

الحمد لله الذي
خلقنا من
الحمم



دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

شناسایی دوربین منبع بر اساس یادگیری عمیق

نگارنده: مسعود بهاری

استاد راهنما

دکتر منصور فاتح

استاد مشاور

دکتر محسن رضوانی

تیر ۱۴۰۰

در این صفحه صورت جلسه دفاع را قرار دهید. لازم است پس از صحافی این صفحه مجدداً توسط دانشکده مهر گردد و استاد راهنما با امضای خود اصلاحات پایان نامه را تایید کند.

سپاس خدایی را که به من این توفیق را عطا فرمود تا بتوانم قدمی در راه علم و
دانش بردارم و به گردآوری اطلاعات مربوط به این تحقیق پردازم.

تقدیم به:

پدرم بابوسه بردستانش که وجودش مایه می دگر می ام است.

و

مادرم بابوسه بردستانش که وجودش برایم همه مهر است.

به نام آفریدگار پاک که انسان را از خاک آفرید و به واسطه عقل بر تمام موجودات

رحمان داد و آنگاه دانش را وسیله تکامل عقل قرار داد.

بشکر از اساتید محترم آقایان دکتر منصور فتح و دکتر محسن رضوانی که

این جانب را در تهیه این پروژه یاری نمودند.

از خداوند متعال پیروزی و موفقیت روز افزون ایشان را خواستارم.

مسعود بهاری

تیرماه ۱۴۰۰

تهدنامه

اینجانب مسعود بهاری دانشجوی دوره کارشناسی ارشد رشته کامپیوتر گرایش هوش مصنوعی دانشکده کامپیوتر دانشگاه صنعتی شاهرود نویسنده پایان نامه "شناسایی دوربین منبع بر اساس یادگیری عمیق" تحت راهنمایی دکتر منصور فاتح متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « **Shahrood University of Technology** » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود . استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

با پیشرفت فناوری دیجیتال، رسانه‌های دیجیتالی را می‌توان از دوربین دستگاه‌های مختلف ثبت کرد و به راحتی در فضای مجازی و شبکه‌های اجتماعی به اشتراک گذاشت. در این زمینه، دستکاری‌های تصویری و ویدئویی بسیار رایج شده و برای افراد و کل جامعه به طور فزاینده‌ای خطرناک است. شناسایی مدل دوربین به عنوان اولین مرحله از ارزیابی اصالت تصاویر در تحقیقات پزشکی قانونی در مورد تصاویر دیجیتال است. شناسایی دوربین منبع، موضوع مورد بحث این پایان‌نامه است. در اکثر مقالات مرتبط، معماری‌های مختلف مبتنی بر یادگیری عمیق برای استخراج الگوی نویز به عنوان اثر انگشت هر دوربین، طراحی شد. هر تصویر شامل نویزهایی می‌باشد که بر اساس طراحی سیستم لنز این نویزها برای هر دوربین، منحصر به فرد است. این الگوی نویز بر اثر ویرایش و دستکاری تصویر از بین نخواهد رفت و هدف نهایی استخراج الگوی نویز، شناسایی مدل دوربین هر تصویر است که در مباحث جرم‌شناسی تصویری کاربرد دارد. به طور خاص، در این پایان‌نامه پیش‌پردازش داده‌ها و معماری CNN تحلیل می‌شود.

پیش‌پردازش داده‌ها همیشه به شناسایی دوربین منبع کمک نمی‌کند زیرا ممکن است در روند اعمال پیش‌پردازش‌ها قسمتی از داده‌ها از دست برود و کارایی سیستم شناسایی دوربین منبع کاهش یابد. در نتیجه انتخاب پیش‌پردازش‌های مناسب امری مهم است. روش پیشنهادی این پایان‌نامه بر روی پیش‌پردازش داده‌ها با استفاده از توابع افزایش وضوح تصویر، افزایش تضاد در تصویر و تشخیص لبه در تصویر تمرکز دارد. در ادامه، تعیین ساختار شبکه عصبی عمیق برای استخراج ویژگی مورد بررسی قرار گرفت. ساختار شبکه عصبی عمیق شامل دو لایه کانولوشنی، یک لایه جمع‌کننده حداکثری و دو لایه تمام متصل مورد استفاده قرار گرفت. برای طبقه‌بندی مجموعه داده‌گان از تابع softmax و ماشین بردار پشتیبان SVM در انتهای ساختارهای شبکه عصبی عمیق استفاده شد. در این پایان‌نامه برای جلوگیری از بیش‌برازش از روش DropConnect به جای روش Dropout استفاده

شد تا با حفظ نورون‌های بیشتر نسبت به روش Dropout الگوی نویز دقیق‌تر استخراج شود و کارایی شناسایی دوربین منبع بهبود پیدا کند. در سطح مدل و سنسور ساختار شبکه عصبی عمیق با تابع softmax توانست به ترتیب ۹۸,۵ درصد و ۹۵,۴ درصد از داده‌های آزمون را به درستی شناسایی کند. همچنین ساختار شبکه عصبی عمیق همراه با ماشین بردار پشتیبان به ترتیب ۹۸,۳ درصد و ۹۳,۲ درصد از داده‌های آزمون را به درستی شناسایی کرد. به این ترتیب، روش‌های پیشنهادی دقت‌های بالاتری را نسبت به مراجع دیگر داشته است.

کلمات کلیدی : جرم‌شناسی تصویری، شناسایی دوربین منبع، پیش‌پردازش، یادگیری

عمیق، ماشین بردار پشتیبان

فهرست مطالب

د	فهرست جداول
ه	فهرست اشکال
۱	فصل ۱: کلیات تحقیق
۲	۱-۱ مقدمه
۳	۲-۱ تعریف مسأله
۵	۳-۱ اهمیت و ضرورت انجام تحقیق
۶	۴-۱ اهداف تحقیق
۶	۵-۱ چالش‌های تحقیق
۷	۶-۱ جنبه نوآوری پژوهش
۸	۷-۱ ساختار پایان‌نامه
۹	فصل ۲ مطالعات انجام شده
۱۰	۱-۲ مقدمه
۱۱	۲-۲ اجزای یک دوربین دیجیتال
۱۳	۳-۲ انواع رویکرد در تشخیص جعل
۱۴	۱-۳-۲ رویکرد فعال
۱۴	۲-۳-۲ رویکرد غیرفعال
۲۱	۴-۲ شناسایی دوربین دیجیتال منبع
۲۱	۱-۴-۲ عدم تطابق قانونی لنز

۲۲	۲-۴-۲ انحراف رنگی
۲۳	۳-۴-۲ نقصان و خرابی حسگر
۲۴	۴-۴-۲ با استفاده از الحاق CFA
۲۶	۵-۴-۲ با استفاده از ویژگیهای تصویر
۲۶	۵-۲ استفاده از شبکه عصبی عمیق
۲۹	۱-۵-۲ شبکه های عصبی کانولوشن سنتی
۳۱	۲-۵-۲ پیشرفت داده
۳۳	۳-۵-۲ ادغام و ترکیب
۳۴	۴-۵-۲ انتخاب پیچ
۳۴	۵-۵-۲ طبقه بندی های مختلف
۳۵	۶-۲ جمع بندی
۳۹	فصل ۳: ارائه روش پیشنهادی
۴۰	۱-۳ مقدمه
۴۰	۲-۳ آماده سازی داده ها
۴۱	۱-۲-۳ پیش پردازش داده ها
۴۳	۳-۳ مدل CNN پیشنهادی
۴۵	۱-۳-۳ لایه کانولوشن
۴۶	۲-۳-۳ لایه جمع کننده
۴۶	۳-۳-۳ لایه تمام متصل
۴۷	۴-۳-۳ شناسایی دوربین منبع
۴۷	۵-۳-۳ انتخاب پارامترها
۴۸	۴-۳ جمع بندی

فصل ۴: نتایج و آزمایش‌ها

۴۹

۴-۱ مقدمه	۵۰
۴ معرفی مجموعه دادگان MICHE-I	۵۰
۴-۳ خلاصه آزمایش‌ها	۵۱
۴-۴ نتایج پیش‌پردازش داده‌ها	۵۲
۴-۵ نتایج طبقه‌بندی ساختارهای پیشنهادی	۵۴
۴-۶ ارزیابی ساختارهای پیشنهادی	۵۷
۴-۷ جمع‌بندی	۶۱

فصل ۵: نتیجه‌گیری

۶۳

۵-۱ جمع‌بندی و نتیجه‌گیری	۶۴
۵-۲ کارهای آتی	۶۵

۶۶

مراجع

۷۵

لغت‌نامه

به ترتیب حروف الفبای فارسی	۷۶
به ترتیب حروف الفبای انگلیسی	۸۰

فهرست جداول

- جدول ۱-۲) جدول شامل معماری‌های طراحی شده در زمینه شناسایی دوربین منبع [۱۲۵]..... ۳۶
- جدول ۲-۲) جدول شامل پارامترها، مجموعه داده‌ها و نتایج تجربی معماری‌های طراحی شده [۱۲۵].. ۳۷
- جدول ۱-۳) ساختار پیشنهاد شده پایان‌نامه (N تعداد کلاس‌های موجود)..... ۴۸
- جدول ۱-۴) مقایسه عملکرد ساختارها..... ۵۵

فهرست اشکال

شکل ۱-۱) فرآیند کلی شناسایی دوربین منبع	۵
شکل ۱-۲) اجزای اصلی یک دوربین دیجیتال	۱۱
شکل ۲-۲) طبقه‌بندی شناسایی جعل تصویر	۱۴
شکل ۳-۲) ساختار تعمیم یافته تشخیص جعل تصویر	۱۶
شکل ۴-۲) نمایش عملیات رتوش [۷۱]	۱۹
شکل ۵-۲) الف تصویر واقعی، ب تصویر شبیه‌سازی شده	۱۹
شکل ۶-۲) نمونه‌ای از پیوند تصویر، الف و ب تصاویر اصلی، ج تصویر پیوند خورده [۷۲]	۲۰
شکل ۷-۲) چارچوب کلی الگوریتم‌های مبتنی بر یادگیری عمیق برای شناسایی دوربین منبع [۱۲۵]	۲۸
شکل ۱-۳) فرآیند آماده‌سازی و پیش‌پردازش داده‌ها	۴۳
شکل ۲-۳) مدل‌های CNN پیشنهادی	۴۴
شکل ۱-۴) تصویر سه مورد ثبت شده با دوربین‌های پشتی سه دستگاه موبایل در فضای داخلی	۵۱
شکل ۲-۴) تصاویر ب حاصل اعمال پیش‌پردازش افزایش کنتراست تصاویر الف می‌باشد.	۵۲
شکل ۳-۴) تصاویر ب حاصل اعمال پیش‌پردازش افزایش وضوح تصویر بر روی تصاویر قسمت ب شکل ۴-۲ می‌باشد. تصاویر الف نیز تصاویر موجود در مجموعه دادگان بدون پیش‌پردازش می‌باشد.	۵۳
شکل ۴-۴) قسمت ب حاصل اعمال تمامی فیلترهای فرآیند پیش‌پردازش این پایان‌نامه می‌باشد.	۵۴
شکل ۵-۴) نمودار دقت بر اساس تعداد مراحل آموزش برای دو ساختار DCNN و DLSVM با اعمال پیش پردازش	۵۷
شکل ۶-۴) نمونه ماتریس درهم‌ریختگی 2×2	۵۸
شکل ۷-۴) ماتریس درهم‌ریختگی برای نتایج ساختار DCNN در سطح مدل	۵۹
شکل ۸-۴) ماتریس درهم‌ریختگی برای نتایج ساختار DLSVM در سطح مدل	۵۹
شکل ۹-۴) ماتریس درهم‌ریختگی برای نتایج ساختار DCNN در سطح سنسور	۶۰
شکل ۱۰-۴) ماتریس درهم‌ریختگی برای نتایج ساختار DLSVM در سطح سنسور	۶۰

فصل ۱: کلیات تحقیق

۱-۱ مقدمه

با استفاده سریع از دوربین‌های دیجیتال و استفاده گسترده از نرم‌افزارهای پردازش تصویر، مانند فتوشاپ، iPhoto و ACD See ویرایش و دستکاری عکس‌های دیجیتال بسیار آسان شده است. هم زمان، مسائل مربوط به امنیت اطلاعات بیشتر و بیشتر برجسته می‌شوند. استفاده‌ی گسترده‌ی تصاویر جعلی در رسانه‌های رسمی، تحقیقات علمی و شواهد بیمه به جد بر سیستم‌های اجتماعی وابسته به اطلاعات بصری قابل اعتماد، تاثیر می‌گذارد. به همین دلیل تکنولوژی تصویر دیجیتال پزشکی قانونی برای اطمینان از اصالت و یکپارچگی تصویر به وجود آمد [۱، ۲]. پزشکی قانونی تصویر، عمدتاً شامل شناسایی منبع [۳] و کشف جعل [۴] است. به عنوان یک هدف مهم تحقیقاتی در زمینه تکنولوژی پزشکی قانونی تصاویر دیجیتال، هدف از شناسایی دوربین منبع، شناسایی دستگاه تصویربرداری از تصویر ضبط شده است که به طور خاص به دو شکل تقسیم می‌شود:

۱. طبقه‌بندی تصاویر یا فیلم‌ها بر اساس نوع دستگاهی که آن‌ها را تولید می‌کند [۵، ۶]؛
۲. تشخیص اینکه کدام دستگاه برای ضبط یک تصویر یا فیلم دیجیتالی خاص استفاده شده است [۷، ۸].

شناسایی دوربین منبع در پزشکی قانونی تصویر می‌تواند برای شناسایی اصالت و مشروعیت تصاویر در سخنرانی‌های دادگاه و همچنین ردیابی منبع فیلم‌ها و تصاویر غیرقانونی استفاده شود. دستگاه تصویربرداری در طی فرآیند تولید، برخی از آثار یا نشانه‌ها را در تصویر باقی می‌گذارد. شناسایی برخی از دستگاه‌ها با ویژگی‌های منحصر به فرد دشوار است. از این رو، ردیابی این نشانه‌ها توسط ویژگی‌های فیزیکی دستگاه انجام می‌شود. ما می‌توانیم از این ویژگی‌های دستگاه‌ها به عنوان اثر انگشت دستگاه استفاده کنیم. در حال حاضر، شناسایی منبع رسانه‌های دیجیتال را می‌توان به دو دسته تقسیم کرد: شناسایی منبع براساس تشخیص همبستگی [۹-۱۲] و شناسایی منبع بر اساس طبقه‌بندی الگو [۱۳-۱۸].

شناسایی منبع بر اساس تشخیص همبستگی برای تعیین منبع با استخراج "اثر انگشت دستگاه" از

تصاویر مختلف ایجاد شده توسط یک دوربین و مقایسه همبستگی آن‌ها استفاده می‌شود. لوکاس و همکاران در سال ۲۰۰۵ روشی را با استفاده از نویز الگوی حسگر به عنوان اثر انگشت منحصر به فرد برای شناسایی دوربین پیشنهاد کردند [۹] و در سال ۲۰۱۳ آن روش را گسترش دادند [۱۰-۱۲].

از شناسایی دوربین منبع بر اساس طبقه‌بندی الگو می‌توان برای طبقه‌بندی تصاویر با توجه به برخی ویژگی‌های دوربین منبع با استفاده از تکنیک‌های یادگیری ماشین استفاده کرد. در این روش، استخراج ویژگی‌های مناسب تصویر مهم است. در [۱۴] و [۱۵]، از ویژگی‌های استخراج شده در یک دامنه فرکانس برای طبقه‌بندی مدل استفاده می‌شود. تواما و همکاران [۱۷] ابتدا یک روش طبقه‌بندی مدل دوربین برای Patch^۱ تصویر مبتنی بر شبکه‌های عصبی کانولوشن (CNN)^۲ پیشنهاد دادند. همچنین، روش‌های طبقه‌بندی مدل مبتنی بر CNN برای تصاویر بزرگ در [۵] و [۱۸] ارائه شده است. در این رویکردها، یک مدل CNN می‌تواند در مرحله آموزش به طور خودکار یاد بگیرد که چه نوع ویژگی‌هایی برای طبقه‌بندی مدل منبع مناسب است. همانطور که در [۵]، [۱۷] و [۱۸] نشان داده شده است، CNN پتانسیل بالایی برای دستیابی به طبقه‌بندی مدل بسیار دقیق دارد. در [۱۷]، نویز از تصاویر ورودی استخراج شده و به CNN وارد می‌شود. در [۵] و [۱۸] تصاویر مستقیماً به یک مدل CNN وارد می‌شوند. هر دو روش ورودی‌ها را به مدل‌های CNN خورنده و آنها را در یک حوزه مکانی آموزش می‌دهند. از طرف دیگر، همانطور که در [۱۴] و [۱۵] نشان داده شده است، ویژگی‌های استخراج شده در یک دامنه فرکانس نقش مهمی در طبقه‌بندی مدل یا نام تجاری دارند. به عنوان مثال، ویژگی دوره‌ای نویز تصویر یکی از انواع مهم اطلاعات برای طبقه‌بندی مدل است. چنین ویژگی‌هایی را می‌توان به عنوان سیگنال‌های اوج یک دامنه فرکانس مشاهده کرد.

۱-۲ تعریف مسأله

با توسعه فناوری دیجیتال، تصاویر دیجیتال را می‌توان به راحتی از دوربین دستگاه‌های مختلف به دست

^۱ پچ قطعه ای کوچک (به طور کلی مستطیل شکل) از تصویر است.

^۲ Convolutional Neural Network

آورد و به طور گسترده در سیستم عامل‌های شبکه‌های اجتماعی گسترش داد. در این زمینه، دستکاری‌های تصویری و ویدئویی بسیار رایج شده و برای افراد و کل جامعه به طور فزاینده‌ای خطرناک است. در دهه گذشته، با استفاده از این پدیده‌ها، روش‌های زیادی برای تشخیص جعلی و بومی‌سازی یا شناسایی دوربین پیشنهاد شده است [۱۹-۲۱]. تحقیقات پزشکی قانونی در مورد تصاویر دیجیتال برای ارزیابی اصالت تصاویر است. شناسایی مدل دوربین اولین مرحله برای تحقیقات پزشکی قانونی تصویر است. شناسایی دوربین منبع (SCI)^۱، موضوع مورد بحث این پایان‌نامه است. در اکثر مقالات مرتبط با موضوع شناسایی دوربین منبع، از ساختار شبکه‌های عصبی عمیق کانولوشنی CNN استفاده شده است. معماری CNN، جهت استخراج الگوی نویز سنسورهای (SPN)^۲ دوربین استفاده شده است. این الگوی نویز به عنوان اثر انگشت هر دوربین شناخته می‌شود. هدف نهایی استخراج الگوی نویز، شناسایی مدل دوربین هر تصویر است که در مباحث جرم‌شناسی تصویری^۳ کاربرد دارد. در مقاله پایه [۳۲] به طور خاص، تنها معماری CNN تحلیل می‌شود. ولی امکان استفاده از شبکه‌های عصبی عمیق دیگر نیز وجود دارد که به فراخور کاربرد می‌توان از آن‌ها استفاده کرد. تحقیقات بسیاری وجود دارد که نویسندگان از روش‌های یادگیری عمیق در پزشکی قانونی استفاده می‌کنند. به عنوان مثال، در مقاله [۲۲]، نویسندگان از CNN برای تشخیص فشرده‌سازی دو برابر (JPEG)^۴ استفاده کردند، در حالی که در [۵]، نویسندگان با استفاده از CNN، ویژگی‌ها و طبقه‌بندی (SVM)^۵ را در بالای ویژگی‌های استخراج شده استخراج کردند. روش‌ها و رویکردهای مختلفی توسط محققان برای شناسایی دوربین منبع پیش‌بینی شده است که تاکنون شامل ویژگی‌های استخراج تصویر [۲۳-۲۵]، استفاده از فراداده [۲۶]، آرایه فیلتر رنگ (CFA)^۶ درهم آمیختن [۱۷، ۲۴]، اعوجاج شعاعی لنز ذاتی [۲۷]، استخراج نویز غیر یکنواختی پاسخ عکس [۱۷] [۲۸]، الگوی نویز اکتساب و غیره است. اما مشکل فراداده این است که به راحتی می‌توان آن‌را دستکاری کرد. بنابراین، استفاده از روش فراداده برای شناسایی دوربین منبع مناسب نیست. باز هم، استخراج ویژگی دست‌ساز از یک تصویر برای تعداد زیادی از تصاویر کار ساده‌ای

¹ Source Camera Identification

² Sensor Pattern Noise

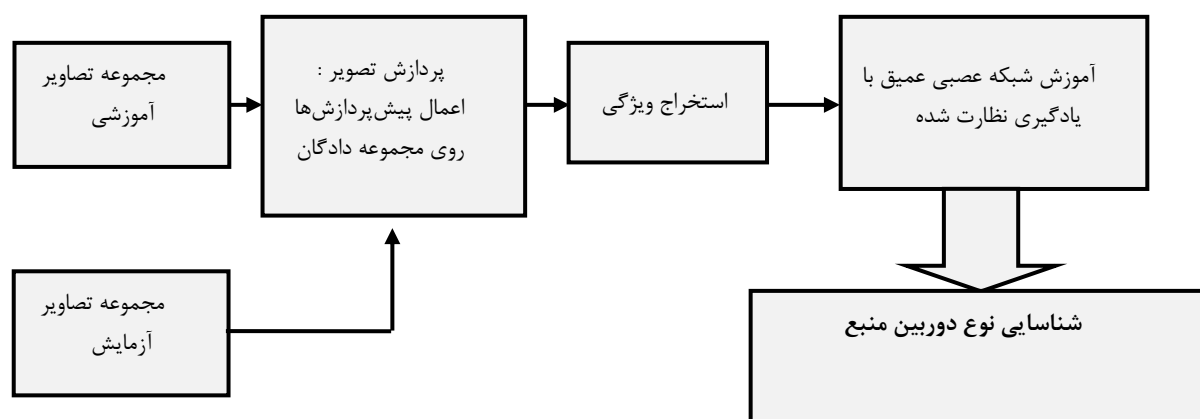
³ Image Forensic

⁴ Joining Photographic Experts Group

⁵ Support Vector Machines

⁶ Color Filter Fray

نیست. در این پژوهش، ما از روش یادگیری عمیق با استفاده از CNN برای شناسایی دوربین منبع از یک تصویر داده شده استفاده کرده‌ایم. فرآیند کلی شناسایی دوربین منبع روش پیشنهادی را در شکل زیر به وضوح نشان داده‌ایم.



شکل ۱-۱) فرآیند کلی شناسایی دوربین منبع

۱-۳ اهمیت و ضرورت انجام تحقیق

با توجه به در دسترس بودن روزافزون دستگاه‌های کسب تصویر و سیستم عامل‌های اشتراک چندرسانه‌ای، عکس‌ها به بخشی گسترده از زندگی روزمره ما تبدیل شده‌اند. از آنجا که بارگیری، کپی، جعل و توزیع مجدد رسانه‌های دیجیتال در طول سال‌ها آسان‌تر شده است، برخی از اشکال تنظیم و بررسی صحت به یک ضرورت فوری تبدیل شده‌اند. به همین منظور، جامعه پزشکی قانونی مجموعه وسیعی از ابزارها را برای مهندسی معکوس تاریخ اشیا چندرسانه‌ای ایجاد کرده است [۲۰، ۲۱]. از جمله مشکلاتی که انجمن پزشکی قانونی تصاویر با آن روبرو شده، شناسایی مدل دوربین است [۲۹، ۳۰]. با استفاده از یک تصویر، مشخص می‌شود که کدام نوع دوربین برای فیلم‌برداری یا تصویربرداری استفاده شده است. حل این مشکل می‌تواند به یک تحلیلگر در اشاره به صاحب مطالب غیرقانونی و بحث برانگیز (به عنوان مثال عکس-های پدو پورنوگرافی، صحنه‌های اقدام تروریستی و غیره) کمک کند. علاوه بر این، تجزیه و تحلیل انجام شده بر روی Patch‌های تصویری می‌تواند برای افشای جعل‌های متصل شده بین تصاویر از دوربین‌های مختلف استفاده شود [۳، ۳۱].

۴-۱ اهداف تحقیق

هنگامی که ما روشی برای پردازش تصاویر دیجیتال انتخاب می‌کنیم، وجود تصاویر دیجیتالی در تلفن‌های همراه گرفته تا صفحات سایت‌های خبری آنلاین و غیره مجموعه‌ای شگفت‌آور از پرسش‌های قانونی و اخلاقی را ایجاد می‌کند که باید به آن‌ها بپردازیم. به عنوان مثال، پیامدهای دستکاری داده‌ها در تصویر یک فرد بی‌گناه چیست؟ آیا این نوع پردازش یک عمل امنیتی عمدی است که به طور مشروع استفاده می‌شود، یا یک فریب عمدی است؟ آیا دستکاری در تصویر یا جعل تصاویر در مواردی که تصویر ممکن است بر رفتار و افکار عمومی تأثیر بگذارد، مناسب است؟ حتی قبل از اینکه اقداماتی بر روی یک تصویر سوال برانگیز انجام شود، باید بتوانیم مواردی را در مورد آن تصویر تشخیص دهیم، مثل نوع دوربین تصویری آن عکس.

هدف این پایان‌نامه پیاده‌سازی الگوریتمی است که بتواند با استفاده از شبکه عصبی مصنوعی عمیق، دوربین منبع تصاویر دیجیتال را با بالاترین دقت شناسایی و ردیابی کند. در شناسایی دوربین منبع، ما به دنبال شناسایی مدل خاص یک دوربین یا شناسایی دقیق دوربینی هستیم که یک تصویر را تولید کرده است.

به عبارت دیگر، هدف دیگر از پیاده‌سازی الگوریتم شناسایی دوربین دیجیتال، شناسایی و تشخیص متمایز انواع حسگرهای دوربین بر اساس تصاویر تولید شده است. این موضوع از سال‌های گذشته به خصوص در جامعه پزشکی قانونی بسیار مورد توجه قرار گرفته است.

۵-۱ چالش‌های تحقیق

با ظهور اینترنت و دوربین‌های دیجیتال ارزان قیمت و همچنین نرم‌افزارهای قدرتمند ویرایش تصویر (مانند Adobe Photoshop و Illustrator و GNU Gimp)، کاربران عادی بیش از گذشته به ابزارهای دیجیتال دسترسی دارند. هم‌زمان، علم ما از پیامدهای فنی، اخلاقی و حقوقی مرتبط با ویرایش تصویر بسیار عقب است. گاهی ممکن است دستکاری یک تصویر (مثلاً تصویر از یک فرد بی‌گناه) تهدیدهای قانونی را برای آن فرد و جامعه به همراه داشته باشد. از این رو طراحی و بکارگیری رویکردهای کارآمد و موثر برای

شناسایی چنین فعالیت‌هایی از اهمیت بالاتری برخوردار است. در این راستا روش‌ها و چالش‌های بسیاری وجود دارد. در این بخش ما این چالش‌ها را به صورت دسته‌بندی شده بیان می‌کنیم و سپس در فصل دوم به بررسی و تحلیل این چالش‌ها می‌پردازیم.

- تشخیص گرافیک رایانه‌ای از عکاسی دیجیتال
- شناسایی و کشف تکنیک‌های ضد پزشکی قانونی در برابر شناسایی دوربین
- تشخیص تصویر شبیه‌سازی شده
- ناسازگاری‌های روشنایی و نور

۱-۶ جنبه نوآوری پژوهش

اکثر دوربین‌ها از لنز برای گرفتن تصویر استفاده می‌کنند. هر لنز برای هر دوربین دارای اعوجاج شعاعی لنز ویژه است و این ویژگی به عنوان یک اثر انگشت منحصر به فرد برای شناسایی دوربین منبع در نظر گرفته می‌شود. در مقاله‌ی منبع برای شناسایی دوربین منبع تنها بر روی شبکه‌های عصبی مصنوعی CNN متمرکز شده است. در این پژوهش، ما از روش یادگیری عمیق با استفاده از CNN برای شناسایی دوربین منبع، استفاده کرده‌ایم. روش‌های مختلفی برای طبقه‌بندی تصاویر، مانند رویکرد یادگیری ماشین، رویکرد یادگیری عمیق و غیره وجود دارد [۳۲]. بسیاری از محققان از روش یادگیری ماشین برای شناسایی دوربین منبع استفاده کرده‌اند، این روش برای حجم زیاد داده‌های ورودی مناسب نیست. اما در صورت وجود مقدار زیادی داده، رویکرد یادگیری عمیق به جای روش یادگیری ماشین، از دقت بالایی برخوردار است. بنابراین، امروزه، یادگیری عمیق برای مقادیر زیادی از داده محبوبیت بیشتری دارد و به خصوص CNN عمیق در بین محققان برای طبقه‌بندی تصویر و تشخیص اشیا رواج دارد.

در این پژوهش ما از CNN عمیق برای دو مفهوم از شناسایی دوربین استفاده می‌کنیم: شناسایی دوربین منبع منفرد (ISCI)^۱ و شناسایی مدل دوربین منبع (SCMI)^۲. ISCI با هدف تشخیص یک دوربین خاص در

¹Individual Source Camera Identification

²Source Camera Model Identification

میان دوربین‌های با مدل‌های یکسان و مدل‌های مختلف است، در حالی که SCMI یک مدل دوربین خاص را در میان سایر دوربین‌ها متمایز می‌کند اما نمی‌تواند یک دوربین خاص را از بین همان مدل‌های دوربین تشخیص دهد.

۷-۱ ساختار پایان‌نامه

ساختار این پایان‌نامه به صورت زیر سازمان‌دهی شده است: در فصل دوم پیش‌زمینه‌ای از کارهای انجام‌شده در خصوص انواع روش‌های شناسایی دوربین منبع بررسی خواهد شد. فصل سوم شامل جزئیات روش پیشنهاد شده می‌باشد. فصل چهارم نیز نتایج شبیه‌سازی و ارزیابی‌های انجام‌شده را نشان خواهد داد. در فصل پنجم نتیجه‌گیری کلی و چالش‌های اصلی در زمینه شناسایی دوربین منبع را بیان خواهیم کرد همچنین کارهای آینده را نیز در این فصل بیان خواهیم کرد.

فصل ۲ : مطالعات انجام شده

۲-۱ مقدمه

با توسعه فناوری دیجیتال، تصاویر دیجیتال را می‌توان به راحتی از دستگاه‌های دیجیتال مجهز به دوربین به‌دست آورد و به طور گسترده در سیستم عامل‌های شبکه‌های اجتماعی گسترش داد. در همین حال، با استفاده از نرم‌افزارهای ویرایش تصویر می‌توان تصاویر دیجیتالی را به راحتی دستکاری کرد و اطلاعات نسبی مشترک بین تصویر و دوربین دیجیتال توسط مجرمان غیرقانونی یا سازمان‌های غیرمجاز حذف و منجر به نقض مالکیت شود. از این رو، سازمان قضایی نیاز فوری به مطالعه روش‌های قانونی و قابل اعتماد پزشکی قانونی دارد. در همین راستا جامعه پزشکی قانونی، صحت تصاویر دیجیتال را مورد توجه قرار داده است. در چند سال گذشته، پزشکی قانونی چندرسانه‌ای بیشتر در دو زمینه مورد توجه قرار گرفته است: ۱- شناسایی دوربین منبع و ۲- کشف جعل.

شناسایی دوربین منبع در زمینه شناسایی دستگاه‌های دیجیتال منبع (دوربین‌ها، تلفن‌های همراه، دوربین‌های فیلمبرداری و غیره) با استفاده از رسانه‌های تولید شده به وسیله آن‌ها متمرکز است. در حالی که شناسایی جعل تلاش می‌کند تا اصالت رسانه‌های دیجیتال (کلیپ‌های صوتی، کلیپ‌های ویدیویی، تصاویر و غیره) را ارزیابی کند.

در این فصل، ما ابتدا چندین روش در پزشکی قانونی تصویر مربوط به شناسایی دوربین منبع و تشخیص جعل را بیان کرده و در ادامه کارهای انجام شده در این حوزه را بررسی می‌کنیم. شناسایی دوربین منبع مراحل مختلف پردازش تصاویر دوربین دیجیتال را برای ویژگی‌های منحصر به فرد دستگاه کاوش می‌کند و از وجود اعوجاج شعاعی لنز [۲۷]، نقص سنسور [۱۰، ۳۳]، آرایه فیلتر رنگ (CFA) درهم‌آمیزی [۲۴، ۳۴]، [۳۵]، ویژگی‌های ذاتی تصویر [۲] و غیره بهره می‌برد. جعل تصویر شامل انشعاب از تصاویر برای ساختن یک تصویر جدید، استفاده از تکثیر منطقه یا تغییر مکان برای پنهان کردن یا جابه‌جایی برخی از اشیا در تصویر و اعمال ویرایش تصویر برای حذف یا اضافه کردن اشیا جدید از تصویر یا به تصویر است. برخی از روش‌ها

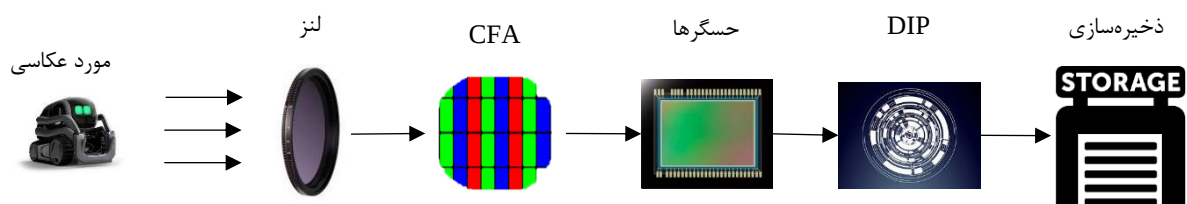
برای تشخیص جعل، از ناسازگاری در انحراف رنگی [۳۶]، روشنایی [۳۷] و عملکرد پاسخ دوربین (CRF^۱) [۳۸] به عنوان نشانه‌های جعل استفاده می‌کنند.

۲-۲ اجزای یک دوربین دیجیتال

ساختار کلی دوربین دیجیتال در شکل ۲-۱ نشان داده شده است. دوربین‌های دیجیتال از سیستم لنز، فیلترها، آرایه فیلتر رنگ (CFA^۲)، حسگر تصویر و پردازشگر تصویر دیجیتال (DIP^۳) تشکیل شده‌اند [۳۹]. انحراف ناشی از لنزها، مانند انحراف رنگی و انحراف کروی ممکن است تصاویر رنگی ثبت شده را دچار اختلال کند.

انحراف رنگی، عدم همگرایی طول موج‌های مختلف در همان موقعیت بر روی سنسور است، درحالی‌که انحراف کروی باعث می‌شود تا نور عبوری از محیط پیرامون عدسی کروی نسبت به نوری که از مرکز عدسی عبور می‌کند، در نقطه نزدیک‌تر به عدسی جمع شود.

در سیستم‌های لنز می‌توان با استفاده از ترکیبات خاص لنزهای محدب و مقعر و همچنین استفاده از لنزهای کروی، این اثرات را به حداقل رساند.



شکل ۲-۱ اجزای اصلی یک دوربین دیجیتال

سیستم لنز همچنین شامل کنترل نوردهی خودکار، تنظیم فاصله کانونی خودکار و واحد تثبیت‌کننده تصویر است. دیافراگم خودکار و سرعت شاتر به همراه یک کنترل‌کننده خودکار کالیبره شده،

¹ Camera Response Function

² Color Filter Array

³ Digital Image Processing

کیفیت تصاویر را با تغییر نوردهی افزایش می‌دهد. تنظیم فاصله کانونی خودکار، لنزها را با یک موتور مینیاتوری به داخل و خارج حرکت داده تا زمانی که واضح‌ترین تصویر ممکن از سوژه به دست آید. تثبیت کننده تصویر با خنثی کردن لرزش دوربین به ارائه تصاویر واضح‌تر کمک می‌کند.

پس از عبور از لنزها، نور از مجموعه‌ای از فیلترها عبور می‌کند. فیلتر مادون قرمز یک فیلتر جاذب یا بازتابنده است که اجازه می‌دهد فقط قسمت قابل مشاهده طیف عبور کند، مسدود کردن اشعه مادون قرمز می‌تواند وضوح تصویر تشکیل شده را کاهش دهد. وقتی فاصله بین پیکسل‌های سنسور نمی‌تواند فرکانس فضایی دقیق‌تر از اشیای هدف مانند الگوهای تزئینی را پشتیبانی کند، فیلتر anti-aliasing میزان نویز را کاهش می‌دهد.

قلب دوربین دیجیتال حسگر تصویر است. سنسور تصویر آرایه‌ای از ردیف‌ها و ستون‌ها از عناصر فوتودیود یا پیکسل است. وقتی نور به آرایه پیکسل برخورد می‌کند، هر پیکسل متناسب با شدت نور سیگنال آنالوگ تولید می‌کند، سپس به یک سیگنال دیجیتال تبدیل می‌شود و توسط DIP پردازش می‌شود. بیشتر دوربین‌های دیجیتال از یک دستگاه شارژ همراه^۱ (CCD) به عنوان حسگر تصویر استفاده می‌کنند اگر چه تراشه‌های (CMOS)^۲ یک گزینه محبوب هستند. پیکسل‌های سنسور به رنگ‌ها حساس نیستند، آن‌ها فقط روشنایی نور را ثبت می‌کنند، بنابراین یک خروجی تک رنگ ایجاد می‌کنند. برای تولید یک تصویر رنگی، از آرایه فیلتر رنگ (CFA) در جلوی سنسور استفاده می‌شود تا هر پیکسل شدت نور را فقط برای یک رنگ ثبت کند. بیشتر دوربین‌های دیجیتال از CFA با الگوی بایر سبز-قرمز-سبز-آبی (GRGB) استفاده می‌کنند. خروجی از سنسور با فیلتر Bayer موزاییکی از پیکسل‌های قرمز، سبز و آبی با شدت‌های مختلف است. از آنجا که هر پیکسل فقط یکی از سه رنگ را ضبط می‌کند، تصویر کامل رنگی توسط DIP با استفاده از درون‌یابی مختلف انجام می‌شود. از دیگر فیلترهای CFA جایگزین می‌توان به الگوی فیروزه‌ای-زرد-سبز-سرخابی^۳ (CYGM)، الگوی قرمز-سبز-آبی-زرد^۴ (RGBE) و الگوی فیروزه‌ای-قرمز-زرد

¹ Charge-Coupled Device

² Complementary Metal Oxide Semiconductor

³ Cyan-Yellow-Green-Magenta

⁴ Red-Green-Blue-Emerald

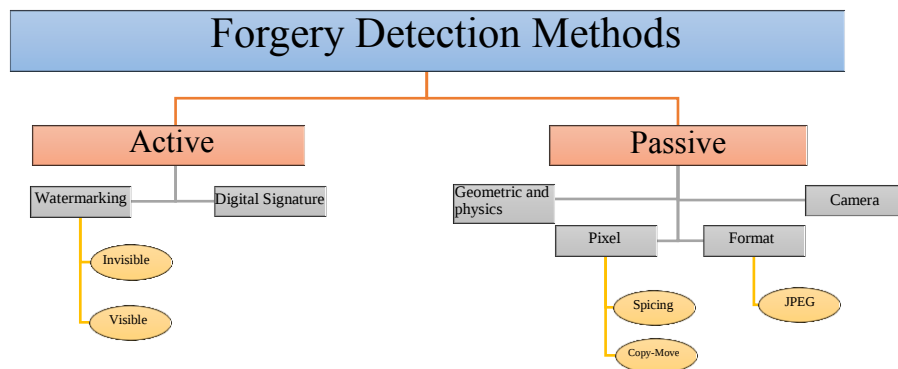
(CMY)^۱ اشاره کرد.

DIP علاوه بر درونیایی، پردازش‌های بیشتری مانند تعادل رنگ سفید، کاهش نویز، دستکاری ماتریس، وضوح تصویر، اصلاح دیافراگم و اصلاح گاما را برای تولید یک تصویر با کیفیت خوب انجام می‌دهد.

۲-۳ انواع رویکرد در تشخیص جعل

تشخیص جعل تصویر با هدف تأیید صحت یک تصویر دیجیتالی انجام می‌شود. راه حل تأیید اعتبار تصویر به دو گروه (۱) فعال و (۲) غیرفعال یا منفعل طبقه‌بندی می‌شود. در شکل ۲-۲ این طبقه‌بندی نشان داده شده است. تکنیک‌های فعال جعل تصویر، مانند علامت‌گذاری دیجیتال [۴۰] یا علامت‌های دیجیتالی [۴۱]، قبل از ارسال تصاویر در شبکه‌هایی که ممکن است تصویر مورد دستکاری قرار بگیرد، از کد احراز هویت شناخته شده‌ای که در محتوای تصویر جاسازی شده است، استفاده می‌کند. با تأیید حضور چنین کد احراز هویتی، و با مقایسه با کد درج شده اصلی، تأیید اعتبار امکان پذیر است. با این حال این روش برای درج کد احراز هویت در داخل تصویر قبل از توزیع تصویر، به سخت افزار یا نرم افزار خاصی نیاز دارد. روش تشخیص جعل غیرفعال، بدون هیچگونه امضا یا علامت در تصویر اصلی، از تصویر دریافت شده از فرستنده فقط برای ارزیابی اصالت یا یکپارچگی آن، استفاده می‌کند. فرض بر این است که اگر چه جعل‌های دیجیتالی ممکن است هیچ سرنخی از دستکاری در تصویر ایجاد نکند، اما به احتمال زیاد ویژگی‌های اصلی آماری یک تصویر را به نحوی مختل می‌کنند. برای مثال اگر تصویر متشکل از یک منظره طبیعی است یک نشانه یا ناسازگاری جدیدی در تصویر از خود باقی می‌گذارد. از این ناسازگاری‌ها می‌توان برای تشخیص جعل استفاده کرد. این روش به این دلیل مورد توجه است که به هیچگونه اطلاعات قبلی در مورد تصویر نیاز ندارد.

¹ Cya- Magenta-Yellow



شکل ۲-۲ طبقه‌بندی شناسایی جعل تصویر

۲-۳-۱ رویکرد فعال

در رویکرد فعال، سخت‌افزار یا نرم‌افزار ثبت‌کننده‌ی تصویر اطلاعاتی را در داخل تصویر تعبیه می‌کند که از نظر چشم انسان تصویر تغییری نکند. این اطلاعات تعبیه شده، مانند یک امضاء روی تصویر هستند و اگر شخصی تصویر را دستکاری کند، اطلاعات گنجانده شده را نیز به ناچار دستکاری کرده است. در نتیجه نظم داده‌های تعبیه شده را بر هم زده و به راحتی می‌توان تشخیص داد که تصویر دستکاری شده است. از جمله روش‌های معروف که از این رویکرد پیروی می‌کند، واترمارکینگ هستند. این روش، ساده‌ترین و مؤثرترین روش محافظت از حقوق کپی رایت تصویر است [۴۲]. اما مشکل بزرگ این رویکرد محدود بودن به دوربین‌های خاص است که عمل تعبیه کردن اطلاعات را انجام می‌دهند و اگر دستگاهی این کار را انجام ندهد، نمی‌توان در مورد صحت تصاویرش نظری داد.

۲-۳-۲ رویکرد غیرفعال

در تکنیک تشخیص جعل غیرفعال، برای ارزیابی اصالت، تصویر را بدون علامت مشخصه یا امضای تصویر اصلی از فرستنده بررسی می‌کند. در این تکنیک فرض می‌کنیم هیچ سرنخ تصویری از دستکاری در تصاویر دیجیتال وجود نداشته باشند. این تکنیک محبوب است زیرا به هیچ اطلاعات قبلی در مورد تصویر نیاز ندارد. تکنیک‌های موجود ردیابی‌های مختلف دستکاری را شناسایی می‌کند و آن‌ها را بطور جداگانه با محلی‌سازی منطقه دستکاری شده شناسایی می‌کند. در این زمینه از چهار تکنیک مبتنی بر پیکسل [۴۳]، مبتنی بر

قالب، مبتنی بر هندسه و فیزیک و مبتنی بر دوربین استفاده می‌شود.

هدف از شناسایی کور یا غیرفعال طبقه‌بندی تصاویر داده شده به دو کلاس تصاویر اصلی (یا معتبر) و جعلی است. رویکردهای غیرفعال تشخیص جعل تصویر، ابتدا ویژگی‌ها را از تصاویر استخراج می‌کنند، سپس یک طبقه‌بندی کننده را انتخاب کرده و با استفاده از ویژگی‌های استخراج شده از مجموعه‌های آموزش تصویر، طبقه‌بندی را آموزش می‌دهند و در نهایت ویژگی‌ها را طبقه‌بندی می‌کنند. تعداد کمی از این رویکردها در مراجع [۴۴-۴۹] پیشنهاد شده است.

در بخش زیر یک چارچوب کلی از رویکرد غیرفعال تشخیص جعل تصویر را بررسی می‌کنیم:

۱. پیش‌پردازش تصویر: قبل از فرآیند استخراج ویژگی، برخی عملیات بر روی تصاویر مورد بررسی مانند برش، تبدیل تصویر RGB به مقیاس خاکستری، تغییر شکل (DCT)^۱ یا (DWT)^۲ برای بهبود عملکرد طبقه‌بندی، انجام می‌شود.

۲. استخراج ویژگی‌ها: مجموعه‌ای از ویژگی‌ها برای هر کلاس استخراج می‌شود که به تمایز آن از کلاس‌های دیگر کمک می‌کند و در عین حال نسبت به تفاوت‌های مشخصه درون کلاس از داده‌های جعلی ورودی ثابت است. به طور خاص، ویژگی‌های آموزنده و ویژگی‌های انتخابی را که باید به دستکاری تصویر حساس باشند، استخراج می‌کنیم. یکی از مشخصه‌های مطلوب ویژگی‌های انتخاب شده و بردار ویژگی‌های ساخته شده باید با ابعاد کم باشد، که از پیچیدگی محاسباتی آموزش و طبقه‌بندی می‌کاهد.

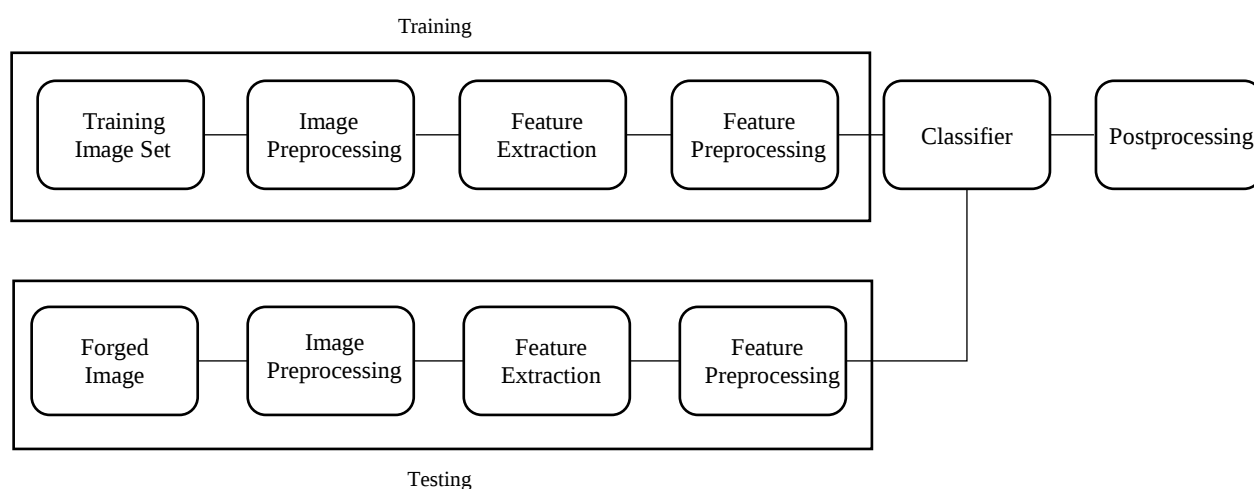
۳. انتخاب طبقه‌بندی و پیش‌پردازش ویژگی‌ها: بر اساس مجموعه ویژگی‌های استخراج شده، طبقه‌بندهای مناسب را طراحی کرده و مجموعه بزرگی از تصاویر را برای آموزش طبقه‌بندها انتخاب می‌کنیم. برخی از پارامترهای مهم طبقه‌بندی را برای طبقه‌بند به دست می‌آوریم. پیش‌پردازش ویژگی‌ها برای کاهش ابعاد ویژگی‌ها بدون کاهش عملکرد طبقه‌بند مبتنی بر یادگیری ماشین و کاهش هم‌زمان پیچیدگی محاسباتی استفاده می‌شود [۵۰].

¹ Discrete Cosine Transform

² Discrete Wavelet Transform

۴. طبقه‌بندی: هدف از طبقه‌بندی، ایجاد تمایز در تصاویر ارائه شده و طبقه‌بندی آن‌ها به دو دسته تصاویر اصلی و جعلی است. طبقه‌بندهای مختلفی مانند SVM در مراجع [۵۱-۵۶] استفاده شده است.

۵. پس پردازش: در برخی از مواردی مانند کپی-انتقال و تقسیم، عملیات پس پردازش شامل محلی‌سازی منطقه جعلی به عنوان موضوع اصلی مورد تحقیق در مراجع [۵۷-۶۰] است. با توجه به مراحل شرح داده شده، ساختار تشخیص جعل منفعل تصویر در شکل ۲-۳ ارائه شده است.



شکل ۲-۳ ساختار تعمیم یافته تشخیص جعل تصویر

۲-۳-۱-۲ تشخیص جعل تصاویر مبتنی بر قالب ذخیره سازی
در این تکنیک تصاویر با فرمت‌های مختلفی ذخیره‌سازی می‌شوند که یکی از این فرمت‌ها JPEG می‌باشد. صحت تصاویر JPEG به کمک ویژگی‌های آماری مرتبط با فشرده‌سازی، مورد بررسی قرار می‌گیرد [۶۱].

این روش کارایی بالایی ندارد ولی می‌تواند به تشخیص جعل کمک کند.

۲-۳-۲-۲ تشخیص جعل تصاویر مبتنی بر قوانین فیزیک و هندسی
در هر تصویری که ثبت می‌شود سه عامل نور، دوربین و سوژه مورد عکاسی حضور دارند، بررسی نور در

این سه عامل می‌تواند صحت تصویر را بررسی کند. تصویری را که در فضای باز و هوای آفتابی گرفته شده و مورد دستکاری قرار گرفته را در نظر بگیرید بدلیل اینکه نمی‌توان نور را به گونه‌ای روی اجسام و افراد حاضر در تصویر تنظیم کرد که همه‌ی سایه‌ها یکسان باشد، براحتی صحت تصویر مشخص می‌شود. نورپردازی مهم‌ترین عامل در ثبت تصاویر است. در این تکنیک جهت و شدت نور در قسمت‌های مختلف تصویر اندازه‌گیری شده و هماهنگی‌های بین آن‌ها مورد بررسی قرار می‌گیرد و در نتیجه جعل تشخیص داده می‌شود [۶۲].

تناقضات هندسی موجود در تصاویر نیز به کمک تخمین فاصله‌ها و اندازه اجسام داخل تصویر مشخص می‌شود. تناقضات هندسی با بررسی بازتاب نور و سایه اجسام قابل تشخیص هستند. اگر شی بازتاب کننده مانند آینه، سطح فلزی براق، سطح آب و غیره در تصویر موجود باشد، با بررسی هم‌خوانی اجسام حاضر در تصویر و شی بازتاب کننده می‌توان جعل را تشخیص داد. شکل سایه اجسام به‌راحتی قابل جعل نیست و معمولاً در اندازه و جهت سایه اجسام در تصویر جعلی تناقض وجود دارد، با پیدا کردن این تناقضات می‌توان تصاویر جعلی را متمایز کرد.

۳-۲-۳-۲ تشخیص جعل تصویر مبتنی بر پیکسل

این روش با تمرکز بر پیکسل‌های تصویر دیجیتال، تضادهای به وجود آمده در تصویر را بر اساس ویژگی‌ها و روابط آماری موجود در پیکسل‌های تصاویر طبیعی بررسی می‌کند. از جمله این روش‌ها می‌توان به نمونه‌برداری مجدد [۶۳، ۶۴] و بهبود تضاد در تصویر [۶۵] و توصیف‌گرهای محلی [۶۶، ۶۷] اشاره کرد. روش مبتنی بر پیکسل به طور کلی به چهار دسته کلی- انتقال، پیوندزنی، رتوش کردن و ویژگی‌های آماری تقسیم می‌شود [۶۳].

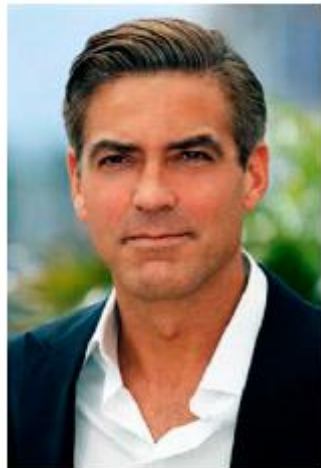
عملیات پس‌پردازش روی تصاویر می‌تواند استفاده از تکنیک‌های تشخیص جعل را دچار محدودیت کند. تغییر اندازه تصویر و پنهان کردن مرزها می‌تواند تشخیص جعل را دچار اختلال کند ولی میزان نور و عدم تطبیق سایه و نویز الگوی سنسور عواملی هستند که با عمل پس‌پردازش از بین نمی‌رود. با این حال، ناهماهنگی‌های روشنایی تنها در موارد وجود تناقض در نور مفید هستند. عدم تطابق در اندازه هنگامی قابل

استفاده است که اندازه واقعی اشیاء موجود در تصویر موجود باشد و نویز الگوی سنسور نیز نیازمند اطلاعات اضافی از دوربین است.

کپی رایج ترین تکنیک دستکاری تصویر است که در آن قسمت‌هایی از تصویر اصلی کپی شده، به مکان دلخواه منتقل شده و جای‌گذاری می‌شود. این کار معمولاً به منظور پنهان کردن جزئیات خاص یا کپی برداری از برخی جنبه‌های تصویر انجام می‌شود. مناطق بافتی به عنوان قطعات ایده‌آل برای جعل کپی-انتقال استفاده می‌شود، زیرا مناطق دارای بافت دارای ویژگی‌های تغییر رنگ و نویز مشابه تصویر هستند که برای چشم انسان غیرقابل تصور است. تار شدن معمولاً در امتداد مرز منطقه اصلاح شده استفاده می‌شود تا اثر بی‌نظمی بین منطقه اصلی و چسبیده را کاهش دهد. اولین تلاش در شناسایی مناطق دستکاری شده توسط فریدریش و همکاران انجام شد [۶۸]. نویسندگان روش جلوگیری از جعل کپی-انتقال را با استفاده از تبدیل DCT بلوک‌های هم‌پوشانی پیشنهاد کردند. بهترین تعادل بین عملکرد و پیچیدگی با استفاده از الگوریتم تطبیق بلوک بدست آمده است. در مرجع [۶۹] روشی با استفاده از تجزیه و تحلیل ترکیب اصلی (PCA)^۱ برای نمایش تقسیم‌بندی‌های تصویر یعنی بلوک‌های مربع هم‌پوشانی ارائه شده است. در PCA، نمایش تقسیمات تصویر یعنی بلوک‌های مربع با هم تداخل دارند. نتایج مبتنی بر PCA، نتایج در کاهش هزینه محاسبات و تعداد محاسبات مورد نیاز $O(N_t N \log N)$ است. جایی که N_t ابعاد نمایش PCA کوتاه شده و N تعداد پیکسل‌های تصویر است.

یکی از مشهورترین نوع جعل تصاویر رتوش تصویر است که در آن نکات برجسته تصاویر با ایجاد تغییرات ریاضی مانند چرخش، مقیاس گذاری، گسترش، شیب، تلنگر و غیره برای افزایش جذابیت و کیفیت تصاویر، ارتقا یا کاهش می‌یابد. این روش، یکی از قدیمی‌ترین انواع جعل تصویر است و از نوع جعل کم خطر محسوب می‌شود. برای مثال از رتوش تصویر برای صفحه جلد یک مجله توسط عکاسان برای بهبود ویژگی‌های خاص تصویر استفاده می‌شود تا مجله جلوه بهتری داشته باشد. در شکل ۲-۴ یک نمونه از عملیات رتوش تصویر نشان داده شده است.

1 Principal Component Analysis



ب: تصویر رتوش شده



الف: تصویر اصلی

شکل ۲-۴) نمایش عملیات رتوش [۷۱]

شبیه‌سازی تصویر (کپی- انتقال) چرخه‌ای است که برخی از ویژگی‌های تصویر با هر اندازه و شکلی در برخی از مناطق مشابه دوباره جاگذاری می‌شوند تا اطلاعات قابل توجهی را پوشش دهد. در شکل ۲-۵ شکل تصویر ب با کمک عمل شبیه‌سازی جعل شده است.



الف



ب

شکل ۲-۵) الف تصویر واقعی، ب تصویر شبیه‌سازی شده

در روش پیوندزنی دو تصویر با یکدیگر ترکیب شده یا به اصطلاح دو تصویر پیوند می‌خورند و تصویر حاصل یک تصویر جعلی می‌باشد یک نمونه از این کار در شکل ۲-۶ آمده است.



شکل ۲-۶) نمونه‌ای از پیوند تصویر، الف و ب تصاویر اصلی، ج تصویر پیوند خورده [۷۲]

۲-۳-۴ تشخیص جعل تصاویر مبتنی بر دوربین عکاسی

برای ثبت هر تصویر دیجیتال، تصویر از سنسور دوربین به حافظه منتقل می‌شود و طی این انتقال، مراحلی مانند چندی‌سازی، همبستگی رنگ، اصلاح گاما، تنظیم روشنایی، فیلتر شدن و فشرده‌سازی را طی می‌کند [۷۳، ۷۴]. این روش‌ها بر اساس خواص آماری که لنز، پردازنده یا نرم افزار دوربین‌های مختلف عکاسی، در تصاویر ایجاد می‌کنند، صحت تصاویر را بررسی می‌کنند. هرگونه تناقض در این الگوها یا خواص، می‌تواند نشان دهنده‌ی دستکاری تصویر باشد. روش‌هایی مانند تابع پاسخ دوربین [۷۵] و نویز الگوهای سنسور [۷۶] برای تشخیص جعل از دوربین کمک می‌گیرند. از آنجا که برخی تصاویر جعلی از ترکیب چند تصویر مختلف ایجاد می‌شوند، در این دسته از روش‌ها اگر قسمت‌های مختلف تصویر جعلی با دوربین‌های عکاسی متفاوتی گرفته شده باشند، این امر نشانه‌ای بر ترکیبی بودن تصویر و لذا جعلی بودن آن است. برای بررسی صحت تصاویر، یک روش خودکار در مرجع [۷۵] پیشنهاد شد که از اصل هماهنگی ویژگی‌های فیزیکی در میان مناطق مختلف تصویر استفاده می‌کرد. در این روش تصویر به صورت خودکار به نواحی مجزا از هم تقسیم و سپس تابع پاسخ دوربین تخمین زده می‌شد. در ادامه براساس تابع پاسخ دوربین، ویژگی‌هایی استخراج و برای کلاس‌بندی به svm داده می‌شد.

۲-۴ شناسایی دوربین دیجیتال منبع

موضوع اصلی این پایان‌نامه شناسایی دوربین منبع می‌باشد که در دسته رویکردهای غیر فعال قرار می‌گیرد.

همانطور که گفته شد امروزه فراوانی نرم افزارهای ویرایش و انتقال تصویر باعث افزایش روزافزون جعل در رسانه‌های دیجیتال شده است و شناسایی دوربین منبع می‌تواند صحت تصاویر ثبت شده را مورد تایید قرار داده و در مراجع قضایی مورد استناد دهد.

در ادامه، ما مجموعه‌ای از روش‌های شناسایی دوربین دیجیتال منبع را مورد مطالعه قرار داده و کارهای انجام شده در این زمینه را بررسی می‌کنیم.

۲-۴-۱ عدم تطابق کانونی^۱ لنز

جوی و همکاران [۲۷] برای شناسایی دوربین منبع اعوجاج شعاعی لنز را به عنوان اثر انگشت پیشنهاد می‌دهند. اعوجاج شعاعی باعث می‌شود که خطوط مستقیم به صورت خطوط منحنی بر روی تصاویر خروجی ظاهر شوند و این امر زمانی اتفاق می‌افتد که بزرگ‌نمایی عرضی M_T (نسبت فاصله تصویر به فاصله جسم) ثابت نباشد بلکه تابعی از فاصله تصویر خارج از محور r باشد. نویسندگان ادعا می‌کنند که تولیدکنندگان مختلف برای جبران اعوجاج شعاعی از سیستم مختلف لنز استفاده می‌کنند و فاصله کانونی لنز بر میزان اعوجاج شعاعی تأثیر می‌گذارد. بنابراین، هر مدل دوربین الگوی اعوجاج شعاعی منحصر به فردی را بیان می‌کند که به شناسایی آن کمک می‌کند. دو آزمایش بر روی ۳ مدل مختلف دوربین با به دست آوردن دقت طبقه‌بندی متوسط به ترتیب ۹۱/۵۳ و ۹۱/۳۹ درصد انجام شد. اگرچه این روش برای دو دوربین از یک مدل آزمایش نشده است، اما بر اساس استدلال نویسندگان در مورد اختلاف اعوجاج شعاعی، می‌توان انتظار دقت کمی داشت. از آنجایی که اعوجاج با استفاده از روش خط مستقیم اندازه‌گیری می‌شود اگر خط مستقیمی در تصویر وجود نداشته باشد، این روش ناموفق خواهد بود.

¹ Aberration

۲-۴-۲ انحراف رنگی

نقص در سیستم‌های تصویربرداری نوری منجر به انواع مختلف انحراف در تصاویر گرفته شده می‌شود. انحراف رنگی ناشی از نقص در لنز برای تمرکز کامل نور در طول موج‌های مختلف در یک دوربین دیجیتال است که باعث ایجاد اختلاف در محلی می‌شود که سنسور نور با طول موج‌های مختلف را دریافت می‌کند. انحراف رنگی دو نوع است: طولی و جانبی. انحراف طولی باعث می‌شود که طول موج‌های مختلف در فواصل مختلف از لنز متمرکز شوند در حالی که انحراف جانبی به طول موج‌های مختلفی متمرکز می‌شود که در موقعیت‌های مختلف حسگر متمرکز می‌شوند. دستکاری تصویر باعث می‌شود انحراف در تصویر ناسازگار باشد و جعل آشکار شود.

جانسون و فرید [۷۷] مدلی را برای انحراف رنگی جانبی و تکنیکی خودکار برای تخمین پارامترهای مدل، پیاده‌سازی کردند. این روش برای تشخیص دستکاری زمانی موثر است که منطقه دستکاری شده نسبتاً کوچک باشد. لان و همکاران [۷۸] با به حداکثر رساندن اطلاعات متقابل بین کانال‌های R و B اصلاح شده با کانال G ، پارامترهای انحراف رنگی جانبی را تخمین زدند. سپس پارامترهای استخراج شده به عنوان ویژگی ورودی برای طبقه‌بند SVM برای شناسایی تلفن همراه منبع از تصاویر استفاده می‌شود که منجر به میانگین دقت ۹۲,۲۲ درصد می‌شود. گلو و همکاران [۷۹] رویکرد جدیدی برای تخمین انحراف رنگی جانبی با پیچیدگی محاسباتی کم به دست آوردند و برای اولین بار، نتایج استفاده از انحراف رنگی جانبی را در سناریوهای دنیای واقعی بر اساس پایگاه داده تصویر "درسدن" ارائه دادند. پایگاه داده درسدن بیش از ۱۴۰۰۰ تصویر از صحنه‌های مختلف داخلی و خارجی تحت شرایط کنترل شده و در نتیجه قابل مقایسه از ۷۳ دوربین دیجیتال به دست آمده است. دوربین‌ها فقط از ۲۵ مدل مختلف تهیه شده‌اند.

۲-۴-۳ نقصان و خرابی^۱ حسگر

۲-۴-۳-۱ نقص پیکسل

گردتز و همکاران [۳۳] نقص پیکسل‌های CCD را بررسی کرده و از آن‌ها برای تطبیق تصاویر هدف برای کشف دوربین منبع دیجیتال استفاده کردند. نقص پیکسل شامل نقص نقطه، نقص نقطه داغ، پیکسل مرده، تله پیکسل و نقص خوشه است. برای یافتن پیکسل‌های ناقص، توسط هر ۱۲ دوربین آزمایش شده، چند تصویر با پس زمینه سیاه گرفته شده و با نقاط معیوب که سفید به نظر می‌رسند، مقایسه می‌شود. نتیجه نشان می‌دهد که هر دوربین الگوی مشخصی از نقص پیکسل دارد. با این حال، همچنین نشان داده شده است که تعداد پیکسل‌های ناقص قابل مشاهده برای دوربین مشابه بین تصاویر متفاوت است و بستگی زیادی به محتوای تصویر دارد. همچنین نشان داده شده است که تعداد پیکسل‌های ناقص قابل مشاهده بر روی تصاویر با همان محتوای گرفته شده توسط یک دوربین در دمای مختلف متفاوت است. علاوه بر این، برای دوربین‌های دارای CCD سطح بالا، نویسندگان نمی‌توانند هیچ پیکسل ناقص قابل مشاهده‌ای پیدا کنند، به این معنی که لزوماً همه دوربین‌ها نقص پیکسل ندارند. علاوه بر این، بیشتر دوربین‌ها مکانیسم داخلی برای جبران نقص پیکسل‌ها دارند. بنابراین، این روش با اطمینان نمی‌تواند برای همه دوربین‌های دیجیتال اعمال شود.

۲-۴-۳-۲ الگوی نویز سنسور

یک روش مطمئن برای شناسایی دوربین منبع بر اساس الگوی نویز حسگر توسط لوکاس و همکاران در [۱۰] ارائه شده است. عدم یکنواختی پیکسل (PNU)^۲، که در آن پیکسل‌های مختلف به دلیل نقص در فرآیندهای تولید سنسور، حساسیت‌های نوری مختلفی دارند، منبع اصلی الگوی نویز است. این امر PNU را به یک ویژگی طبیعی برای شناسایی منحصر به فرد سنسورها تبدیل می‌کند. نویسندگان ۹ مدل دوربین را مطالعه کردند که ۲ مورد از آن‌ها CCD مشابه دارند و ۲ مدل دقیقاً یکسان هستند. شناسایی دوربین حتی

¹ Imperfections

² Pixel Non-Uniformity

برای دوربین‌های همان مدل نیز ۱۰۰ درصد دقیق است. همچنین نتیجه برای شناسایی تصاویر فشرده شده قابل قبول است. یک مشکل در آزمایشات انجام شده این است که نویسندگان از مجموعه تصویر یکسانی برای محاسبه الگوی مرجع دوربین و همبستگی تصاویر استفاده کردند. ما چندین آزمایش با این مدل برای تصاویر برش خورده انجام داده‌ایم. به نظر می‌رسد که این مدل نتواند دوربین منبع تصاویر برش خورده را پیش‌بینی کند. علاوه بر این، برای کارکرد مدل، اندازه تصاویر استفاده شده برای محاسبه الگوی مرجع دوربین باید همان اندازه تصویر آزمون باشد.

۲-۴-۴ با استفاده از الحاق CFA

۲-۴-۴-۱ تأثیر درون‌یابی رنگ^۱ در نوارهای رنگ

بایرام و همکاران [۲۴] برای تعیین ساختار همبستگی موجود در هر باند رنگی که می‌توان برای طبقه‌بندی تصویر استفاده کرد، روند درون‌یابی CFA را بررسی کردند. فرض اصلی این است که الگوریتم درون‌یابی و طراحی الگوی فیلتر CFA هر سازنده (یا حتی هر مدل دوربین) تا حدودی با سایرین متفاوت است، که منجر به ایجاد ساختارهای همبستگی غیرقابل تشخیص در تصاویر گرفته شده می‌شود. با استفاده از الگوریتم تکرار حداکثر انتظار^۲ (EM)، ۲ مجموعه ویژگی ضرایب درون‌یابی از تصاویر و محل اوج و اندازه‌ها در طیف فرکانسی نقشه‌های احتمال، برای طبقه‌بندی بدست می‌آیند.

هنگام استفاده از هسته درون‌یابی 5×5 ، دقت طبقه‌بندی برای دو دوربین مختلف ۹۵٫۷۱ درصد است اما با مقایسه سه دوربین به ۸۳٫۳۳ درصد کاهش می‌یابد. برای تعیین تأثیر درون‌یابی بر دقت طبقه‌بندی باید از مجموعه بزرگتری از دوربین‌ها استفاده می‌شد. هیچ آزمایشی برای دوربین‌های همان مدل اجرا نمی‌شود، دوربین‌هایی با مدل یکسان به طور معمول از الگوی فیلتر CFA و الگوریتم درون‌یابی یکسان برخوردار هستند و به همین دلیل ما انتظار داریم که این روش منسوخ شود. علاوه بر این، نویسندگان اشاره کرده‌اند که این روش برای تصاویر فشرده به خوبی کار نمی‌کند.

¹ Interpolation

² Expectation-Maximization

۲-۴-۴-۲ مدل همبستگی پیکسل درجه دوم

لانگ و هوانگ [۳۴] یک ماتریس ضریب را از یک مدل همبستگی پیکسل درجه دوم به دست آوردند که در آن همبستگی فضایی دوره‌ای بین پیکسل از فرم درجه دوم پیروی می‌کند. برای آزمایشات از چهار دوربین همراه با تصاویر کارتونی استفاده شده است که دقت ۹۵ درصد برای دوربین اول و ۹۸ درصد برای دوربین دوم حاصل شده است. نویسندگان همچنین آزمایش را با تصاویر اصلاح شده (فشرده سازی، افزودن نویز گوسی، اصلاح گاما و صاف کردن) انجام داده‌اند. هنگامی که با کیفیت ۸۰ فشرده می‌شود، دقت به ۸۰ درصد می‌رسد. دقت در تصاویر با تغییرات دیگر حتی کمتر است. از آنجایی که دوربین‌های مشابه از الگوریتم demosaick استفاده می‌کنند، بنابراین انتظار داریم که این مدل، دوربین‌های مشابه را به درستی از یکدیگر تفکیک نکند. علاوه بر این، مدل برای تصاویر اصلاح شده عملکرد ضعیفی دارد.

۲-۴-۴-۳ معیار باینری شباهت^۱

سلیکتوتان و همکاران [۳۵] از مجموعه‌ای از اقدامات شباهت باینری برای شناسایی دوربین منبع تلفن همراه استفاده کرده‌اند. فرض اصلی این است که الگوریتم درونیابی اختصاصی CFA، همبستگی‌هایی را در بین صفحات بیت‌ی مجاور یک تصویر ایجاد می‌کند که می‌تواند توسط این معیارها نشان داده شود. معیارهای باینری شباهت، معیارهایی است که برای اندازه‌گیری شباهت بین تصاویر باینری، یعنی بین صفحات بیت یک تصویر استفاده می‌شود. ۱۰۸ معیار شباهت باینری بدست آمده است و مانند [۸۰]، مجموعه‌ای از ۱۰ معیار کیفیت تصویر به عنوان ویژگی‌های اضافی برای طبقه‌بندی استفاده می‌شود. بیشترین دقت متوسط برای طبقه‌بندی ۳ گروه دوربین ۹۸,۷ درصد است، در حالی که کمترین دقت متوسط ۸۱,۳ درصد است. هنگام طبقه‌بندی ۹ دوربین، دقت طبقه‌بند به ۶۲,۳ درصد کاهش یافته است. نتایج نشان می‌دهد که این روش به دوربین‌های هدف و تعداد دوربین‌های استفاده شده بستگی دارد.

¹ Similarity

۲-۴-۵ با استفاده از ویژگی‌های تصویر

خرازی و همکاران [۲] مجموعه ای از ویژگی‌های تصویر را که می‌توانند در شناسایی دوربین منبع استفاده شوند، شناسایی می‌کنند. ۳۴ ویژگی پیشنهادی در ۳ گروه ویژگی‌های رنگ، معیارهای کیفیت تصویر و آمار دامنه موجک^۱ دسته‌بندی می‌شوند. ویژگی‌ها از تصاویر دو دوربین استخراج می‌شوند، سپس برای آموزش و آزمایش طبقه‌بند استفاده می‌شوند. نتیجه برای تصاویر غیر فشرده به ۹۸,۷۳ درصد و برای تصاویر فشرده با ضریب JPEG ۹۳,۷۵ درصد است. با بررسی پنج دوربین، میزان دقت به ۸۸ درصد کاهش می‌یابد. تسای و همکاران [۸۱] نیز با استفاده از مجموعه دوربین‌های مختلف، مطالعه مشابهی را برای این روش انجام داده‌اند. از این رو، این روش برای دوربین‌های دارای CCD مشابه خوب کار نمی‌کند و برای شناسایی دوربین‌های منبع از همان مدل، مناسب نیست. علاوه بر این، لازم است که همه دوربین‌ها از محتوا و وضوح یکسان برخوردار باشند، که در عمل آسان نیست.

۲-۵ استفاده از شبکه عصبی عمیق

در گذشته، با استفاده از الگوریتم‌های متعدد مبتنی بر تجزیه و تحلیل آماری و تشخیص الگو، به پزشکی قانونی تصویر منبع پرداخته شده است. اخیراً، پیشرفت در توانایی‌های محاسباتی باعث ایجاد علاقه مجدد در فنون مبتنی بر یادگیری ماشین شده است. به‌طور خاص، طرح‌های مبتنی بر یادگیری عمیق با موفقیت در زمینه پزشکی قانونی تصویر منبع به کار گرفته شده‌اند و اثربخشی آن‌ها را در رقابت‌های مختلف اثبات کرده‌اند [۸۲، ۸۳].

تکنیک‌های مبتنی بر شبکه‌های عصبی مصنوعی، سال‌هاست که شناخته شده‌اند. با این وجود، در سال ۲۰۰۶ [۸۴-۸۶] استفاده از شبکه‌های عصبی عمیق موفقیت‌هایی را عملی ساخته و زمینه‌ای را ایجاد کرد که به عنوان یادگیری عمیق شناخته می‌شود [۸۷، ۸۸]. پس از آن، روش‌های مبتنی بر یادگیری عمیق به طور مداوم در زمینه‌هایی مانند تشخیص دیجیتال متون دست نویس و طبقه‌بندی تصویر، نتایج قابل توجهی کسب کرده‌اند. تعدادی از تکنیک‌های تخصصی در زمینه بینایی ماشین مانند شبکه‌های عصبی کانولوشن

¹Wavelet

CNN، شبکه‌های عصبی بازگشتی (RNN^1) و شبکه‌های مولد تخاصمی (GAN^2) ایجاد شده است. در این میان، نشان داده شده است که استفاده از CNN در کارهای مربوط به تصویر، موثر بوده و به عنوان مبنای روش‌های پزشکی قانونی تصویر دیجیتال متعددی پذیرفته شده است.

اجزای اساسی CNN ها عمدتاً از لایه‌های کانولوشن^۳، لایه‌های جمع‌کننده^۴ و توابع فعال‌ساز^۵ تشکیل شده‌اند که برای ساخت معماری CNN ها در امتداد هم قرار گرفته‌اند. به گفته خان و همکاران [۸۹]، می‌توان نوآوری‌های اخیر در ساخت و آموزش CNN را به عنوان اصلاح ساختار، بهینه‌سازی پارامترها، منظم‌سازی و عملکرد تابع ضرر دسته‌بندی کرد. در این میان، اصلاح ساختار مهم‌ترین نقش را در بهبود عملکرد ایفا می‌کند. معماری معمولی CNN مانند AlexNet [۹۰]، VGGNet [۹۱]، GoogleNet [۹۲]، ResNe [۹۳]، DenseNet [۹۴]، Xception [۹۵]، SENet [۹۶] و Siamese Network [۹۷]، کاملاً شناخته شده هستند. در مرجع [۸۹] خان و همکاران به بررسی اجمالی انواع معماری‌های شبکه‌های عصبی کانولوشنی عمیق پرداخته‌اند. گو و همکاران نیز در مرجع [۹۸] جزئیات بیشتری را در مورد یادگیری عمیق بیان کرده‌اند.

رویکردهای مبتنی بر یادگیری عمیق در پزشکی قانونی تصویر منبع نتایج قابل توجهی را بدست آورده است، این امر توجه هر دانشجو یا محقق را برای کار در این زمینه جلب می‌کند. مقالاتی که در ادامه به مقایسه روش‌ها و نتایج آن‌ها می‌پردازیم، همگی دارای چند طرح مشترک هستند، برای اینکه مقایسه ساده‌تر انجام شود همه‌ی طرح‌ها در یک چارچوب ارایه می‌شود، در ابتدا ورودی که تصویری با وضوح کامل است در وصله‌های کوچک پیکسلی با اندازه ثابت برش داده می‌شود. سپس همه‌ی وصله‌ها پردازش شده و به کمک یک استراتژی مناسب، وصله‌های مفیدتر انتخاب می‌شوند. در مرحله بعدی می‌توان روی وصله‌ها به

¹ Recurrent Neural Network

² Generative Adversarial Network

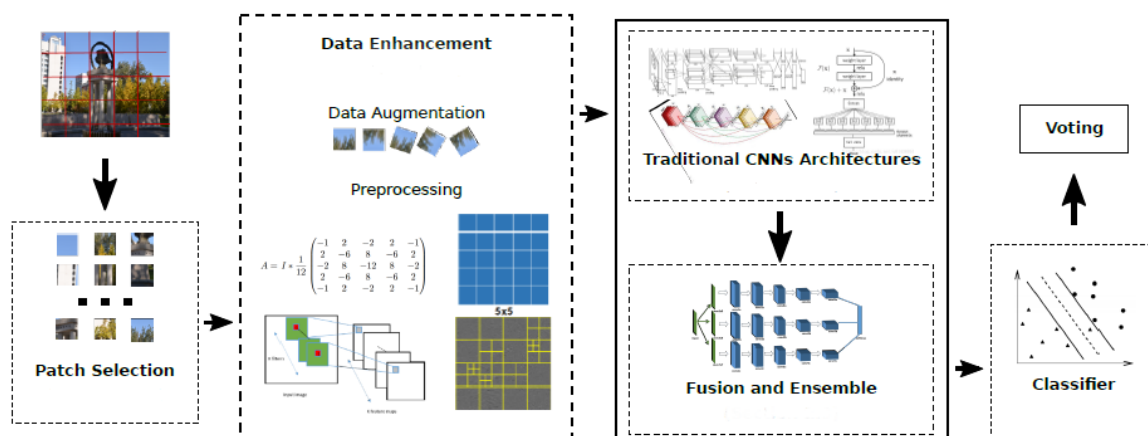
³ Convolutional Layers

⁴ Pooling Layers

⁵ Activation Functions

منظور بهبود نسبت سیگنال به نویز (SNR^1) توسط یک فیلتر فضایی پیش‌پردازش انجام داد. در ادامه وصله‌های پیش‌پردازش شده به یک شبکه عصبی کانولوشنی وارد می‌شوند. طبقه‌بندی یک وصله را می‌توان با داشتن یک لایه softmax در انتهای شبکه، یا با آموزش یک طبقه‌بندی جداگانه در مورد ویژگی‌های استخراج شده توسط آخرین لایه CNN، به دست آورد و در پایان با رأی دادن در میان تکه‌های پیکسل، نتیجه طبقه‌بندی برای تصویر اصلی را بدست می‌آید. الگوریتم‌های بررسی شده به ۵ زیربخش زیر تقسیم می‌شوند:

۱. اتخاذ شبکه‌های عصبی کانولوشن سنتی ($T.CNN^2$) برای کارهای شناسایی دوربین منبع.
۲. بهبود عملکرد با استفاده از پیشرفت داده ($D.A^3$)، از جمله افزایش داده‌ها و پیش‌پردازش داده‌ها.
۳. بهبود عملکرد از طریق ادغام و ترکیب (F/E^4)؛
۴. بهبود عملکرد با استفاده از انتخاب پیچ ($P.S^5$)؛
۵. اتخاذ طبقه‌بندی‌های مختلف (C^6).



شکل ۲-۷ چارچوب کلی الگوریتم‌های مبتنی بر یادگیری عمیق برای شناسایی دوربین منبع [۱۲۵]

¹ Signal-To-Noise Ratio
² Patch
³ Traditional CNNs Architectures
⁴ Data Augmentation
⁵ Fusion and Ensemble
⁶ Patch Selection
⁷ Classifier

۲-۵-۱ شبکه های عصبی کانولوشن سنتی

کارهای اولیه در زمینه شناسایی دوربین منبع با استفاده از شبکه عصبی کانولوشن سنتی متشکل از لایه های کانولوشن انباشته بود. ابتدا طرح های مبتنی بر یادگیری عمیق برای شناسایی دوربین منبع توسط بوندی و همکاران ارائه شد [۹۹]. این روش از معماری ساده با پنج لایه شامل سه مرحله کانولوشن و دو لایه کاملاً متصل استفاده می کند که روی وصله هایی با سایز ۴۸ در ۴۸ کار می کند. نویسندگان شبکه خود را برای شناسایی دوربین در سطح نمونه و مدل آزمایش کردند و به ترتیب ۲۹٫۸ و ۷۲٫۹ درصد دقت بدست آوردند. علاوه بر این، شناسایی دوربین در سطح مدل با استراتژی رأی گیری در وصله های مربوطه، در تصاویر با وضوح کامل به ۹۴٫۱ درصد افزایش یافت. لیو یونسیا و همکاران [۱۰۰] یک روش شناسایی دوربین منبع وصله کارآمد بر اساس یک شبکه عصبی کانولوشن عمیق پیشنهاد دادند. اول، برای دستیابی به استحکام بیشتر با کاهش هزینه آموزش، تکه های نماینده با توجه به معیارهای متعدد برای افزایش تنوع در داده های آموزشی انتخاب می شوند. دوم، یک پیمانۀ پیش بینی باقی مانده چند مقیاس کوچک برای کاهش تأثیر محتوای صحنه پیشنهاد شده است. در نهایت، یک شبکه VGG اصلاح شده برای شناسایی دوربین منبع در سطح نام تجاری، مدل و نمونه پیشنهاد می شود. فریر-اوبرگون و همکاران [۳۲] با تمرکز بر تصاویر گرفته شده توسط تلفن های همراه، معماری شش لایه CNN، شامل دو لایه کانولوشن، یک لایه حداکثر ادغام و سه لایه کاملاً متصل را پیشنهاد داده اند. تابع فعال ساز مورد استفاده در این کار واحد خطی اصلاح شده (L-ReLU) بوده است که عملکرد نسبتاً بهتری در مقایسه با ReLU داشته است. سپس، هوانگ و همکاران [۱۰۱] معماری مشابه معماری ارائه شده توسط بوندی و همکاران [۹۹] را ارائه دادند. نویسندگان مرجع [۳۲] توانستند صحت به دست آمده توسط فریر-اوبرگون و همکاران را با استفاده از نرمال سازی دسته ای و لایه های کانولوشن بیشتر بهبود بخشند. یائو و همکاران [۱۰۲]، یک شبکه عصبی کانولوشن ۱۳ لایه را ارائه می دهد. روش پیشنهادی در برابر فشردگی JPEG و افزودن نویز قوی تر است. با این حال، در برابر عملیات مقیاس سازی مجدد مقاوم نیست.

چن و همکاران [۱۰۳] استفاده از یک شبکه عصبی باقی مانده (ResNet) با ۲۶ لایه برای شناسایی

دوربین منبع را بررسی کرد و اثربخشی آن را با آزمایش‌های متعدد اثبات کردند. دقت‌های بدست آمده برای شناسایی سطح نشان تجاری، مدل و دستگاه به ترتیب ۹۹٫۱۲ درصد، ۹۴٫۷۳ درصد و ۴۵٫۸۱ درصد بوده است. با توجه به بررسی‌های انجام شده، ResNet عملکرد بهتری نسبت به AlexNet و GoogleNet در شناسایی منبع دوربین دارد [۹۹]. دینگ و همکاران [۱۰۴] روش مرجع [۱۰۳] را با ترکیب معماری ResNet با یک استراتژی یادگیری چند وظیفه گسترش دادند و عملکرد آن را بهبود بخشیدند. سه وظیفه (طبقه‌بندی در سطح نام تجاری، سطح مدل و سنسور) در یک چارچوب ادغام شده است.

مراجع [۱۰۵-۱۰۹] برای شناسایی دوربین از طریق معماری جدید مانند DenseNet و XceptionNet ارائه شده است. ماررا و همکاران [۱۰۵] از XceptionNet برای بدست آوردن دقت کلی ۹۵٫۱۵ درصد برای وصله‌های پیکسل و ۹۹٫۳۱ درصد برای تصویر با وضوح کامل با استفاده از استراتژی رأی‌گیری استفاده کرده است.

تمام آثار ذکر شده تاکنون برای حل شناسایی دوربین منبع در یک سناریوی بسته در نظر گرفته شده است. این فرض به گونه‌ای در نظر گرفته شده که از همه دستگاه‌هایی که احتمالاً می‌توانند تصاویر را ثبت کنند دانش کاملی داریم، این بدان معنی است که دستگاه‌های ثبت تصویر برای داده‌های آزمون همان دستگاه‌هایی است که برای به دست آوردن داده‌های آموزش استفاده می‌شود. با این حال، بحث سناریوی باز مطرح است، جایی که اطلاعات مربوط به تصاویر به طور کامل مشخص نیست و این امکان وجود دارد که داده‌های آزمون با دوربین‌هایی ثبت شده باشد که در ثبت داده‌های آموزش از آن دوربین‌ها استفاده نشده باشد. اخیراً، تعدادی روش یادگیری عمیق برای این سناریوی چالش برانگیز ارائه شده است. بایار و استام [۱۱۰] دو طرح مختلف را برای رفع مشکل تنظیم معماری سناریوی باز برای شناسایی مدل دوربین ارائه دادند که هدف آن قضاوت در مورد مشخص بودن یا نامشخص بودن دستگاه ضبط کننده تصویر است. اولین مورد به جای یک لایه طبقه‌بندی، از نگاشت فازی با یک استراتژی تعیین حد آستانه استفاده می‌کند تا ارزیابی کند که آیا مدل دوربین منبع شناخته شده است یا ناشناخته است. رویکرد دیگر، از طبقه‌بندی ویژگی‌های استخراج شده توسط CNN استفاده می‌کند. مایر و استام [۱۱۱] سعی کردند با استفاده از یک

شبکه عصبی سیامی که بسته به میزان شباهت ورودی‌ها، از راهکار متمایزی برای به کارگیری آن‌ها استفاده می‌کند، شباهت میان تصاویر را بسنجند. ویژگی‌ها از آخرین لایه CNN استخراج شده و با انتقال به یک شبکه سیامی، اندازه‌گیری شباهت منبع را می‌آموزند، با این روش بررسی اینکه آیا دو تصویر توسط دستگاه مشابه گرفته شده است یا خیر امکان‌پذیر می‌شود.

۲-۵-۲ پیشرفت داده

بهبود داده‌ها، از جمله افزایش داده‌ها و پردازش‌های قبلی، به طور گسترده‌ای به عنوان یک روش موثر برای بهبود عملکرد هنگام برخورد با کارهای بینایی رایانه استفاده شده است. این تکنیک‌ها برای طرح‌های شناسایی دوربین منبع نیز به کار گرفته شده‌اند. بوندی و همکاران [۱۱۲، ۱۱۳] نرمال‌سازی تصاویر را با کم کردن میانگین اندازه پیکسل‌ها انجام دادند که روش محبوبی برای متمرکز کردن داده هاست و به شبکه برای یادگیری سریع‌تر کمک می‌کند. کمال و همکاران [۱۰۷] چهار عملیات مختلف افزایش داده شامل محصولات تصادفی، چرخش‌های تصادفی، دستکاری تصاویر (از جمله فشردن‌سازی JPEG، همبستگی گاما و تغییر اندازه) و افزودن تصاویر را اعمال کردند. لازم به ذکر است که نویسندگان با جمع‌آوری تصاویر بیشتر اندازه مجموعه آموزش را افزایش دادند. نتایج نشان داد که افزودن تصاویر بیشتر تأثیر زیادی در عملکرد دارد. با توجه به اینکه روش‌های مبتنی بر یادگیری عمیق مبتنی بر داده هستند، افزایش تعداد نمونه‌های آموزشی قطعاً منجر به دقت بالاتر تشخیص می‌شود. تیم GPU_muscle [۱۰۶، ۱۰۸] بیش از ۵۰۰ گیگابایت عکس از منابع مختلف (Flickr، Yandex.Fotki، Wikipedia Commons و سایر موارد) جمع‌آوری کرد و با آموزش CNN سنتی دقت بیش از ۹۸ درصد را بدست آوردند. علاوه بر این، استفاده از تصاویر دستکاری شده مانند تصاویر پیشرفته با تصحیح گاما، تصاویر فشرده شده با JPEG و تصاویر تبدیل شده توسط عملیات نمونه‌برداری، تا حدی کارایی مدل CNN را افزایش می‌دهند.

تکنیک‌های پیش‌پردازش برای بهبود نسبت سیگنال به نویز (SNR) داده‌های ورودی، با توجه به اهمیت محتوای تصویر که تفاوت اصلی بین وظایف بینایی رایانه و شناسایی منبع می‌باشد، معرفی شدند. معمولاً وظایف بینایی رایانه به‌طور جدی به محتوای تصویر بستگی دارد، در حالی که در هنگام کار با شناسایی

دوربین منبع نیاز به بررسی محتوای تصویر نیست. در حالت دوم، کلاس صحیحی که به یک نمونه نسبت داده می شود، بستگی زیادی به مولفه نویز دوربین دارد. بر اساس این مشاهده، برخی از تحقیقات [۱۱۴-۱۱۸] پیشنهاد کردند که تداخل محتوای تصویر را با استفاده از دو نوع پیش پردازش کاهش دهند. تکنیک اول بر اساس ایده استفاده از فیلتر denoising F بر روی تصویر ورودی I و سپس کسر نتیجه آن عملیات از I و بدست آوردن نویز N است که در رابطه (۱-۲) نشان داده شده است.

$$N = I - F(I) \quad (1-2)$$

تواما و همکاران [۱۱۵] یک کاهنده نویز مبتنی بر موجک را انتخاب کرد، زیرا فیلترهای این نوع به طور گسترده‌ای در طرح‌های مبتنی بر مدل بر اساس پاسخ عکس غیر یکنواخت (PRNU)^۱ برای شناسایی دوربین منبع استفاده می‌شود. بایار و استام [۱۱۴] تأثیر یک فیلتر متوسط را با پنجره ۳در۳ (MFR)^۲ ارزیابی کردند. روش دوم پیش‌پردازش مبتنی بر این ایده است که با استفاده از فیلتر فضایی G می‌توان به راحتی نویز را از رابطه (۲-۲) بدست آورد.

$$N = I * G \quad (2-2)$$

تواما و همکاران [۱۱۵] همچنین اثربخشی فیلتر بالاگذر (HP^۳) را آزمایش کردند. با توجه به نتایج این تحقیق، استفاده از فیلترهای HP در طرح‌های مبتنی بر CNN نتایج بهتر از denoiser مبتنی بر موجک حاصل می‌کنند. دینگ و همکاران [۱۰۴] با اعمال فیلتر گاوسی با سه مقیاس (۳در۳، ۵در۵، ۷در۷) عملکرد شناسایی دوربین منبع را ارزیابی کردند.

به دنبال ایده فیلترهای مکانی، یانگ و همکاران [۱۱۸، ۱۱۹] فیلترهای خودآموز را به عنوان راهی برای بهبود بیشتر SNR ارائه داده‌اند. بایر و و استام [۱۲۱] با الهام از ویژگی‌های اسپم [۱۲۰]، یک پیچیدگی محدود را ارائه داده‌اند تا اطمینان حاصل کنند که فیلترهای بالاگذر آموخته شده در یک محدوده مشخص قرار دارند. وزن مرکزی هسته کانولوشن برابر با ۱- است و مجموع وزن های دیگر برابر با ۱ است. به جای

¹ Photo Response Non Uniformity

² Median Filter Residual

³ High Pass

طراحی فیلتر، وانگ و همکاران [۱۱۷] از الگوهای باینری محلی (LBP^1) برای کدگذاری تصویر استفاده کردند. تصاویر ابتدا توسط عملیات کدگذاری LBP در مرحله پیش‌پردازش، پردازش و سپس وارد معماری CNN می‌شوند. نتایج مراجع [۱۱۷-۱۲۱] نشان داد که فیلترهای خودآموز، پیچیدگی‌های محدود و کدگذاری LBP عملکرد بهتری نسبت به فیلترهای بالاگذر دارند.

اخیراً، ژائو [۱۱۶] نحوه تأثیرگذاری بر عملکرد مدل‌های CNN را با استفاده از دو تکنیک پیش‌پردازش فیلترهای صاف‌کننده گاوسی لاپلاس و فیلترهای کاهنده نویز ارزیابی کرده است. نتایج نشان می‌دهد که مدل CNN بدون پیش‌پردازش عملکرد بهتری را ارائه می‌دهد. نویسنده این نتیجه را از منظر داده‌ها و استراتژی آموزش بیان کرده است. البته لازم به ذکر است که مجموعه داده مورد استفاده فقط از تصاویر گرفته شده توسط سه مدل دوربین تشکیل شده است و تعداد دوره‌ها در مرحله آموزش به دلیل محدودیت در توان محاسباتی موجود، اندک بوده است.

۲-۵-۳ ادغام و ترکیب

استراتژی‌های ادغام و ترکیب با هدف تلفیق چندین مدل و ویژگی در کنار یکدیگر، عملکرد شناسایی دوربین منبع را افزایش می‌دهند. یانگ و همکاران [۱۱۸] با ادغام سه مدل با هم، یک شبکه ترکیبی سازگار با محتوا ایجاد کردند که به طور قابل توجهی دقت کلی را با توجه به مدل‌های واحد افزایش می‌دهد. بایار و استام [۱۱۴] در اولین لایه از ساختار CNN پیچش‌های محذب و MFR را با هم ترکیب کردند و نتایج نسبت به CNN محدود، اندکی بهبود یافتند. کمال و همکاران [۱۰۷] از ویژگی‌های گروه DenseNet201 استفاده کرده‌اند که با استفاده از سه مقیاس تصویر (۶۴در۶۴، ۱۲۸در۱۲۸ و ۲۵۶در۲۵۶) آموزش دیده‌اند که برای مدل CNN مفیدند. فریرا و همکاران [۱۰۹] پیشنهاد داده‌اند برای تقویت عملکرد، معماری InceptionNet و XceptionNet ادغام شود.

¹ Local Binary Patterns

۲-۵-۴ انتخاب پیچ

نویز PRNU قابل استخراج از یک تصویر با محتوای آن مرتبط است. مناطق صاف و غیر اشباع با درخشندگی بالا برای تخمین PRNU مناسب هستند. علاوه بر این، فرآیندهای سازگار با محتوا مانند کاهش دما CFA و فشرده‌سازی JPEG می‌توانند در حین دستیابی به تصویر استفاده شوند. بنابراین، انتخاب مناطق مختلف در تصویر می‌تواند در عملکرد تأثیرگذار باشد. بر اساس این واقعیت، یک استراتژی خوب برای انتخاب بهترین وصله‌های پیکسل برای آموزش CNN می‌تواند برای دستیابی به عملکرد بالاتر ضروری باشد. بوندی و همکاران [۱۱۲] فقط برای آموزش، وصله‌های پیکسلی که مقادیر متوسط آن‌ها نزدیک به نیمی از دامنه پویایی تصویر است، را انتخاب می‌کنند. معیار دیگری برای انتخاب پیچ با هدف یافتن وصله‌های پیکسلی با بافت بهتر، توسط کمال و همکاران ارائه شد [۱۰۷]. بوندی و دیگران [۱۱۳] در این روش مقدار کیفیت برای یک وصله پیکسل از واریانس و میانگین آن محاسبه می‌شود. برای آموزش مدل CNN از وصله‌های پیکسل با مقدار اندازه‌گیری بالاتر استفاده می‌شود. یانگ و همکاران [۱۱۹] رویکرد متفاوتی را در نظر گرفتند که در آن وصله‌های پیکسل با توجه به میانگین و واریانس آن‌ها به سه زیر مجموعه تقسیم می‌شوند. سپس، یک مدل متفاوت CNN در مورد هر زیر مجموعه آموزش داده می‌شود. سرانجام، وصله‌های پیکسل آزمایش با استفاده از مدل مربوط به ویژگی‌های آن‌ها طبقه‌بندی می‌شوند.

۲-۵-۵ طبقه‌بندی‌های مختلف

برخی از کارهای بینایی رایانه [۱۲۲-۱۲۴] ادعا می‌کنند که استفاده از طبقه‌بندی جداگانه SVM در ویژگی‌های آموخته شده توسط CNN به جای یک لایه softmax می‌تواند عملکرد طبقه‌بندی را بهبود بخشد. به همین ترتیب، برخی از محققان پزشکی قانونی تصویر، اخیراً کشف کرده‌اند که طبقه‌بندی‌های مختلف می‌تواند عملکرد را بهبود بخشد. به طور خاص، هوانگ و همکاران [۱۰۱]، کمال و همکاران [۱۰۷]، بوندی و دیگران [۱۱۲، ۱۱۳] و بایر و استام [۱۱۰، ۱۱۴] استراتژی‌های یادگیری دو مرحله‌ای را با تغذیه ویژگی‌های استخراج شده توسط مدل CNN با طبقه‌بندی‌های متفاوت مانند ماشین بردار پشتیبانی (SVM)،

درختان بسیار تصادفی (ERT)^۱، اندازه‌گیری تشابه کسینوس و نزدیک‌ترین نمره انجام داده‌اند. نتایج تایید می‌کند که این طبقه‌بندها می‌توانند عملکرد بهتری با توجه به لایه‌های ساده softmax داشته باشند.

۲-۶ جمع‌بندی

مهم‌ترین پارامترهای معماری و تنظیمات پایه آثار بررسی شده بخش ۲-۵ در جدول ۲-۱ خلاصه شده است، جایی که دوازده معماری اصلی را بررسی کرده‌ایم. سپس، در جدول ۲-۲ تنظیمات آزمایشی و عملکرد بدست آمده توسط آن معماری‌ها و همچنین برخی CNNهای معمولی را خلاصه کرده‌ایم. استفاده از معماری‌های مختلف یادگیری عمیق در زمینه شناسایی دوربین منبع پیشرفت چشم‌گیری را در این زمینه پدید آورده است.

¹ Extremely Randomized Trees

قسمت کانولوشن				قسمت تمام متصل			
مقاله	اندازه ورودی	تعداد لایه‌ها	تابع فعال‌ساز	ادغام (pooling)	تعداد لایه‌ها	تابع فعال‌ساز	حذف تصادفی (dropout)
[99]	48×48×3	3	ReLU	Max	1	ReLU	✓
[32]	32×32×3	2	L-ReLU	Max	2	L-ReLU	✓
[101]	36×36×3	3	ReLU	Avg	1	ReLU	✓
[102]	64×64×3	13	ReLU	Max	2	-	✓
[103]	256×256×3	1 Conv, 12 Residual	ReLU	-	-	-	-
[112]	64×64×3	4	-	Max	1	ReLU	-
[113]	64×64×3	10	-	Max	1	ReLU	-
[114]	256×256×2	4	TanH	Max, Avg	2	TanH	-
[115]	256×256	3	ReLU	Max	2	ReLU	✓
[117]	256×256×3	3	ReLU	Max	2	ReLU	✓
[118]	64×64×3	6	ReLU	Avg	-	-	-
[119]	64×64×3	1 Conv, 3 Residual	ReLU	Avg	-	-	-

جدول ۱-۲) جدول شامل معماری‌های طراحی شده در زمینه شناسایی دوربین منبع [۱۲۵].

جدول ۲-۲) جدول شامل پارامترها، مجموعه داده‌ها و نتایج تجربی معماری‌های طراحی شده [۱۲۵].

مقاله	اندازه ورودی	پیشرفت داده	ترکیب و ادغام	انتخاب پیچ	طبقه‌بند	مجموعه داده	عملکرد بر اساس پیچ		عملکرد بر اساس رای‌گیری	
							مدل	سنسور	مدل	سنسور
[99]	48×48×3	-	-	-	Softmax	Dresden [132]	72.9% (27)	29.8% (74)	94.1% (27)	-
[32]	32×32×3	-	-	-	Softmax	MICHE-I [133]	98.1% (3)	91.1% (5)	-	-
[101]	36×36×3	-	-	-	SVM	Dresden [132]	-	-	-	99.9% (10)
[102]	64×64×3	-	-	✓	Softmax	Dresden [132]	93% (25)	-	>98% (25)	-
[103]	256×256×3	-	-	-	Softmax	Dresden [132]	94.7% (27)	45.8% (74)	-	-
[105]	64×64×3	32×32×3 DenseNet-40 224×224×3 DenseNet-121 299×299×3 XceptionNet	-	-	Softmax	VISION [135]	-	80.77% (35)	-	97.47% (35)
							-	87.96% (35)	-	95.06% (35)
							-	93.88% (35)	-	99.10% (35)
							-	95.15% (35)	-	99.31% (35)
[107]	256×256×1 Inception-Xception	✓	✓	✓	Softmax	SPC2018 [134]	98.37% (10, weighted)	-	-	-
[112]	64×64×3	-	-	✓	SVM	Dresden [132]	93% (18)	-	>95 % (18)	-
[113]	64×64×3	-	-	✓	Softmax	Dresden [132]	94.93% (18)	-	-	-
[114]	256×256×2	✓	✓	-	ET	Dresden [132]	98.58% (26)	-	-	-
[115]	256×256	-	-	-	Softmax	Dresden [132]	98.99% (12) 98.01% (14)	-	-	-
[117]	256×256×3	✓	-	-	Softmax	Dresden [132]	98.78% (12) 97.41% (14)	-	-	-
[119]	64×64×3	✓	✓	✓	Softmax	Dresden [132]	-	97.03% (9)	-	-
[108]	480×480×3 Inception-Xception	✓	-	-	Softmax	SPC2018 [134]	98% (10, weighted)	-	-	-
[118]	64×64×3	✓	✓	-	Softmax	Dresden [132]	-	94.14% (9)	-	-
[109]	299×299 Inception-Xception	-	✓	✓	Softmax	SPC2018 [134]	93.29% (10, weighted)	-	-	-
[104]	48×48×3	✓	-	-	Softmax	Dresden [132]	-	-	79.71% (27)	53.4% (74)

فصل ۳ : ارائه روش پیشنهادی

۳-۱ مقدمه

در این فصل، به توضیح روش پیشنهادی برای پیش‌پردازش داده‌ها پرداخته شده است. همچنین توضیحات شبکه عصبی مبتنی بر یادگیری عمیق جهت شناسایی دوربین منبع آورده شده است. سیستم SCI را می‌توان به چهار مرحله اصلی شامل جمع‌آوری داده برای تهیه دادگان، پیش‌پردازش داده‌ها، استخراج ویژگی و شناسایی دوربین منبع تقسیم کرد. در برخی از مراجع مرحله پیش‌پردازش داده‌ها به‌طور خاص حذف می‌گردد که علت آن، از دست رفتن اطلاعات مفید در مرحله پیش‌پردازش داده‌ها است. در این فصل با استفاده از برخی تکنیک‌های پردازش تصویر، داده‌های موجود در مجموعه داده مورد پیش‌پردازش قرار می‌گیرند. روش پیش‌پردازش پیشنهادی، شامل استفاده از توابعی است که باعث بهبود کیفیت تصاویر می‌شود. در ادامه تمرکز این پایان‌نامه بر روی پیدا کردن معماری یادگیری عمیق مناسب برای استخراج ویژگی و شناسایی دوربین منبع می‌باشد. با کمک این معماری، می‌توان داده‌های تصویری را مورد بررسی قرار داد و دوربین منبع را با دقت بالا شناسایی کرد.

۳-۲ آماده‌سازی داده‌ها

در این بخش به توضیح پردازش‌هایی می‌پردازیم که روی داده‌ها اعمال می‌شود تا داده‌ها را برای تحلیل‌های بعدی آماده کند.

داده‌های تصویری موجود در مجموعه داده بصورت رنگی ثبت شده اند. برای افزایش سرعت پردازش و کاهش زمان اجرا، تصاویر بصورت سیاه و سفید بارگذاری می‌شود.

همه‌ی داده‌ها باید اندازه یکسان داشته باشند. پس از انجام پیش‌پردازش‌های توضیح داده شده در بخش ۳-۱، یکسان‌سازی اندازه اعمال می‌شود و همه تصاویر با اندازه ۳۲۰ در ۳۲۰ بارگذاری می‌شوند. همچنین داده‌ها، به یک ماتریس ۳۲۰ در ۳۲۰ یک لایه تغییر شکل داده می‌شوند تا بعنوان داده ورودی برای آموزش یک شبکه عصبی عمیق مناسب باشند.

۳-۲-۱ پیش پردازش داده‌ها

اغلب مواقع، اعمال توابع پیش‌پردازشی روی داده‌های تصویری برای تجزیه و تحلیل بهتر داده‌ها مناسب نیستند. حتی در برخی مواقع اعمال پیش‌پردازش با از دست دادن اطلاعات مفید، باعث کاهش کارایی سیستم می‌شود. در این بخش ابتدا به توضیح آماده‌سازی داده‌ها می‌پردازیم و سپس توابعی را برای پیش‌پردازش داده‌ها معرفی می‌کنیم که نتایج شناسایی دوربین منبع را بهبود ببخشد. برای شناسایی دوربین منبع به کمک الگوی نویز سنسور ترجیحا نباید از پیش‌پردازش‌هایی استفاده کرد که موجب از بین رفتن یا کاهش شدید نویز تصویر می‌شوند. با این تفسیر، توابعی که کیفیت و تضاد تصویر را بهبود ببخشند یا لبه‌های تصویر را شناسایی کنند در شناسایی دوربین منبع موثر خواهند بود.

۳-۲-۱ استفاده از عملگر Prewitt

از عملگر Prewitt برای تشخیص لبه در پردازش تصویر استفاده می‌شود. این عملگر شیب شدت تصویر را در هر نقطه از طریق محاسبه شیب‌ها در جهت x و y شناسایی می‌کند. عملگر Sobel نیز تشخیص لبه را انجام می‌دهد با این تفاوت که فیلتر تشخیص لبه Sobel به مقادیر نزدیک‌تر سلول در هسته^۱ وزن بیشتری می‌بخشد، فیلتر Prewitt شیب‌های در جهت x و y را با ترکیب تصویر ورودی (A) با هسته‌های ۳ در ۳ رابطه ۱-۱ تخمین می‌زند:

$$G_x = \begin{pmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{pmatrix} * A \quad G_y = \begin{pmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix} * A \quad (1-3)$$

سپس طبق رابطه ۲-۳ به هر سلول شبکه در تصویر خروجی، ریشه مربع حاصل از مربع شیب‌های x و y اختصاص داده می‌شود.

$$G = \sqrt{G_x^2 + G_y^2} \quad (2-3)$$

¹ Kernel

کتابخانه SciPy یکی از کتابخانه‌های کلیدی پایتون است. همانند Numpy از این کتابخانه مهم نیز می‌توان برای پیاده‌سازی برخی از کاربردهای پردازش تصویر با پایتون استفاده کرد. کتابخانه SciPy حاوی توابع و الگوریتم‌هایی است که به راحتی می‌توان از آن‌ها برای انجام عملیات (ساده) دستکاری و پردازش تصویر با پایتون استفاده کرد. از این کتابخانه برای استفاده از فیلتر Prewitt در کد برنامه مربوط به پایان‌نامه استفاده شده است.

ابزار SciKit-Image نیز بعنوان بسته برنامه‌نویسی منبع باز با ساختار آرایه‌ای تعریف شده توسط کتابخانه NumPy کار می‌کند. به کمک این ابزار که یکی از مهم‌ترین کتابخانه‌های پردازش تصویر در پایتون است، فیلتر Sobel را روی داده‌ها اعمال می‌کنیم.

در این پایان‌نامه فیلترهای تشخیص لبه Sobel و Prewitt بصورت جداگانه روی داده‌های ورودی اعمال شد، با توجه به نتایج ارائه شده در فصل ۴، فیلتر Prewitt بعنوان فیلتر مناسب تشخیص لبه در قسمت پیش‌پردازش داده‌ها مورد استفاده قرار گرفت.

۲-۱-۲-۳ بهبود کیفیت تصاویر

یکی دیگر از کتابخانه‌های مفید در پردازش تصویر کتابخانه Pillow نسخه گسترش‌یافته کتابخانه

PIL

است. این کتابخانه شامل مجموعه‌ی قابل توجهی از قابلیت‌های پردازش تصویر است که در این پایان‌نامه از دو تابع Sharpness^۱ و Contrast^۲ برای بهبود کیفیت تصاویر استفاده شده است.

با اعمال فیلتر Sharpness روی داده‌های ورودی میزان وضوح تصاویر افزایش یافته و این کار بر شناسایی الگوی نویز سنسور موجود در تصاویر موثر می‌باشد.

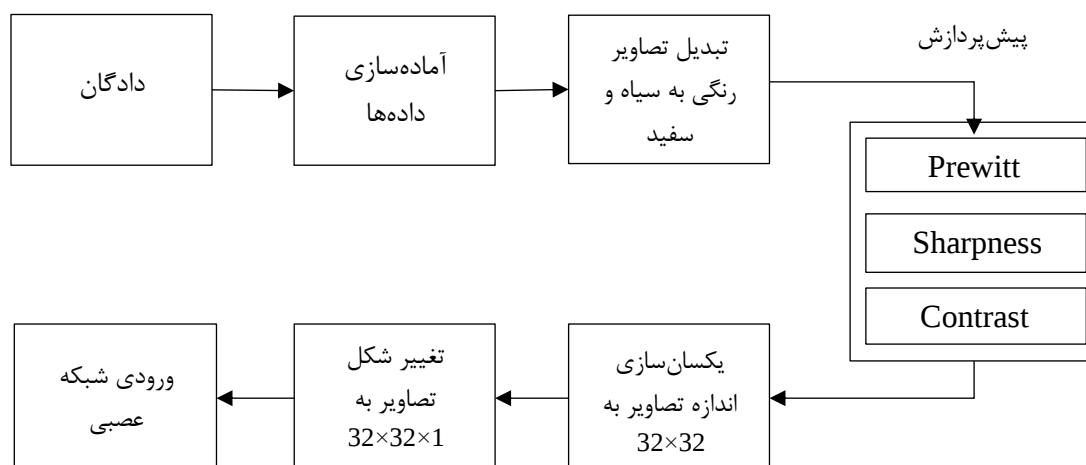
یکی دیگر از فیلترهای مناسب برای پیش‌پردازش داده‌ها، فیلتر Contrast است. این فیلتر تفاوت بین روشن‌ترین مناطق سفید و تیره‌ترین مناطق سیاه را افزایش داده و باعث افزایش کارایی سیستم شناسایی دوربین منبع می‌شود.

^۱ میزان وضوح تصاویر

^۲ تضاد

فیلترهای تنظیم روشنایی و رنگ تصویر نیز در این کتابخانه وجود دارد که در فرآیند پیش‌پردازش بدلیل عدم کارایی مناسب در زمینه شناسایی دوربین منبع از آن‌ها استفاده نمی‌شود.

در شکل ۳-۱ فرآیند کلی آماده‌سازی و پیش‌پردازش پیشنهادی داده‌ها نشان داده شده است.



شکل ۳-۱) فرآیند آماده‌سازی و پیش‌پردازش داده‌ها

۳-۳ مدل CNN پیشنهادی

در این بخش، مدل CNN پیشنهادی مبتنی بر یادگیری عمیق بر اساس شبکه‌های عصبی کانولوشنی را ارائه می‌دهیم. از این مدل برای شناسایی دوربین منبع استفاده می‌کنیم و در ادامه به شرح کامل جزئیات این مدل می‌پردازیم. CNN به عنوان دنباله‌ای از لایه‌های کانولوشن، جمع‌کننده و تمام متصل^۱ تعریف می‌شود که در آن هر لایه از طریق توابع مختلف، یک حجم ورودی را به حجم جدیدی از خروجی تبدیل می‌کند. معمولاً فرآیند محاسبات شبکه عصبی با یک عملیات غیر خطی دنبال می‌شود. در پایان فرآیند، نتیجه یک تابع ضرر است که نقش اساسی در هر به روزرسانی وزن نورون‌های شبکه عصبی دارد. معمولاً انتخاب پیکربندی معماری به صورت آزمایشی انجام می‌شود. ساختار این پایان‌نامه از تصاویر آماده‌سازی شده و پیش‌پردازش شده به عنوان ورودی استفاده می‌کند. داده‌ها برای استخراج ویژگی به ساختار پیشنهادی با دو لایه کانولوشنی، یک لایه جمع‌کننده حداکثری و دو لایه تمام متصل

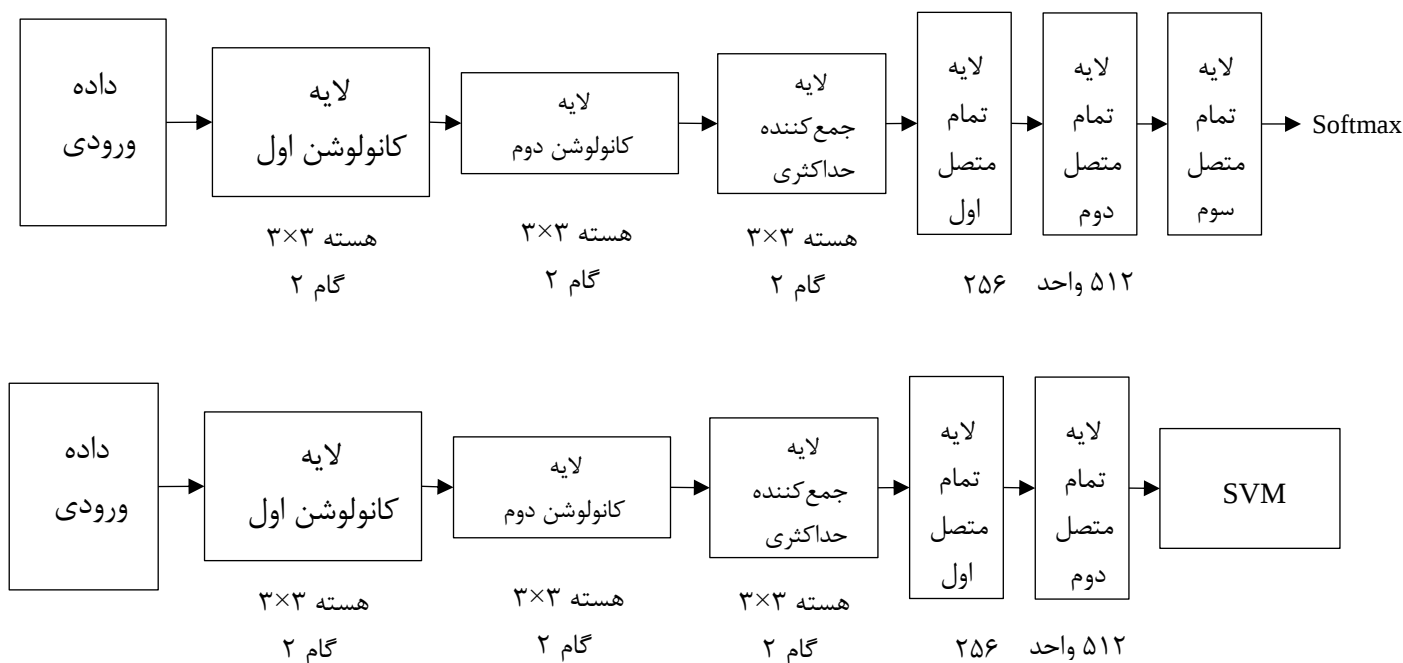
¹ Fully Connected

وارد می‌شوند. در نهایت برای شناسایی دوربین منبع از دو روش استفاده می‌شود : ۱- استفاده از یک لایه تمام متصل با تابع فعال‌ساز Softmax^۱ و ۲- استفاده از ماشین بردار پشتیبان (SVM).

تابع بیشینه هموار (Softmax)، به عنوان تابع نمایی نرمال شده نیز شناخته می‌شود. معمولاً از تابع بیشینه هموار برای نرمال‌سازی خروجی شبکه عصبی و تبدیل آن به توزیع نرمال در آخرین لایه شبکه عصبی به عنوان تابع فعال‌سازی استفاده می‌شود. در شبکه های عصبی نرمال‌سازی نسبت به کلاس خروجی پیش‌بینی شده صورت می‌گیرد.

ماشین بردار پشتیبان یک الگوریتم دسته‌بندی همراه با یادگیری ماشین می‌تواند مدلی بسیار قابل توجه برای دسته‌بندی داده‌ها ارائه کند. الگوریتم ماشین بردار پشتیبان هنگامی که قدرت پیش‌بینی بالا مورد نیاز باشد یک گزینه بسیار عالی است. استفاده از ماشین بردار پشتیبان همراه با یادگیری عمیق نتایج خوبی را در این زمینه به وجود آورد.

ساختار مدل‌های پیشنهادی این پایان‌نامه در شکل ۲-۳ آورده شده است.



شکل ۲-۳ (مدل‌های CNN پیشنهادی

^۱ تابع بیشینه هموار

هر دو ساختار پیشنهادی دارای لایه‌های ورودی و میانی یکسان می‌باشند که در ادامه توضیح داده می‌شوند. لایه انتهایی وظیفه شناسایی دوربین منبع را بر عهده دارد که در ساختار اول از تابع Softmax و در ساختار دوم از ماشین بردار پشتیبان استفاده شده است.

۳-۳-۱ لایه کانولوشن

پارامترهای هر لایه کانولوشن را مجموعه‌ای از هسته‌های قابل آموزش تشکیل می‌دهند (که به آن‌ها فیلتر نیز می‌گویند). همانطور که در شکل ۳-۲ مشاهده می‌شود، اندازه هر یک از این هسته‌ها برای لایه‌های کانولوشن 3×3 می‌باشد. بر اساس مرجع [۱۲۶]، با توجه به تصویر دو بعدی I به عنوان ورودی و یک هسته دو بعدی K ، می‌توان محاسبه لایه کانولوشن را در معادله ۳-۱ توصیف کرد.

$$S_t(i, j) = (K * I)(i, j) = \sum_{m=0}^2 \sum_{n=0}^2 I(i - m, j - n) K(m, n) \quad (3-1)$$

نقشه ویژگی $(S_t(i, j))$ تولید شده، حاصل خروجی اعمال هسته (K) به لایه قبلی یا تصویر ورودی است. لایه کانولوشن اول دارای ۳۲ نقشه ویژگی و لایه کانولوشن دوم دارای ۶۴ نقشه ویژگی می‌باشد. گام حرکت هسته را برای ساختار پیشنهادی در لایه کانولوشن برابر با ۲ قرار می‌دهیم.

تابع فعال‌ساز واحد خطی اصلاح شده نشتی (Leaky ReLU^1) برای این ساختار انتخاب شده است [۱۲۷]. دلیل این انتخاب حساسیت بسیار بالای سیستم شناسایی دوربین منبع به اطلاعات نویز سطح پایین است. در معادله ۳-۲ نشان داده شده است که Leaky ReLU بر خلاف تابع فعال‌ساز واحد خطی اصلاح شده (ReLU^2) [۱۲۸] یک شیب خطی غیر صفر (α) را به نوروں غیر فعال اختصاص می‌دهد تا آن نوروں بصورت کامل از فرآیند آموزش شبکه حذف نگردد.

¹ Leaky Rectified Linear Unit

² Rectified Linear Unit

$$f(x) = \mathbb{I}(x < 0)(\alpha x) + \mathbb{I}(x \geq 0)(x) \quad (2-3)$$

با استفاده از تابع فعال‌ساز Leaky ReLU تمامی نورون‌ها تا آخرین مرحله آموزش شبکه عصبی در فرآیند حضور خواهند داشت که باعث افزایش کارایی عملکرد شناسایی دوربین منبع خواهد شد.

۳-۳-۲ لایه جمع‌کننده

معمولاً لایه‌های جمع‌کننده در بین سری لایه‌های کانولوشن قرار می‌گیرند. وظیفه این نوع لایه‌ها کاهش تعداد پارامترها و کاهش پیچیدگی محاسبات مدل است. این عمل با کاهش تدریجی اندازه فضایی داده ورودی انجام می‌شود. این کار مشابه مرحله گام برداشتن در لایه کانولوشن انجام می‌شود. در این پایان‌نامه از لایه جمع‌کننده حداکثری استفاده شده است. در این لایه، با قرار گرفتن هسته بر روی ورودی حداکثر مقدار حجم در هر منطقه زیر هسته استخراج می‌شود. همانطور که در شکل ۳-۲ مشاهده می‌شود، لایه جمع‌کننده بین هر جفت لایه کانولوشن و لایه تمام متصل قرار دارد. در ساختار پیشنهادی، اندازه هسته 3×3 با گام حرکت ۲ در نظر گرفته شده است.

۳-۳-۳ لایه تمام متصل

سرانجام سه لایه تمام متصل در پایان فرآیند در نظر گرفته شده است. شبکه عصبی عمیق حجم ورودی را می‌گیرد و یک بردار بعدی N را به عنوان خروجی می‌دهد. N تعداد کلاس‌هایی است که معماری برای آن آموزش داده شده است. لایه‌های تمام متصل اول و دوم به ترتیب ۲۵۶ و ۵۱۲ نورون دارند. در این حالت، تابع فعال‌ساز Leaky ReLU نیز بر روی خروجی لایه تمام متصل اعمال می‌شود. برای جلوگیری از بیش‌برازش^۱، استفاده از Dropout [۱۲۹] پیشنهاد شده است. در هر مرحله آموزش، نورون‌های بین لایه‌ها با احتمال $1-p$ از CNN خارج می‌شوند یا با احتمال p نگه داشته می‌شوند. این روش هم‌چنین به طور قابل توجهی سرعت آموزش را بهبود می‌بخشد و به نظر می‌رسد تعاملات نورون‌ها را کاهش می‌دهد و آن‌ها را به سمت یادگیری ویژگی‌های قوی‌تر هدایت می‌کند که تعمیم بهتر به

¹ Overfitting

داده‌های جدید است. در این پایان‌نامه از DropConnect [۱۳۰] که تعمیم یافته Dropout می‌باشد، استفاده شده است. هنگام آموزش با استفاده از Dropout، یک زیر مجموعه خروجی از نورون‌های فعال به طور تصادفی در هر لایه روی صفر تنظیم می‌شود. در عوض، DropConnect یک زیرمجموعه از ورودی نورون‌ها را که به طور تصادفی انتخاب شده، روی صفر تنظیم می‌کند. بنابراین هر واحد ورودی، زیر مجموعه تصادفی واحدهای موجود در لایه قبلی را دریافت می‌کند که کار باعث می‌شود همه نورون‌ها در فرآیند آموزش حضور داشته باشند. بعد از لایه‌های تمام متصل اول و دوم DropConnect اعمال می‌شود تا از بیش‌برازش جلوگیری شود.

۳-۳-۴ شناسایی دوربین منبع

۳-۳-۴-۱ استفاده از Softmax

در آخرین لایه شبکه عصبی از تابع Softmax برای نرمال‌سازی خروجی شبکه و تبدیل آن به توزیع احتمال استفاده می‌کنیم. نرمال‌سازی در این حالت نسبت به کلاس‌های خروجی پیش‌بینی شده، صورت می‌پذیرد. در این ساختار پیشنهادی، خروجی آخرین لایه تمام متصل سوم عملکرد Softmax را تغذیه می‌کند.

۳-۳-۴-۲ استفاده از SVM

برای وظایف طبقه‌بندی، بسیاری از مدل‌های یادگیری عمیق از تابع فعال‌ساز softmax برای یادگیری برچسب‌های خروجی استفاده می‌کنند. در این پایان‌نامه، یک ماشین بردار پشتیبان خطی را جایگزین لایه softmax می‌کنیم. ویژگی‌ها استخراج شده به عنوان ورودی به شبکه SVM وارد می‌شوند [۱۳۱].

۳-۳-۵ انتخاب پارامترها

در کنار طراحی مناسب ساختار شبکه عصبی، انتخاب پارامترهای بهینه نیز اهمیت خاصی در

شناسایی دوربین منبع خواهند داشت. بعد از بررسی عملکرد توابع بهینه‌ساز مختلف در این پایان‌نامه از تابع بهینه‌ساز Adam با نرخ یادگیری 0.001 استفاده می‌کنیم. در جدول ۱-۳ ساختارهای پیشنهاد شده در این پایان‌نامه را می‌توان مشاهده کرد. قبل از اولین لایه Dense یا تمام متصل از روش flatten استفاده می‌شود تا نورون‌ها به صورت یک بردار نمایان شوند.

جدول ۱-۳) ساختار پیشنهاد شده پایان‌نامه (N تعداد کلاس‌های موجود)

لایه	نام لایه	اندازه هسته	تابع فعال‌ساز	گام	تعداد خروجی
۱	Conv2D	3×3	LeakyReLU	۲	۳۲
۲	Conv2D	3×3	LeakyReLU	۲	۶۴
۳	MaxPooling2D	3×3	-	۲	۶۴
۴	Flatten	-	-	-	۵۷۶
۵	Dense	-	LeakyReLU	-	۲۵۶
۶	Dense	-	LeakyReLU	-	۵۱۲
۷	Dense/SVM	-	Softmax/-	-	N

۳-۴ جمع‌بندی

در این بخش روش پیشنهادی برای شناسایی دوربین منبع به همراه پیش‌پردازش‌های اعمال شده روی داده‌ها شرح داده شد. از افزایش وضوح تصویر، افزایش تضاد و تشخیص لبه در بخش پیش‌پردازش داده‌ها استفاده شد. در این بخش به شرح روش یادگیری عمیق استفاده شده در ساختارهای پیشنهادی پرداخته‌ایم.

پارامترهای تاثیرگذار شبکه عصبی عمیق در فرآیند شناسایی دوربین منبع با در نظر گرفتن کارایی بهتر شبکه تعیین شدند. در جدول ۱-۳ ساختار پیشنهادی این پایان‌نامه نشان داده شده است.

فصل ۴ : نتایج و آزمایش‌ها

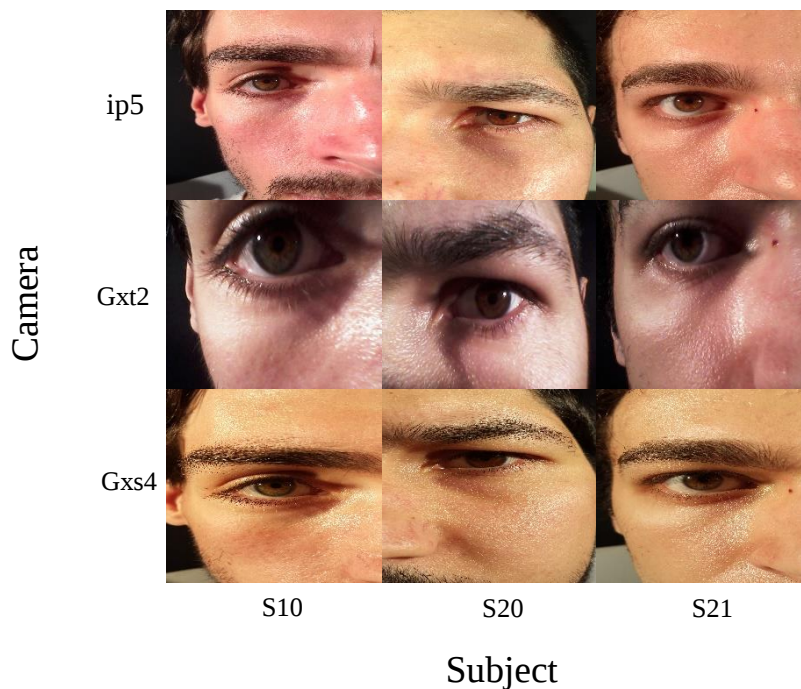
۴-۱ مقدمه

در این فصل، به توضیح مجموعه دادگان مورد استفاده (MICHE-I) پرداخته شده است. همچنین به بررسی عملکرد روش پیشنهادی پایان نامه می پردازیم. پیش پردازش های توضیح داده شده در بخش ۳-۳-۱ را روی مجموعه داده MICHE-I اعمال می کنیم و نتایج حاصل از این پیش پردازش ها با کمک نمایش گرافیکی نشان می دهیم. همچنین کارایی روش پیشنهادی با برخی از روش های دیگر در زمینه شناسایی دوربین منبع مقایسه شده و نتایج در ادامه آورده شده است. عملکرد روش پیشنهادی با انجام پیش پردازش ها و عدم انجام آن ها بررسی خواهد شد.

۴-۲ معرفی مجموعه دادگان MICHE-I

این پایگاه داده شامل بیش از ۳۷۳۲ تصویر می باشد که با سه دستگاه موبایل iPhone5، Samsung galaxyIV و Samsung galaxy tabletII در محیط های داخلی ساختمان و فضای آزاد با شرایط نوری متفاوت برای مقایسه بهتر ثبت شده است. تصاویر در دو سطح مدل دستگاه و دوربین جلو و عقب جمع آوری شده است. این امر شناسایی منبع دوربین را به دو دسته شناسایی مدل دستگاه موبایل و شناسایی نوع دوربین جلو یا عقب تقسیم می کند.

شکل ۴-۱ در این بخش به نمایش نمونه هایی از این مجموعه دادگان قبل از مرحله پیش پردازش می پردازیم.



شکل ۴۴-۱) تصویر سه مورد ثبت شده با دوربین‌های پشتی سه دستگاه موبایل در فضای داخلی

۳-۴ خلاصه آزمایش‌ها

در این بخش، آزمایش‌ها را مورد بررسی قرار می‌دهیم. برای ارزیابی تاثیر پیش‌پردازش‌های اعمال شده روی مجموعه داده، عملکرد ساختارهای پیشنهادی با داده‌ها در مراحل مختلف پیش‌پردازش مورد بررسی قرار می‌گیرد و کارایی نهایی ساختارهای پیشنهادی با پیش‌پردازش‌های مختلف گزارش می‌شود. برای نمایش عملکرد ساختارها در سطح مدل گوشی موبایل استفاده شده برای ثبت تصویر و نوع دوربین استفاده شده در گوشی موبایل از ماتریس درهم ریختگی استفاده کرده و به تحلیل آن می‌پردازیم. تاثیر پارامترها روی عملکرد ساختار نیز مورد بررسی قرار می‌گیرد و نتایج بصورت نموداری برای مقادیر مختلف پارامتر مورد بررسی نمایش داده می‌شود. در آخر نیز به مقایسه عملکرد روش پیشنهادی با روش‌های دیگر مراجع می‌پردازیم.

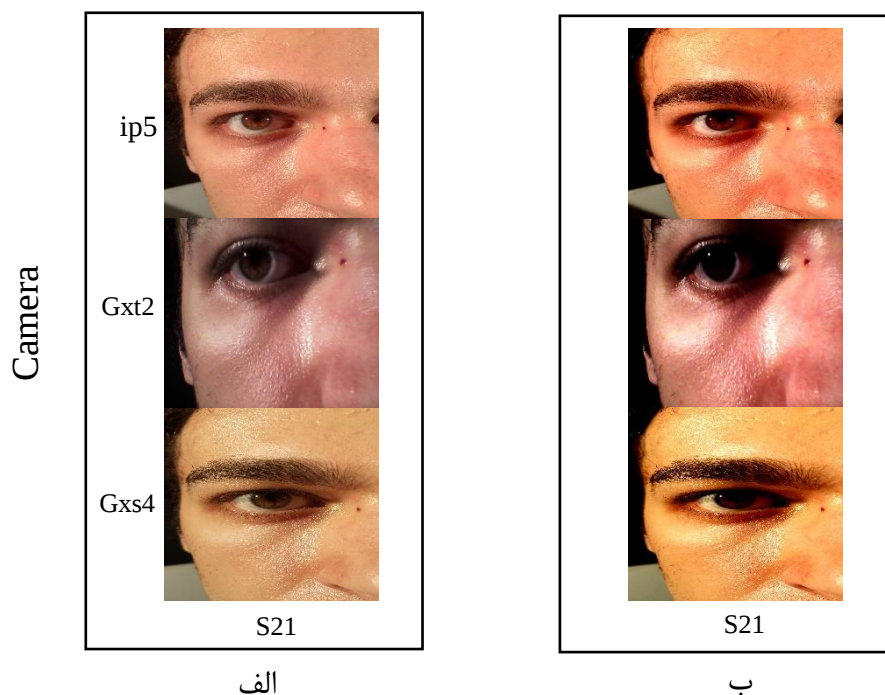
۴-۴ نتایج پیش‌پردازش داده‌ها

همانطور که در بخش ۳-۳-۱ روش پیش‌پردازش مورد استفاده در این پایان‌نامه توضیح داده شد، در

این بخش، به شرح نتایج پیش‌پردازش روی تصاویر مجموعه داده MICHE-I می‌پردازیم.

ابتدا، پیش‌پردازش تصاویر موجود در مجموعه داده را انجام می‌دهیم، سپس به یکسان‌سازی اندازه تصاویر و تغییر شکل آن‌ها می‌پردازیم.

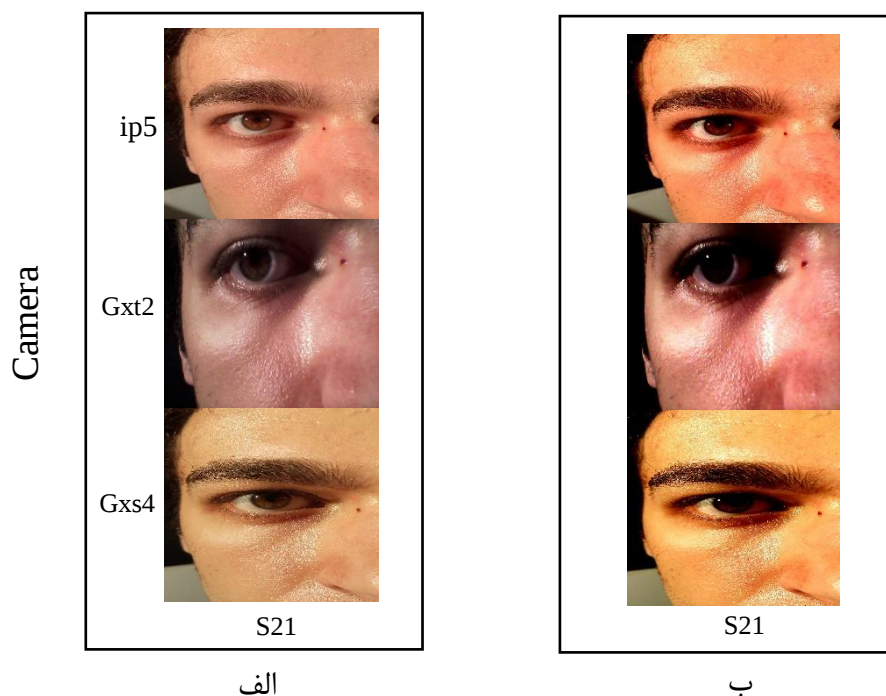
به دلیل حساس بودن سیستم شناسایی دوربین منبع به نویزهای سطح پایین تصویر، پیش‌پردازش‌هایی را انتخاب کردیم که بتواند در افزایش کارایی سیستم موثر باشد. اولین مورد، بهبود کیفیت تصاویر است، از این بخش ابتدا افزایش تضاد تصویر را روی داده‌ها اعمال می‌کنیم و نتایج حاصل از این پیش‌پردازش برای سه مورد از تصاویر موجود در شکل ۴-۱ را در شکل ۴-۲ به نمایش می‌گذاریم.



شکل (۴-۴) تصاویر ب حاصل اعمال پیش‌پردازش افزایش کنتراست تصاویر الف می‌باشد.

ضریب افزایش کنتراست تصویر را روی ۱,۵ تنظیم می‌کنیم، همانطور که در شکل ۴-۲ مشاهده می‌شود پس از اعمال این عملگر تصاویر با کیفیت بالاتری خواهیم داشت و این افزایش کیفیت موجب استخراج ویژگی‌های مفیدتر و در نتیجه افزایش کارایی سیستم شناسایی دوربین منبع خواهد شد.

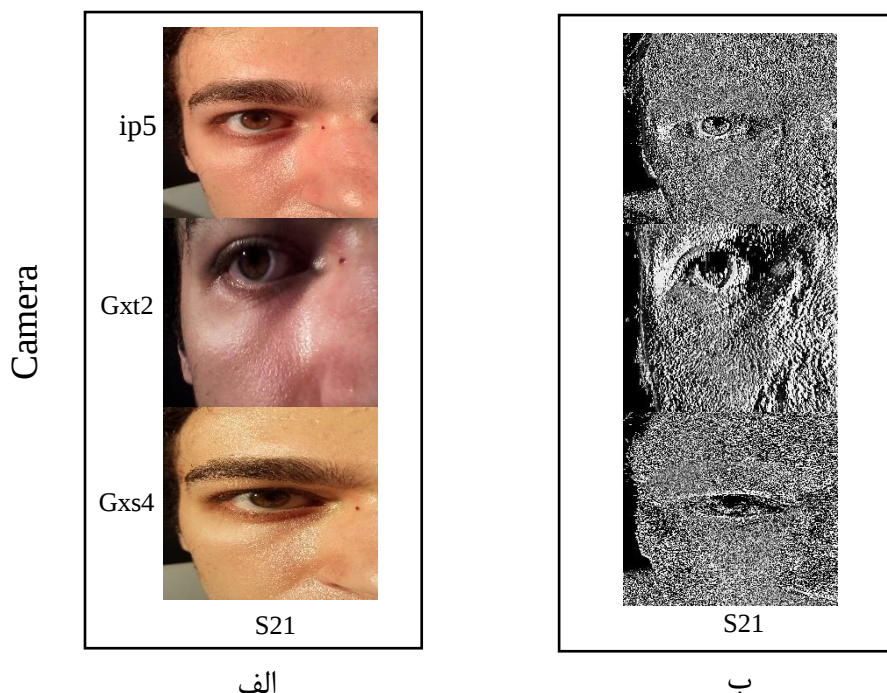
تصاویر قسمت ب شکل ۲-۴ را با عملگر افزایش وضوح تصویر پیش پردازش کرده و نتایج حاصل را در شکل ۳-۴ نمایش می دهیم.



شکل ۳-۴ تصاویر ب حاصل اعمال پیش پردازش افزایش وضوح تصویر بر روی تصاویر قسمت ب شکل ۲-۴ می باشد. تصاویر الف نیز تصاویر موجود در مجموعه دادگان بدون پیش پردازش می باشد.

هم اکنون، با اعمال فیلتر تشخیص لبه Prewitt روی تصاویر حاصل از پیش پردازش بهبود کیفیت، تصاویر را مورد پردازش قرار داده سپس به عنوان داده ورودی برای شروع فرآیند آموزش تغییر اندازه و تغییر شکل می دهیم.

در شکل ۴-۴ نتایج حاصل از اعمال فیلتر تشخیص لبه به عنوان آخرین مورد از فرآیند پیش پردازش آورده شده است. در این قسمت تصاویر از کانال رنگی RGB به کانال سیاه و سفید (GrayScale) تغییر می یابد.



شکل ۴-۴) قسمت ب حاصل اعمال تمامی فیلترهای فرآیند پیش پردازش این پایان نامه می باشد.

در این پایان نامه ۷۰ درصد مجموعه دادگان بصورت تصادفی در بسته های ۱۲۸ تایی به عنوان داده آموزشی، ۲۰ درصد مجموعه دادگان به عنوان داده آزمون و ۱۰ درصد نیز به عنوان داده ارزیابی در نظر گرفته شده است. به دلیل حجم کوچک مجموعه دادگان نیاز به سخت افزار خیلی قوی نمی باشد، به همین دلیل همه ساختارها در بستر برنامه نویسی پایتون با پس زمینه تنسورفلو روی لپ تاپ MSI cx62 با گرافیک NVIDIA 940MX که دارای ۳۸۴ هسته cuda می باشد پیاده سازی شده است. استفاده از کارت گرافیک به عنوان واحد پردازش، باعث افزایش سرعت اجرا و کاهش زمان اجرا خواهد شد.

۴-۵ نتایج طبقه بندی ساختارهای پیشنهادی

در این بخش نتایج ساختارهای پیشنهادی را بررسی می کنیم. در ساختار پیشنهادی این پایان نامه از یادگیری عمیق مانند بسیاری دیگر از مقالات در زمینه شناسایی دوربین استفاده شده است. به طور کلی استفاده از یادگیری عمیق در بحث یادگیری ماشین به ویژه پردازش تصویر باعث افزایش کارایی مدل ها شده است.

در شبکه‌های عصبی عمیق رعایت تعادل بین عمق، دقت و سرعت شبکه مهم است. هر چه لایه‌های بیشتری استفاده شوند، دقت کلی بالاتر می‌رود، اما باعث می‌شود که شبکه کندتر اجرا شود، زیرا تعداد محاسبات افزایش می‌یابد. به طور کلی، با بررسی عمق‌های مختلف می‌توان بهترین عمق که بالاترین نتیجه را به دست می‌آورد پیدا کرد.

در جدول ۱-۴ عملکرد دو ساختار پیشنهادی پایان‌نامه و نتایج مقالات گذشته آورده شده است.

جدول ۱-۴) مقایسه عملکرد ساختارها

دقت ساختار				دقت ساختار				پیش‌پردازش	ساختار	مرجع
DO	سنسور	DO ^۲	مدل	DC	سنسور	DC ^۱	مدل			
✓	۹۱,۱ درصد	✓	۹۸,۱ درصد	-	-	-	-	-	DCNN ^۳	[۳۲]
✓	۹۴,۸ درصد	✓	۹۸,۵ درصد	✓	۹۵,۴ درصد	✓	۹۸,۵ درصد	✓	DCNN	روش پیشنهادی
✓	۹۲,۴ درصد	✓	۹۸,۳ درصد	✓	۹۳,۲ درصد	✓	۹۸,۳ درصد	✓	DLSVM ^۴	روش پیشنهادی

با بررسی جدول ۱-۴ مشاهده می‌شود دقت روش پیشنهادی در سطح سنسور با توجه به انجام پیش‌پردازش داده‌ها نسبت به مرجع [۳۲] افزایش یافته است. همچنین دقت ساختارها با حضور DropConnect مورد بررسی قرار گرفت، در سطح سنسور بهبود کارایی بیشتری را نسبت به شناسایی در سطح مدل شاهد بودیم. استفاده از ساختار DLSVM نیز بهبود دقت را به همراه دارد در ادامه روش‌های پیشنهادی با دو ساختار DCNN و DLSVM از نظر زمان اجرا مقایسه خواهند شد.

^۱ DropConnect

^۲ DropOut

^۳ Deep Convolutional Neural Network

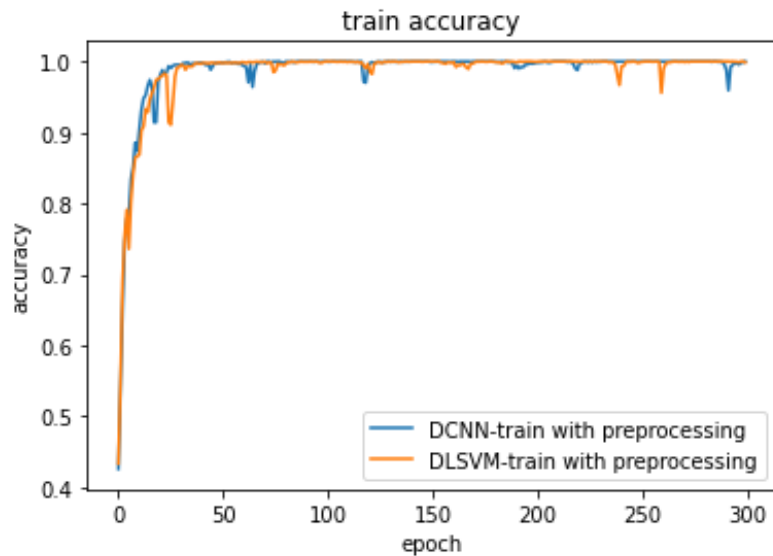
^۴ Deep Learning-Support Vector Machine

مهم‌ترین تفاوت بین روش پیشنهادی و مرجع [۳۲] اعمال پیش‌پردازش روی داده‌ها است. پیش‌پردازش داده‌ها معمولاً برای حذف نویزهای موجود در داده‌ها انجام می‌شود، نتایج روش‌های پیشنهادی این پایان‌نامه همراه با پیش‌پردازش نشان می‌دهد بعضی از پیش‌پردازش‌ها مانند افزایش کیفیت تصویر و تشخیص لبه باعث شده است کارایی شناسایی دوربین منبع در سطح سنسور در ساختار DCNN حدود ۴ درصد و در ساختار DLSVM حدود ۳ درصد افزایش پیدا کند. در سطح مدل با اعمال پیش‌پردازش‌ها برای هر یک از دو ساختار شاهد افزایش ناچیز زیر یک درصد هستیم ولی با این حال کارایی بهبود پیدا کرده است.

علاوه بر دقت روش پیشنهادی می‌توان به زمان اجرا اشاره کرد، پیش‌پردازش داده‌ها باعث ساده‌تر شدن آن‌ها در عین حال برجسته‌تر شدن ویژگی‌های موثر در فرآیند آموزش شده است. در روش‌های پیشنهادی سرعت اجرای برنامه به طور چشم‌گیری افزایش یافته است، در ادامه به مقایسه زمان اجرای روش‌ها می‌پردازیم.

زمان اجرا، رابطه مستقیم با تعداد مراحل آموزش دارد. تعیین مناسب تعداد مراحل آموزش به عنوان یکی از مهم‌ترین پارامترهای شبکه عصبی عمیق در کاهش زمان اجرا موثر خواهد بود. مشخص است هر چه تعداد مراحل آموزش بالاتر باشد دقت آموزش افزایش پیدا می‌کند ولی ممکن است باعث ایجاد بیش‌برازش شود.

در شکل ۴-۵ نمودار دقت بر اساس آموزش روش پیشنهادی با دو ساختار DCNN و DLSVM همراه با اعمال پیش‌پردازش آورده شده است.



شکل ۴-۵) نمودار دقت بر اساس تعداد مراحل آموزش برای دو ساختار DCNN و DLSVM با اعمال پیش‌پردازش

همانطور که در نمودار شکل ۴-۵ مشخص است، حضور فرآیند پیش‌پردازش باعث افزایش دقت شناسایی دوربین منبع در سطح مدل می‌شود. از مرحله ۵۰ام فرآیند آموزش، دقت آموزش به حداکثر مقدار خود می‌رسد از این رو تعداد مراحل آموزش برای دو ساختار روش‌های پیشنهادی این پایان‌نامه را برابر ۵۰ در نظر می‌گیریم.

افزایش کارایی حاصل از دو ساختار در صورت عدم اعمال پیش‌پردازش داده‌ها طبق شکل ۴-۵ مشخص است. با بررسی نمودار، کاهش زمان و افزایش سرعت اجرا نمایان است. در صورت عدم اعمال پیش‌پردازش نیاز به حداقل ۱۰۰ مرحله آموزش داریم که این مقدار با حضور پیش‌پردازش‌ها به ۵۰ مرحله کاهش پیدا کرده است. نمودار روند دقت مرحله آموزش در سطح سنسور نیز مشابه شکل ۴-۵ به دست آمد، همچنین تعداد مراحل آموزش برای شناسایی دوربین منبع در سطح مدل نیز برابر ۵۰ در نظر گرفته می‌شود.

۴-۶ ارزیابی ساختارهای پیشنهادی

در این بخش، برای ارزیابی ساختارهای پیشنهادی از ماتریس درهم‌ریختگی استفاده می‌کنیم. در بررسی عملکرد الگوریتم‌های یادگیری با ناظر در مبحث مدل‌های طبقه‌بندی از ماتریس درهم‌ریختگی و

معیارهای مرتبط برای ارزیابی مدل استفاده می‌شود. ماتریس درهم‌ریختگی به عنوان ماتریس خطا جدول خلاصه شده از ارزیابی عملکرد مدل است. تعداد پیش‌بینی‌های درست و نادرست شمارش شده و در جدولی بر اساس تعداد کلاس‌های موجود در طبقه‌بندی قرار می‌گیرند.

در شکل ۴-۶ یک نمونه ماتریس درهم‌ریختگی 2×2 آورده شده است.

		Actual Values	
		Yes	No
Predicted Values	Yes	True Positive	False Positive
	No	False Negative	True Negative

شکل ۴-۶) نمونه ماتریس درهم‌ریختگی 2×2

Predicted Values : مقادیر پیش‌بینی شده

Actual Values : مقادیر واقعی

True Positive (TP) : مدل به درستی کلاس مثبت را پیش‌بینی کرده است.

True Negative (TN) : مدل به درستی کلاس منفی را پیش‌بینی کرده است.

False Positive (FP) : مدل به اشتباه داده مربوط به کلاس مثبت را به عنوان داده کلاس منفی

پیش‌بینی کرده است. (خطای نوع ۱)

False Negative (FN) : مدل به اشتباه داده مربوط به کلاس منفی را به عنوان داده کلاس مثبت

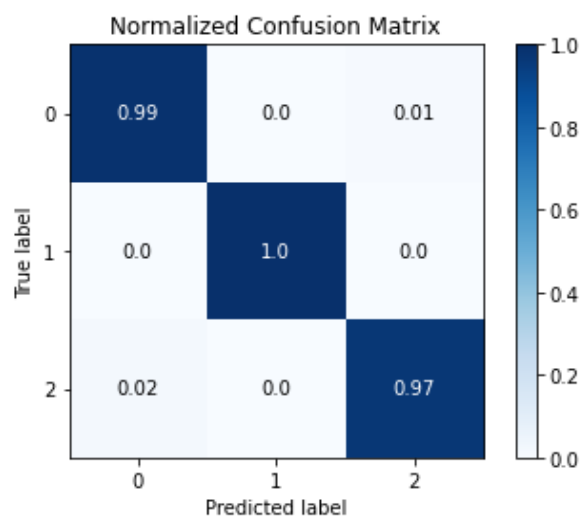
پیش‌بینی کرده است. (خطای نوع ۲)

با داشتن ماتریس درهم‌ریختگی و محاسبه معیار دقت عملکرد مدل طبقه‌بند را می‌توان مورد بررسی

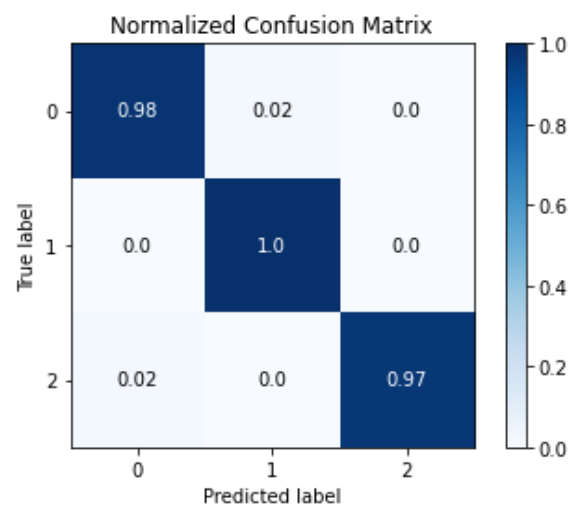
قرار داد. معیار دقت^۱ : همانطور که در رابطه ۴-۱ آمده است مقدار این معیار برابر با نسبت پیش‌بینی‌های درست به کل پیش‌بینی‌ها است. این معیار میزان تشخیص صحیح طبقه‌بند را نشان می‌دهد.

$$Accuracy = \frac{(1-4)}{TP+TN+FP+FN}$$

ماتریس درهم‌ریختگی نرمال‌سازی شده حاوی اطلاعات طبقه‌بندی در سطح مدل برای دو ساختار پیشنهادی در شکل ۴-۷



شکل ۴-۷) ماتریس درهم‌ریختگی برای نتایج ساختار DCNN در سطح مدل

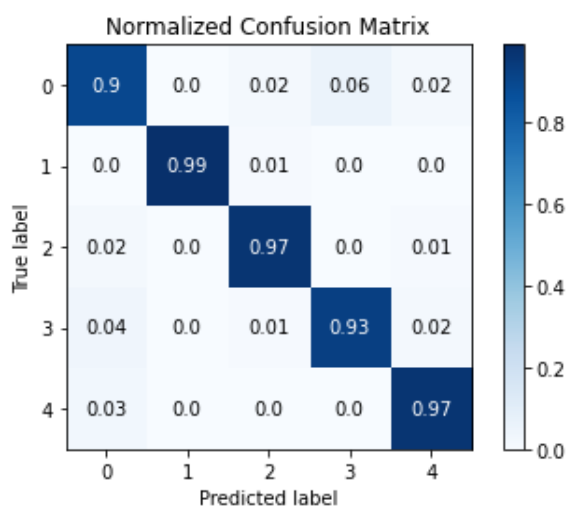


شکل ۴-۸) ماتریس درهم‌ریختگی برای نتایج ساختار DLSVM در سطح مدل

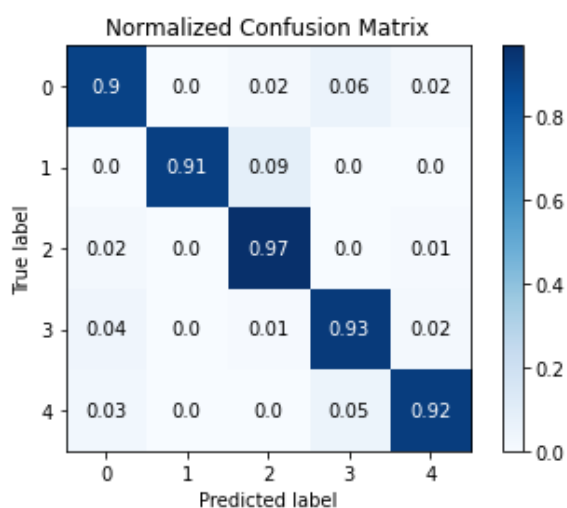
¹ Accuracy

با جایگذاری مقادیر ماتریس درهم‌ریختگی در رابطه ۴-۱ دقت ساختارهای پیشنهادی در سطح مدل محاسبه شده است. برای ساختار DCNN دقت ۹۸,۵ درصد و برای ساختار DLSVM دقت ۹۸,۳ درصد بدست آمد.

علاوه بر بررسی ساختارها در سطح مدل، به بررسی ساختارها در سطح سنسور نیز می‌پردازیم. در این قسمت داده‌ها در ۵ کلاس $GS4_F(0)$ ، $GS4_R(1)$ ، $GST2_F(2)$ ، $IP5_F(3)$ و $IP5_R(4)$ طبقه‌بندی می‌شوند. ماتریس درهم‌ریختگی نرمال‌سازی شده حاصل از تست این طبقه‌بندی‌ها در شکل ۴-۹ و ۴-۱۰ نشان داده شده است.



شکل ۴-۹) ماتریس درهم‌ریختگی برای نتایج ساختار DCNN در سطح سنسور



شکل ۴-۱۰) ماتریس درهم‌ریختگی برای نتایج ساختار DLSVM در سطح سنسور

ساختارهای پیشنهادی برای شناسایی دوربین منبع در سطح سنسور عملکرد خوبی را از خود نشان

دادند. طبق رابطه ۴-۱ برای ساختار DCNN دقت تست ۹۵,۴ درصد و برای ساختار DLSVM دقت تست ۹۳,۲ درصد بدست آمد.

مقدار معیار دقت نشان از تاثیر پیش‌پردازش داده‌ها و بهبود عملکرد ساختارهای پیشنهادی نسبت به ساختار مرجع [۳۲] دارد.

هم

۴-۷ جمع‌بندی

در این بخش، به بررسی نتایج ساختارهای پیشنهادی و تاثیر پیش‌پردازش داده‌ها پرداخته شد. کیفیت وضوح و تضاد داده‌های مجموعه دادگان با اعمال پیش‌پردازش‌ها افزایش یافت و به کمک فیلتر Prewitt لبه‌های تصاویر تشخیص داده شد. اعمال پیش‌پردازش باعث بهبود عملکرد ساختارها شدند. بررسی عملکرد ساختارهای پیشنهادی به کمک ماتریس درهم‌ریختگی نشان از تاثیر فرآیند پیش‌پردازش بر افزایش دقت شناسایی دوربین منبع دارد. در زمینه شناسایی دوربین منبع ساختارهای مبتنی بر یادگیری عمیق مختلفی طراحی شده است، برای مجموعه دادگان این پایان نامه ساختارهای پیشنهادی نتایج بهتری را نسبت به ساختارهای قبلی داشتند.

فصل ۵ : نتیجه گیری

در این فصل، به جمع‌بندی روش‌های پیشنهادی شرح داده شده در این پایان‌نامه می‌پردازیم. همچنین در قسمتی از این فصل پیشنهادهایی برای بهبود بیشتر عملکرد این روش‌ها در شناسایی دوربین منبع ارائه می‌دهیم.

۵-۱ جمع‌بندی و نتیجه‌گیری

تمرکز این پایان‌نامه بر روی شناسایی دوربین منبع به کمک یادگیری عمیق است. ساختار شبکه عصبی مناسب مبتنی بر عمیق شدند.

افزایش وضوح تصویر، افزایش تضاد و مشخص کردن لبه‌ها به عنوان پیش‌پردازش به فرآیند آموزش کمک می‌کند تا با سرعت و دقت بیشتری به اتمام برسد و الگوی نویز سنسور به عنوان یک ویژگی مهم در دوربین‌های دیجیتال برای هر کلاس از داده‌های موجود در مجموعه دادگان به بهترین شکل استخراج شود.

در ادامه دو ساختار مبتنی بر یادگیری عمیق ارائه شد که برای طبقه‌بندی کلاس‌ها پس از استخراج ویژگی از یک لایه تمام متصل با تابع فعال‌ساز Softmax و یا از ماشین بردار پشتیبان استفاده شد. هر دو روش پیشنهادی عملکرد مناسبی در شناسایی دوربین منبع داشتند.

روش‌های پیشنهادی نسبت به کارهای پیشین شرح داده شده در فصل ۲ دقت بالاتری را به دست آوردند. نتایج نشان داد اعمال پیش‌پردازش بر روی داده‌ها می‌تواند در بهبود کارایی موثر باشد. همچنین از نظر زمان اجرا، روش‌های پیشنهادی زمان کمتری را برای فرآیند اجرا ثبت کردند.

اعمال پیش‌پردازش‌های مناسب روی مجموعه دادگان مزیت مهم روش پیشنهادی این رساله نسبت به مقاله پایه [۳۲] می‌باشد. اعمال پیش‌پردازش روی داده‌ها به روند یادگیری در شناسایی بهتر ویژگی‌های خاص هر دوربین کمک کرد تا نتایج بدست آمده گواه این مطلب باشد.

در کنار پیاده‌سازی روش پیشنهادی رساله، الگوریتم مرجع پایه [۳۲] نیز با شرایط ذکر شده در مقاله پیاده‌سازی شد و با شرایط یکسان داده‌های ورودی به بررسی نتایج پرداختیم.

نتایج بدست آمده از ساختار مرجع [۳۲] با شرایط یکسان داده‌های ورودی با نتایج ذکر شده در مقاله

اختلاف مقدار داشت. الگوریتم مقاله در سطح مدل ۸۲,۴ درصد و در سطح سنسور ۷۳,۸ درصد دوربین منبع را به درستی شناسایی کرد در صورتی که مقاله شناسایی در سطح مدل و سنسور را به ترتیب ۹۸,۱ درصد و ۹۱,۱ درصد گزارش کرده است.

۵-۲ کارهای آتی

در این پایان‌نامه، تعدادی از الگوریتم‌های پیش‌پردازش داده‌ها بکار گرفته شد. نتایج نشان داد پیش‌پردازش داده‌ها می‌تواند عملکرد شناسایی دوربین منبع را بهبود ببخشد. در کارهای آتی می‌توان از پیش‌پردازش‌های دیگری نیز استفاده کرد. مجموعه دادگان مورد استفاده مجموعه‌ی کوچکی از تصاویر ثبت شده به‌وسیله موبایل‌های موجود در جامعه را در خود دارد. با وجود تنوع دستگاه‌های تلفن همراه و گسترش روز افزون این دستگاه‌ها، لازم است مجموعه دادگان مورد استفاده همواره به روز رسانی شود. وجود مجموعه دادگان مناسب در این زمینه می‌تواند نتایج بهتری را داشته باشد.

برای حل مشکل مجموعه دادگان در کارهای آتی می‌توان از شبکه‌های مولد تخاصمی (GAN) استفاده کرد. در سال‌های اخیر، پیشرفت چشم‌گیر شبکه عصبی GAN باعث شده تا این نوع شبکه عصبی مورد توجه قرار بگیرد. در کارهای آتی با داشتن تعداد کمی از داده‌های تصویری از هر دستگاه موبایل و استفاده از شبکه عصبی GAN می‌توان حجم مجموعه دادگان را افزایش داد و شاهد بهبود عملکرد شناسایی دوربین منبع باشیم.

- [1] R. Bohme and M. Kirchner, "Digital Image Forensics: There is More to a Picture Than Meets the Eye, chapter Counter-forensics: Attacking Image Forensics," ed: Springer-Verlag, 2013.
- [2] M. Kharrazi, H. T. Sencar, and N. Memon, "Blind source camera identification," in 2004 International Conference on Image Processing, 2004. ICIP'04., 2004, vol. 1: IEEE, pp. 709-712 .
- [3] D. Cozzolino, D. Gagnaniello, and L. Verdoliva, "Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques," in 2014 IEEE International Conference on Image Processing (ICIP), 2014: IEEE, pp. 5302-5306 .
- [4] P. Korus, "Digital image integrity—a survey of protection and verification techniques," Digital Signal Processing, vol. 71, pp. 1-26. ٢٠١٧ ,
- [5] L. Bondi, L. Baroffio, D. Güera, P. Bestagini, E. J. Delp, and S. Tubaro, "First steps toward camera model identification with convolutional neural networks," IEEE Signal Processing Letters, vol. 24, no. 3, pp. 259-263, 2016.
- [6] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," IEEE Transactions on Information Forensics and Security, vol. 5, no. 2, pp. 280-287, 2010.
- [7] K. S. Choi, E. Y. Lam, and K. K. Wong, "Source camera identification using footprints from lens aberration," Digital photography II, vol. 6069, pp. 172-179, 2006.
- [8] T. Qiao, F. Retraint, R. Cogranne, and T. H. Thai, "Source camera device identification based on raw images," in 2015 IEEE international conference on image processing (ICIP), 2015: IEEE, pp. 3812-3816 .
- [9] J. Lukas, J. Fridrich, and M. Goljan, "Determining digital image origin using sensor imperfections," in Image and Video Communications and Processing 2005, 2005, vol. 5685: International Society for Optics and Photonics, pp. 249. ٢٠٠٥ -
- [10] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, pp. 205-214, 2006.
- [11] M. Chen, J. Fridrich, M. Goljan, and J. Lukás, "Determining image origin and integrity using sensor noise," IEEE Transactions on information forensics and security, vol. 3, no. 1, pp. 74-90, 2008.
- [12] Y. Tomioka, Y. Ito, and H. Kitazawa, "Robust digital camera identification based on pairwise magnitude relations of clustered sensor pattern noise," IEEE transactions on information forensics and security, vol. 8, no. 12, pp. 1986-1995, 2013.
- [13] C. Chen and M. C. Stamm, "Camera model identification framework using an ensemble of demosaicing features," in ٢٠١٥ IEEE international workshop on information forensics and security (WIFS), 2015: IEEE, pp. 1-6 .
- [14] A. Roy, R. S. Chakraborty, V. U. Sameer, and R. Naskar, "Camera Source Identification Using Discrete Cosine Transform Residue Features and Ensemble Classifier," in CVPR Workshops, 2017, pp. 1848-1854 .

- [15] Y. Hu, C.-T. Li, and C. Zhou, "Selecting forensic features for robust source camera identification," in 2010 International Computer Symposium (ICS2010), 2010: IEEE, pp. 506-511 .
- [16] B. Wang ,Y. Guo, X. Kong, and F. Meng, "Source camera identification forensics based on wavelet features," in 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009: IEEE, pp. 702-705 .
- [17] A. Tuama, F. Comby ,and M. Chaumont, "Camera model identification with the use of deep convolutional neural networks," in 2016 IEEE International workshop on information forensics and security (WIFS), 2016: IEEE, pp. 1-6 .
- [18] H. Yao, T. Qiao, M. Xu, and N. Zheng, "Robust multi-classifier for camera model identification based on convolution neural network," IEEE Access, vol. 6, pp. 24973-24982, 2018.
- [19] H. Farid, Photo forensics. MIT press, 2016.
- [20] M. C. Stamm, M. Wu, and K. R. Liu, "Information forensics: An overview of the first decade," IEEE access, vol. 1, pp. 167-200, 2013.
- [21] A. Piva, "An overview on image forensics," International Scholarly Research Notices, vol. 2013, 2013.
- [22] M. Barni et al., "Aligned and non-aligned double JPEG detection using convolutional neural networks," Journal of Visual Communication and Image Representation, vol. 49, pp. 153-163, 2017.
- [23] A. Tuama, F. Comby, and M. Chaumont, "Source camera model identification using features from contaminated sensor noise," in International Workshop on Digital Watermarking, 2015: Springer, pp. 83-93 .
- [24] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on CFA interpolation," in IEEE International Conference on Image Processing 2005, 2005, vol. 3: IEEE, pp .III-69 .
- [25] T. Gloe, "Feature-based forensic camera model identification," in Transactions on data hiding and multimedia security VIII: Springer, 2012, pp. 42-62.
- [26] M. Boutell and J. Luo, "Photo classification by integrating image content and camera metadata," in Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., 2004, vol. 4: IEEE, pp. 901-904 .
- [27] K. San Choi, E. Y. Lam, and K. K. Wong, "Automatic source camera identification using the intrinsic lens radial distortion," Optics express, vol. 14, no. 24, pp. 11551-11565, 2006.
- [28] T. Filler, J. Fridrich, and M. Goljan, "Using sensor pattern noise for camera model identification," in 2008 15th IEEE International Conference on Image Processing, 2008: IEEE ,pp. 1296-1299 .
- [29] A. Rocha, W. Scheirer, T. Boulton, and S. Goldenstein, "Vision of the unseen: Current trends and challenges in digital image and video forensics," ACM Computing Surveys (CSUR), vol. 43, no. 4, pp. 1-42, 2011.
- [30] A. T. Ho and S. Li, "Forensic camera model identification," 2015.
- [31] L. Gaborini, P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Multi-clue image tampering localization," in 2014 IEEE International Workshop on Information Forensics and Security (WIFS), 2014: IEEE ,pp. 125-130 .
- [32] D. Freire-Obregon, F. Narducci, S. Barra, and M. Castrillon-Santana, "Deep learning for source camera identification on mobile devices," Pattern Recognition Letters, vol. 126, pp. 86-91, 2019.
- [33] Z. J. Geradts, J. Bijhold, M. Kieft ,K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," in Enabling

- technologies for law enforcement and security, 2001, vol. 4232: International Society for Optics and Photonics, pp. 505-512 .
- [34] Y. Long and Y. Huang, "Image based source camera identification using demosaicking," in 2006 IEEE Workshop on Multimedia Signal Processing, 2006: IEEE, pp. 419-424 .
 - [35] O. Celiktutan, I. Avcibas, B. Sankur, and N. Memon, "Source cell-phone identification," IEEE Signal Processing and Communications Applications, pp. 1-3, 2005.
 - [36] M. K. Johnson and H. Farid, "Exposing digital forgeries through chromatic aberration," in Proceedings of the 8th workshop on Multimedia and security, 2006, pp. 48-55 .
 - [37] M. K. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting," in Proceedings of the 7th workshop on Multimedia and security, 2005, pp. 1-10 .
 - [38] Y.-F. Hsu and S.-F. Chang, "Detecting image splicing using geometry invariants and camera characteristics consistency," in 2006 IEEE International Conference on Multimedia and Expo, 2006: IEEE, pp. 549-552 .
 - [39] J. Adams, K. Parulski, and K. Spaulding, "Color processing in digital cameras," IEEE micro, vol. 18, no. 6, pp. 1998, 20-20 .
 - [40] E. Alkim, P. S. Barreto, N. Bindel, J. Krämer, P. Longa, and J. E. Ricardini, "The lattice-based digital signature scheme qtesla," in International Conference on Applied Cryptography and Network Security, 2020: Springer, pp. 441-460 .
 - [41] X. Zhao, J. Li, S. Li, and S. Wang, "Detecting digital image splicing in chroma spaces," in International workshop on digital watermarking, 2010: Springer, pp. 12-22 .
 - [42] Bharti, C.N. and P. Tandel. A survey of image forgery detection techniques. in Wireless Communications, Signal Processing and Networking (WiSPNET), International Conference on. 2016. IEEE.
 - [43] S. Bravo-Solorio and A. K. Nandi, "Automated detection and localisation of duplicated regions affected by reflection, rotation and scaling in image forensics," Signal Processing ,vol. 91, no. 8, pp. 1759-1770, 2011.
 - [44] Luo W, Huang J, Qiu G. Robust detection of region-duplication forgery in digital image. In: Proc. of the 18th International conference on pattern recognition 2006. p. 746–9.
 - [45] Mahdian B, Saic S. Detection of near-duplicated image regions. In: Computer recognition systems 2Advances in soft computing, vol. 45. 2007.p. 187–95.
 - [46] Myna A, Venkateshmurthy M, Patil C. Detection of region duplication forgery in digital images using wavelets and log-polar mapping. In:Proc. of the International conference on computational intelligence and multimedia applications (ICCIMA 2007) 2007. p.371–7.
 - [47] Kirchner M, Fridrich J. On detection of median filtering in digital images.In: Proc. of the SPIE, electronic imaging, media forensics and security XII, vol. 7541. 2010. p. 754110–754110-12.
 - [48] Cao G, Zhao Y, Ni R. Edge-based blur metric for tamper detection. J Inf Hiding Multimed Signal Process 2010a;1(1):20–7.
 - [49] Gul G, Avcibas I, Kurugollu F. SVD based image manipulation detection. In:Proc. International conference on image processing (ICIP) 2010. p.1765–8.
 - [50] Mahalakshmi DS, Vijayalakshmi K, Priyadharsini S. Digital image forgery detection and estimation by exploring basic image manipulations.Digital Invest 2012;8(34):215 25.

- [51] Fang Y, Dirik AE, Sun X, Memon N. Source class identification for DSLR and compact cameras. In: Proc. IEEE workshop on multimedia signal processing 2009b. p. 1–5.
- [52] Khanna N, Chiu GT-C, Allebach JP, Delp EJ. Forensic techniques for classifying scanner, computer generated and digital camera images. In: Proc. IEEE International conference on acoustics, speech and signal processing 2008. p. 1653–6.
- [53] Chen W, Shi YQ, Xuan GR, Su W. Computer graphics identification using genetic algorithm. In: Proc. IEEE International conference on pattern recognition 2008b. p. 14.
- [54] Wang W, Dong J, Tan T. Effective image splicing detection based on image chroma. In: Proc. IEEE International conference on image processing 2009. p. 1257–60.
- [55] Chen W, Shi Y, Su W. Image splicing detection using 2-d phase congruency and statistical moments of characteristic function. In: Proc. of SPIE electronic imaging: security, steganography, and watermarking of multimedia contents 2007.
- [56] Hsu Y, Chang S. Detecting image splicing using geometry invariants and camera characteristics consistency. In: Proc. IEEE International conference on multimedia and Expo (ICME) 2006. p. 549–52.
- [57] Sergio B, Asoke N. Automated detection and localisation of duplicated regions affected by reflection, rotation and scaling in image forensics. *Signal Process* 2011;91:1759–70.
- [58] Muhammad G, Hussain M, Khawaji K, Bebis G. Blind copy move image forgery detection using dyadic uncedimated wavelet transform. In: Proc. of 17th International conference on digital signal processing 2011. p. 1–6.
- [59] Ghorbani M, Firouzmand M, Faraahi A. DWT-DCT (QCD) based copy-move image forgery detection. In: Proc. of 18th International conference on systems, signals and image processing (IWSSIP) 2011. p. 1–4.
- [60] Gopi E, Lakshmanan N, Gokul T, Ganesh S, Shah P. Digital image forgery detection using artificial neural network and auto regressive coefficients. In: Proc. Canadian conference on electrical and computer engineering 2006. p. 194–7.
- [61] Bianchi, T. and A. Piva, Image forgery localization via block-grained analysis of JPEG artifacts. *IEEE Transactions on Information Forensics and Security*, 2012. 7(3): p. 1003-1017.
- [62] Johnson, M.K. and H. Farid, Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security*, 2007. 2(3): p. 450-461.
- [63] Popescu, A.C. and H. Farid, Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on signal processing*, 2005. 53(2): p. 758-767.
- [64] Mahdian, B. and S. Saic, Blind authentication using periodic properties of interpolation. *IEEE Transactions on Information Forensics and Security*, 2008. 3(3): p. 529-538.
- [65] Stamm, M.C. and K.R. Liu, Forensic detection of image manipulation using statistical intrinsic fingerprints. *IEEE Transactions on Information Forensics and Security*, 2010. 5(3): p. 492-506.
- [66] Shi, Y.Q., et al. Steganalysis versus splicing detection. in *International Workshop on Digital Watermarking*. 2007. Springer.
- [67] Fridrich, J. and J. Kodovsky, Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security*, 2012. 7(3): p. 868-882.

- [68] Fridrich J, Soukal D, Lukas J. Detection of copy-move forgery in digital images. In: Proc. of digital forensic research workshop 2003. p. 55–61.
- [69] Popescu A, Farid H. Exposing digital forgeries by detecting duplicated image regions. Technical Report TR2004-515. Department of Computer Science, Dartmouth College; 2004.
- [70] Mahdian B, Saic S. Detection of near-duplicated image regions. In: Computer recognition systems 2Advances in soft computing, vol. 45. 2007.p. 187–95.
- [71] E. Kee and H. Farid, "A perceptual metric for photo retouching," Proceedings of the National Academy of Sciences, vol. 108, no. 50, pp. 19907-19912, 2011.
- [72] L. Alamro and N. Yusoff" ,Copy-move forgery detection using integrated DWT and SURF," Journal of Telecommunication, Electronic and Computer Engineering (JTEC), vol. 9, no. 1-2, pp. 67-71, 2017.
- [73] Cao, H. and A.C. Kot, Accurate detection of demosaicing regularity for digital image forensics. IEEE Transactions on Information Forensics and Security, 2009. 4(4): p. 899-910.
- [74] Swaminathan, A., M. Wu, and K.R. Liu, Digital image forensics via intrinsic fingerprints. IEEE transactions on information forensics and security, 2008. 3(1): p. 101-117.
- [75] Hsu, Y.-F. and S.-F. Chang, Camera response functions for image forensics: an automatic algorithm for splicing detection. IEEE Transactions on Information Forensics and Security, 2010. 5(4): p. 816-825.
- [76] Chen, M., et al., Determining image origin and integrity using sensor noise. IEEE Transactions on information forensics and security, 2008. 3(1): p. 74-90.
- [77] Johnson M, Farid H. Metric measurements on a plane from a single image. Technical Report TR2006-579. Department of Computer Science, Dartmouth College; 2006a.
- [78] Lanh T, Emmanuel S, Kankanhalli M. Identifying source cell phone using chromatic aberration. In: Proc. IEEE International conference on multimedia and Expo 2007b. p. 883–6.
- [79] Gloe T, Winkler A, Borowka K. Efficient estimation and large-scale evaluation of lateral chromatic aberration for digital image forensics. In:Proc. SPIE conference on media forensics and security 2010. p.754107.
- [80] T.-T. Ng and S.-F. Chang, "A model for image splicing," in 2004 International Conference on Image Processing, 2004. ICIP'04., 2004, vol. 2: IEEE, pp. 1169-1172 .
- [81] M.-J. Tsai and G.-H. Wu, "Using image features to identify camera sources," in 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, 2006, vol. 2: IEEE, pp. II-II.
- [82] Guan, H.; Kozak, M.; Robertson, E.; Lee, Y.; Yates, A.N.; Delgado, A.; Zhou, D.; Kheyrkhah, T.; Smith, J.; Fiscus, J. MFC Datasets: Large-Scale Benchmark Datasets for Media Forensic Challenge Evaluation. In Proceedings of the 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW), Waikoloa Village, HI, USA, 7–11 January 2019; pp. 63–72.
- [83] IEEE. Signal Processing Society - Camera Model Identification; IEEE: Piscataway, NJ, USA, 2018.
- [84] Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. Neural Comput. 2006,18, 1527–1554.
- [85] Bengio, Y.; Lamblin, P.; Popovici, D.; Larochelle, H. Greedy layer-wise training of deep networks. In Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2007; pp. 153–160.

- [86] Ranzato, M.; Poultney, C.; Chopra, S.; LeCun, Y. Efficient learning of sparse representations with an energy-based model. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2007; pp. 1137–1144.
- [87] LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015, 521, 436.
- [88] Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
- [89] Khan, A.; Sohail, A.; Zahoor, U.; Qureshi, A.S. A Survey of the Recent Architectures of Deep Convolutional Neural Networks. *arXiv* 2019, arXiv:1901.06032.
- [90] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2012; pp. 1097–1105.
- [90] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2012; pp. 1097–1105.
- [91] Simonyan, K.; Zisserman, A. Very deep convolutional networks for large scale image recognition. *arXiv* 2014, arXiv:1409.1556.
- [92] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- [93] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- [94] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
- 95. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
- [96] Jie, H.; Li, S.; Albanie, S.; Gang, S.; Wu, E. Squeeze-and-Excitation Networks. *arXiv* 2017, arXiv:1709.01507.
- [97] Chopra, S.; Hadsell, R.; LeCun, Y. Learning a Similarity Metric Discriminatively, with Application to Face Verification. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01*; IEEE Computer Society: Washington, DC, USA, 2005; CVPR '05; pp. 539–546.
- [98] Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* 2018, 77, 354–377.
- [99] Bondi, L.; Baroffio, L.; Güera, D.; Bestagini, P.; Delp, E.J.; Tubaro, S. Camera identification with deep convolutional networks. *arXiv* 2016, arXiv:1603.01068.
- [100] Liu, Y., Zou, Z., Yang, Y., Law, N.F.B. and Bharath, A.A., 2021. Efficient Source Camera Identification with Diversity-Enhanced Patch Selection and Deep Residual Prediction. *Sensors*, 21(14), p.4701.
- [101] Huang, N.; He, J.; Zhu, N.; Xuan, X.; Liu, G.; Chang, C. Identification of the source camera of images based on convolutional neural network. *Digit. Investig.* 2018, 26, 72–80.
- [102] Yao, H.; Qiao, T.; Xu, M.; Zheng, N. Robust multi-classifier for camera model identification based on convolution neural network. *IEEE Access* 2018, 6, 24973–24982.

- [103] Chen, Y.; Huang, Y.; Ding, X. Camera model identification with residual neural network. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 4337–4341.
- [104] Ding, X.; Chen, Y.; Tang, Z.; Huang, Y. Camera Identification based on Domain Knowledge-driven Deep Multi-task Learning. *IEEE Access* 2019, 7, 25878–25890.
- [105] Marra, F.; Gragnaniello, D.; Verdoliva, L. On the vulnerability of deep learning to adversarial attacks for camera model identification. *Signal Process. Image Commun.* 2018, 65, 240–248.
- [106] Stamm, M.; Bestagini, P.; Marcenaro, L.; Campisi, P. Forensic Camera Model Identification: Highlights from the IEEE Signal Processing Cup 2018 Student Competition [SP Competitions]. *IEEE Signal Process. Mag.* 2018, 35, 168–174.
- [107] Kamal, U.; Rafi, A.M.; Hoque, R.; Das, S.; Abrar, A.; Hasan, M. Application of DenseNet in Camera Model Identification and Post-processing Detection. *arXiv* 2018, arXiv:1809.00576.
- [108] Kuzin, A.; Fattakhov, A.; Kibardin, I.; Iglovikov, V.I.; Dautov, R. Camera Model Identification Using Convolutional Neural Networks. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 3107–3110.
- [109] Ferreira, A.; Chen, H.; Li, B.; Huang, J. An Inception-Based Data-Driven Ensemble Approach to Camera Model Identification. In Proceedings of the 2018 IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, 11–13 December 2018; pp. 1–7.
- [110] Bayar, B.; Stamm, M.C. Towards open set camera model identification using a deep learning framework. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2007–2011.
- [111] Mayer, O.; Stamm, M.C. Learned forensic source similarity for unknown camera models. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2012–2016.
- [112] Bondi, L.; Baroffio, L.; Güera, D.; Bestagini, P.; Delp, E.J.; Tubaro, S. First steps toward camera model identification with convolutional neural networks. *IEEE Signal Process. Lett.* 2017, 24, 259–263.
- [113] Bondi, L.; Güera, D.; Baroffio, L.; Bestagini, P.; Delp, E.J.; Tubaro, S. A preliminary study on convolutional neural networks for camera model identification. *Electron. Imaging* 2017, 2017, 67–76.
- [114] Bayar, B.; Stamm, M.C. Augmented convolutional feature maps for robust cnn-based camera model identification. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 4098–4102.
- [115] Tuama, A.; Comby, F.; Chaumont, M. Camera model identification with the use of deep convolutional neural networks. In Proceedings of the 2016 IEEE International workshop on information forensics and security (WIFS), Abu Dhabi, UAE, 4–7 December 2016; pp. 1–6.
- [116] Zuo, Z. Camera Model Identification with Convolutional Neural Networks and Image Noise Pattern, 2018. Available online: <http://hdl.handle.net/2142/100123> (accessed on 2 July 2018).
- [117] Wang, B.; Yin, J.; Tan, S.; Li, Y.; Li, M. Source camera model identification based on convolutional neural networks with local binary patterns coding. *Signal Process. Image Commun.* 2018, 68, 162–168.

- [118] Yang, P.; Zhao, W.; Ni, R.; Zhao, Y. Source camera identification based on content-adaptive fusion network.arXiv 2017, arXiv:1703.04856.
- [119] Yang, P.; Ni, R.; Zhao, Y.; Zhao, W. Source camera identification based on content-adaptive fusion residual networks. *Pattern Recognit. Lett.* 2019, 119, 195–204.
- [121] Pevny, T.; Bas, P.; Fridrich, J. Steganalysis by subtractive pixel adjacency matrix. *IEEE Trans. Inf. Forensics Secur.* 2010, 5, 215–224.
- [120] Bayar, B.; Stamm, M.C. Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection. *IEEE Trans. Inf. Forensics Secur.* 2018, 13, 2691–2706.
- [122] Tang, Y. Deep learning using linear support vector machines. arXiv 2013, arXiv:1306.0239.
- [123] Chen, S.; Wang, Y. Convolutional Neural Network and Convex Optimization; Department of Electrical and Computer Engineering University of California San Diego: San Diego, CA, USA, 2014.
- [124] Abdalnabi, A.H.; Wang, G.; Lu, J.; Jia, K. Multi-task CNN model for attribute prediction. *IEEE Trans.Multimed.* 2015, 17, 1949–1959.
- [125] Yang, Pengpeng, Daniele Baracchi, Rongrong Ni, Yao Zhao, Fabrizio Argenti, and Alessandro Piva. "A survey of deep learning-based source image forensics." *Journal of Imaging* 6, no. 3 (2020): 9
- [126] Goodfellow, Ian, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. Deep learning. Vol. 1, no. 2. Cambridge: MIT press, 2016.
- [127] Maas, A. L., Hannun, A. Y., and Ng, A. Y. (2013). Rectified nonlinearities improve neural network acoustic models. In *ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, pages 13–19.
- [128] Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." In *Icml*. 2010.
- [129] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *CoRR*, abs/1207.0580.
- [130] Wan, Li, Matthew Zeiler, Sixin Zhang, Yann Le Cun, and Rob Fergus. "Regularization of neural networks using dropconnect." In *International conference on machine learning*, pp. 1058–1066. PMLR, 2013.
- [131] Agarap, Abien Fred. "An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification." arXiv preprint arXiv:1712.03541 (2017).
- [132] Gloe, T.; Böhme, R. The Dresden Image Database for benchmarking digital image forensics. In *Proceedings of the 2010 ACM Symposium on Applied Computing*, Sierre, Switzerland, 22–26 March 2010; pp. 1584–1590.
- [133] De Marsico, M.; Nappi, M.; Riccio, D.; Wechsler, H. Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols. *Pattern Recognit. Lett.* 2015, 57, 17–23.
- [134] IEEE. Signal Processing Society - Camera Model Identification; IEEE: Piscataway, NJ, USA, 2018.
- [135] Shullani, D.; Fontani, M.; Iuliani, M.; Al Shaya, O.; Piva, A. VISION: a video and image dataset for source identification. *EURASIP J. Inf. Secur.* 2017, 2017, 15.

نعت نامہ

به ترتیب حروف الفبای فارسی

عنوان فارسی	عنوان انگلیسی
ادغام و ترکیب	Fusion and Ensemble
استخراج ویژگی	Feature Extraction
افزایش داده	Data Augmentation (DA)
امضا دیجیتال	Digital Signature
انحراف	Aberration
آرایه فیلتر رنگ	Color Filter Array (CFA)
بالاگذر	High Pass
بیش‌برازش	Overfitting
پاسخ عکس غیر یکنواخت	Photo Response Non Uniformity (PRNU)
پردازش تصویر	Image Processing
پردازش گر تصویر دیجیتال	Digital Image Processing (DIP)
پس‌پردازش	Postprocessing
پیش‌پردازش	Preprocessing
تابع بیشنیه هموار	Softmax Function

عنوان فارسی	عنوان انگلیسی
تابع فعال ساز	Activation Function
تبدیل کسینوس گسسته	Discrete Cosine Transform (DCT)
تبدیل موجک گسسته	Discrete Wavelet Transform (DWT)
تجزیه و تحلیل اجزای اصلی	Principal Component Analysis (PCA)
تصویر جعلی	Forged Image
تضاد	Contrast
جرم‌شناسی تصویر	Image Forensics
حداکثر انتظار	Expectation-Maximization (EM)
درختان بسیار تصادفی	Extremely Randomized Trees (ERT)
درون‌یابی	Interpolation
دستگاه شارژ همراه	Charge-Coupled Device (CCD)
دقت	Accuracy
روش‌های تشخیص جعل	Forgery Detection Methods
شباهت	Similarity

عنوان فارسی	عنوان انگلیسی
شبکه عصبی بازگشتی	Recurrent Neural Network (RNN)
شبکه عصبی تخصصی	Generative Adversarial Network (GAN)
شبکه عصبی کانولوشنی	Convolutional Neural Network (CNN)
شناسایی دوربین منبع	Source Camera Identification (SCI)
شناسایی دوربین منبع منفرد	Individual Source Camera Identification (ISCI)
شناسایی مدل دوربین منبع	Source Camera Model Identification (SCMI)
طبقه‌بند	Classifier
علامت‌گذاری	Watermarking
عملکرد پاسخ دوربین	Camera Response Function (CRF)
غیرفعال	Passive
غیریکنواختی پیکسل	Pixel Non-Uniformity (PNU)
فعال	Active
قالبی برای ذخیره‌سازی تصویر فشرده	Joint Photographic Experts Group (JPEG)
قطعه‌ای کوچک از تصویر	Patch

عنوان فارسی	عنوان انگلیسی
الگوهای باینری محلی	Local Binary Patterns (LBP)
الگوی نویز سنسور	Sensor Pattern Noise (SPN)
لایه تمام متصل	Fully Connected Layer
لایه جمع کننده	Pooling Layer
لایه کانولوشنی	Convolutional Layer
ماتریس درهم ریختگی نرمال سازی شده	Normalized Confusion Matrix
ماشین بردار پشتیبان	Support Vector Machine (SVM)
مجموعه تصاویر آموزشی	Training Image Set
مقیاس خاکستری	Grayscale
مکمل اکسید فلز نیمه رسانا	Complementary Metal Oxide Semiconductor (CMOS)
میزان وضوح تصاویر	Sharpness
نسبت سیگنال به نویز	Signal-To-Noise Ratio (SNR)
نقصان و خرابی	Imperfections
واحد خطی اصلاح شده	Rectified Linear Unit (ReLU)

عنوان فارسی	عنوان انگلیسی
واحد خطی اصلاح شده نشتی	Leaky Rectified Linear Unit (Leaky ReLU)
هسته	Kernel
یادگیری عمیق	Deep Learning (DL)

به ترتیب حروف الفبای انگلیسی

عنوان فارسی	عنوان انگلیسی
انحراف	Aberration
دقت	Accuracy
فعال	Active
تابع فعال ساز	Activation Function
عملکرد پاسخ دوربین	Camera Response Function (CRF)
دستگاه شارژ همراه	Charge-Coupled Device (CCD)
طبقه‌بند	Classifier
آرایه فیلتر رنگ	Color Filter Array (CFA)
مکمل اکسید فلز نیمه رسانا	Complementary Metal Oxide

عنوان فارسی	عنوان انگلیسی
تضاد	Contrast
لایه کانولوشنی	Convolutional Layer
شبکه عصبی کانولوشنی	Convolutional Neural Network (CNN)
افزایش داده	Data Augmentation (DA)
یادگیری عمیق	Deep Learning (DL)
پردازش گر تصویر دیجیتال	Digital Image Processing (DIP)
امضا دیجیتال	Digital Signature
تبدیل کسینوس گسسته	Discrete Cosine Transform (DCT)
تبدیل موجک گسسته	Discrete Wavelet Transform (DWT)
حداکثر انتظار	Expectation-Maximization (EM)
درختان بسیار تصادفی	Extremely Randomized Trees (ERT)
استخراج ویژگی	Feature Extraction
تصویر جعلی	Forged Image
روش های تشخیص جعل	Forgery Detection Methods

عنوان فارسی	عنوان انگلیسی
لایه تمام متصل	Fully Connected Layer
ادغام و ترکیب	Fusion and Ensemble
شبکه عصبی تخصصی	Generative Adversarial Network (GAN)
مقیاس خاکستری	Grayscale
بالاگذر	High Pass
جرم‌شناسی تصویر	Image Forensics
پردازش تصویر	Image Processing
نقصان و خرابی	Imperfections
شناسایی دوربین منبع منفرد	Individual Source Camera Identification (ISCI)
درون‌یابی	Interpolation
قالبی برای ذخیره‌سازی تصویر فشرده	Joint Photographic Experts Group (JPEG)
هسته	Kernel
واحد خطی اصلاح شده نشتی	Leaky Rectified Linear Unit (Leaky ReLU)
الگوهای باینری محلی	Local Binary Patterns (LBP)

عنوان فارسی	عنوان انگلیسی
ماتریس درهم‌ریختگی نرمال‌سازی شده	Normalized Confusion Matrix
بیش‌برازش	Overfitting
غیرفعال	Passive
قطعه‌ای کوچک از تصویر	Patch
پاسخ عکس غیر یکنواخت	Photo Response Non Uniformity (PRNU)
غیریکنواختی پیکسل	Pixel Non-Uniformity (PNU)
لایه جمع‌کننده	Pooling Layer
پس‌پردازش	Postprocessing
پیش‌پردازش	Preprocessing
تجزیه و تحلیل اجزای اصلی	Principal Component Analysis (PCA)
واحد خطی اصلاح شده	Rectified Linear Unit (ReLU)
شبکه عصبی بازگشتی	Recurrent Neural Network (RNN)
الگوی نویز سنسور	Sensor Pattern Noise (SPN)
میزان وضوح تصاویر	Sharpness

عنوان فارسی	عنوان انگلیسی
نسبت سیگنال به نویز	Signal-To-Noise Ratio (SNR)
شبهت	Similarity
تابع بیشنیه هموار	Softmax Function
شناسایی دوربین منبع	Source Camera Identification (SCI)
شناسایی مدل دوربین منبع	Source Camera Model Identification (SCMI)
ماشین بردار پشتیبان	Support Vector Machine (SVM)
مجموعه تصاویر آموزشی	Training Image Set
علامت گذاری	Watermarking

Abstract

With the advancement of digital technology, digital media can be recorded from the cameras of various devices and easily shared in cyberspace and social networks. In this context, video and video manipulations have become very common and increasingly dangerous for individuals and society as a whole. Identifying the camera model is the first step in assessing the authenticity of images in forensic research on digital images. Source camera identification is the topic of this dissertation. In most related articles, various deep learning architectures were designed to extract the noise pattern as the fingerprint of each camera. Each image contains noises that are unique to each camera based on the design of the lens system. This noise pattern will not be lost due to image editing and manipulation, and the ultimate goal of extracting the noise pattern is to identify the camera model of each image that is used in the discussion of image criminology. Specifically, this dissertation analyzes CNN data processing and architecture.

Data preprocessing does not always help identify the source camera, as some data may be lost during the preprocessing process and the efficiency of the source camera identification system may be reduced. Therefore, choosing the right preprocessors is important. The proposed method of this dissertation focuses on data processing using the functions of increasing the image resolution, increasing the contrast in the image and recognizing the edge in the image. Next, the determination of deep neural network structure for feature extraction was examined. A deep neural network structure consisting of two convolutional layers, a maximal retractor layer and two fully connected bilayers was used. The softmax function and the SVM support vector machine at the end of deep neural network structures were used to classify the datasets. In this dissertation, to prevent overfitting, the DropConnect method was used instead of the DropOut method, in order to extract a more accurate noise pattern by preserving more neurons than the DropOut method, and to improve the detection efficiency of the source camera. At the model and sensor levels, the deep neural network structure with softmax function was able to correctly identify 98.5% and 95.4% of the test data, respectively. Also, the structure of deep neural network with support vector machine correctly identified 98.3% and 93.2% of the test data, respectively. In this way, the proposed methods have been more accurate than other sources.

Keywords: Image Forensics, Source Camera Identification, Preprocessing, Deep Learning, Support Vector Machine.



Faculty Of Computer Engineering
MSc Thesis in Artificial Intelligence Engineering

Source Camera Identification (SCI) Based on Deep Learning

By: masoud Bahaari

Supervisor:
Mansoor Fateh

Advisor:
Mohsen Rezvani

July 2021