



تشخیص آتش با استفاده از شبکه های عصبی کانولوشن

افسانه تقی زاده اردکانی

کارشناس ارشد، دانشکده امنیت، گروه هوش مصنوعی و ریاتیک، تهران،
Taghizadedafsaneh@gmail.com

چکیده

در سال های اخیر روش های تشخیص آتش در زمینه آکادمی و صنعت مورد توجه بیشتری قرار گرفته است. روش های تشخیص مبتنی بر سنسور به شدت بر روی سیگنال های فیزیکی خارجی تأکید دارند، که احتمالاً اگر محیط خارجی به طور چشمگیری تغییر کند دقت تشخیص را کاهش می دهد. با توسعه سریع دانش و تولید دوربین با کیفیت بالا، روش های مبتنی بر قابلیت استخراج ویژگی های تصویری، راه حل دیگری را فراهم می کند تا شناسایی الگو را برای فیلم ویدئو نیز انجام دهد. با این حال، این روش ها به دلایل زری نمی تواند به طور گسترده و موفقیت آمیز برای شناسایی آتش استفاده شود:

۱. آیتم های تداخل بیش از اندازه مانند: لامپ روشنایی و برجسته سازی خودرو در اتاق یا تونل وجود دارد که عمدتاً عملکرد تشخیص را مختل می کند.
۲. ویژگی ها بستگی به اطلاعات قبلی در مورد شعله و دود دارند و روش استخراج اتوماتیک برای صحنه های مختلف آتش سوزی وجود ندارد.

به عنوان یک پیشرفت در شناخت الگو، یادگیری عمیق قادر به کشف اطلاعات مفید از داده های خام است و می تواند به طور خودکار نتیجه های دقیقی را تشخیص دهد. بنابراین، بر اساس ایده یادگیری عمیق، یک روش تشخیص آتش جدید مبتنی بر شبکه عصبی کانولوشن چند کاناله در این مقاله برای غلبه بر کمبودهای ذکر شده در بالا پیشنهاد شده است.

مرحله اول، تصاویر رنگی سه کاناله به عنوان ورودی شبکه عصبی کانولوشن ساخته می شوند؛ مرحله دوم، لایه های پنهان با پیچیدگی لایه های متعدد و ادغام ها ساخته میشوند، و به طور همزمان، پارامترهای مدل را پیدا می کنند و با استفاده از پس انتشار تنظیم می شود. در نهایت، روش softmax برای انجام طبقه بندی در مورد تشخیص آتش استفاده می شود. برای صرفه جویی در زمان آموزش، ما از GPU برای ساخت مدل های آموزش و آزمایش استفاده می کنیم. از مجموعه داده های آتش در اینترنت، تعدادی تصاویر برای آموزش و تعدادی تصاویر برای آزمون جمع آوری می کنیم، و پس از آن اجرا آزمایش با چهار روش پایه از جمله شبکه های عصبی عمیق، ماشین بردار پشتیبانی، پشته خودرو رمزگذار و شبکه عمیق مقایسه می شود.

فهرست مطالب

۱- مقدمه ۵

۲- کارهای مرتبط ۷

۱-۲ تشخیص آتش بر اساس ویژگی های تصویر ۷

۲-۲ شبکه های عصبی کانولوشن ۷

۱-۲-۲ ساختار CNN ۸

۲-۲-۲ لایه ی کانولوشن ۸

۳-۲-۲ لایه های نمونه فرعی CNN (Sub-sampling) ۹

۳- روش پیشنهادی ۹

۹

۱-۳ پیش درمان تصویر ۹

۲-۳ آموزش ۱۰

۳-۳ شتاب دهنده GPU ۱۰

۴- نتایج آزمایش ۱۱

۱۱

۱-۴ مجموعه داده ۱۱

۲-۴ روشهای پایه برای مقایسه ۱۱

۳-۴ نقشه های ویژگی ۱۳

۴-۴ بحث ۱۷

۵- نتیجه گیری ۱۷

۶- منابع ۱۸

۱۸

فهرست اشکال

- شکل (۱) : نقشه های ویژگی. ۸
- شکل (۲) : ساختار LeNet-5. ۸
- شکل (۳) : نمودار جریان تشخیص آتش. ۱۰
- شکل (۴) : ساختار MCCNN. ۱۰
- شکل (۵) : روند کاری جزئیات رویکرد ما. ۱۱
- شکل (۶) : داده های تصویری. ۱۲
- شکل (۷) : نقاط به دست آمده از V-SIFT. ۱۴
- شکل (۸) : ویژگی SAE. ۱۴
- شکل (۹) : تصاویر اصلی و لایه پنهان نقشه های MCCNN. ۱۵
- شکل (۱۰) : عملکرد MCCNN. ۱۵
- شکل (۱۱) : کانال ورودی چهار کانال CNN. ۱۶
- شکل (۱۲) : ماتریس اختلال MCCNN. ۱۶
- شکل (۱۳) : ROC از هر شش روش. ۱۶

فهرست فرمولها

۸	فرمول (1)
۹	فرمول (2)
۹	فرمول (3)
۹	فرمول (4)
۹	فرمول (5)
۹	فرمول (6)
۹	فرمول (7)

مبتنی بر ویژگی، این مقاله یک مدل یادگیری عمیق عمیق، به عنوان مثال، MCCNN، برای تشخیص شعله در تصویر استفاده می کند و دقت تشخیص بیشتر از چهار روش ابتدایی را به دست می آورد. Khan Muhammad و همکاران [۲] در مقاله ی خود سیستم پیشنهادی دقت تشخیص آتش را با حداقل هشدارهای نادرست بهبود داده است، با این وجود اندازه مدل نسبتاً سنگین است یعنی ۲۳۸ مگابایت. همچنین می توان CNN های سبک وزن را برای کاهش اندازه مدل در حین حفظ تعادل بین دقت و آلام های دروغین بررسی کند. علاوه بر این، چارچوب پیشنهادی فریم های مهم را بدون مکانیزم تأیید اعتبار در سیستم مدیریت فاجعه انتشار می دهد. در این زمینه، برای مخفی کردن برخی از اطلاعات در داخل فریم های کلیدی برای اهداف تأیید اعتبار، روش های پنهان سازی روشی مانند استیگانگی و علامت گذاری می توانند مورد استفاده قرار گیرند.

Khan Muhammad و همکاران [۳] در مقاله ی خود به قابلیت های پیشرفته پردازش دستگاه های هوشمند، نتایج قابل توجهی را در سیستم های نظارتی برای شناسایی وقایع غیر طبیعی مختلف مانند آتش سوزی، حوادث و سایر موارد اضطراری نشان داده اند. آتش یکی از حوادث خطرناک است که اگر در زمان کنترل نشده باشد، می تواند باعث زیان های بزرگی شود. این به اهمیت توسعه سیستم های تشخیص آتش سوزی نیاز دارد. بنابراین، در این مقاله پژوهشی، ما پیشنهاد می کنیم از معماری CNN برای مصارف مقرون به صرفه در دوربین های نظارت استفاده شود. Qi-xing ZHANG و همکاران [۴] در مقاله ی خود Faster R-CNN برای شناسایی دود در جنگل استفاده شده است. تصاویر موجود در دود آتش نشانی جنگل برای آموزش مدل های عمیق در مقیاس و تنوع محدود است، با تولید تصاویر دود در سیگنال جنگل و وارد کردن دو نوع دود، دود واقعی و دود شبیه سازی، به پس زمینه جنگل، نتایج آزمایش با تصاویر دود واقعی جنگل اثبات امکان این راه حل را نشان می دهد. این نه تنها مشکل کمبود داده را حل می کند،

در زندگی واقعی، کنترل آتش سوزی سخت است، به ویژه در سازه های بزرگ مانند کارخانه، تونل یا کارگاه و غیره. بنابراین، در این صورت تشخیص آتش سوزی، از اهمیت عالی نظری و واقع گرایانه برخوردار است. روش های تشخیص آتش سوزی عمدتاً بر اساس سیگنال فیزیکی از طریق سنسورهای دود، آشکارساز شعله مادون قرمز، شعاع مادون قرمز، آشکارساز شعله ی ماوراء بنفش و غیره بدست می آید. به عنوان مثال، Wentao Mao و همکاران [۱] پیشنهاد روش تشخیص آتش بر اساس تشخیص آتش بر اساس شبکه عصبی فشرده چند کاناله. Khan Muhammad و همکاران [۲] روش تشخیص زودهنگام آتش با استفاده از شبکه عصبی مصنوعی در حین نظارت بر مدیریت بحران های مؤثر. Khan Muhammad و همکاران [۳] یک روشی ارائه می دهد که آتش در شبکه های عصبی مصنوعی در فیلم های نظارت را تشخیص می دهد. Qi-xing ZHANG و همکاران [۴] تشخیص هوای آتشفشان جنگل Wildland بر اساس Faster R-CNN با استفاده از تصاویر دود مصنوعی. Zhen Zhong^{1,2} و همکاران [۵] روش تشخیص شعله ی مبتنی بر شبکه عصبی کانولوشن در دنباله ی ویدیو را مورد بررسی قرار داده اند. با توجه به دقت تشخیص زمانی محیط اطراف، دما و تغییرات نوردی و با توجه به اثر فعال سنسورها، روش های تشخیص دود معمولاً در یک منطقه کوچک کار می کنند. اخیراً، با توسعه سریع دوربین ها با کیفیت بالا، تئوری تشخیص الگو یک راه حل موثر برای تشخیص آتش است. به طور کلی، روش های معمول تشخیص آتش بر اساس تکنیک پردازش تصویر عبارتند از دو مرحله: ویژگی های استخراج و مدل تشخیص آموزش. ویژگی های اولیه شامل بافت، ویژگی رنگ و آنتروپی و غیره است. مدل های بعدی را می توان با ماشین بردار پشتیبانی (SVM)، شبکه عصبی مصنوعی (ANN) و غیره بدست آورد. Wentao Mao و همکاران [۱] در مقاله ی خود به یادگیری عمیق در مورد مسئله شناسایی آتش پرداخته اند. بر خلاف روش های تشخیص آتش سنتی

بلکه کار را از برجسب نمونه نیز حذف می کند. برای دو روش نسل دود، اگر چه تصاویر تولید شده توسط روش دوم که قرار دادن دود شبیه سازی به پس زمینه جنگل بصری واقع بینانه نیست، عملکرد بهترین دلیل احتمالی این است که مکان دود کردن توسط روش دوم در مقایسه با روش اول دقیق تر است. ممکن است با بهبود فرآیند مصنوعی تصاویر دود جنگل و یا در نظر گرفتن گسترش این راه حل، توالی های ویدئویی، عملکرد را افزایش دهیم. با این حال، در مورد چنین روش هایی سه ضعف وجود دارد. اولاً، لازم است مصنوعی ویژگی های تشخیصی را استخراج کند که بر دانش پیشین تا حد زیادی تکیه دارد. به عنوان مثال، مقادیر آستانه شدت در فضای HSI باید به طور عمده بر اساس تجربه مهندسی تنظیم شود. دوماً، اکثر ANN ها و SVM ها مدل کم عمق هستند، به این معنی که این مدل ها تنها دارای لایه پنهان مخفی با ظرفیت کمتر از استفاده از روابط پیچیده نقشه برداری از تصویر مبدا به حالت آتش هستند، بنابراین آنها نمی توانند به طور موثری شعله را تشخیص دهند، مخصوصاً در زمینه های پیچیده. بسیاری از روش های اخیر تشخیص آتش برای صحنه ثابت و نوع آتش تنها پیشنهاد می شود که اگر نوع آتش، تغییر رنگ و صحنه تغییر کند، این روش ها برای برنامه های کاربردی زندگی عمیقتر نیستند. به طور کلی، برای آموزش یک مدل با قابلیت انعطاف پذیری قوی، داده های آتش سوزی واقعی برای پوشش دادن صحنه های مختلف و نوع آتش مورد نیاز است. با این حال، بهره وری از روش های فعلی به طور عمده با عملکرد نامطلوب کاهش می یابد، زیرا مدل سطحی نمی تواند نیاز به پردازش داده های عظیم را برآورده کند. چگونگی ساخت مدل شناختی موثر با داده های عظیم کلید حل مشکلات تشخیص آتش است. یادگیری عمیق پتانسیل را برای غلبه بر مشکلات موجود در بالا ذکر می کند. توجه داشته باشید که یادگیری عمیق به یک کلاس از الگوریتم های یادگیری ماشین اشاره می کند، جایی که عمیق به معنی معماری عمیق با بسیاری از لایه های مراحل پردازش اطلاعات است. با تقلید از مکانیزم مغز انسان می باشد. جالب

ترین ویژگی یادگیری عمیق این است که می تواند به طور خودکار ویژگی های تشخیصی را از لایه های پایین به لایه بالا استخراج کند. در مقایسه با شبکه عصبی مصنوعی سنتی که دارای معماری کم عمق است، یادگیری عمیقی تا حد زیادی بر عمق مدل تأکید می کند و توجه بیشتری به یادگیری ویژگی های غیر نظارتی دارد. به طور خاص، یادگیری عمیق، نمونه های عظیم را به عنوان ورودی می داند و به طور خودکار ویژگی هایی را با گره های عصبی چند لایه سوق می دهد و سپس یک مدل می گیرد که توانایی تعمیم بهتر را دارد. یادگیری عمیق به طور موفقیت آمیزی در بسیاری از زمینه ها مانند طبقه بندی ImageNet، تشخیص گفتار و Biomedicine به خوبی معرفی شده است، اما بر اساس بهترین اطلاعات ما، تحقیقات زیادی در مورد کاربرد یادگیری عمیق در مورد موضوع شناسایی آتش صورت گرفته است. در این مقاله سعی داریم این مشکل را حل کنیم. به دنبال تجزیه و تحلیل بالا، یکی از تکنیک های یادگیری عمیق معمولی، یعنی شبکه چند عاملی پیچیده شبکه عصبی (MCCNN) را انتخاب می کنیم و یک روش تشخیص آتش جدید مبتنی بر تصویر را پیشنهاد می کنیم. به طور خلاصه، این روش ابتدا ماتریس تصویر سه کاناله (R, G, B) را به عنوان ورودی MCCNN که دارای ۸ لایه است، پردازش می کند و سپس به صورت خودکار از طریق کانولوشن و نمونه برداری نمونه ها را استخراج می کند و پارامترها را از طریق استراتژی BP تنظیم می کند. در نهایت، مدل تشخیص برای پیش بینی اینکه آیا آتش وجود دارد یا خیر، به دست می آید. سه نکته اصلی در مورد تحقیقات ما وجود دارد. اول، تکنیک یادگیری عمیق برای به رسمیت شناختن مشکل آتش با دقت بهتر تشخیص شعله برای تصویر عظیم آتش و داده های ویدئویی استفاده می شود؛ دوم، روش پیشنهادی برای انواع و صحنه های مختلف آتش سوزی بیشتر و عمیق تر است. سوم، ما با استفاده از GPU، تکنولوژی را تسریع می کنیم تا سرعت ساخت مدل را افزایش دهیم، و یک روش عملی تر تشخیص آتش را پیشنهاد می کنیم.

۲- کارهای مرتبط

۱-۲ تشخیص آتش بر اساس ویژگی های تصویر

روش های تشخیص آتش با استفاده از ویژگی های تصویری عمدتاً ویژگی هایی از تصاویر را حذف می کند و از طریق طراحی الگوریتم های استخراج ویژگی که به میزان زیادی بر دانش پیشین متکی هستند، از طریق دستی کنترل می شوند. روش کلی را می توان به صورت زیر خلاصه کرد. ابتدا عناصر تشخیصی را استخراج می کند. ویژگی آتش همیشه تمرکز تحقیقات برای یافتن آتش است. برخی از ویژگی های استاتیک مانند رنگ آتش، فضای رنگ، بافت و غیره اثربخشی خود را در برخی صحنه های آتش نشان داده اند. با این حال، همانطور که پیشتر مورد بحث قرار گرفت، این ویژگی های استاتیک با سر و صدای خارجی آشنا هستند. علاوه بر این، برخی از ویژگی های پویا مانند شکل، ویژگی اختلاف قاب، حرکت دینامیکی، می تواند راه حل قوی برای تشخیص آتش را فراهم کند. دوم، آموزش یک مدل (مانند BP، SVM، رگرسیون لجستیک، روش بیزی، و غیره).

که رفتار بردارهای ویژگی استخراج را به عنوان ورودی مدل در نظر می گیرد. در نهایت، از مدل آموزش یافته برای تصمیم گیری در مورد ویدیو یا تصویر هدف استفاده می کنید. در این روش، اطلاعات دامنه یا دانش زمینه ای در این روش ها نقش کلیدی ایفا می کنند، زیرا ویژگی های نامناسب دقت تشخیص را کاهش می دهند. علاوه بر این، طراحی یک ویژگی نماینده تمایل دارد زمان زیادی را برای تنظیم آستانه و سایر پارامترها صرف کند. لازم به ذکر است که دوربین های IR برای تشخیص زود هنگام آتش با استفاده از ترموگرافی بسیار مفید هستند. دوربین های مادون قرمز می توانند ویژگی های خاصی از اشیاء را از طریق تابش حرارتی تشخیص دهند و درک کنند که دوربین های معمولی قادر به انتخاب آن نیستند. در تصاویر حرارتی، اجسام داغ به راحتی می توانند تعیین شوند و از برخی روش های پردازش سیگنال مانند تحلیل ویولت برای استخراج

ویژگی های معمول استفاده می شود. حتی با این مزایا، سیستم های مونیتورینگ با تعداد زیادی از دوربین های مادون قرمز تنظیم نمی شوند، زیرا دوربین های IR بسیار گران تر از دوربین های معمولی با وضوح بالا هستند. در عین حال، دوربین های IR برای تشخیص شیء مبهم مشکل است، در حالی که دوربین نوری می تواند اطلاعات کمکی در مورد آتش سوزی را بدست آورد، به عنوان مثال، بازتاب دیوار در تونل. توسعه یک روش ساده، موثر و راحت پخش تشخیص آتش بصری این پژوهش است.

۲-۲ شبکه های عصبی کانولوشن

شبکه های عصبی کانولوشن (CNN) یک شبکه عصبی جریان مستقیم است. به عنوان شاخه ای از یادگیری عمیق، شش ویژگی های جذاب در مورد (CNN) وجود دارد:

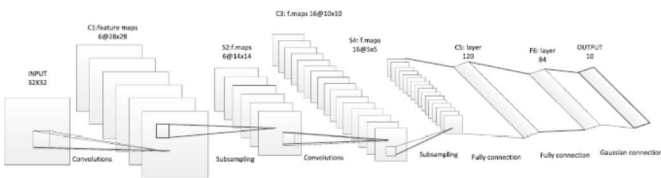
۱. ادراک محلی: در فضای تصویر، پیکسل های محلی همبستگی نزدیک دارند. در مقابل، همبستگی پیکسل های دور مشخص نیست، بنابراین نیازی نیست که هر گره نوری تصویر، تصویری را درک کند، اما هر گره می تواند منطقه محلی را درک کند و سپس به تدریج اطلاعات تصویر محلی را در لایه بالاتر نمایش دهد؛

۲. اشتراک گذاری وزن: در مناطق مختلف کانولوشن (CNN)، وزنه های مشترک را برای کاهش محاسبات استفاده می کنند. با استفاده از روش استخراج ویژگی مشابه در مناطق مختلف کانولوشن، به این معنی است که خواص آماری مشابه در مناطق مختلف کانولوشن محلی وجود دارد؛

۳. فیلترها: CNN فیلترهای زیادی را می سازد که انواع مختلفی از ویژگی ها را استخراج می کنند. این استراتژی می تواند به کمبود وزن های مشترک که در روش استخراج ویژگی های تک (یک فیلتر مشابه برای مناطق مختلف پیچش) است، غلبه کند.

در شکل ۱ نقشه های مشخصی با پنج فیلتر مختلف کانولوشن برای نشان دادن ویژگی های مختلف تصویر را نشان می دهد.

ویژگی های با اندازه $14 * 14$ ؛ $C3$ خروجی $S2$ را با شانزده فیلتر که اندازه $5 * 5$ است نشان می دهد، بنابراین خروجی 16 نقشه ویژگی $C3$ با اندازه $10 * 10$ است؛ $S4$ شبیه به $S2$ است با حجم پلات $2 * 2$ است و خروجی 16 نقشه ویژگی $S4$ با اندازه $5 * 5$ است؛ $C5$ یک لایه متقاطع است، و خروجی 120 نقشه ویژگی $C5$ با اندازه $1 * 1$ است؛ $F6$ (۷) به طور کامل ارتباط با خروجی $C5$ را نشان می دهد، و تعداد پارامترهای کل 10164 است؛ $S8$ خروجی LeNet شامل واحد تابع پایه ای شعاع اقلیدسی است و اندازه خروجی ده کلاس است که به ترتیب با عدد ۹-۰ مشخص می شود.



شکل (۲): ساختار LeNet-5 را نمایش می دهد، C نشان دهنده لایه کنولوشن است، S نشان دهنده زیر نمونه F برای اتصال به طور کامل می باشد.

۲-۲-۲ لایه ی کنولوشن

لایه کنولوشن به عنوان بخش مهمی از CNN نقش حیاتی دارد. در لایه کنولوشن، نقشه های ویژگی از لایه قبلی با هسته قابل یادگیری سازش می کنند و سپس از طریق تابع فعال سازی برای ایجاد نقشه های ویژگی خروجی قرار می گیرند. هر نقشه ی خروجی ممکن است با کنولوشن و ورودی های متعدد را ترکیب شود. به طور کلی، ما:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * kernel_{ij}^l + b_j^l \right) \quad (1)$$

در معادله (۱)، l نشان دهنده تعداد لایه ها، $kernel_{ij}^l$ نشان دهنده فیلتر، M_j یک انتخاب از نقشه های ورودی است. b_j^l نشان دهنده تعصب هر لایه است. هر نقشه خروجی یک تعصب افزایشی دارد، اما برای یک نقشه خروجی خاص، نقشه های ورودی با هسته های متمایز

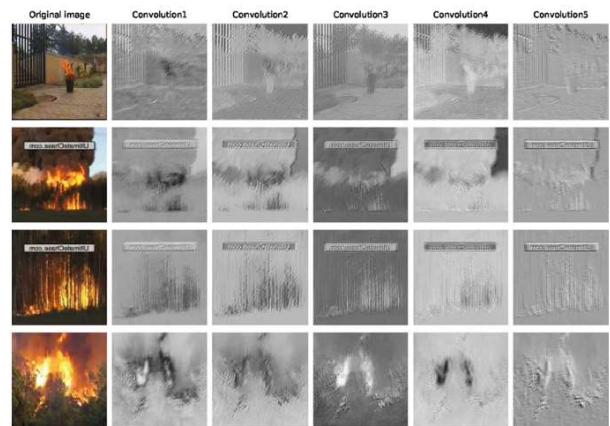
۴. زیر نمونه برداری: برای کاهش پیچیدگی محاسبات، CNN نمونه گیری را (همچنین به عنوان نمونه برداری کم نامیده می شود) برای استخراج ویژگی هایی که ترجمه غیرمنتظره هستند، استفاده می کند؛

۵. اتصال به طور کامل: CNN در لایه بالا به تمام ویژگی های انتزاعی استخراج شده متصل است و از طریق بسیاری از کانولوشن ها آن را جمع آوری می کند؛

۶. استراتژی Softmax: CNN طبقه بندی با استفاده از Softmax انجام می دهد، که می تواند قابل اجرا به صورت طبقه بندی و چند طبقه بندی باشد.

۲-۲-۱ ساختار CNN

LeNet5 نوعی از CNN با ۸ لایه است. کل شبکه شامل لایه ورودی، سه لایه کنولوشن، سه لایه جمع کننده، یک لایه کاملاً متصل و یک لایه خروجی است.



شکل (۱): نقشه های ویژگی استخراج شده توسط پنج فیلتر مختلف را نمایش می دهد. ستون اول تصاویر اصلی است، و پنج ستون های دیگر تصاویر خاکستری پیچیده به دست آمده توسط پنج فیلتر های مختلف کنولوشن می باشد.

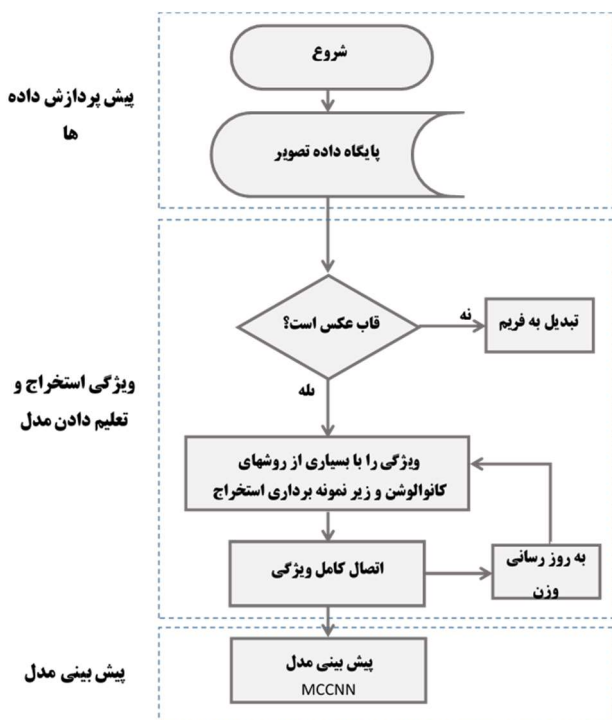
شکل ۲ چارچوب LeNet5 را نشان می دهد. فرآیند خاصی است: (۱) ورودی LeNet یک تصویر خاکستری با ابعاد $32 * 32$ است؛ (۲) $C1$ نشان دهنده کنولوشن ورودی است و $5 * 5$ فیلتر است که پنجره های اسلاید با ابعاد $5 * 5$ هستند. خروجی $C1$ دارای ۶ نقشه ویژگی با ابعاد $28 * 28$ است. (۳) $S2$ downsamples خروجی $C1$ با حداکثر تجمع که اندازه $2 * 2$ است، و خروجی $S2$ شش نقشه از

۳- روش پیشنهادی

در این مقاله، به عنوان الگوریتم آموزشی برای بهبود دقت تشخیص آتش، از شبکه عصبی کانولار چند کاناله (MCCNN) استفاده می کنیم. در مقایسه با LeNet5، MCCNN بیشتر برای تشخیص آتش کاربرد دارد، زیرا توجه بیشتری به رنگ شعله می دهد تا بتواند شعله خوبی را در نظریه تشخیص دهد. روش را می توان خلاصه کرد. اول، قبل از درمان تصویر منبع. دوم، مدل تشخیص را بسازید و پارامترهای مدل را به صورت خودکار به روز کنید. شایان ذکر است که این گام در محیط GPU اجرا می شود. در نهایت، نتیجه را بر روی مجموعه تست پیش بینی کنید. شکل ۳ نمودار جریان را نشان می دهد.

۳-۱-۳ پیش درمان تصویر

داده های تصویری که توسط MCCANN مورد استفاده قرار می گیرند نیاز به تنوع زیادی دارند (صحنه ها و نوع آتش). پس از جمع آوری داده ها، ابتدا همه فیلم ها را به تصاویر منتقل می کنیم و سپس تمام تصاویر را به اندازه $۳۲ * ۳۲ * ۳۲$ اندازه می کنیم. در این کار، برای تغییر اندازه تصاویر از درون یابی نزدیکترین همسایه استفاده می کنیم. علاوه بر این، ما تمام این ماتریس نمونه را نرمالیزه می کنیم.



پیچیده خواهند شد. ما فرض می کنیم که هر لایه کانولوشن l توسط یک لایه $l + 1$ downsampling دنبال می شود. بر اساس شبکه BP، مشتق خطا برای هر وزن لایه l برابر با ورودی این لایه است، حساسیت این لایه را چند برابر می کند، و سپس وزن را به روز می کنید. برای به روز رسانی وزن گره در لایه کانولوشن l ما نیاز به محاسبه حساسیت δ^l گره در لایه بعدی $l + 1$ و معادله (۲) فرایند محاسبه حساسیت را نشان می دهد، که در آن β معادلات جاری مربوطه را نشان می دهد، و $up(\delta_j^{l+1})$ یک عملیات upsampling را نشان می دهد که می تواند با معادله اجرا شود. (۳). در نهایت، گرادیان هسته را دریافت و وزن را به روز رسانی می کند، معادله (۲) به شرح زیر است:

$$\delta^l = \beta_j^{l+1} (f'(\mu_j^l) \circ up(\delta_j^{l+1})) \quad (2)$$

$$up(x) = x \otimes 1_{n \times n} \quad (3)$$

$$\frac{\partial E}{\partial b_j} = \sum_{\mu, v} (\delta_j^l)_{\mu v} \quad (4)$$

۳-۲-۲ لایه های نمونه فرعی CNN (Sub-sampling)

در لایه های نمونه فرعی، تعدادی از نقشه های ویژگی ورودی همانند تعدادی از نقشه های ویژگی خروجی است. با این حال، اندازه نقشه های ویژگی خروجی به نصف کاهش می یابد. معادله (۵) فرایند نمونه برداری را نشان می دهد، در حالی که $down(x_j^{l-1})$ نشان دهنده تابع نمونه برداری پایین است. روند به روز رسانی وزن نیز به BP متکی است، همچنین معادله (۶) و (۷) فرایند محاسبه حساسیت را نشان می دهد و سپس وزن لایه های نمونه فرعی برداری را به روز می کند.

$$x_j = f(\beta_j^l down(x_j^{l-1}) + b_j^l) \quad (5)$$

$$d_j^l = down(x_j^{l-1}) \quad (6)$$

$$\frac{\partial E}{\partial b_j} = \sum_{\mu, v} (\delta_j^l \circ d_j^l)_{\mu v} \quad (7)$$

شکل (۳): نمودار جریان تشخیص آتش بر اساس شبکه چند عاملی کانولوشنال عصبی.

۲-۳ آموزش

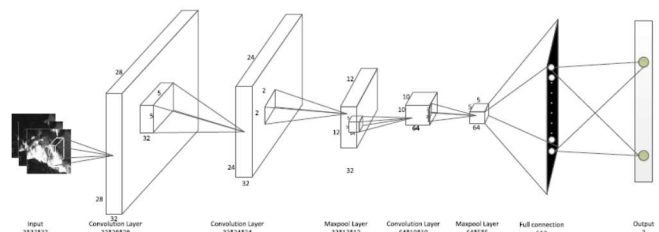
همانطور که در شکل ۴ نشان داده شده است، مدل MCCNN مورد استفاده در این مقاله دارای ۸ لایه است که خروجی هر لایه را وارد لایه بعدی می کند به شرح زیر:

۱. لایه اول یک لایه ورودی است که ماتریس تصویر رنگ را با اندازه $3 \times 32 \times 32$ به عنوان ورودی در نظر می گیرد.

۲. لایه دوم یک لایه کانولوشن است که ورودی را از طریق 32 فیلتر مختلف کانولوشن با ابعاد 5×5 شکل می دهد و سپس 32 نقشه ویژگی را با ابعاد 28×28 به دست می آورد.

۳. لایه سوم نیز یک لایه کانولوشن است، در حالیکه تعداد و اندازه فیلترهای کانولوشن شبیه لایه دوم است و همچنین 32 نقشه ویژگی با ابعاد 24×24 دارد.

۴. لایه چهارم یک لایه زیر نمونه برداری است که در آن Max-pooling را انتخاب می کنیم، خروجی دارای 32 نقشه ویژگی با اندازه 12×12 است.



شکل (۴): ساختار MCCNN.

۵. لایه پنجم یک لایه کانولوشن است که دارای 64 فیلتر کانولو با ابعاد 3×3 است و سپس 64 نقشه ویژگی با اندازه 10×10 دریافت می کند.

۶. لایه ششم یک لایه زیر نمونه برداری است که شبیه لایه چهارم است و پس از آن 64 نقشه ویژگی با ابعاد 5×5 می شود.

۷. لایه هفتم یک لایه کامل متصل است که شامل 512 گره عصبی است و این لایه تمام نقشه های محلی را

ادغام می کند و اطلاعات جهانی تصویر را به دست می آورد.

۸. لایه هشتم یک لایه خروجی است که دارای دو گره عصبی است از آنجا که برچسب ما شامل تصاویر آتش و تصاویر بدون آتش است.

همه پارامترهای به روز رسانی بر BP متکی هستند و به روز رسانی در هر دسته اجرا می شود تا به طور خودکار پارامترهای بهینه را انتخاب و ذخیره کند.

۳-۳ شتاب دهنده GPU

با استفاده از عصر جدید داده های بزرگ در سال های اخیر، محاسبات سریع موازی به طور فزاینده ای محبوب و مهم است، به خصوص در منطقه یادگیری عمیق. تکنیک GPU یک راه حل رقابتی است. در مقایسه با پردازنده، GPU با استفاده از بسیاری از ترانزیستورها توانایی محاسبات بهبود می بخشد. شبکه عصبی می تواند به راحتی محاسبات GPU را برای آموزش استفاده کند، زیرا فرایند ارسال خورده شبکه عصبی که شامل مجموعه های از مجموع وزنی گره ورودی است می تواند توسط ضرب ماتریس انجام شود. علاوه بر این، در طی روند به روز رسانی وزن ها، خطاهای باقی مانده و شیب ها نیز می تواند با استفاده از ضرب ماتریس انجام شود. در این عملیات ها خوب است. برای کاهش زمان آموزش، ما شبکه عصبی پیچیده چند کانالی خود را با استفاده از Keras، یک کتابخانه یادگیری عمیق که در اطراف نرم افزار یادگیری عمیق در Theano پیچیده است، اجرا می کنیم. Theano یک کتابخانه پایتون است که به کاربران اجازه می دهد تا به طور موثر تعریف، بهینه سازی و ارزیابی عبارات ریاضی شامل آرایه های چند بعدی انجام دهند. کاربران فقط نیاز به نصب پایتون، Theano و CUDA NVIDIA، و سپس می توانند MCCNN را آموزش دهند. تکنیک شتاب GPU مورد استفاده در این مقاله می تواند در شکل ۵ نشان داده شود.

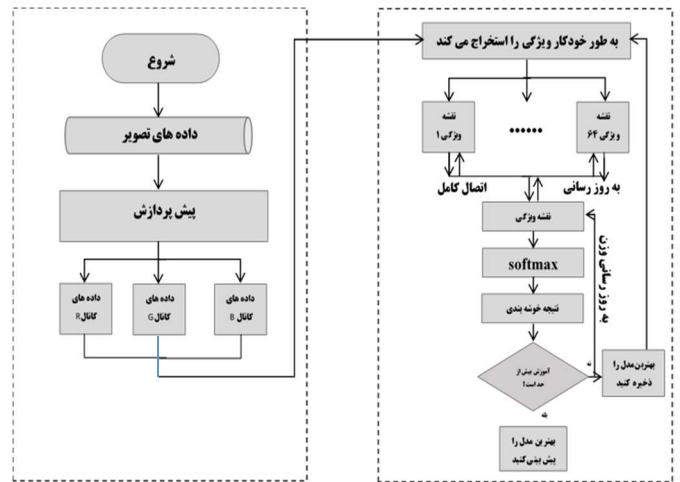
Web Crawler را برای جمع آوری خودکار تصویر آتش در اینترنت توسعه می دهیم. جدول ۱ اطلاعات دقیق در مورد داده های آتش شده تصویر آتش نشان می دهد. در کل، این پایگاه داده دارای ۱۱۴۹۴ تصویر از جمله ۹۰۷۴ تصویر آتش و ۶۲۲۷ تصویر بدون آتش است. این تصاویر صحنه های مختلف و انواع آتش را پوشش می دهد. به طور تصادفی این تصاویر را به دو مجموعه تقسیم می کنیم: مجموعه قطار دارای ۷۰۰۰ تصویر آتش است و مجموعه تست شامل ۴۴۹۴ تصویر آتش است. شکل ۶ برخی از تصاویر نمونه در پایگاه داده آتش نشانی ما را نشان می دهد، و تمام تصاویر از صحنه های واقعی و زندگی آمده است. این تصاویر شامل صحنه های مختلف و انواع شعله های مختلف، نتیجه تجربی از اهمیت عملی تر و معنایی متقاعد کننده خواهد بود.

۴-۲ روشهای پایه برای مقایسه

ما چهار روش ابتدایی را برای مقایسه انتخاب می کنیم. اطلاعات جزئی این چهار روش پایه به شرح زیر است:

۱. روش اول، شبکه عصبی عمیق (DNN) با ساختار شبکه ۱۰۲۴-۷۸۰-۲۵۰-۲ است، جایی که هر شخصیت تعداد گره های عصبی را در لایه مربوطه نشان می دهد. در این روش، ماتریس تصاویر خاکستری به عنوان ورودی در نظر گرفته می شود. استراتژی پیشنهادی تغذیه بین واحدهای مخفی و خروجی ها استفاده می شود. واحدهای مخفی معمولاً از تابع لجستیک استفاده می کنند. و هر واحد پنهان ورودی کل آن را از لایه زیر به حالت اسکالر نشان می دهد. این روش یک شبکه معمولی اما ساده است. این موفقیت در بسیاری از زمینه ها مانند مدل سازی آکوستیک، تشخیص گفتار و غیره مورد استفاده قرار گرفته است.

۲. روش دوم، یک نوع از تبدیل ویژگی غیر مارگریتی مقیاس (SIFT) است، که در آن خواص آتش



مرحله پردازنده (داده های پیشنهادی)

مرحله GPU (آموزش و پیش بینی)

شکل (۵): روند کاری جزئیات رویکرد ما. این شامل دو مرحله است و مرحله اول در CPU کار می کند و مرحله دوم در GPU کار می کند.

همانطور که در شکل ۵ نشان داده شده است، GPU پس از بارگیری تمام داده ها و پیش پردازش، کار را آغاز می کند. کار GPU، ساخت مدل MCCNN است. اولاً GPU با ماتریس چند کاناله با استفاده از چند کانولوشن و زیر نمونه برداری برای به دست آوردن ۶۴ ویژگی نقشه ها، و سپس تمام نقشه های ویژگی را به طور کامل متصل می کند. دوم، طبقه بندی Softmax نتایج طبقه بندی را به دست می دهد، و برای انتشار وزن استفاده می شود. سپس MCCNN به طور خودکار بهترین مدل را ذخیره می کند. در نهایت، پس از تمرین فرآیند، MCCNN به طور خودکار بهترین مدل ذخیره شده را بارگیری می کند و نتیجه را در مجموعه تست پیش بینی می کند.

۴- نتایج آزمایش

۴-۱ مجموعه داده

در این مقاله، آزمایش های کامپیوتری را برای تست عملکرد روش پیشنهادی در تصویر آتش و داده های ویدئویی با انواع مختلف ممکن انجام می دهیم. داده های تصویر عمدتاً از پایگاه داده های تصویری عمومی VisFire می آیند. برای بهبود تنوع تصویر آتش، ما یک نرم افزار

جدول ۱

مجموعه داده تصویر شعله

نام	شماره	یادداشت
تصویر کل	۱۱۴۹۴	از جمله ۵۲۶۷ تصویر آتش و ۶۲۲۷ تصویر بدون آتش، و همه تصاویر از صحنه های مختلف آمده است.
مجموعه آموزش	۷۰۰۰	از جمله ۴۰۰۰ تصویر آتش و ۳۰۰۰ عکس بدون آتش، و همه تصاویر از صحنه های مختلف آمده است.
مجموعه تست	۴۴۹۴	از جمله ۱۲۶۷ تصویر آتش و ۳۲۲۷ تصویر بدون آتش، و همه تصاویر از صحنه های مختلف آمده است.



شکل ۶ داده های تصویری با (a) نشان دهنده آتش سوزی و (b) نمایندگی از تصاویر آتش نشانی.

۴. روش چهارم همچنين نوعی از روش یادگیری عمیق است، یعنی شبکه اعتقاد عمیق (DBN) (ساختار شبکه: ۱۰۲۴-۷۸۴-۱۰۰-۲). DBN از الگوریتم یادگیری بی نظیر استفاده می کند تا وزن را از طریق جمع کردن بسیاری از ابزارهای تشخیصی (ماشین های بولتزمن محدود) استخراج کند. هنگامی که بر روی نمونه های بدون نظارت آموزش داده می شود، DBN می تواند به ورودی ورودی خود را بازنگری کند. بعد از مرحله یادگیری حریص، DBN از نظارت برای انجام طبقه بندی استفاده می کند. با توجه به ساختار ساده آن، DBN به طور گسترده ای برای حل برخی از مشکلات واقعی مانند پیش بینی سری زمانی، یادگیری ویژگی های نادر و غیره استفاده شده است.

با توجه به محدودیت فضا، ما ارائه نظریه مفصل این چهار روش را ارائه نمی دهیم. محیط آزمایش i7- CPU 4790، RAM 8G، سیستم عامل اوبونتو

نشان داده شده در فضای رنگی معرفی شده است. به طور خاص، ما ابتدا توصیف ویژگی را با استفاده از SIFT استخراج می کنیم و سپس ویژگی توصیف را از طریق ویژگی شعله ای استاتیک در فضای رنگ B, G, R، فیلتر می کنیم. در نهایت، بردارهای ویژگی به دست آمده برای آموزش مدل SVM و پیش بینی مجموعه های آزمون استفاده می شود

۳. روش سوم نوع دیگری از روش یادگیری عمیق است، پیک خودکار رمزگذار (SAE) (ساختار شبکه: ۱۰۲۴-۷۸۴-۳۸۶-۱۰۰-۲). ورودی نیز ماتریس تصویر خاکستری است. اول، ما فرایند پیش آموزش را اجرا می کنیم که شامل سه خودکار رمزگذاری می شود تا بازسازی ورودی و به دست آوردن ویژگی پنهان. دوم، شبکه های عصبی عمیق را با استفاده از ویژگی های پنهان که از مرحله اول ساخته شده است، بسازید. سرانجام، تنظیم دقیق را برای بهینه سازی شبکه به دست آمده از مرحله دوم انجام می دهیم. در این روش، خودکار رمزگذار برای استخراج ویژگی هایی است که می تواند خروجی یک لایه را به ورودی حفظ کند. تمام کانکتورهای خودکار به ترتیب به شکل یک شبکه عمیق ت همانطور که استراتژی ریاضی برای آموزش لایه به لایه معرفی شده است، SAE توانایی بهینه سازی جهانی را ندارد. اما به علت ساختار ساده و واضح بودن آن، SAE به طور گسترده در زمینه های مختلف مانند تشخیص بیماری آلزایمر، تشخیص گسل و غیره به طور گسترده ای مورد استفاده قرار گرفته است. شکیل می شوند.

NVIDIA GeForce ۱۵,۰۴ و کارت گرافیکی 705 با ۴۸ CUDA, 1G DDR3 vRAM و ۶۴ بیت گسترده.

علاوه بر دقت طبقه بندی، ما همچنین نرخ یادآوری، نرخ دقیق و نمره F1 را برای نشان دادن عملکرد روش پیشنهادی معرفی می کنیم. دقت، همچنین ارزش پیش بینی مثبت نامیده می شود، کسری از نمونه های واقعی مثبت در میان نمونه های مثبت طبقه بندی شده است، به شرح زیر است:

$$precision = \frac{TP}{TP + FP} \quad (8)$$

جایی که TP به معنی تعداد نمونه های واقعی مثبت است، FP به معنی تعداد نمونه های مثبت کاذب است. به یاد بیاورید، همچنین حساسیت نامیده می شود، کسری از نمونه های مثبت است که به درستی بر تعداد کل نمونه های مثبت طبقه بندی شده است:

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

جایی که FN تعداد نمونه های منفی کاذب است. به یاد بیاورید و دقیق معمولاً برای ارزیابی نتیجه طبقه بندی بد استفاده می شود. از تعریف بالا، فراخوانی می تواند برای محاسبه زنگ هشدار از دست رفته، به عنوان (1-Recall) استفاده شود، که نشان می دهد که چند نمونه مثبت مرتب طبقه بندی نشده اند. برای مثال، Fire را به عنوان کلاس مثبت و NoFire به عنوان کلاس منفی تنظیم می کنیم.

۳ نمونه از آتش و ۹۷ نمونه NoFire وجود دارد. یک مدل شناخت می تواند تمام ۹۷ نمونه منفی را به درستی طبقه بندی کند، اما فقط ۲ نمونه مثبت، یعنی یک نمونه آتش به عنوان NoFire شناخته شده است. در این مورد، $Recall = 2/3 = 66.67\%$

دقت $= 2/3 = 66.67\%$. اگرچه هشدار نادرستی رخ نداده است، یک هشدار آتش نشانی از دست رفته است که پیامد جدی خواهد بود. در بعضی موارد شدید، Recall و Precision با یکدیگر اختلاف دارند. بنابراین، ما همچنین F1 نمره را برای محاسبه میانگین هارمونیک Recall و دقت، به عنوان:

$$F1\ score = \frac{2Recall \times Precision}{Recall + Precision} \quad (10)$$

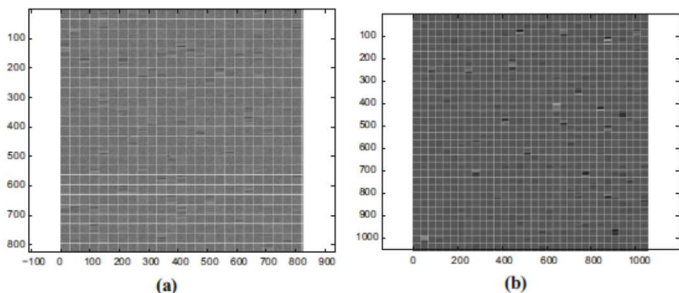
نمره $F1$ بیشتر به معنای عملکرد بهتر مدل است.

۴-۳ نقشه های ویژگی

برای مقایسه بهتر، ما می خواهیم در ابتدا عملکرد روش سنتی را همانطور که در شکل ۷ نشان می دهد، ارائه دهیم. همانطور که در شکل ۷ نشان داده شده است، شش تصویر آتش با صحنه های مختلف و انواع برای استخراج ویژگی انتخاب شده اند. ستون سمت چپ هر تصویر تصویر اصلی را نشان می دهد و ستون سمت راست تصاویر مربوط به VSIFT را نشان می دهد. ما مشاهده می کنیم که V-SIFT می تواند به طور موثر منطقه شعله در شکل a۷، b را تشخیص دهد، با این حال، تصویر پس زمینه پیچیده شامل بسیاری از کلیدهای پر سر و صدا است که ممکن است دقت تشخیص آتش را کاهش دهد. در این سناریو، V-SIFT نمیتواند کلیدهای منطقه آتش را دقیقاً شناسایی کند. درست مانند شکل c-f۷، V-SIFT به دلیل روشنایی، دود شعله و دیگر اشیاء تداخل (ماشین یا کارگر) به خوبی کار نخواهد کرد. بنابراین، واضح است که V-SIFT تا حد زیادی بر دانش پیشین متکی است. شکل ۸ نقشه های ویژگی های استخراج شده توسط SAE و DBN را نشان می دهد.

تصویر اصلی در شکل a۷ نشان داده شده است. شکل a۸ نشان می دهد نقشه های آموخته شده توسط اولین خودکار رمزگذار در SAE، و در شکل b۸ نمایش ویژگی

است و ستون سمت راست به وسیله V-SIFT نقشه ویژگی استخراج شده را می دهد. خطوط آبی نشان دهنده نکات کلیدی و بردارهای جهت آنها هستند (شکل رنگی آنلاین).



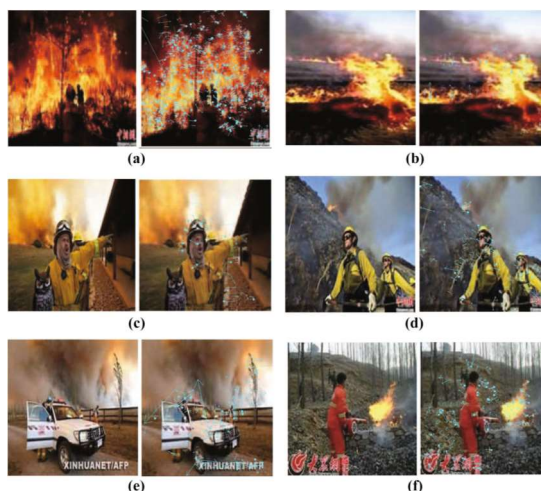
شکل ۸. ویژگی SAE با استفاده از (a) و ویژگی DBN با استفاده از (b) را نشان می دهد.

تغییر MCCNN از دست دادن در مجموعه آموزش و تست با آموزش های مختلف دوره ها روشن است که تلفات مجموعه های آموزشی و تست در ابتدا به شدت کاهش می یابد ۲۰ دوره، و به تدریج در ۲۰ تا ۵۰ دوره کاهش می یابد. پس از آن، از دست دادن کوچک است تغییر، و از دست دادن تمرین در حدود ۰,۰۱ با افت نوسان تست مجموعه تغییر می کند. در حدود ۰,۰۷۵. توجه داشته باشید که از بین رفتن آزمون، کمی بعد از ۲۰ دوره افزایش می یابد. از دست دادن آموزش می افتد، که به این معنی است که بیش از حد به وقوع می پیوندد. شکل ۱۰b نشان دهنده است. تغییر مقدار دقت در مجموعه آموزش و تست با مجموعه های مختلف epochs آشکار است که دقت به طور صعودی در ۲۰ اکتبر اول، و به تدریج صعود در ۲۰ تا ۵۰ دوره. سپس دقت مجموعه آموزش و تست مجموعه پایدار هستند دقت مجموعه آموزش و تست مجموعه در حدود ۹۹,۸۵ و ۹۸,۸۶٪ بود. از شکل ۱۰، می توان نتیجه گرفت که MCCNN است مناسب برای تشخیص آتش از دیدگاه آماری. به عنوان یکی از ویژگی های اصلی از آتش، رفتار زمانی آن است، به عنوان مثال، شکل، جریان و غیره، لازم است برای اثربخشی روش پیشنهاد شده بر روی اطلاعات زمانی. از ویدیو آتش جمع آوری شده از پایگاه Visfire، ما محاسبه تفاوت قاب فریم های متوالی و این اطلاعات زمانی را به عنوان ورودی

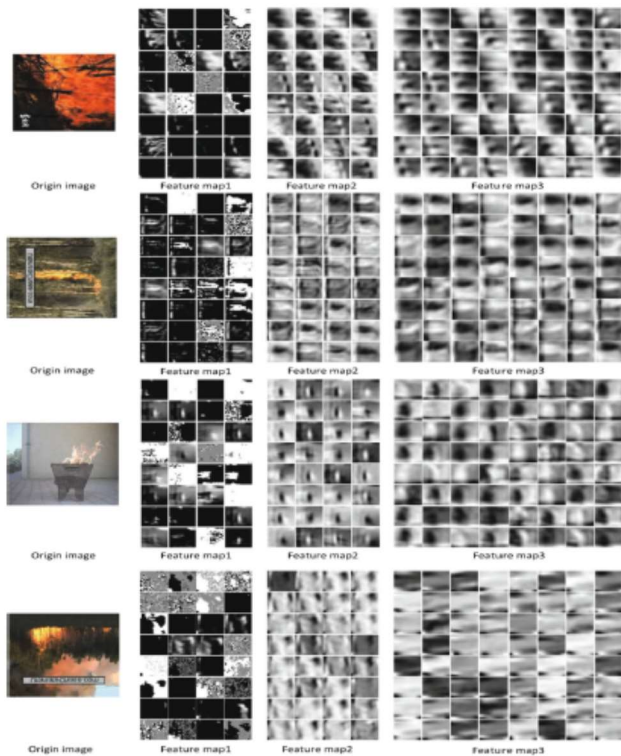
های آموخته شده توسط DBN. توجه داشته باشید که این نقشه تجسم ماتریس وزنه آموزش دیده است. مقدار بالاتر وزن است، رنگ عمیق تر است. واضح است که هر دو نقشه در بسیاری از نقاط دارای همان رنگ عمیق هستند و هیچ تفاوت آشکاری وجود ندارد. به این معناست که هیچ ساختار واضح مورد بهره برداری قرار نگرفته است (در مقایسه با شکل ۹، این نتیجه گیری ساده خواهد بود). در نتیجه، تشخیص آتش با استفاده از SAE و DBN در صحنه های مختلف محدود می شود. شکل ۹ نقشه های ویژگی های استخراج شده توسط MCCNN را نشان می دهد.

ستون اول نشان دهنده تصاویر اصلی است. ستون دوم نشان دهنده ۳۲ نقشه ویژگی است که خروجی لایه اول کانولیت است. ستون سوم ۳۲ نقشه ویژگی را نشان می دهد و خروجی های لایه دوم کانولیت هستند. و ستون چهارم نشان دهنده ۶۴ نقشه ویژگی است که چهارمین خروجی لایه کانولوشن هستند. شایان ذکر است که اندازه نقشه های ویژگی در ستون های مختلف برابر نیست، اما نقشه های ویژگی در همان ستون یکسان هستند. به خاطر تصویر، این نقشه ها را همان اندازه نشان می دهیم. از شکل ۹، ما می توانیم ویژگی های پنهان را مشاهده کنیم می تواند تصاویر منشا را به خوبی نشان دهد.

شکل ۱۰a نشان می دهد :

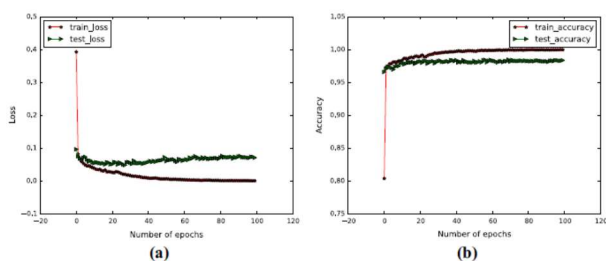


شکل ۷. نقاط نقاط به دست آمده از V-SIFT برای شش نمونه تصاویر آتش سوزی. در هر تصویر، ستون سمت چپ تصویر اصلی



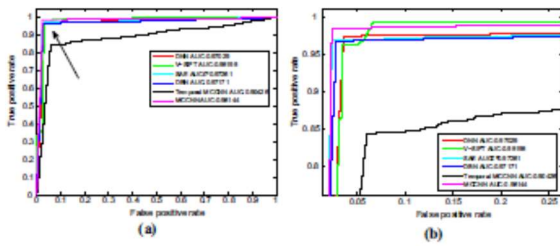
شکل ۹. تصاویر اصلی و لایه پنهان نقشه های به دست آمده توسط MCCNN.

ویژگی VIDEO در فضای پنهان بالا و این قابلیت برای طبقه بندی سودمند است. با مقایسه با SAE، NN، DBN و MCCNN، V-SIFT دارای کارایی پایین است، زیرا قابلیت دستیابی استخراج ممکن است برخی از اطلاعات مهم دامنه را نادیده بگیرد. برعکس، نتایج همچنین بهره وری از یادگیری عمیق را در استخراج ویژگی های خودکار نشان می دهد. از نظر برخی از شاخص های کلیدی مانند Precision، Recall.



شکل ۱۰. عملکرد MCCNN با دوره های آموزشی مختلف با (a) از دست دادن متقابل آنتروپی و (b) دقت طبقه بندی.

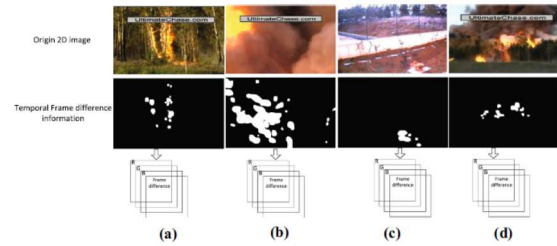
تعیین می کند. MCCNN، همانطور که در شکل ۱۱ نشان داده شده است. برای سادگی، ما این روش را Temporal می نامیم. MCCNN بدیهی است، اطلاعات اختلاط قاب میتواند خطوط را به طور چشمگیری ترسیم کند. از آتش، به خصوص در شکل ۱۱. اما به نظر می رسد به راحتی توسط سر و صدا اختلال مانند دود و درخت، مانند شکل ۱۱، b، d. ما یک کانال را در شبکه MCCNN اضافه می کنیم. نشان داده شده در شکل ۴، یعنی کانال چهارم علاوه بر کانال های R، G، B. اینجا برای محاسبه ساده تر، ما تصاویر اصلی را به $32 * 32 * 3$ تغییر می دهیم، زیرا $32 * 32$ نیز بیش از حد است. برای تفاوت فریم های درشت و مناسب برای نادیده گرفتن اطلاعات محلی. دیگری قسمت های شبکه هم همینطور جدول ۲ نتایج تطبیقی همه شش را ارائه می دهد. روش ها، که نشان می دهد که روش پیشنهادی (بدون اطلاعات زمانی) بهتر عمل می کند. چهار روش ابتدایی دیگر و MCCNN Temporal از نظر شناخت دقت. به طور خاص، MCCNN بالاترین دقت را دریافت می کند 99.86% در مجموعه آموزش، و همچنین دارای بالاترین سطح شناختی از 98.06% می باشد. از این نتایج می توانیم پیدا کنیم که MCCNN توانایی خوبی در استخراج ویژگی دارد که همچنین در شکل ۹ نشان داده شده است که MCCNN به خوبی می تواند اصل را بیان کند.



شکل ۱۳: ROC از هر شش روش با (a) منحنی کل و (b) نقاشی بزرگ در قسمت نشانگر فلش نشان داده شده است.

MCCNN ما در حال جستجو و آزمایش دیگر اطلاعات آتش سوزی برای MCCNN هستیم. ما همچنین مشاهده می کنیم که MCCNN نیاز به زمان بسیار محاسباتی بیشتری دارد، به ویژه در زمان آزمایش. دلیل این است که MCCNN دارای بسیاری از هسته کانولاسیون [لطفاً به معادله (۱)] که غیرخطی هستند. به این معنی است که داده های تست باید محاسبات را از طریق کل شبکه تکرار کنند. V-SIFT بسیار شبیه است. در VSIFT SVM برای مدل سازی استفاده می شود و هسته RBF برای آموزش و آزمایش انتخاب شده است. برعکس، DNN، SAE و DBN شبکه نسبتاً کم عمق دارند (که در توضیح بالا ذکر شده است) و هیچ عملی غیر خطی وجود ندارد. بنابراین، زمان آزمون بسیار کوتاه تر از MCCNN است. لطفاً توجه داشته باشید که این زمان آزمون برای کل ۴۴۹۴ تصویر است. بنابراین برنامه واقعی واقعی را حفظ نخواهد کرد.

به منظور ایجاد نتایج بیشتر، ما ماتریس سردرگمی MCCNN را نشان می دهد، همانطور که در شکل ۱۲ نشان داده شده است. در شکل ۱۲، مقادیر نمونه خاص در شبکه ها مشخص شده است. به این ترتیب، میزان مثبت واقعی (غلط) و نرخ منفی واقعی (غلط) در ردیف سوم و ستون نشان داده شده است. ما همچنین تصاویر نادرست طبقه بندی شده را بررسی می کنیم و سعی می کنیم اثر بصری مشترک آنها را پیدا کنیم. این تصاویر از یک



شکل ۱۱. کانال ورودی چهار کانال CNN با استفاده از اطلاعات اختلاف فریم زمانی.

و نمره F1، MCCANN نیز بهتر از چهار روش دیگر و Baseline MCCANN است.

بسیار عجیب و غریب است که MCCNN Temporal از لحاظ شاخص های خطا کمترین عملکرد را به دست می آورد. با تجزیه و تحلیل بصری، اطلاعات موقتی اضافی برای مدل شناختن مطمئناً مفید است. اما از نتایج مقایسه ای، این فرض ثابت نشده است. با توجه به این پدیده، ما دو دلیل را حدس می زنیم: (۱) فیلم های آتش نشانی جمع آوری شده از VisiFire با تنوع کمتر محدود می شوند، در حالی که اطلاعات تفاوت فریم تکرار داده های مشابه را افزایش می دهد. (۲) مدل یادگیری عمیق نسبتاً به داده های مکرر حساس است، که باعث می شود مدل بر روی بخشی از داده ها ایجاد شود و پس از آن باعث ایجاد بیش از حد می شود. توجه داشته باشید که دقت آموزش MCCNN Temporal، مسلماً بالاتر از دقت آزمون است. البته احتمال دیگری وجود دارد: اطلاعات اختلافی قاب ممکن است ورودی مناسب نباشد.

Confusion Matrix

	Fire	No Fire	
Fire	TP:2194 48.8%	FP:54 1.2%	97.6% 2.4%
No Fire	FN:33 0.7%	TN:2213 49.2%	98.5% 1.5%
	Fire	No Fire	98.1% 1.9%
			Target Class

شکل ۱۲. ماتریس اختلال MCCNN بدون اطلاعات زمانی.

وضعیت یا ویدیو نیستند، اما همه آنها دارای پسزمینی یا تداخل آشکاری هستند.

شکل ۱۳ ROC تمام شش روش را نشان می دهد. به منظور نشان دادن، ما همچنین نقاشی بزرگ و بزرگ را برای منطقه مشخص شده در شکل a۱۳ نشان می دهیم، همانطور که در شکل b۱۳ نشان داده شده است. در شکل ۱۳، منحنی MCCNN دارای بزرگترین AUC (۰,۹۸۱۴۴) است در حالی که منطقه AUC NN V-SIFT, SAE و DBN به ترتیب ۰,۹۷۰۲۹، ۰,۹۵۱۹۸، ۰,۹۷۲۶۱ و ۰,۹۷۱۷۱ است. توجه داشته باشید که MCCNN Temporal کوچکترین منطقه AUC (۰,۹۰۴۲۶) را می گیرد، که مطابق با نتایج در جدول ۲ است. بنابراین، شکل ۱۳ همچنین ثابت می کند که روش پیشنهاد شده در زمینه تشخیص آتش می تواند به خوبی مورد استفاده قرار گیرد.

۴-۴ بحث

از نقشه های ویژگی که در شکل ها نشان داده شده است. ۷، ۸ و ۹، می توانیم V-SIFT را به راحتی تحت تاثیر قرار دهیم. نقشه های ویژگی SAE دارای ساختار واضح نیستند زیرا SAE ویژگی های پنهان را با استفاده از خطاهای ورودی بازپرداخت به دست می آورد که بیشتر به هر یک از پیکسل ها توجه می کند و اطلاعات کلی تصویر و ساختار را نادیده می گیرد. علاوه بر این، نقشه های مشخصی از کمبود اطلاعات ساختاری DBN، بسیاری از ابزارهای ویژه برای یادگیری وزن، که بیشتر به هر یک از پیکسل ها توجه بیشتری نسبت به اطلاعات کلی داده می شود. در مقایسه با چهار روش پایه، نقشه های ویژگی از روش پیشنهادی دارای اطلاعات بیشتر در مورد لبه، بافت و ساختار تصویر و غیره است. همانطور که در شکل ۹ نشان داده شده است، نقشه های ویژگی های روش پیشنهادی به وضوح تصویر اصلی را ارائه می دهند و وزن های مختلف را برای مناطق مختلف. همه این مزایا به تشخیص آتش کمک می کند.

از جدول ۲، روش پیشنهادی دقت برابر با ۹۸,۰۶٪ است که بالاترین در همه روش ها است. در مقایسه با V-SIFT که از ویژگی های دستی استخراج شده است، روش پیشنهادی عملکرد رضایت بخش را به دست می دهد، زیرا MCCNN به طور خودکار از تصویر اصلی استخراج می کند و توجه بیشتری به خود داده و عمق مدل می دهد. در مقایسه با DNN, SAE و DBN که به طور خودکار ویژگی را استخراج می کنند، MCCNN توانایی استخراج ویژگی های خوب را دارد که بر اساس فیلد پذیرفته شده زیست شناسی است و می تواند ویژگی های موجود در لایه های مخفی را نیز داشته باشد.

در همین حال، روش پیشنهادی نیز از لحاظ یادآوری و دقت عملکرد خوبی دارد، که ثابت می کند که روش پیشنهادی توانایی جدایی پذیری در تشخیص آتش است. از دیدگاه هزینه محاسبات، روش پیشنهادی بدتر از چهار روش ابتدایی است که می تواند از زمان آموزش در جدول ۲ مشاهده شود. جوهر یادگیری عمیق، اجرای هوش مصنوعی با تقلید مغز انسان است، بنابراین لایه های زیادی وجود دارد و گره در مدل یادگیری عمیق، به عنوان MCCNN به همان شیوه کار می کند. به همین دلیل هزینه محاسبات MCCNN نسبتا بالا است. با این حال، می توان از برخی ترفندها اجتناب کرد. به عنوان مثال، مدل MCCNN را می توان به صورت آفلاین آموزش داد با بهترین وزن های ذخیره شده ذخیره شده، و سپس این مدل را در کاربرد واقعی برای پیش بینی نمونه های آزمایش بارگذاری می کند. علاوه بر این، سرورهای توزیع شده بزرگ با چند گرافیک و چندین موضوع می توانند برای حل این مشکل معرفی شوند.

۵- نتیجه گیری

در سال های اخیر، با توسعه سریع هوش مصنوعی، آموزش عمیق توجه بیشتری به خود جلب کرده است. این مقاله با موفقیت به یادگیری عمیق در مورد مشکل تشخیص

همیشه لازم است که انواع داده های آتش را بهبود ببخشیم.

هنوز مشکلات زیادی در آینده حل می شود. اول، چگونگی تشخیص شعله در هوای بد، مانند آب و هوای مبهم، یک کار قابل حل است. آب و هوای مه آلودگی نادرست تشخیص دود را افزایش می دهد. دوم، همانطور که در شرایط واقعی، ممکن است دود باشد، اما در واقع هیچ آتش سوزی وجود ندارد. برای انجام چندین طبقه بندی در تشخیص آتش، به عنوان مثال، اضافه کردن برجسب دود علاوه بر آتش و عدم برجسب آتش، سوال دیگری است. یادگیری عمیق می تواند به خوبی در مسئله چند طبقه بندی خوب کار کند، اما استفاده از تصاویر خام به عنوان ورودی مستقیم باعث می شود تا پایداری مدل یادگیری عمیق کاهش یابد. شاید لازم باشد پیش از پردازش و یا استخراج ویژگی ها در پیشبرد، که دانش میدان را به مدل عمیق افزایش می دهد. چگونه برای ادغام با سنسورهای دیگر مانند دوربین های IR و سنسور تشخیص دود برای ایجاد یک سیستم نظارت کامل یک مشکل واقعی است. و در نهایت، چگونه باید هزینه های محاسباتی را کاهش داد همچنین باید در نظر گرفته شود.

۶- منابع

[1] Ali Farhat, Omar Hommos, Ali Al-Zawqari, Abdulhadi Al-Qahtani, Faycal Bensaali, Abbas Amira and Xiaojun Zhai , **Optical character recognition on heterogeneous SoC for HD automatic number plate recognition system**, EURASIP Journal on Image and Video Processing 26 June 2018, <https://doi.org/10.1186/s13640-018-0298-2> .

[2] Gundram Leifert , Tobias Strauß , Tobias Grüning , Welf Wustlich , Roger Labahn , **Cells in Multidimensional Recurrent Neural Networks** , Journal of Machine Learning Research 17 (2016) 1-37 Submitted 5/14; Revised 12/15; Published 7/

[3] Israr Ud Din, Imran Siddiqi, Shehzad Khalid and Tahir Azam, **Segmentation-free optical character recognition for printed Urdu text** , EURASIP Journal on Image and Video Processing 2017 2017:62, <https://doi.org/10.1186/s13640-017-0208-z>

آتش پرداخته است. بر خلاف روش های تشخیص آتش سنتی مبتنی بر ویژگی، این مقاله یک مدل یادگیری عمیق عمیق، به عنوان مثال، MCCNN، برای تشخیص شعله در تصویر استفاده می کند و دقت تشخیص بیشتر از چهار روش ابتدایی را به دست می آورد. بر اساس نتایج تجربی، ما یافته های زیر را داریم:

۱. با روشهای سنتی مبتنی بر ویژگی، روش پیشنهادی می تواند ویژگی های خود را به طور خودکار از تصاویر خام استخراج کند (البته برای آموزش آسان، ما نیاز به تغییرات پیش فرض تصاویر خام داریم). از ویژگی نقشه برداری به دست آمده، روش پیشنهاد شده توانایی بهتر ساختن ساختار درونی تصاویر خام را از دیگر روش های یادگیری عمیق عمیق دارد که نشان می دهد ویژگی های به دست آمده از روش پیشنهادی بیشتر تبعیض آمیز است.

۲. با استفاده از ویژگی های به دست آمده، روش پیشنهاد شده می تواند نتایج دقیق تر و شناختی قوی را از نظر شاخص های خطای مختلف بدست آورد. و از نتایج با دوره های مختلف آموزش، پیشنهاد شده است که زمان آموزش مناسب برای جلوگیری از overfitting نگه داشته شود.

۳. برای روش شناختی مبتنی بر یادگیری عمیق، تنوع داده ها مهم است. از نتایج مقایسه ای، روش پیشنهاد شده به خوبی در تصاویر آتش نشانی با صحنه های مختلف و انواع شعله های مختلف کار می کند. برای ساختن یک مدل عمیق عمیق برای شناختن، پیشنهاد شده است که اطلاعات آتش را با تعداد زیادی صحنه و انواع مختلف جمع آوری کنیم. صحنه های مختلف و انواع می تواند اطلاعات دامنه کافی از آتش را فراهم کند. اما از نتایج اضافه شدن اطلاعات موقتی، تعداد کمی از اطلاعات ویدئویی تکراری، توانایی تعمیم مدل مبتنی بر یادگیری عمیق را پایین می آورد. برای روش مبتنی بر یادگیری عمیق، بدون توجه به تصاویر و ویدیوها،

[4] Taras Iakymchuk, Alfredo Rosado-Muñoz, Juan F Guerrero-Martínez, Manuel Bataller-Mompeán and Jose V Francés-Víllora, **Simplified spiking neural network architecture and STDP learning algorithm applied to image classification**, EURASIP Journal on Image and Video Processing 2015, <https://doi.org/10.1186/s13640-015-0059-4>

[5] **Gauri Katiyar and Shabana Mehfuz**, A hybrid recognition system for off-line handwritten characters, **Katiyar and Mehfuz SpringerPlus (2016) 5:357 DOI 10.1186/s40064-016-1775-7**, <https://doi.org/10.1186/s40064-016-1775-7>