Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال اول، شماره اول، بهار و تابستان ۱۴۰۲، صفحات ۲۱-۱۴

روش غیرمداخله کننده و مقاوم در مقابل پوشش جهت کشف جعل در شناسایی چهره بر اساس یادگیری عمیق

سید ابراهیم حسینی [۱]، حمید حسن‌پور [۲*]

[۱] دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود، ایران

[۲] دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود، ایران

مقاله پژوهشی

تاریخ دریافت:
۱۴۰۱/۰۶/۰۵

تاریخ پذیرش:
۱۴۰۱/۱۰/۲۱

کلیدواژه‌ها:
شناسایی چهره، کشف تقلب، یادگیری عمیق، الگوی دودویی محلی، نویز متناوب

نویسنده مسئول:
h.hassanpour@shahroodut.ac.ir

چکیده

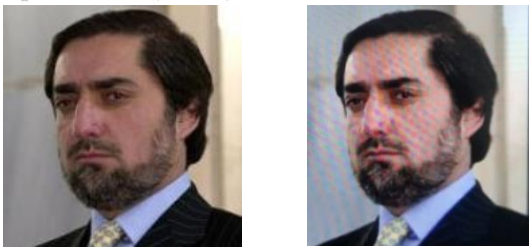
در به‌کارگیری سیستم‌های شناسایی چهره روش‌های مختلف تقلب نظیر استفاده از ماسک پوششی و به‌کارگیری عکس شخص معتبر دو مشکل اساسی هستند که کاربردهای آن‌ها را محدود می‌کنند. براساس بررسی‌های انجام‌شده روش‌هایی برای تشخیص تقلب در شناسایی چهره معرفی‌شده‌اند که در بعضی موارد مداخله کننده هستند، یعنی شخص را وادار به انجام حرکتی می‌کنند تا بتوانند چهره واقعی را از تقلبی تمیز دهند. استفاده از روش‌های مداخله کننده اغلب نارضایتی کاربران را به همراه دارد. در این مقاله با ارائه روشی غیرمداخله کننده و براساس ویژگی‌هایی مانند انعکاس نور یا وجود نویز متناوب اقدام به شناسایی تصاویر واقعی از تقلبی می‌کنیم. در این روش ابتدا با بهره‌گیری از الگوی دودویی محلی لبه‌ها و بافت تصویر برجسته می‌شوند. سپس جهت طبقه‌بندی تصاویر واقعی و غیرواقعی، ویژگی‌های تصویر توسط مدل یادگیری عمیق متشکل از سه لایه پیچش استخراج می‌شوند. نتایج نشان‌دهنده مقاومت روش پیشنهادی در برابر پوشش چشم است. به‌منظور ارزیابی روش پیشنهادی مجموعه داده CASIA در این تحقیق مورداستفاده قرارگرفته است. نتایج حاکی از دقت ۹۸ درصدی روش پیشنهادی در این مجموعه داده است که در مقایسه با روش‌های موجود دقت بالاتری دارد.

doi : 10.22034/ABMIR.2023.18914.1015

۱- مقدمه

وجود سیستم‌های تجاری متنوع جهت شناسایی افراد، به‌عنوان‌مثال استفاده از سیستم‌های شناسایی چهره و اثرانگشت، گواهی بر پیشرفت قابل‌توجه در این زمینه است. باوجوداین دستاوردها، شناسایی چهره همچنان یک مبحث فعال درزمینه بینایی ماشین محسوب می‌شود [۱]. مهم‌ترین مزیت روش شناسایی چهره بین سایر روش‌های شناسایی، دسترسی آسان به تصویر چهره فرد است. عباسپور و همکاران روشی برای شناسایی چهره با دقت بالا ارائه کرده‌اند که قادر است در یک مجموعه داده بزرگ با استفاده از طبقه‌بندی سلسله مراتبی شناسایی را انجام دهد [۲]. نقطه‌ضعف عمده این روش و روش‌های مشابه باوجود راندمان بالا در دقت شناسایی، که استفاده از آن¬ها را برای برخی کاربردها محدود می‌کند، عدم توانایی در تمایز بین چهره زنده با تصویر یا ویدئو غیرزنده انسان است. درواقع تقلب یک مشکل جدی در این روش‌های شناسایی چهره محسوب می‌شود [۳]. روش پیشنهادی در این مقاله عملیات شناسایی تقلب را با استفاده از اختلاف ویژگی‌هایی از قبیل انعکاس نور یا وجود نویز متناوب که میان تصاویر واقعی و تقلبی وجود دارد، به‌وسیله شبکه یادگیری عمیق انجام می‌دهد. نویز متناوب به‌عنوان یکی از انواع نویز در تصویر، بخصوص تصاویری که از روی مانیتور گرفته می‌شوند در اثر تداخل الکتریکی یا مغناطیسی ایجاد می‌شود. این نویز در تصاویر تهیه‌شده در برخی از کاربردهای بصری مانند تلویزیون و نمایشگرهای نوری دیده می‌شود [٤].

نویز متناوب به‌عنوان یک الگوی تکرارشونده روی تصویر خود را نشان می‌دهد. به‌عنوان‌مثال شکل ۱(الف) تصویری را نشان می‌دهد که از روی نمایشگر گرفته‌شده است. این تصویر دارای مقداری نویز متناوب نسبت به چهره همان فرد در شکل ۱ (ب) است، درصورتی‌که در شکل ۱ (ب) که تصویر مستقیم از چهره انسان گرفته‌شده است چنین نویزی مشاهده نمی‌شود.



(الف) (ب)

شکل ۱: (الف) چهره عکس‌برداری شده از روی نمایشگر (حاوی نویز متناوب)، (ب) چهره عکس‌برداری شده از شخص واقعی (عاری از نویز متناوب).

روش‌های دیگری نیز تاکنون برای شناسایی تقلب معرفی‌شده‌اند ازجمله فعالیتی که توسط پن و همکاران انجام‌شده است [۳]. در این مقاله یک روش شناسایی در مقابل جعل تصویر در شناسایی چهره پیشنهادشده است که تشخیص آن از طریق پلک زدن چشمان انسان است. این روش به هیچ سخت‌افزار اضافه‌ای به‌جز یک دوربین عادی نیاز ندارد. در این روش و روش‌هایی که شخص مجبور به انجام فعالیتی جهت شناسایی تقلب است (روش مداخله جویانه)، سیستم جهت تشخیص تقلب معطل می‌ماند. لذا دارای ویژگی‌های مناسبی جهت شناسایی تقلب نیستند.

پژوهش‌هایی که اخیرا برای کشف تقلب در شناسایی چهره انجام شد نظیر [۵]، از تبدیل موجک گسسته (DWT[1]) جهت بررسی ویژگی‌های تصویر در مقیاس‌های متفاوت استفاده می‌کند. اگرچه تبدیل موجک گسسته یک ابزار قدرتمند برای پردازش سیگنال و تصویر است، اما به علت تنوع بالای تقلب‌ها و محدودیت تبدیل موجک گسسته در تشخیص نویز متناوب و لبه در تشخیص تقلب موفق عمل نمی‌کند. بر همین اساس ویژگی مورداستفاده در این روش قادر به تمییز قائل شدن میان چهره واقعی و تقلبی در مجموعه داده CASIA با دقت قابل‌قبول نیست.

در همین راستا پژوهش‌های دیگری نیز نظیر [٦] توسط پینتو و همکاران انجام‌شده است. این روش از ویژگی بازتابش نور جهت تشخیص تقلب بهره می‌برد و در فرایند تشخیص، شخص وادار به انجام حرکتی نمی‌شود بلکه صرفا از طریق تفاوت ویژگی انتشار نور که در تصویر شخص واقعی و تقلبی وجود دارد اقدام به تشخیص جعل می‌کند. این‌گونه روش‌ها که صرفا تمرکز خود را

[1] Discrete Wavelet Transform

بر روی یک ویژگی از تصویر (مانند بازتابش نور) قرار می‌دهند، در شناسایی انواع تقلب موفق عمل نمی‌کنند. روش‌هایی مانند [۳] که برپایه تشخیص از روی چشم کار می‌کنند زمانی دچار مشکل اساسی می‌شوند که شخص چشمان خود را بپوشاند. به‌عنوان‌مثال اگر شخص از عینک آفتابی یا عینک با شیشه‌های رنگی استفاده کند یا در دوران همه‌گیری کوید-۱۹ از ماسک استفاده کند، ممکن است این روش‌ها را دچار خطا کند.

روش دیگر کار زو و همکاران است که از شبکه با حافظه طولانی کوتاه‌مدت (LSTM[1]) استفاده می‌کند [۷]. این نوع شبکه‌ها جهت رفع مشکلات ناپدید شدن گرادیان در شبکه‌های عصبی بازگشتی طراحی‌شده‌اند و دارای محدودیت‌هایی از قبیل زمان آموزش طولانی بوده و نیاز به حافظه زیاد برای آموزش شبکه‌دارند. همین عوامل سبب برتری استفاده از شیوه یادگیری عمیق می‌شود که سرعت آن نیز بالاتر بوده و نیاز به حافظه کمتری دارد.

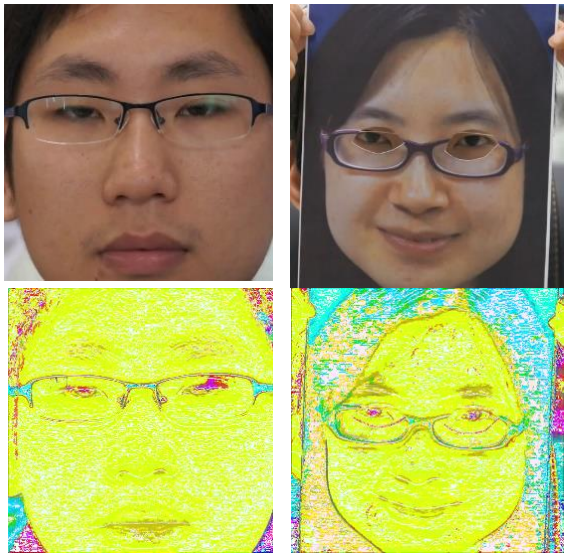
بررسی‌های ما در این تحقیق نشان می‌دهد که روش‌های رایج تقلب شامل پوشیدن ماسک چهره، عکس چاپ‌شده حاوی چهره و بازپخش تصویر چهره از روی نمایشگرهای مختلف است. به‌عنوان نمونه در شکل ۲ (الف) چهره چاپ‌شده بر روی کاغذ و در شکل ۲ (ب) ماسک چهره فردی را می‌توان مشاهده کرد که قسمت چشمان آن برش داده‌شده است. بر این اساس ارائه روشی که بتواند شناسایی لبه‌های ناشی از برش در تصویر (ناحیه چشم) و بافت چهره چاپی در تصویر را انجام دهد به بهبود عملکرد در تشخیص این نوع تقلب‌ها کمک می‌کند. رویکرد پیشنهادی در این مقاله استفاده از روشی برپایه یادگیری عمیق است. به‌علاوه از الگوی دودویی محلی جهت شناسایی بافت و لبه‌ها بهره می‌برد. با توجه به تنوع تقلب در شناسایی چهره، استفاده از الگوی دودویی محلی و برجسته کردن جزئیات فرکانس بالا، شناسایی تقلب را تسهیل می‌کند. در شکل ۳ می‌توان تصاویری از نتیجه اعمال الگوی دودویی محلی را مشاهده نمود.

۲- پیش‌پردازش تصاویر و استخراج ویژگی

با توجه به شکل ۳، به‌کارگیری الگوی دودویی محلی در تصاویر حاوی تقلب ناشی از ماسک سبب برجسته شدن لبه‌ها در اطراف چشم می‌شود. به همین ترتیب، اعمال الگوی دودویی محلی بر روی تصاویر چهره چاپ‌شده روی کاغذ نیز سبب برجسته شدن لبه‌ها و بافت تصویر می‌شود. برجسته‌سازی این نشانه‌ها سبب تسهیل در تشخیص این نوع تقلب‌ها شده و استفاده از الگوی دودویی محلی در شناسایی تقلب را توجیه می‌کند.

(الف) (ب)

شکل ۲: دو نوع از تصاویر حاوی تقلب، (الف) تصویر چاپ‌شده، (ب) ماسک چهره برش داده‌شده در قسمت چشم‌ها

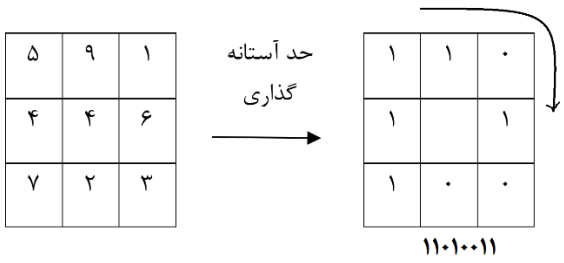


شکل ۳: نتیجه اعمال الگوی دودویی محلی (تصاویر ردیف پایین)، تصاویر سمت راست تقلب با ماسک (وجود حفره اطراف چشم) و سمت چپ چهره واقعی است

ایده اصلی الگوی دودویی محلی مقایسه مقدار خاکستری هر پیکسل مجاور با پیکسل مرکزی است. به‌عنوان‌مثال با در نظر گرفتن

[1] Long Short-Term Memory

شعاع ۱ و تعداد پیکسل‌های ۸ با در نظر گرفتن نقطه مرکزی به‌عنوان نقطه پایه، مقدار خاکستری نقطه مرکزی با مقادیر خاکستری ۸ پیکسل در همسایگی مقایسه می‌شود. اگر مقدار خاکستری پیکسل‌های مجاور بزرگ‌تر از پیکسل مرکزی باشد، مقادیر خاکستری همه پیکسل‌های همسایه روی ۱ تنظیم می‌شود، و اگر مقدار خاکستری پیکسل‌های مجاور کوچک‌تر از پیکسل مرکزی باشد مقادیر خاکستری همه پیکسل‌های همسایه روی صفر تنظیم می‌شوند. سپس می‌توان از این الگو به‌عنوان توصیف‌کننده بافت استفاده کرد. به‌عنوان نمونه در شکل ٤ نحوه انجام این عملیات را می‌توان مشاهده نمود [۸].



شکل ٤: نحوه عملکرد الگوی دودویی محلی

در این مقاله، اندازه پارامترهای شعاع و همسایگی در استفاده از الگوی دودویی محلی به ترتیب ۱ و ۸ در نظر گرفته‌شده است. پس از استخراج الگوی دودویی محلیُ در مرحله بعد این الگو به تصویر اصلی اضافه می‌شود تا هیچ‌کدام از جزئیات تصویر را از دست ندهیم. شکل ۵ مراحل کلی انجام این فرایند را به تصویر می‌کشد. هدف از استفاده الگوی دودویی محلی به دست آوردن سه نوع زیر–الگو است، شامل: لبه‌ها، بافت و نواحی هموار. همان‌گونه که در شکل ۵ مرحله ۳ قابل‌مشاهده است، تصویر حاصل از اعمال الگوی دودویی محلی سبب برجسته شدن لبه‌ها می‌شود. همین عامل سبب بهبود دقت شبکه می‌شود.

در مرحله بعد ادامه عملیات استخراج ویژگی انجام می‌شود. این مرحله ازآنجایی‌که داده ورودی را برای شبکه آماده می‌سازد، اهمیت پیدا می‌کند. پیدا کردن پارامترهای مناسب برای اعمال داده–ها به شبکه بسیار اهمیت دارد. برای اعمال داده‌ها ابتدا تصویر چهره به ابعاد ٦٤×٦٤ تبدیل می‌شود. بررسی‌های ما نشان می‌دهد که این مقدار از تفکیک‌پذیری تصویر برای آموزش شبکه با توجه به داده‌ها کافی است. افزایش ابعاد تصویر نه‌تنها سبب افزایش دقت نمی‌شود بلکه هزینه محاسبات را نیز بالا می‌برد.

در قدم بعدی با عملیات نرمال‌سازی مقدار پیکسل‌های تصاویر در یک بازه‌ی ثابت (به‌عنوان‌مثال بین ۰ تا ۱) قرار می‌گیرند. زیرا برخی از پیکسل‌ها (در اثر اضافه شدن الگوی دودویی محلی) ارزش عددی بالاتری (یا پایین‌تری) نسبت به سایرین دارند. عمل نرمال‌سازی سبب یادگیری بهتر توسط شبکه نیز می‌شود [۹].

۳ – شناسایی تقلب با شبکه عصبی پیچشی

در طراحی شبکه عصبی پیچشی (CNN[1]) از ساختار سیستم بینایی الهام گرفته‌شده است. برخلاف شبکه‌های کاملاً متصل، CNN ها از اتصالات محلی برای استخراج ویژگی‌های فضایی دوبعدی بافت تصاویر استفاده می‌کنند. علاوه بر این، پارامترهای شبکه را می‌توان از طریق مکانیسم ادغام[2] به میزان قابل‌توجهی کاهش داد [۱۰]. این عوامل سبب افزایش دقت در تشخیص می‌شود، به‌علاوه ادغام سبب افزایش سرعت در شناسایی می‌شود. ساختار شبکه به‌کاررفته در این تحقیق در شکل ٦ نشان داده‌شده است.

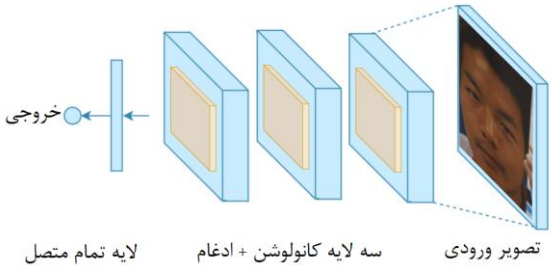
عنوان مرحله	تصویر
تصویر ورودی (مرحله ۱)	
الگوی دودویی محلی (مرحله ۲)	

[1] Convolutional Neural Network

[2] Pooling

جمع تصاویر دو مرحله قبل
(مرحله ۳)

شکل ۵: مراحل اعمال الگوی دودویی محلی بر روی تصویر



شکل ٦: معماری CNN حاوی سه لایه پیچش، ادغام و تمام متصل

به‌طورکلی، لایه‌های پیچش مهم‌ترین بخش CNN ها هستند. در هر لایه پیچش، مکعب (ابعاد تصویر) ورودی با چندین فیلتر قابل یادگیری پیچیده (Convolve) می‌شود و درنتیجه آن چندین نقشه ویژگی ایجاد می‌شود. درروش پیشنهادی از سه لایه با مقدار ۱٦ نقشه ویژگی استفاده‌شده است. بعد از هر لایه پیچش، یک‌لایه ادغام قرار دارد که در ادامه به جزئیات آن می‌پردازیم. لازم به ذکر است که تابع فعال‌ساز استفاده‌شده از نوع ReLU است. استفاده از این تابع دارای دو مزیت است: همگرایی سریع و استحکام در مقابل ناپدید شدن گرادیان [۱۱]. این تکنیک علاوه بااینکه ارتباط میان نمونه‌های پیکسل‌های ورودی را یاد می‌گیرد، در هر نقشه ویژگی باعث کاهش تعداد پارامترها می‌شود.

به دلیل وجود اطلاعات مازاد در نقشه‌های ویژگی، لایه‌های ادغام به‌صورت دوره‌ای پس از چندلایه پیچش در CNN ها استفاده می‌شوند. از طریق عملیات ادغام، اندازه نقشه‌های ویژگی کوچک‌تر شده و نمایش ویژگی‌های استخراج‌شده انتزاعی‌تر می‌شود [۱۰]. درواقع لایه‌های تراکم برای کاهش تعداد پارامترهای موردنیاز برای توصیف لایه‌ها در عمق شبکه توسعه داده‌شده است [۱۲].

پس از عملیات ادغام لایه‌ها، نقشه‌های ویژگی لایه قبلی به‌صورت بردار درآمده و به یک‌لایه کاملاً متصل اعمال می‌شوند. تعداد نرون‌ها در این لایه ۱۲۸ عدد است. از مزیت‌های شبکه عصبی عمیق استخراج ویژگی از داده‌ها به‌صورت خودکار است.

۳-۱ ساختار شبکه عصبی عمیق

مدل مورداستفاده تصاویر رنگی با ابعاد ٦٤×٦٤ را به‌عنوان ورودی دریافت می‌کند. در ادامه سه لایه میانی پیچش قرار داده‌شده است. تعداد نقشه‌های ویژگی در هر سه لایه پیچش برابر ۱٦ در نظر گرفته‌شده است. هدف از این کار این است که ابتدا تمرکز بر روی ویژگی‌های بزرگ‌تر باشد و در ادامه به جزئیات به‌صورت دقیق‌تر توجه شود. پس از هر لایه پیچش از یک‌لایه ادغام با اندازه ۲×۲ استفاده‌شده است که پیشتر جزئیات آن ذکر شد. پس از بررسی‌های انجام‌شده از یک تکنیک دیگر در شبکه‌های یادگیری عمیق به نام حذف (Dropout) استفاده کردیم. این تکنیک توسط سریواستاوا معرفی و گسترش پیدا کرد [۱۳]. این تکنیک در مرحله آموزش پس از هر لایه پیچش ۲۰ درصد از نورون‌ها را از چرخه فعالیت خارج می‌کند. این کار شبکه را مجبور به یادگیری با نرون‌های کمتر کرده و مانع از بیش‌برازش شبکه می‌شود، همچنین در زمان محاسبات نیز صرفه‌جویی می‌شود [۱۲].

٤ – مجموعه داده‌های آزمون

داده استفاده‌شده در این مقاله از مجموعه CASIA است [۱٤]. در جمع‌آوری این مجموعه داده از ۵۰ شخص متفاوت استفاده‌شده است. داده‌ها با سه درجه وضوح پایین، معمولی و زیاد تصویربرداری شده‌اند. در این مجموعه، داده‌های مرتبط با چندین نوع تقلب از قبیل ماسک چهره، تصویر چهره چاپ‌شده، و بازپخش تصویر چهره از روی نمایشگر گنجانده‌شده است. به‌صورت کلی تصاویر واقعی شامل سه درجه کیفی، و تصاویر تقلبی نیز سه نوع تقلب با سه درجه کیفی در این مجموعه داده وجود دارد. پس از استخراج تصاویر از ویدیو تعداد ۳۸۰۰ تصویر از حالات مختلف چهره تهیه‌شده است. در روش‌هایی که از یادگیری عمیق استفاده می‌شود به داده‌های زیاد جهت آموزش نیاز دارند. تکنیک‌های

افزایش داده[1] با ایجاد تغییراتی در تصاویر و افزودن آن به مجموعه تصاویر آموزشی باعث افزایش داده‌ها می‌شود که سبب آموزش بهتر شبکه و همچنین عدم رخ دادن بیش برازش در آن می‌شود. این تغییرات شامل عناصری از قبیل برش تصویر، بزرگنمایی و ایجاد وارونگی جانبی به همراه کشیدن لبه‌ها در تصویر است. در شکل ۷ سه نمونه از تصاویر ایجادشده با تکنیک افزایش داده را می‌توان مشاهده کرد. پس از به‌کارگیری تکنیک افزایش داده تعداد تصاویر آموزش به ۷۵۰۰ عدد رسید.

شکل ۷: تصویر اصلی و سه نمونه از تصاویر ایجادشده با تکنیک افزایش داده

برای انجام آزمایش بیشتر با اضافه کردن ۲۰ نفر به مجموعه داده CASIA، مجموعه داده دیگری به‌نام CASIA+شامل ۷۰ نفر تهیه شد. جزئیات آن بدین شرح است که از چهره هر شخص با چشمان پوشیده شده توسط عینک آفتابی تصاویری تهیه شد. تصاویر تقلب نیز با عکس‌برداری از بازپخش تصاویر اصلی در صفحه نمایشگرهایی باکیفیت‌های مختلف انجام پذیرفت. چند نمونه از این تصاویر واقعی و تقلبی را می‌توان در شکل ۸ مشاهده کرد.

شکل ۸: نمونه‌هایی از تصاویر موجود در مجموعه داده CASIA+ (تصاویر سمت راست از چهره شخص (تصاویر واقعی) و تصاویر سمت چپ از روی بازپخش تصاویر (تصاویر تقلبی) تهیه‌شده‌اند.)

۵ – نتایج آزمایش‌ها

سیستم مورداستفاده جهت آموزش و آزمون دارای مشخصات پردازنده Core i5 و مقدار حافظه موقت ۸ گیگابایت بوده است. مدت‌زمان آموزش حدود ۳۰ دقیقه به طول انجامید و شبکه آموزش‌دیده شده به‌صورت برخط قابلیت تشخیص تقلب را دارا است.

داده‌های موجود از مجموعه‌های CASIA و CASIA+ را به دو مجموعه آموزش و آزمون تقسیم کردیم. تقسیم‌بندی به‌صورت تصادفی و به‌گونه‌ای انجام گرفت که هیچ‌گونه اشتراکی میان مجموعه آموزش و آزمون وجود نداشته باشد. به‌بیان‌دیگر داده‌های آزمون در آموزش‌دیده نشده‌اند. نسبت تقسیم داده‌ها در مجموعه آموزش و آزمون به ترتیب برابر ٪۷۰ و ٪۳۰ کل داده‌ها است. نسبت تقسیم داده‌های آزمون در کلاس‌های تقلب و واقعی به ترتیب با نسبت ٪۶۰ و ٪۴۰ است. در مرحله آموزش داده‌ها را به تعداد ۱۵ مرتبه به الگوریتم اعمال کرده و به‌دقت ۹۹ درصد در داده‌های اعتبار سنجی[2] رسیدیم و آموزش در این مرحله متوقف شد. داده‌های اعتبار سنجی نیز به میزان ٪۱۵ درصد از مجموعه داده‌های آموزش به‌صورت تصادفی انتخاب‌شده است.

در داده‌های آزمون نیز پس از استخراج الگوی دودویی محلی و اضافه کردن آن به تصویر اصلی، داده‌ها را به شبکه آموزش داده‌شده اعمال کردیم و درنهایت دقت کلی ۹۸ درصد حاصل شد. با انجام آزمایش دیگری بر روی مجموعه داده CASIA+، دقت ۹۷ درصد را دریافت کردیم. جزئیات مربوط به تعاریف موجود در ماتریس درهم‌ریختگی در جدول ۱ ذکرشده است.

[1] Data Augmentation Techniques

[2] Validation

با توجه (۱) به جدول ارزیابی روش پیشنهادی شامل صحت (Precision) و فراخوانی (Recall) می‌رسیم (جدول ۲).

$$\text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN} \quad (۱)$$

جدول ۱: ماتریس درهم‌ریختگی.

	شخص واقعی	**تصویر تقلبی**
تصویر واقعی را تقلبی تشخیص دهد (FP)	شخص واقعی را به‌درستی کشف کند (TP)	شخص اصیل (واقعی)
تصویر تقلبی را به‌درستی کشف کند (TN)	شخص تقلبی را واقعی تشخیص دهد (FN)	تصاویر تقلبی (واقعی)

جدول ۲: میزان صحت و دقت روش پیشنهادی بر روی مجموعه داده CASIA و +CASIA.

	Precision		Recall		F-measure		**دقت**	
	CASIA	**CASIA+**	**CASIA**	**CASIA+**	**CASIA**	**CASIA+**	CASIA	**CASIA+**
تصویر واقعی	۹۷	۹۵	۹۸	۹۷	۹۷	۹۶	**۹۸**	۹۷
تصویر تقلبی	۹۸	۹۸	۹۹	۹۷	۹۹	۹۸		

جدول ۳ عملکرد روش پیشنهادی را با روش‌های موجود بر روی یک مجموعه داده مشترک، براساس نتایج ارایه شده در مقالات، مقایسه می‌کند. با توجه به نتایج این جدول نکته قابل‌توجه درروش [۵] است که در مجموعه داده CASIA به علت تنوع زیاد تقلب‌ها دقت پایین‌تری به نسبت روش‌های [۷]، [۱۵] و روش پیشنهادی در این مقاله دریافت کرده است. نمونه‌هایی از داده‌ها که توسط روش پیشنهادی به‌اشتباه تشخیص داده‌شده‌اند را می‌توان در شکل ۹ مشاهده نمود.

جدول ۳: مقایسه میزان دقت با سایر روش‌ها در مجموعه داده CASIA.

روش	*دقت*
Boulkenafet et al [16].	۹۳
LSTM-CNN [7]	۹۴/۸
Yang et al [15].	۹۵
Zhang et al [5].	۹۴/۵
روش پیشنهادی	۹۸

شکل ۹: چند نمونه تشخیص اشتباه درروش پیشنهادی (تصاویر تقلبی از نوع چاپ‌شده است.)

یکی از علت‌های خطا در تشخیص صحیح این تصاویر توسط روش پیشنهادی پایین بودن کیفیت آن‌ها است. این نتایج نشان می‌دهد که یک عملیات پیش‌پردازش مناسب دیگر ممکن است سبب بهبود در تشخیص شود که در کارهای آینده انجام خواهد شد.

۶- نتیجه‌گیری

در این مقاله یک روش برای تشخیص تقلب در شناسایی چهره براساس شبکه عصبی عمیق و با برجسته‌سازی ویژگی‌های تصویر

توسط الگوی دودویی محلی پیشنهاد داده‌شده است. براساس بررسی‌های انجام‌شده روش‌هایی که مداخله کننده هستند و بر اساس تکان دادن سر و پلک زدن چشم عملیات شناسایی تقلب را انجام می‌دهند، با پوشیده شدن قسمتی از صورت و چشم دچار مشکل اساسی شده و دقت در عملکرد خود را از دست می‌دهند. با توجه به اهمیتی که شناسایی چهره در سیستم‌های تجاری مختلف دارد، تشخیص تقلب براساس روش غیرمداخله‌کننده ضروری است. نتایج آزمایش‌های ارائه‌شده در این مقاله نشان می‌دهد که روش پیشنهادی قادر به تشخیص تقلب باوجود پوشش چشم است.

References

[1] Ahonen, Timo, Abdenour Hadid, and Matti Pietikäinen. (2004, May) "Face Recognition with Local Binary Patterns." European conference on computer vision.

[2] Abbaspoor, N., & Hassanpour, H. (2022). "Face Recognition in A Large Dataset Using a Hierarchical Classifier". Multimedia Tools and Applications, 1-19.

[3] Pan, Gang. "Eyeblink-based Anti-spoofing In Face Recognition from A Generic Webcamera." IEEE 11th international conference on computer vision. IEEE, 2007.

[4] Alibabaie, N., & Latif, A. (2021). Adaptive Periodic Noise Reduction in Digital Images Using Fuzzy Transform. Journal of Mathematical Imaging and Vision, 63(4), 503-527.

[5] Zhang, Wanling, and Shijun Xiang. "Face Anti-spoofing Detection Based On DWT-LBP-DCT Features." Signal Processing: Image Communication 89 (2020): 115990.

[6] Pinto, A., Schwartz, W. R., Pedrini, H., & de Rezende Rocha, A. (2015). Using Visual Rhythms for Detecting Video-based Facial Spoof Attacks. IEEE Transactions on Information Forensics and Security, 10(5), 1025-1038.

[7] Z. Xu, S. Li, and W. Deng. Learning Temporal Features Using LSTM-CNN Architecture for Face Anti-spoofing. In Pattern Recognition (ACPR), 3rd IAPR Asian Conference on, pages 141–145. IEEE, 2015. 2, 7, 8

[8] Tang, J., Su, Q., Su, B., Fong, S., Cao, W., & Gong, X. (2020). Parallel Ensemble Learning of Convolutional Neural Networks and Local Binary Patterns for Face Recognition. Computer Methods and Programs in Biomedicine, 197, 105622.

[9] Patro, S., and Kishore Kumar Sahu. "Normalization: A Preprocessing Stage. 1503.06462 (2015).

[10] Li, Shutao, et al. "Deep Learning for Hyperspectral Image Classification: An overview." IEEE Transactions on Geoscience and Remote Sensing 57.9 (2019): 6690-6709.

[11] A.Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Ne ural Networks,” in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.

[12] Mathew, Amitha, P. Amudha, and S. Sivakumari. "Deep Learning Techniques: An Overview." International Conference on Advanced Machine Learning Technologies and Applications. Springer, Singapore, 2020.

[13] Nitish Srivastava, Geoffrey Hinton, and Alex and Krizhevsky. Dropout: A Simple Way to Prevent Neural Networks from Overftting. Journal of Machine Learning Research, 15:1929-1958, 2014.

[14] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Li, “A Face Antispoofing Database with Diverse Attacks,” in IAPR Intl. Conference on Biometrics, 29 2012-april 1, pp. 26 –31

[15] J. Yang, Z. Lei, and S. Z. Li. "Learn Convolutional Neural Network for Face Anti-spoofing". CoRR, abs/1408.5601, 2014. 2, 7, 8

[16] Z. Boulkenafet, J. Komulainen, and A. Hadid. Face Antispoofing Based on Color Texture Analysis. In Image Processing (ICIP), IEEE International Conference on, pages 2636–2640. IEEE, 2015.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 1, No. 1, Spring & Summer 2023, 14-21

A Non-intrusive and Cover Resistant Method for Detecting Forgery in Face Recognition Using Deep Learning

Seyed Ebrahim Hosseini [1], Hamid Hassanpour [2*]

[1] Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.

[2] Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.

Article Information

Original Research Paper

Received:
17 September 2022

Accepted:
11 January 2023

Keywords:
face recognition, forgery detection, deep learning, local binary pattern, periodic noise.

Corresponding Author*:
h.hassanpour@shahroodut.ac.ir

Abstract

In the use of face recognition systems, various fraud, such as the use of a mask and a photo of a genuine person, are two major problems that limit their applications. Studies have shown a number of methods for detecting fraud in face recognition, which are sometimes intrusive, enforcing the person to make a move in order to distinguish the real face from the fake one. The use of intrusive methods often leads to user dissatisfaction. In this article, we present a non-intrusive method using features such as light reflection or the presence of periodic noise to distinguish real images from the fake one. In this method, the edges and texture of the image are highlighted by a local binary pattern to better detect fraud. Then, by extracting the image feature using a deep learning technique with three layers of convolution, it will be able to distinguish between real and fake face images. This method is resistant to covering the eyes and face. In order to evaluate the proposed method, the CASIA dataset was used in this research. The results show 98% accuracy of the proposed method on this dataset. Among the existing methods, we see an increase in accuracy.

: 10.22034/ABMIR.2023.18914.1015