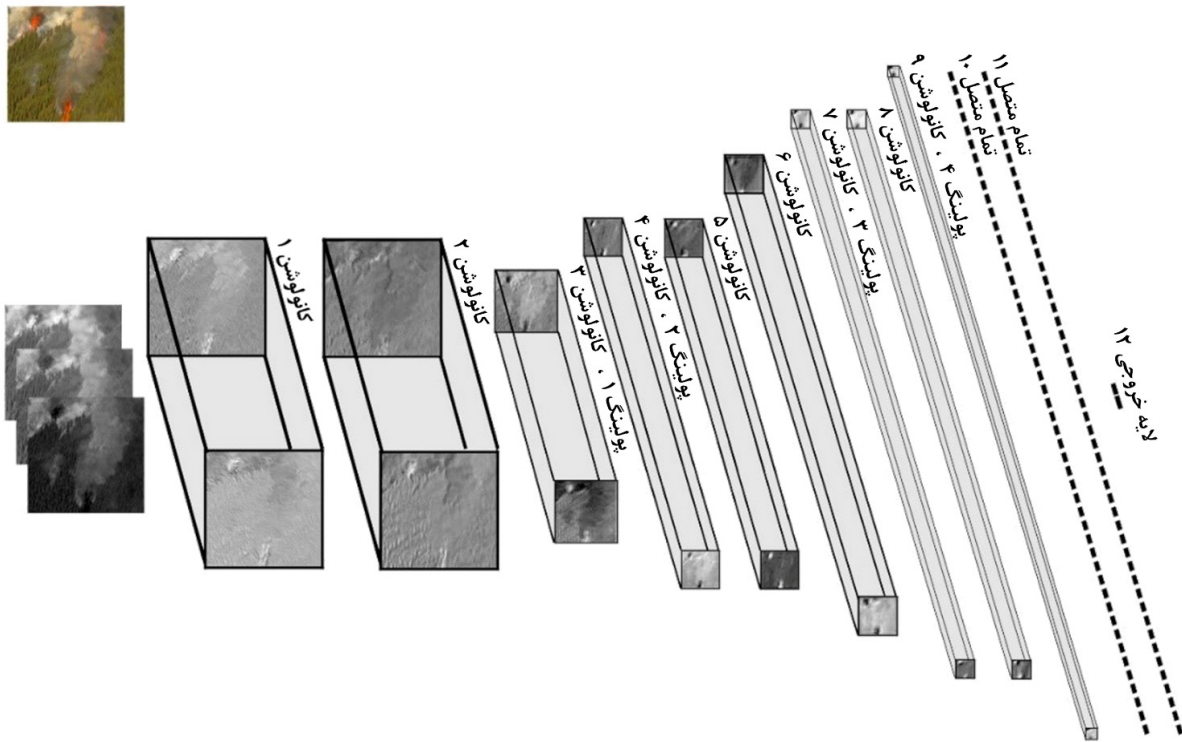
دو معیار سرعت و میانگین متوسط دقت (mAP)^۳ که در واقع ارائه شده برای درک حساسیت شبکه به اشیای مطلوب و میزان دقت در اجتناب از شناسایی موارد نادرست و برابر با ضرب دقت در صحت است، مدل «mask_rcnn_inception_v2_coco» [۱۹] را با سرعت بسیار کمتر و دقت کمی کمتر از میان ۴ مدل پیشنهادی دیگر انتخاب کردیم که در بخش بعد نتایج به دست آمده نشان داده شده است.

عملاً اتفاقی که در شبکه‌های عصبی کانولوشنی رخ می‌دهد، همان محاسبات ساده ریاضی و به طور خاص ماتریسی در مقیاس زیاد است و بنابراین استفاده از پردازنده‌ای که بتواند عملیات نه چندان پیچیده و ساده اما زیاد را در مدت زمان کمتری انجام دهد درخور اهمیت است. واحد پردازش گرافیکی^۴ (GPU) نسبت به واحد پردازش مرکزی^۵ (CPU) از

واقع در این روش وزن‌ها و بایاس‌ها یا پارامترهای اولیه نه به صورت تصادفی و با مقدار کم که از قبل در شبکه دیگر آموزش داده شده و با استخراج ویژگی‌های مختلف، این بار با داده‌های جدید فرایند آموزش را می‌سازند. API به کار رفته تنسورفلو بر روی مجموعه داده اشیای رایج در زمینه^۱ (COCO) تمرین داده شده است. مجموعه داده COCO، تصاویر ۳۰۰ کیلوبایتی از ۹۰ شیء پر استفاده در تصاویر روزمره است. همچنین چارچوب شبکه‌های به کار رفته در API تنسورفلو، مناطق ویژگی CNN^۲ (R-CNN) با ایده ایجاد پیشنهادهای چندگانه و استخراج ویژگی‌ها از هر یک با استفاده از یک CNN و سپس دسته‌بندی هر ناحیه کاندیدا است. در واقع R-CNN ابتدا تصویر را برای یافتن کادرهای محصورکننده مربوط بالقوه قطعه‌بندی می‌کند و سپس الگوریتم شناسایی را برای یافتن محتمل‌ترین اشیا در این کادرهای محصورکننده اجرا می‌کند. با توجه به

3. Mean Average Precision
4. Graphics Processing Unit
5. Central Processing Unit

1. Common Object in Context
2. Regions with CNN Features



شکل ۳: ساختار کلی شبکه ۱۲ لایه پیشنهادی برای مرحله دسته‌بندی.

کرده است. میکروکنترلرها دستگاه‌های محاسباتی کوچک و ارزانی هستند که از سخت‌افزارهای مورد نیاز برای انجام کار مشخص تشکیل شده‌اند. با توجه به مزیت امکان جدابودن فرایند آموزش و آزمایش در یادگیری ماشین می‌توان ابتدا در یک سخت‌افزار مناسب به فرایند زمان‌بر آموزش پرداخت و سپس فرایند آزمایش را از روی مدل به دست آمده بر روی میکروکنترلر متصل به پهنپا با حداقل سخت‌افزارهای مورد نیاز انجام داد.

۴- نتایج تجربی

۴-۱ دسته‌بندی

مدل پیشنهادی این تحقیق برای مرحله دسته‌بندی، شبکه ۱۲ لایه تشکیل شده از ۹ لایه کانولوشنی، ۵ لایه ادغام و ۳ لایه تمام‌متصل با ورودی تصاویر با اندازه‌های $3 \times 128 \times 128$ می‌باشد که از تکرار و آزمایش‌های مختلف به صورت سعی و خطا روی دیگر شبکه‌ها با در نظر گرفتن ۲ مقدار دقت و سرعت انتخاب شده است. در جدول ۲ تأثیر مقدار دورریزی یا حذف تصادفی برای نورون‌های لایه‌های تمام‌متصل نشان داده شده است. همان طور که دیده می‌شود، مقدار 0.6 حذف تصادفی، دقت بیشتر داده‌های تأیید را به دنبال دارد. همچنین در جدول ۳ مشخصات مدل پیشنهادی با جزئیات بیشتر آمده است.

در شکل ۴ نمودار دقت و هزینه داده‌های آموزشی و اعتبار برای مدل پیشنهادی در ۱۰۰ دوره اول آورده شده است. همان طور که ملاحظه می‌شود دقت بالا و نبود اختلاف زیاد هزینه داده‌های آموزشی با داده‌های اعتبار، نشان از برآزش مناسب و آموزش خوب شبکه است.

ماتریس درهم‌ریختگی^۷ آورده شده در جدول ۴ نتایج حاصل را بر اساس اطلاعات واقعی موجود نمایش می‌دهد. پارامتر صحت^۸ متداول‌ترین، اساسی‌ترین و ساده‌ترین معیار اندازه‌گیری کیفیت یک

تعداد هسته بیشتر اما کندتر تشکیل شده و به همین منظور برای انجام عملیات‌های موازی نسبت به سری، گزینه مناسب‌تری در یادگیری عمیق به حساب می‌آید. در میان کارت‌های گرافیک مختلف نیز کارت‌های گرافیک شرکت انویدیا^۱ به دلیل پشتیبانی تمام کتابخانه‌ها و چارچوب‌ها از آن و مبتنی بودن بر کودا^۲ (CUDA) و سی‌بودی‌ان‌ان^۳ (cuDNN) گزینه اصلی می‌باشند. cuDNN یک کتابخانه سطح بالا برای کار با شبکه‌های عصبی مصنوعی عمیق است که به عنوان یک شتاب‌دهنده استفاده می‌شود و در بستر نرم‌افزاری CUDA که توسط شرکت انویدیا برای محاسبات موازی در کارت‌های گرافیک ایجاد شده است، توسعه یافته است. در این تحقیق از کارت گرافیک GTX 960M شرکت انویدیا با قابلیت مبتنی‌بودن بر CUDA و در برنامه‌نویسی‌ها از چارچوب تنسورفلو در برنامه پایتون^۴ و محیط ژوپیتر^۵ استفاده کردیم. چارچوب تنسورفلو یک کتابخانه نرم‌افزاری متن‌باز برای یادگیری ماشین است که توسط تیم تحقیقاتی و محصولات مختلف گوگل استفاده می‌شود. تنسورفلو در آغاز توسط تیم پژوهشی هوش مصنوعی گوگل به صورت داخلی استفاده می‌شد ولی بعدها در اواخر سال ۲۰۱۵ منتشر گردید. دلایل زیادی نظیر سهولت مدیریت وزن‌ها و مقادیر نورون‌ها در لایه‌های مختلف و ایجاد داده‌های چندبعدی موجب شدند تا در طراحی تنسورفلو از ساختار تنسور برای نگهداری مقادیر و نوع داده متغیرها استفاده شود. از دلایل استفاده از تنسورفلو به جای چارچوب‌های دیگر، تفاوت معنادار میزان امتیاز و استفاده توسعه‌دهندگان دیگر در گیت‌هاب^۶ از این چارچوب است. همچنین چارچوب تنسورفلو، امکان استفاده بر روی میکروکنترلرها را نیز فراهم

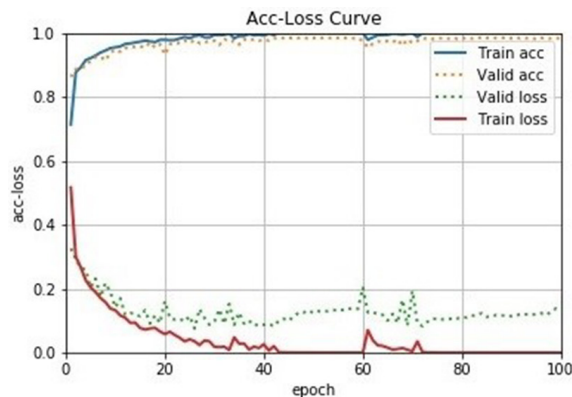
1. NVIDIA
2. Compute Unified Device Architecture
3. CUDA Deep Neural Network
4. Python
5. Jupyter
6. Github

7. Confusion Matrix

8. Accuracy

جدول ۲: تأثیر حذف تصادفی.

حذف تصادفی	دقت داده‌های تأیید
۰٫۲	۹۸٫۲
۰٫۳	۹۷٫۷
۰٫۴	۹۸٫۳
۰٫۵	۹۸٫۰
۰٫۶	۹۸٫۵
۰٫۷	۹۸٫۱
۰٫۸	۹۸٫۲



شکل ۴: نمودار دقت هزینه مدل پیشنهادی.

جدول ۳: شبکه کانولوشنی پیشنهادی برای دسته‌بندی تصاویر.

نوع لایه	تعداد ویژگی‌ها	اندازه کرنل ^۱	اندازه لغزش	تابع فعال‌سازی
کانولوشن ۱	۳۲	۳×۳	۱	رلو ^۲ (ReLU)
کانولوشن ۲	۳۲	۳×۳	۱	ReLU
پولینگ ^۳ ۱	۳۲	۲×۲	۲	-
کانولوشن ۳	۳۲	۳×۳	۱	ReLU
پولینگ ^۴ ۲	۳۲	۲×۲	۲	-
کانولوشن ۴	۶۴	۳×۳	۱	ReLU
کانولوشن ۵	۶۴	۳×۳	۱	ReLU
کانولوشن ۶	۱۲۸	۳×۳	۱	ReLU
پولینگ ^۵ ۳	۱۲۸	۲×۲	۲	-
کانولوشن ۷	۲۵۶	۳×۳	۱	ReLU
کانولوشن ۸	۲۵۶	۳×۳	۱	ReLU
پولینگ ^۶ ۴	۲۵۶	۲×۲	۲	-
کانولوشن ۹	۵۱۲	۳×۳	۱	ReLU
پولینگ ^۷ ۵	۵۱۲	۲×۲	۲	-
تمام‌متصل ۱۰	۵۱۲	حذف تصادفی: ۰٫۶	-	ReLU
تمام‌متصل ۱۱	۵۱۲	حذف تصادفی: ۰٫۶	-	ReLU
لایه خروجی ۱۲	۲	-	-	سافت‌مکس ^۴

تابع هزینه: کراس انتروپی^۵
 نرخ یادگیری: ۰٫۰۰۰۱
 مجموع پارامترها:
 زمان اجرای هر نمونه:
 بیشترین دقت داده‌های اعتبار در ۱۰۰ دوره^۶

- Kernel
- Rectified Linear Unit
- Pooling
- Softmax
- Cross Entropy
- Epoch

مثبت درست (TPR)^۳، به معنی نسبتی از موارد مثبت است که آزمایش، آنها را به درستی به عنوان نمونه مثبت تشخیص داده و معیاری است که مشخص می‌کند دسته‌بندی، به چه اندازه در تشخیص آتش‌سوزی‌ها موفق بوده است. همان گونه که مشخص است، تعداد شرایط عادی‌ای که به اشتباه به عنوان آتش‌سوزی تشخیص داده شده‌اند، هیچ تأثیری در محاسبه این پارامتر ندارد و در واقع زمانی از این پارامتر به عنوان پارامتر

دسته‌بندی است و عبارت است از میزان تشخیص صحیح در مجموع دو دسته. این پارامتر در واقع نشانگر میزان الگوهایی است که درست تشخیص داده شده‌اند. پارامتر دقت^۱، مقدار پیش‌بینی شده آتش‌سوزی است که برای بیان نسبت پاسخ درست در دسته آتش و دود استفاده می‌شود و بیان‌کننده این است که چند درصد از آتش‌سوزی‌ها در واقعیت هم آتش‌سوزی هستند. معیار حساسیت و فراخوانی^۲ یا نرخ پاسخ‌های

- Precision
- Recall



شکل ۵: نمونه‌ای از تصاویر، (الف) مثبت صحیح با وجود هر دو ویژگی آتش و دود، (ب) مثبت صحیح با وجود ویژگی دود، (ج) منفی صحیح و (د) اشتباه تشخیص داده شده.

جدول ۴: عملکرد نهایی بر روی تصاویر آزمایشی.

مثبت صحیح	منفی کاذب	مثبت کاذب	منفی صحیح	صحت	دقت	حساسیت	خاصیت
۲۴۱	۴	۹	۲۴۶	۹۷/۴	۹۶/۴	۹۸/۴	۹۶/۵

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$Specificity = \frac{TN}{TN + FP} \quad (9)$$

ارزیابی استفاده می‌شود که هدف، دستیابی به نهایت صحت در تشخیص نمونه‌های کلاس آتش‌سوزی است. در نقطه مقابل این پارامتر هم که شرایط عادی اهمیت دارد از پارامتر خاصیت و ویژگی^۱ یا نرخ پاسخ‌های منفی درست^۲ (TNR) استفاده می‌شود [۲۰] تا [۲۲]

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

در (۶) تا (۹) مقادیر صحت، دقت و حساسیت و در جدول ۴ نیز مقادیر ماتریس درهم‌ریختگی و معیارهای بررسی شده آمده است. در شکل ۵ چند نمونه از تصاویر صحیح و اشتباه تشخیص داده شده از تصاویر آزمایشی آورده شده است.

1. Specificity
2. True Negative Rate

جدول ۵: مقایسه شبکه پیشنهادی با دیگر شبکه‌ها.

نام شبکه	معماری شبکه
زمان اجرای هر نمونه بیشترین دقت داده‌های اعتبار در ۱۰۰ دوره	ک: کانولوشن، پ: پولینگ و ت: تمام متصل
شبکه ۸ لایه	ورودی $128 \times 128 \times 3$ ک ۱پ ۱ک ۲پ ۲ک ۲پ
۲ میلی ثانیه	ک ۳پ ۴ک ۳پ ۵ک ۴پ ۴ت ۶ت ۷لایه خروجی ۸
۹۷,۵	بهینه‌ساز Adam و نرخ یادگیری ۰,۰۰۰۱
شبکه مشابه با شبکه الکسننت ^[۲۳]	ورودی $113 \times 113 \times 3$ ک ۱پ ۱ک ۱پ ۱ک ۲پ ۲ک
۴ میلی ثانیه	ک ۳پ ۴ک ۳پ ۵ک ۴پ ۴ت ۶ت ۷لایه خروجی ۸
۹۶,۷	بهینه‌ساز SGD و نرخ یادگیری ۰,۰۱ مومنوم ۰,۹ و کاهش وزن ۰,۰۰۰۵
شبکه پیشنهادی	ورودی $128 \times 128 \times 3$ ک ۱پ ۲ک ۲پ ۳ک ۲پ ۳ک ۴ک ۵ک
۵ میلی ثانیه	ک ۳پ ۴ک ۳پ ۵ک ۴پ ۴ت ۵ت ۱۰ت ۱۱لایه خروجی ۱۲
۹۸,۵	بهینه‌ساز Adam و نرخ یادگیری ۰,۰۰۰۱
شبکه فایرننت [۱۶]	ورودی $128 \times 128 \times 3$ ک ۱پ ۲ک ۲پ ۳ک ۳پ ۴ک
۴ میلی ثانیه	ک ۳پ ۴ک ۳پ ۵ک ۴پ ۴ت ۹ت ۱۰لایه خروجی ۱۱
۹۸,۵	بهینه‌ساز گرادیان کاشی تصادفی ^۲ (SGD) نرخ یادگیری ۰,۰۰۰۱ و مومنوم ۰,۹
(در ۱۰۰۰ دوره)	برای بهینه‌ساز Adam
۹۷,۵	ورودی $128 \times 128 \times 3$ ک ۱پ ۲ک ۲پ ۳ک ۳پ ۴ک
شبکه ۱۴ لایه	ک ۳پ ۴ک ۳پ ۵ک ۴پ ۴ت ۷ت ۸ک ۹ک ۱۰ک ۱۱پ ۱۱ک
۸ میلی ثانیه	ک ۳پ ۴ک ۳پ ۵ک ۴پ ۴ت ۱۲ت ۱۳لایه خروجی ۱۴
۹۸,۷	بهینه‌ساز Adam و نرخ یادگیری ۰,۰۰۰۱

- AlexNet
- Stochastic Gradient Descent

جدول ۷: عملکرد نهایی قطعه‌بندی روی تصاویر آزمایشی با جدا در نظر گرفتن آتش و دود.

صحت	دقت	حساسیت
۷۷,۴	۸۱,۵	۵۵,۱
۷۲,۷	۹۲,۷	۵۴,۸

حذف می‌گردند. سپس از میان گزینه‌های موجود، موارد تکراری حذف و بیشترین IoU انتخاب می‌شود و نهایتاً از تطابق گزینه‌های به دست آمده و مورد انتظار، ماتریس درهم‌ریختگی تشکیل می‌شود.
در شکل ۶ تصاویر نشان داده شده در شکل ۵، با اعمال شبکه قطعه‌بندی پیشنهادگر دیده، آمده است.

۵- نتیجه‌گیری

در این تحقیق ابتدا به تشکیل یک پایگاه داده از تصاویر آتش سوزی جنگل‌ها با زاویه دید پهچاد و آماده‌سازی آنها پرداخته شده و در ادامه با پیشنهاد دو شبکه، یکی شبکه ۱۲ لایه برای دسته‌بندی و دیگری استفاده از روش یادگیری انتقالی برای قطعه‌بندی تصاویر ورودی، دسته و ناحیه آتش و دود تصاویر ورودی تشخیص داده شده است. با توجه به نتایج و دقت‌های به دست آمده از شبکه‌های پیشنهادی می‌توان نتیجه گرفت که روش آماده‌سازی یا داده‌افزایی به کار برده شده از کاهش شدید ویژگی به دلیل تغییر اندازه مستقیم جلوگیری کرده و استفاده از شبکه‌های عصبی کانولوشنی می‌تواند به خوبی ویژگی‌های آتش و دود موجود در تصاویر را جدا و به تشخیص دسته و یا محلی‌سازی ناحیه‌های موجود بپردازد. در مرحله دسته‌بندی نشان داده شد که می‌توان به دقت‌های بالاتر در شبکه‌های عمیق‌تر نیز دسترسی پیدا کرد اما این کار باعث صرف هزینه

جدول ۶: ماتریس درهم‌ریختگی قطعه‌بندی بر روی تصاویر آزمایشی.

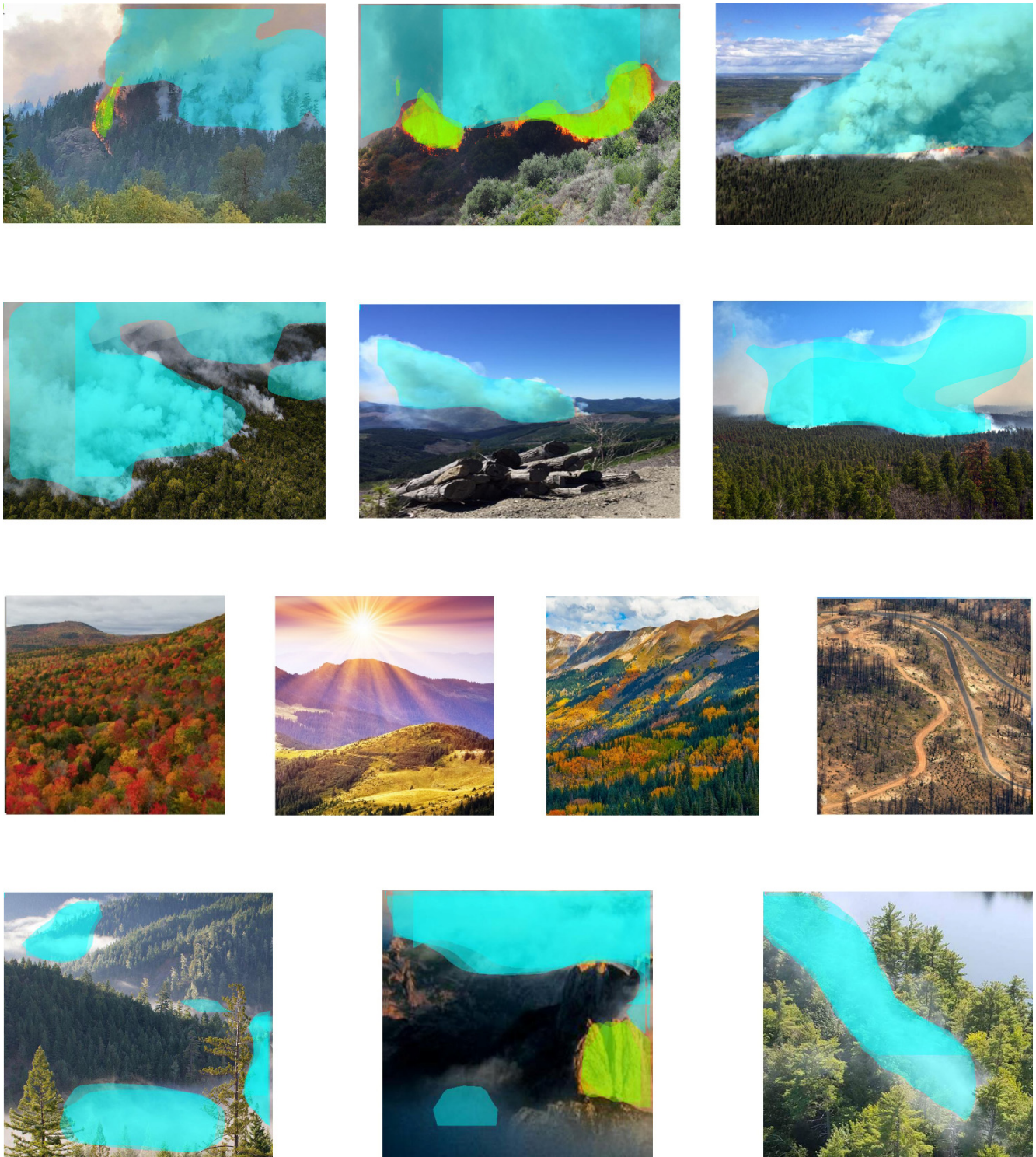
	واقعی		
	آتش	دود	زمینه
پیش‌بینی شده	آتش	۹	۲۸
	دود	۴	۱۴
	زمینه	۲۷۹	۱۷۸

در جدول ۵ مقایسه‌ای بین نتایج حاصل از شبکه کانولوشنی پیشنهادی با دیگر شبکه‌ها آورده شده و در تمام شبکه‌ها از تابع هزینه کراس آن‌تروپی و اندازه دسته ۵۰ استفاده گردیده است.

۴-۲ تشخیص شیء و قطعه‌بندی

همان طور که عنوان شد، برای مرحله قطعه‌بندی از یادگیری انتقالی و آموزش داده‌های آموزشی روی «mask_rcnn_inception_v2_coco» استفاده می‌کنیم. ماتریس درهم‌ریختگی در جدول ۶ و عملکرد نهایی شبکه در جدول ۷ با اعمال همان ابرپارامترهای شبکه انتخابی و انتخاب دسته ۱ به دلیل عمق زیاد شبکه، ورودی اندازه تصاویر دلخواه و بزرگ و محدودیت سخت‌افزار مورد استفاده نشان داده شده است. همچنین سرعت اجرای هر نمونه تصویر در شبکه با توجه به اندازه آن بین ۰,۵ تا ۱ ثانیه متغیر بود و آموزش شبکه تا بیستمین دوره کامل از اعمال کل تصاویر آموزشی ادامه یافت. نحوه محاسبه ماتریس درهم‌ریختگی به این صورت است که ابتدا نواحی استخراج شده با ناحیه‌های مورد انتظار تطابق داده می‌شوند و مقادیر با میزان هم‌پوشانی میان دو ناحیه (IoU)^۱ کمتر از ۰,۵

1. Intersection over Union



شکل ۶: تصاویر شکل ۵ با اعمال بر روی شبکه قطعه‌بندی پیشنهادشده.

تشخیص وجود آتش و دود پردازد و در صورت تشخیص، تصویر مورد نیاز را به نواحی آتش و دود قطعه‌بندی کند. در مواردی که تفاوت قابل‌شدن بین ابر و مه با دود حتی توسط چشم انسانی هم دشوار به نظر می‌رسد، روی آوردن به استفاده از حسگرهای تشخیص دود یا در کل برای کاهش میزان خطای تشخیص، استفاده از روش‌های دیگر نیز در کنار روش‌های پیشنهادی منطقی به نظر می‌رسد.

مراجع

- [1] C. Yuan, Y. M. Zhang, and Z. X. Liu, "A survey on technologies for automatic forest fire monitoring, detection, and fighting using

سرعت کمتر نسبی با توجه به سخت‌افزار استفاده‌شده می‌گردد. همچنین از آنجایی که تشخیص به موقع آتش و یا دود یا تشخیص تمام آتش‌سوزی‌ها از اهمیت بالایی برخوردار است، می‌توان نرخ حساسیت بالاتر تشخیص آتش و دود نسبت به دقت در دسته‌بندی را از مزیت‌های روش پیشنهادی دانست.

با وجود محدودیت داده‌های به کار برده شده و تنوع مناطق جنگلی می‌توان نتایج بهتری را از اعمال داده‌های بومی مناطق پایش‌شده در مرحله آموزش شبکه انتظار داشت. همچنین با توجه به نتایج بهتر دسته‌بندی و اهمیت ابتدایی تشخیص درست نسبت به محلی‌سازی آتش و دود می‌توان دو شبکه را به صورت سری نیز به کار برد که ابتدا به

- [14] G. W. Lindsay, "Convolutional neural networks as a model of the visual system: past, present, and future," *J. of Cognitive Neuroscience*, vol. 33, no. 10, pp. 1-15, Sept. 2020.
- [15] S. Frizzi, et al., "Convolutional neural network for video fire and smoke detection," in *Proc. IECON 42nd Annual Conf. of the IEEE Industrial Electronics Society*, pp. 877-882, Florence, Italy, 23-26 Oct. 2016.
- [16] Y. Zhao, J. Ma, X. Li, and J. Zhang, "Saliency detection and deep learning-based wildfire identification in UAV imagery," *Sensors*, vol. 18, no. 3, Article ID: 712, Feb. 2018.
- [17] H. Pan, D. Badawi, and A. E. Cetin, "Computationally efficient wildfire detection method using a deep convolutional network pruned via Fourier analysis," *Sensors*, vol. 20, no. 10, Article ID: 2891, May 2020.
- [18] https://github.com/tensorflow/models/tree/master/research/object_detection
- [19] K. He, G. Gkioxari, P. Dollár, and R. Girshick, Mask R-CNN, In *ICCV*, 2017.
- [20] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, Jun. 2006.
- [21] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation," *J. of Machine Learning Technologies*, vol. 2, no. 1, pp. 37-63, 2011.
- [22] K. M. Ting, C. Sammut, and G. I. Webb, *Encyclopedia of Machine Learning*, Boston: Springer, 2011.
- [23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. of the 25th Int. Conf. on Neural Information Processing Systems*, pp. 1097-1105, Lake Tahoe, NA, USA, 3-6 Dec. 2012.
- unmanned aerial vehicles and remote sensing techniques," *Canadian J. of Forest Research*, vol. 45, no. 7, pp. 783-792, Jul. 2015.
- [2] J. R. Martinez-de Dios, et al., "Computer vision techniques for forest fire perception," *Image Vision Comput*, vol. 26, no. 4, pp. 550-562, Apr. 2008.
- [3] K. G. Hirsch and P. Fuglem, *Canadian Wildland Fire Strategy: Background Syntheses, Analyses, and Perspectives*, Canadian Council for Forest Ministries, Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta, 2006.
- [4] H. Lin, Z. Liu, T. Zhao, and Y. Zhang, "Early warning system of forest fire detection based on video technology," in *Proc. of the 9th Int. Symp. on Linear Drives for Industry Applications*, vol. 3, pp. 751-758, Hangzhou, China, 7-10 Jul. 2013.
- [5] E. Den Breejen, et al., "Autonomous forest fire detection," in *Proc. of the 3rd Int. Conf. on Forest Fire Research*, pp. 2003-2012, Luso, Portugal, 16-20 Nov. 1998.
- [6] R. W. Beard, T. W. McLain, D. B. Nelson, D. Kingston, and D. Johanson, "Decentralized cooperative aerial surveillance using fixed-wing miniature UAVs," *Proceedings of the IEEE*, vol. 94, no. 7, pp. 1306-1324, Jul. 2006.
- [7] H. Olsson, et al., "Current and emerging operational uses of remote sensing in Swedish forestry," in *Proc. of the 5th Annual Forest Inventory and Analysis Symp. USDA Forest Service*, pp. 39-46, Portland, ONn, USA, 3-6 Oct. 2005.
- [8] J. A. J. Berni, P. J. Zarco-Tejada, L. Surez, and E. Fereres, "Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 722-738, Mar. 2009.
- [9] J. Everaerts, "The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping," *Proc. Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, Pt. B1, pp. 1187-1192, Beijing, China, 3-11 Jul. 2008.
- [10] V. G. Ambrosia and T. Zajkowski, "Selection of appropriate class UAS/sensors to support fire monitoring: experiences in the United States," in *Handbook of Unmanned Aerial Vehicles*, Edited by K. P. Valavanis and G. J. Vachtsevano, Springer, Netherlands, pp. 2723-2754, 2015.
- [11] R. Sporea, *Color Space and Color Model*, (<http://www.opticalimits.com>)
- [12] M. Pietikainen, A. Hadid, G. Zhao, and T. Ahonen, "Local binary pattern for still images," In: *Computer Vision Using Local Binary Patterns. Computational Imaging and Vision*, vol 40, pp. 13-47, London: Springer, 2011.
- [13] T. X. Truong and J. M. Kim, "Fire flame detection in video sequences using multi-stage pattern recognition techniques," *Engineering Applications of Artificial Intelligence J.*, vol. 25, no. 7, pp. 1365-1372, Oct. 2012.

محمدصادق کیهان پناه تحصیلات خود را در مقطع کارشناسی در رشته مهندسی مکانیک در دانشگاه آزاد اسلامی واحد تهران شرق در سال ۱۳۹۴ به پایان رسانده است. همچنین مدرک کارشناسی ارشد خود در رشته مهندسی مکاترونیک را در سال ۱۳۹۹ از دانشگاه تبریز دریافت نمود. زمینه‌های تحقیقاتی مورد علاقه ایشان شامل یادگیری ماشین و شبکه‌های عصبی است.

بهروز کوهستانی مدارک کارشناسی ارشد فناوری اطلاعات - سیستم‌های نرم افزاری و دکتری علوم کامپیوتر - سیستم‌های هوشمند خود را به ترتیب در سال‌های ۱۳۸۷ و ۱۳۹۲ از دانشگاه‌های هریوت وات و اسکس در انگلستان دریافت نمود و هم‌اکنون استادیار دانشکده مهندسی برق و کامپیوتر دانشگاه تبریز می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: هوش مصنوعی، رایانش تکاملی، بهینه‌سازی ترکیبیاتی، سیستم‌های خبره و شبکه‌های عصبی.