



دانشکده علوم و فنون نوین

کروه بین رشته ای فناوری (بنجش علوم و فناوری شبکه)

گزارش مرحله اول پروژه
احراز هویت غیرحضوری متقارضیان خدمات الکترونیک
انتظامی بر مبنای سنجه های بیومتریکی

توسط:

هادی ویسی

فهرست

٤	خلاصه اجرایی
٧	فصل ۱ مبانی زیست‌سنگی و احراز هویت غیرحضوری
٧	۱-۱ احراز هویت غیرحضوری
٨	۱-۲ سامانه‌های زیست‌سنگی
۱۰	۱-۲-۱ عملیات اصلی در یک سامانه‌ی زیست‌سنگی
۱۴	۱-۲-۲ تشخیص زنده بودن و اهمیت آن
۱۵	۱-۲-۳ ارزیابی سامانه‌های زیست‌سنگی و تشخیص زنده بودن
۲۱	فصل ۲ بررسی سامانه‌های بازشناسی چهره و رویکردهای آن
۲۳	۲-۱ مراحل اصلی در سیستم بازشناسی چهره
۲۴	۲-۲ روش‌های دریافت چهره
۲۴	۲-۲-۱ تصاویر دو بعدی
۲۹	۲-۲-۲ تصاویر سه بعدی
۳۲	۲-۳ بازشناسی چهره دو بعدی
۳۸	۲-۳-۲ تشخیص هویت به کمک چهره
۴۷	۲-۴ بازشناسی چهره سه بعدی
۴۷	۴-۱ پیش‌پردازش
۴۸	۴-۲ رویکردهای سنتی بازشناسی چهره سه بعدی
۵۳	۴-۳ رویکردهای مبتنی بر یادگیری عمیق در بازشناسی چهره سه بعدی
۵۶	۴-۴ جمع‌بندی
۵۷	۵-۱ تحلیل و مقایسه رویکردهای بازشناسی چهره
۶۰	۵-۲ بررسی سامانه تشخیص زنده بودن و رویکردهای آن
۶۱	۵-۳-۱ ساختار کلی سامانه تشخیص زنده بودن
۶۳	۵-۳-۲ تشخیص زنده بودن در سامانه بازشناسی چهره
۶۴	۵-۳-۳ پایگاه داده‌های رایج تشخیص زنده بودن برای چهره
۶۶	۵-۳-۴ رویکردهای تشخیص زنده بودن در سامانه بازشناسی چهره
۷۰	۵-۴-۱ رویکردهای غیر تعاملی
۷۵	۵-۴-۲ رویکردهای تعاملی
۷۶	۵-۴-۳ رویکردهای مبتنی بر ترکیب
۷۹	۵-۴-۴ رویکرد شبکه عصبی عمیق در تشخیص زنده بودن

۸۱	۳-۵ جمع‌بندی
۸۳	۳-۵-۱ تحلیل و مقایسه روبکردهای تشخیص زنده بودن
۸۴	فصل ۴ تحلیل نیازهای پروژه
۸۵	۴-۱ دفاتر پلیس + ۱۰+
۸۶	۴-۱-۱ فهرست خدمات قابل انجام در دفاتر پلیس + ۱۰
۸۶	۴-۱-۲ اپلیکیشن پلیس + ۱۰
۸۷	۴-۲-۱ اپلیکیشن پلیس من
۸۸	۴-۲-۲ امکانات بخش راهور در اپلیکیشن پلیس من
۸۹	۴-۲-۳ امکانات بخش گذرنامه در اپلیکیشن پلیس من
۹۰	۴-۳ راهور ۱۲۰
۹۰	۴-۴ ضرورت احراز هویت الکترونیکی
۹۲	۴-۵ بیان مزايا و فواید عملیاتی احراز هویت غیرحضوری
۹۲	۴-۵-۱ منافع برای ارائه دهنده خدمات (پلیس)
۹۳	۴-۵-۲ منافع برای دریافت کننده خدمات (مردم)
۹۳	۴-۵-۳ تحلیل راهبردی طرح احراز هویت غیرحضوری
۹۴	۴-۵-۴ نگرانی‌های رایج در احراز هویت الکترونیکی
۹۵	۴-۶ جمع‌بندی
۹۶	فصل ۵ راهکار پیشنهادی پروژه
۹۷	۵-۱ راه حل شماره یک: استفاده از رابطهای برنامه نویسی API و SDK
۱۰۰	۵-۲ راه حل شماره دو: استفاده از درگاه احراز هویت
۱۰۳	مراجع

خلاصه اجرایی

با گسترش علم و فناوری بسیاری از خدمات نهادها و سازمان به صورت الکترونیکی و هوشمند در بستر اینترنت ارائه می‌شود. پلیس نیز به عنوان یک نهاد ارائه‌دهنده خدمات به مردم و سایر نهادهای مهم در کشور، به دنبال هوشمندسازی خدمات خود به ویژه برای مردم بوده و در همین راستا سامانه‌های الکترونیکی و هوشمند مختلفی را ارائه کرده است. با توجه به مطالعات صورت گرفته که در بخش تحلیل نیازها به طور مفصل آورده شده است، به دلیل عدم احراز هویت غیرحضوری کاربران، بسیاری از خدماتی که در حضور دارند، همچنان فقط با مراجعه‌ی حضوری به مراکز مرتبط مانند دفاتر پلیس $10+$ ارائه می‌شوند. محدودیت بودجه و تجهیزات برای پاسخگویی حضوری به تعداد زیادی از مراجعه‌کنندگان، محدودیت نیروهای پلیس و تمرکز آن‌ها بر روی موضوعات مهم، ارائه خدمات به نیروهای انسانی خود پلیس و کنترل آن‌ها در رسته‌های مختلف این نهاد، محدودیت تعداد دفاتر خدماتی در شهرستان‌ها و عدم دسترسی روستاها به این دفاتر، ازدحام مراجعه به دفاتر در ایام خاصی از سال، پیدایش نیازهای جدید و بیشترین سازگاری برای رفع این نیازها از نظر اقتصادی و تجهیزاتی، و افزایش تقاضای عمومی برای استفاده از خدمات دیجیتال و غیرحضوری سبب شده است تا نیاز به احراز هویت الکترونیکی بسیار مورد توجه قرار بگیرد. با توجه به موارد بیان شده، ضرورت ارائه خدمات غیرحضوری با رشد روزافزون خدمات برخط و افزایش تقاضای مردم برای آن، به ویژه در شرایطی مانند بحران بیماری کرونا، موضوعی بدیهی است که همه سازمان‌ها و نهادهای ارائه‌دهنده خدمات را به سمت بهره‌گیری از آن سوق داده است و مورد تأکید نهادهای بالادستی کشور شامل قانون‌گذاران و سیاست‌گذاران است. خلاصه ضرورت‌های انجام این طرح عبارتنداز:

- تقاضای روزافزون ارائه خدمات الکترونیکی و غیرحضوری از سمت مردم و لزوم بهبود تجربه مشتری^۱ به دلیل سادگی و سرعت کار
- تأکید نهادهای قانون‌گذار بر ارائه خدمات الکترونیکی و غیرحضوری به مردم به ویژه با تشدید موضوع در شرایط بحران کرونا

¹ User Experience

• نیاز به کاهش مراجعات حضوری افراد (از نظر سلامتی، ترافیک)

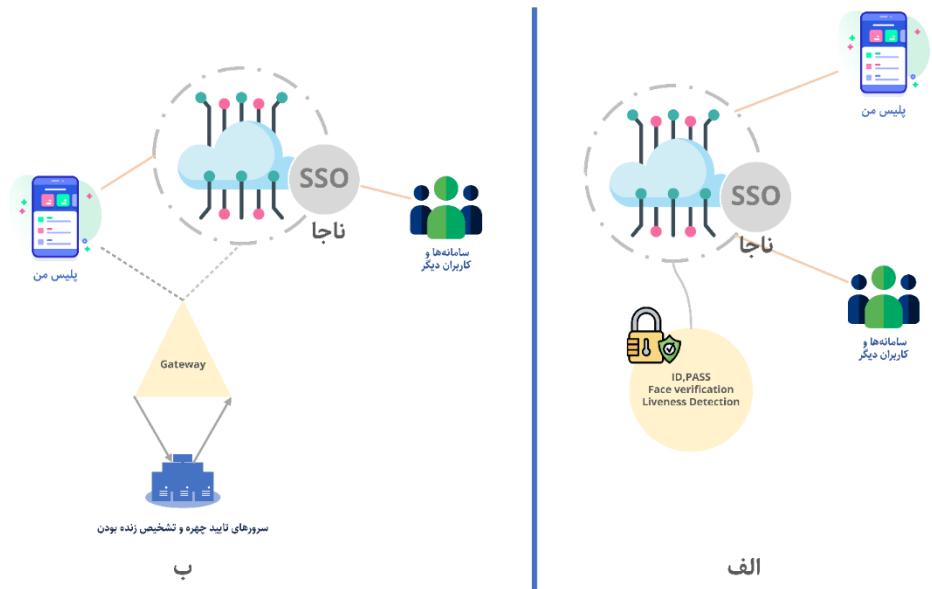
• ضرورت افزایش امنیت و اشراف اطلاعاتی پلیس با تکمیل پایگاههای داده افراد به ویژه در تکمیل اطلاعات زیست‌سنگی، تهیه زیرساخت‌های استفاده از آن‌ها، یکپارچه‌سازی و پیگیری (شفافیت)

• لروم کاهش خطاهای انسانی و سواستفاده افراد از اطلاعات و اسناد

احراز هویت الکترونیکی (از راه دور) به دلیل امنیت و قابل اعتماد بودن و همچنین سازگاری با استانداردهای جهانی، به عنوان یک راه حل مناسب برای پاسخگویی به نیاز بیان شده، ارائه شده است. به دلیل گسترش بیش از پیش تجهیزات دریافت تصویر در جهان و رشد فناوری دوربین‌های دیجیتال در ساختار گوشی‌های هوشمند امروزی و رایانه‌ها و همچنین افزایش استفاده از فناوری زیست‌سنگی چهره در بازارهای تجاری و نیز وجود الگوریتم‌های قابل اطمینان در این حوزه، روش احراز هویت الکترونیکی از طریق تایید چهره به طور روزافزونی توسعه یافته است. در راستای استفاده از فناوری بازشناسی چهره برای احراز هویت، مسائلی مانند تشخیص زنده بودن که بیانگر واقعی بودن تصویر ارسال شده و به معنی اینکه مربوط به یک فرد زنده است، به وجود می‌آید. تشخیص زنده بودن به این معناست که تصویر و یا ویدیوی ارسال شده واقعی باشد و یک مصنوع تقلیبی مانند نمایش یک عکس چهره‌ی چاپ شده و یا نمایش از روی صفحه‌ی تلفن همراه و یا ماسک نباشد. در همین راستا، در این فاز از پروژه، روش‌های مختلف بازشناسی چهره و تشخیص زنده بودن در ابتدا بررسی شده است و در فصل‌های دوم و سوم تلاش شده است تا به روزترین روش‌ها و نتایج آن‌ها بر روی مجموعه داده‌های رایج در این دو حوزه که جزو تعهدات اصلی پروژه هستند، بررسی و ارائه گردد.

بعد از بررسی روش‌های بازشناسی چهره و تشخیص زنده بودن و معرفی روش‌های این دو فناوری، در فصل‌های بعد تحلیل نیاز کارفرما در نحوه استفاده از احراز هویت غیرحضوری انجام شده و روش‌های پیشنهادی پاسخ به این نیازها آورده شده است. همان‌طور که پیش‌تر نیز به آن اشاره شد، نیروی انتظامی نیز با بهره‌گیری از تحول دیجیتال، سعی در ارتقای خدمات خود و بهبود شرایط شهروندان، کسب‌وکارها و دولت داشته و دارد. در راستای تحقق این هدف، بسیاری از خدمات خود را به صورت الکترونیکی و از طریق اپلیکیشن و سامانه‌های وبی مختلف در اختیار مردم قرار داده است و برای توسعه‌ی هرچه بیشتر و برداشتن گام‌های کلان در این راه، نیاز به رفع چالش احراز هویت از راه دور دارد.

برای پاسخ به این نیاز نیروی انتظامی، بعد از همفکری با کارشناسان فنی سازمان و بررسی وضع موجود سامانه‌های مرتبط با موضوع، به ویژه زیرساخت فعلی SSO سازمان، دو راهکار ارائه شده است: ۱- اتصال به سامانه‌ی تایید هویت چهره و تشخیص زنده بودن از طریق SDK و API (شکل ۲ (الف)) و ۲- ایجاد یک درگاه برای احراز هویت مبتنی بر چهره و اتصال به سرورهای ارائه‌دهنده‌ی این خدمات از طریق فراخوانی این درگاه (شکل ۲ (ب)). در روش اول، تایید چهره و تشخیص زنده بودن به عنوان یک روش احراز هویت در SSO خود ناجا قرار می‌گیرد و کلیه عملیات سمت مشتری (کلاینت) توسط تیم فنی مرتبط با سامانه‌های بهره‌بردار انجام می‌شود اما روشی منعطف برای استفاده و انجام تغییرات لازم است. از طرف دیگر، در روش دوم، سامانه‌های بهره‌بردار به یک درگاه که ارائه‌دهنده خدمات احراز هویت مبتنی بر چهره است، متصل شده و کلیه فرایند آن مانند دریافت داده‌های لازم از کاربران برای احراز هویت و پیاده‌سازی رابط کاربری آن توسط درگاه تامین می‌شود و پس از پردازش تایید نهایی به SSO ناجا برای ادامه‌ی فرایند ارسال می‌شود. این روش از نظر فنی هزینه پیاده‌سازی پایین‌تری دارد اما انعطاف‌کمتری هم ارائه می‌دهد. جزئیات این دو روش و نمودار ترتیب آنها در فصل پایانی تشریح شده است و ضروری است کارفرما یکی از این دو روش را برای پیاده‌سازی در ادامه مسیر انتخاب و اعلام کند.



شکل ۱-۰۰ راه حل های ارائه شده برای اتصال سامانه هی احراز هویت غیر حضوری به سایر سامانه های موجود در سازمان

فصل ۱ مبانی زیست‌سنگی و احراز هویت غیرحضوری

با توسعه روزافزون فضای مجازی و رشد خدمات برخط (آنلاین)، گرایش عمومی مردم نیز به عدم مراجعه حضوری و انجام کارها از راه دور، روز به روز بیشتر می‌شود. این کار می‌تواند توسط خود افراد و در منزل یا محل کار آن‌ها و به صورت برخط صورت پذیرد. خدماتی که در حوزه‌های حساسی مانند امور بانکی انجام می‌شود و روز به روز بر تعداد خدمات غیرحضوری و مجازی مالی و بانکی افزوده می‌شود و بسیاری از کارها که قبلاً توسط کارمندان بانک انجام می‌شد، امروزه توسط خود مشتریان انجام می‌شود. این موضوع، علاوه بر کاهش هزینه‌های مختلف سازمانی برای ارائه دهنده خدمت، آسانی و راحتی بیشتری را برای گیرنده خدمت هم فراهم می‌کند، به ویژه در شرایطی مانند بحران کرونا (Covid 19) که گرایش به عدم حضور و تجمع افراد در مکان‌های سرپوشیده مانند دفاتر خدمات الکترونیکی در حال افزایش است.

۱-۱ احراز هویت غیرحضوری

برای عملیاتی کردن خدمات غیرحضوری، یکی از اصلی‌ترین چالش‌های پیش‌رو، موضوع امنیت و اعتبارسنگی هویت مشتریان^۱ (KYC) است. این مساله در رویکرد سنتی، با مراجعه حضوری افراد به دفاتر و پیشخوان‌های پلیس حل می‌شود، این در حالی است که برای ارائه خدمات مجازی باید از احراز هویت الکترونیکی^۲ (eKYC) بهره گرفت.

به طور خلاصه eKYC عبارتی است که برای توصیف دیجیتالی و الکترونیکی شدن فرایندهای KYC استفاده می‌شود. eKYC (مشتری خود را الکترونیکی بشناسید) فراینده‌ای از راه دور و بدون کاغذ است که هزینه‌ها و بوروکراسی سنتی مورد نیاز در فرآیندهای KYC را به حداقل می‌رساند. فرآیند eKYC دیجیتالی و از راه دور شدن فرآیند KYC سنتی است. شناسایی و تأیید هویت مشتری

¹ Know Your Customer (KYC)

² electronic Know Your Customer (eKYC)

در زمان واقعی و بلافاصله اتفاق می‌افتد و به همین دلیل فرآیندهای KYC^۱ رضایت مشتری را افزایش می‌دهند. در همین راستا برای اطمینان از اینکه فرایندهای KYC^۲ دارای استانداردهای ایمنی شناسایی هستند، راه حل‌های به کار رفته باید فرایندهای شناسایی الکترونیکی را با سطح بالایی از ایمنی و قابلیت اطمینان و مطابق با قوانین تعیین شده پیاده‌سازی کنند.

طیف گسترده‌ای از راه حل‌های KYC^۳ مبتنی بر هوش مصنوعی و یادگیری ماشین ارائه شده است که سامانه‌های زیست‌سنجدی و مفاهیم مربوط به آن از جمله‌ی مهم‌ترین این راه حل‌ها می‌باشد.

برای ارائه خدمات غیرحضوری و از راه دور به افراد جامعه، لازم است امنیت ارائه خدمات به ویژه احراز هویت افراد با اطمینان مطلوب تامین شود که برای این کار از ویژگی‌های زیست‌سنجدی چهره آن‌ها به عنوان معیار شناسایی استفاده می‌شود. بدین صورت که فرد متقارضی با بیان یکی از شناسه‌های هویتی خود مانند کد ملی، شماره گواهینامه یا گذرنامه و همچنین ارائه تصویری از خود به صورت برخط درخواست تایید هویت^۴ می‌کند. علاوه بر تایید هویت مبتنی بر چهره، موضوع مهم دیگر این طرح، تشخیص زنده بودن^۵ است که در آن زنده بودن ویدئویی دریافتی بررسی می‌شود. در ادامه به بررسی این حوزه و علوم مربوط به احراز هویت از راه دور از جمله مفاهیم زیست‌سنجدی به طور خلاصه و رویکردهای سامانه‌های بازشناسی چهره و تشخیص زنده بودن پرداخته می‌شود.

۱-۲ سامانه‌های زیست‌سنجدی

از گذشته انسان‌ها به منظور شناسایی دیگر انسان‌ها از ویژگی‌های ظاهری یا رفتاری یکدیگر (از جمله چهره، صدا و نحوه‌ی راه رفتن) استفاده می‌کنند. چندین دهه است که استفاده از این ویژگی‌ها برای تشخیص هویت انسان‌ها در پرونده‌های جنایی نیز مورد استفاده قرار گرفته است. با گذشت زمان در سال‌های اخیر تشخیص هویت به وسیله‌ی ویژگی‌های زیستی به علت مزایایی که دارد کاربردهای بسیار گسترده‌تری پیدا کرده به طوری که امروزه علاوه بر محیط‌های تجاری و صنعتی در کاربردهای عمومی‌تر، از جمله کنترل دسترسی به گوشی‌های تلفن همراه و رایانه‌ها نیز مورد استفاده قرار می‌گیرد.

زیست‌سنجدی^۶ علم شناسایی افراد از طریق مشخصات انسانی او مانند اثر انگشت^۷، کف دست^۸، چهره^۹، امضاء^{۱۰}، دست خط^{۱۱}، عنبیه^{۱۲}، شبکیه^{۱۳} و صدا^{۱۴} است. در علم زیست‌سنجدی، اعضایی از بدن مورد توجه قرار گرفته که استفاده از آن‌ها راحت‌تر و کم ضرر‌تر باشد. هر کدام از روش‌های مورد استفاده دارای نقاط ضعف و قدرتی هستند که با ترکیب آن‌ها با دیگر روش‌های امنیتی می‌توان ضعف‌های موجود را از بین برد. زیست‌سنجدی به روش‌های خودکار تشخیص یا تأیید هویت یک شخص زنده از طریق اندازه‌گیری مشخصه‌های فیزیولوژیکی یا رفتاری وی اطلاق می‌شود. بدین ترتیب زیست‌سنجدی یک مجموعه فناوری محسوب می‌گردد. واژه زیست‌سنجدی می‌تواند به صورت اسم به کار رود که در این صورت اشاره به یک فناوری منفرد و خاص دارد: «اسکن اثر انگشت متداول‌ترین زیست‌سنجدی است». این واژه را می‌توان به صورت صفت نیز به کار برد و بدین ترتیب بین «مشخصه‌های زیست‌سنجدی» مانند الگوی عنبیه و یا «سامانه‌های زیست‌سنجدی» مانند سامانه‌های عنبیه‌نگاری و سایر مشخصه‌های انسانی و سامانه‌های موجود تمایز قائل شد.

¹ Verification

² Liveness

³ Biometric

⁴ Fingerprint

⁵ Palmprint

⁶ Face

⁷ Signature

⁸ Handwriting

⁹ Iris

¹⁰ Retina

¹¹ Voice

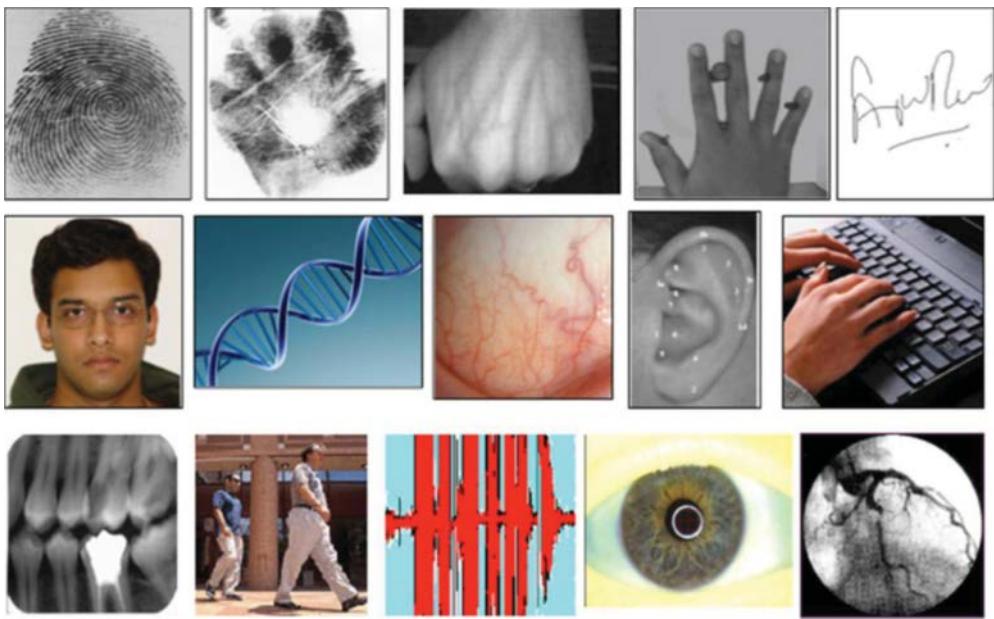
سامانه‌های زیست‌سنگی از ویژگی‌های فیزیولوژیکی، رفتاری و یا روانی یک فرد برای شناسایی وی استفاده می‌کنند. این روش‌های شناسایی از روش‌های سنتی ایمن‌تر هستند و امکان جعل در آن‌ها کاهش یافته است. همین دلایل موجب شده است که این روش‌ها در حوزه‌های امنیتی، حقوقی و تجارت الکترونیکی اهمیت پیدا کند. در جدول ۱-۱ به تفکیک نوع، چند نمونه از ویژگی‌هایی که می‌توان برای زیست‌سنگی مورد استفاده قرار داد ملاحظه می‌شود.

جدول ۱-۱ انواع ویژگی‌های زیست‌سنگی

روانی	رفتاری	فیزیولوژیکی
کارکرد مغز	گفتار	چهره
شناختی	امضا	اثر انگشت
	دست خط	عنبیه‌ی چشم
	ریتم تایپ	خطوط کف دست
	راه رفتن	گوش
	ضریبان قلب	رگ
	الگوی تنفسی	رد پا
	امواج مغزی	شبکیه‌ی چشم
	نوار قلب	DNA

از میان ویژگی‌های یاد شده ویژگی‌های روانی هنوز در مراحل اولیه‌ی توسعه قرار دارند و هنوز به پایداری خوبی نرسیده‌اند و ممکن است چندان دقیق عمل نکند (تا حدی نیز برای ویژگی‌های رفتاری). همچنین نمونه‌گیری ویژگی‌های روانی نیز در حال حاضر بسیار دشوارتر از سایر ویژگی‌ها است. اما ویژگی‌های فیزیولوژیکی در حال حاضر تا حدی توسعه یافته‌اند که می‌توان آن‌ها را برای جلوگیری از جعل و یا برای کاربردهای با سطح امنیت بالا مورد استفاده قرار داد [1].

در شکل ۱-۱ نمونه‌هایی از صفات زیست‌سنگی مشاهده می‌شود. (الف) اثر انگشت، کف دست، عروق دست، شکل دست و امضا. (ب) چهره، ناحیه چشم، شکل گوش و الگوهای تایپ کردن (ضریبه زدن به کلید). (ج) دندان (ادنتولوژی پزشکی قانونی)، راه رفتن، صدا یا گفتار، عنبیه و شبکیه. برخی از این صفات، به عنوان مثال، اثر انگشت، کف دست، صورت، صدا، دندان، شکل گوش و DNA نیز در پزشکی قانونی استفاده می‌شود [2].



شکل ۱-۱ تصویر زیست‌سنجهای رایج در دنیا [2]

۱-۲-۱ عملیات اصلی در یک سامانه‌ی زیست‌سنجهی

عملیات اصلی که در سیستم‌های زیست‌سنجهی صورت می‌گیرد را می‌توان در دو بخش ثبت‌نام^۱ و تطبیق/ مقایسه^۲ قرار داد. هر فرد که از سیستم استفاده می‌کند ابتدا لازم است که در سیستم ثبت‌نام کند. برای این کار، مشخصات زیستی فرد با استفاده از حسگر دریافت می‌شود و با استخراج ویژگی‌های مورد نیاز از آن، الگویی برای آن فرد تولید شده و در پایگاه داده ذخیره می‌شود (شکل ۲-۱).

در مرحله‌ی شناسایی (تطبیق/ مقایسه)، سیستم با دریافت مشخصات زیستی جدید از حسگر، ویژگی‌های مورد نیاز آن را استخراج کرده و با تشکیل الگوی مربوط به آن، الگو را با یک یا چند الگو ذخیره شده در پایگاه داده مقایسه می‌کند و در نهایت پاسخ مورد نظر را به عنوان خروجی بازمی‌گرداند. در شکل ۲-۱ شما کلی عملیات مقایسه ملاحظه می‌شود.

در هر یک از کاربردهای پایه‌ای زیست‌سنجهی، مقایسه به شکلی خاص مورد استفاده قرار می‌گیرد. در «تایید هویت^۳» الگوی مربوط به یک فرد (که کاربر ادعا کرده) از پایگاه داده استخراج می‌شود و الگوی دریافتی تنها با آن مقایسه می‌شود (یک به یک). اما در کاربرد «تعیین هویت^۴» الگوی دریافتی با کل پایگاه داده مقایسه می‌شود تا هویت فرد مورد نظر در صورت ثبت‌نام در سیستم- مشخص شود (یک به N).

حسگر مورد استفاده در این مراحل با توجه به نیاز سیستم و این که چه خصوصیاتی از فرد نیاز است که جمع‌آوری شود، انتخاب می‌شود و یا طراحی می‌گردد. به عنوان مثال معمولاً برای کاربرد اسکن عنبه‌ی چشم دوربین‌های NIR^۵ و برای تشخیص چهره معمولاً دوربین‌های دیجیتال که طول موج مرئی را دریافت می‌کنند استفاده می‌شوند. همچنین برای دریافت صدا نیز معمولاً میکروفون مورد استفاده قرار می‌گیرد.

¹ Enrollment

² Identification

³ Matching

⁴ verification

⁵ Near Infrared

ماژول دریافت داده^۱ بخشی است که وظیفه‌ی تبدیل اطلاعات دریافت شده از حسگر (مانند مقادیر ولتاژ، جریان، دما) به اطلاعاتی که آمده‌ی پردازش باشند را بر عهده دارد. در اکثر حسگرهای مورد استفاده این بخش نیز قرار گرفته است. این ماژول در کاربردهای قانونی نیز که حسگری وجود ندارد نیز می‌تواند مورد استفاده قرار بگیرد.

در بخش پیش‌پردازش، داده‌ها با کاهش نویز و حذف بخش‌های اضافی برای استخراج ویژگی آمده می‌شوند. به عنوان مثال در کاربرد تشخیص صدا در این بخش با دریافت صدای ضبط شده، صدای انسان را جدا کرده و صدای‌های پس‌زمینه و بخش‌های سکوت را حذف می‌کند. پس از آن با نرمال کردن صدا تاثیر عواملی مانند شدت و.. را کاهش می‌دهد. معمولاً کیفیت پیش‌پردازش داده‌ها در دقت سیستم تاثیر زیادی می‌گذارد.

در بخش استخراج ویژگی‌ها، با استفاده از مدل‌های ریاضیاتی و یا یادگیری ماشین^۲ الگوهایی را از داده‌ها استخراج می‌کند. تعداد ویژگی‌های قابل استخراج به صفات مورد زیست‌سنجدی و مدل به کار رفته برای استخراج ویژگی وابسته است. به عنوان مثال اگر ضربان قلب زیست‌سنجدی مورد استفاده باشد، با توجه به این که ویژگی‌های استخراج شده از آن زیاد نیست، دقت تشخیص افراد به کمک آن نیز نسبتاً پایین است. در مقابل در صورتی که از اثر انگشت برای شناسایی افراد استفاده شود، چون تعداد ویژگی‌های قابل استخراج از آن زیادتر است به طبع می‌توان بررسی‌های بیشتری بر روی آن‌ها انجام داد. دقت سسیتم نهایی وابستگی بسیار زیادی به ویژگی‌های استخراج شده توسط مدل دارد.

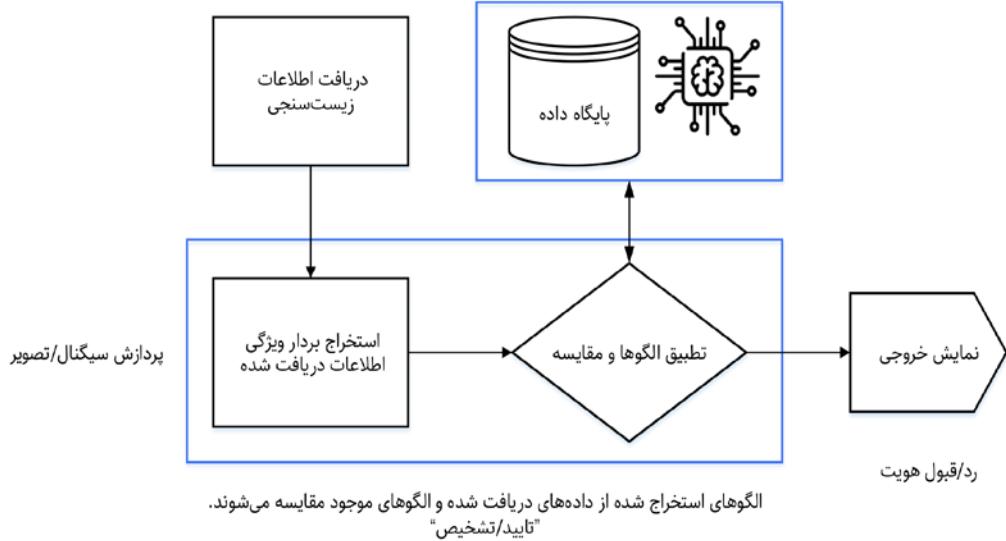
در بخش ساخت الگو با دریافت ویژگی‌های استخراج شده، الگویی از اطلاعات زیستی دریافت شده ساخته می‌شود. هدف از این بخش حذف ویژگی‌های غیر ضروری و تکراری و در نهایت افزایش سرعت و دقت بخش مقایسه می‌باشد. الگوی ساخته شده در هنگام ثبت‌نام برای استفاده‌های بعدی در پایگاه داده ذخیره می‌شود. در برخی کاربردها الگو قبل از ذخیره سازی در پایگاه داده، رمزنگاری نیز می‌شود.

در فرآیند مقایسه، الگوی داده‌ای که به تازگی از حسگر دریافت شده است، با الگوهای ذخیره شده در پایگاه داده مقایسه می‌شود تا مقدار شباهت یا عدم شباهت آن به داده‌های پایگاه داده بررسی شود. مدل‌های بسیاری را می‌توان برای بررسی شباهت دو الگو در این بخش به کار برد که بسته به نوع الگو و ویژگی‌های مورد استفاده می‌توان درست را انتخاب کرد. در این بخش با مقایسه‌ی هریک از دو الگو می‌توان امتیازی به میزان شباهت در الگو به یکدیگر اختصاص داد.

در بخش خروجی با دریافت امتیازهای کسب شده از بخش مقایسه در مورد برابر بودن الگو با الگو یا الگوهایی از پایگاه داده تصمیم‌گیری در ساده‌ترین حالات می‌تواند با اعمال یک آستانه بر روی امتیاز حاصل از مقایسه انجام شود و یا می‌تواند به صورت نسبی صورت بگیرد [3].

¹ Data Acquisition

² Machine Learning



شکل ۱-۲ ساختار یک سیستم زیست-سنجی

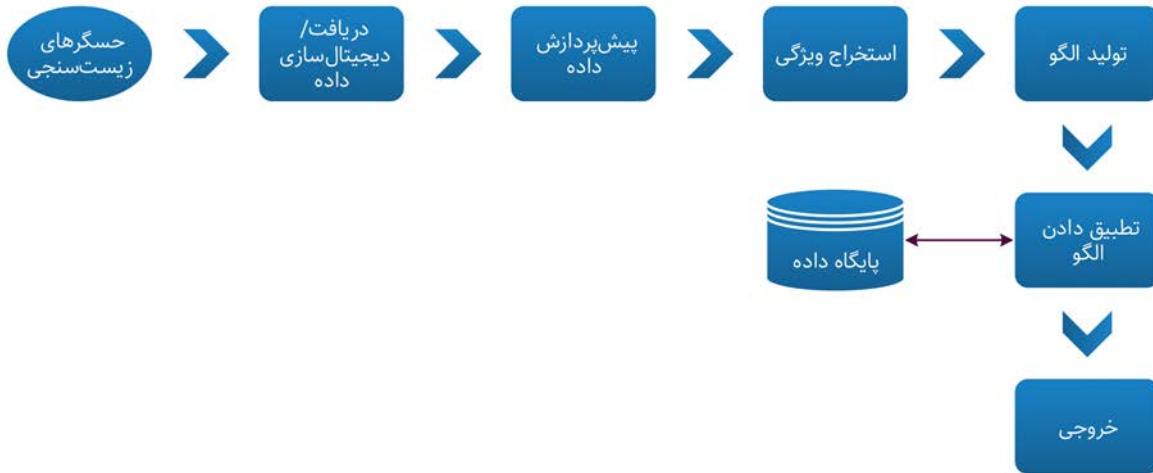
عملیات اصلی که در سیستم‌های زیست-سنجی صورت می‌گیرد را می‌توان در دو بخش ثبت‌نام و تطبیق/مقایسه قرار داد. هر فرد که از سیستم استفاده می‌کند ابتدا نیاز است که در سیستم ثبت‌نام کند. برای این کار، مشخصات زیستی فرد با استفاده از حسگر دریافت می‌شود و با استخراج ویژگی‌های مورد نیاز از آن، الگوبی برای آن فرد تولید شده و در پایگاه داده ذخیره می‌شود (شکل ۱-۳).



شکل ۱-۳ مراحل عملیات ثبت‌نام در سیستم‌های زیست-سنجی

در مرحله‌ی شناسایی (تطبیق/مقایسه)، سیستم با دریافت مشخصات زیستی جدید از حسگر، ویژگی‌های مورد نیاز آن را استخراج کرده و با تشکیل الگوی مربوط به آن، الگو را با یک یا چند الگو ذخیره شده در پایگاه داده مقایسه می‌کند و در نهایت پاسخ مورد نظر را به عنوان خروجی بازمی‌گرداند. در شکل ۴-۱ کلی عملیات مقایسه ملاحظه می‌شود.

ساختار تمامی سیستم‌های زیست-سنجی دارای یک معماری کلی می‌باشد. دریافت داده‌ها، پردازش سیگنال، تطبیق، تصمیم‌گیری و ذخیره‌سازی، زیرسیستم‌های یکسان در یک سیستم زیست-سنجی می‌باشند.



شکل ۴-۱ مراحل عملیات مقایسه در سیستم‌های زیرسنجی

در زیرسیستم دریافت داده، داده‌های خام یک فرد که توسط یک حسگر ویژه اسکن شده است، وارد سیستم می‌شود. فرایندی که در این زیرسیستم انجام می‌شود:

۱. دریافت داده‌ها توسط حسگر
 ۲. تبدیل داده‌های (سیگنال‌ها) دریافتی از حسگرها به فرم مناسب جهت ارسال به زیر سیستم پردازش سیگنال
- عملیات زیرسیستم پیش‌پردازش به شرح ذیل می‌باشد:
۱. دریافت داده‌های خام از زیر سیستم جمع‌آوری داده
 ۲. عملیات فیلترینگ جهت حذف نویز
 ۳. اصلاح داده‌ها
 ۴. تبدیل داده‌های دریافتی به فرم لازم (تولید الگو) برای زیر سیستم تطبیق

لازم به ذکر است که از داده‌های دریافت شده در این زیرسیستم، پس از پردازش، یک الگو از برخی ویژگی‌های موجود تولید و ذخیره می‌شود. در واقع این الگوی تولید شده است که مورد مقایسه و شناسایی قرار می‌گیرد. ماهیت این الگو که از روی یک استاندارد ثابت از پیش تعریف شده تولید می‌شود، با استفاده از مدل‌های ریاضیاتی و یا یادگیری ماشین الگوهایی را از داده‌ها استخراج می‌کند.

خروجی زیرسیستم تطبیق از مقایسه دو الگو بدست می‌آید. فرایند این زیر سیستم شامل:

۱. دریافت داده‌های پردازش شده (الگو) از زیر سیستم قبل
۲. دریافت الگوهای ذخیره شده
۳. مقایسه الگوی تولید شده در زیر سیستم قبل، با الگوهای موجود

زیر سیستم تصمیم‌گیری پس از اجرای زیر سیستم قبل، فرآخوانی می‌شود که وظیفه آن تصمیم‌گیری بر روی تطابق انجام شده‌ی متناسب با درخواست است. در این مرحله یک حد یا آستانه در نظر گرفته شده است. اگر امتیاز بیشتر یا برابر این آستانه باشد، کاربر تأیید می‌شود در غیر اینصورت کاربر پذیرفته نمی‌شود.

زیر سیستم ذخیره‌سازی شامل الگوهایی است که در هنگام ثبت‌نام از کاربران به دست آمده است. ممکن است برای هر کاربر یک یا چند الگو ذخیره شده باشد [3].

۱-۲-۲- تشخیص زنده بودن^۱ و اهمیت آن

امروزه فناوری زیست‌سنگی به سرعت در حال گسترش است و جزء مهمی از زندگی روزمره‌ی انسان‌ها شده است. هدف یک سامانه زیست‌سنگی شناخت خودکار افراد بر اساس ویژگی فیزیولوژیکی و یا رفتاری آن‌ها مانند اثر انگشت، چهره، عنایی و امضا است. با توجه به ماهیت پردازش و ضبط خودکار ویژگی‌ها در این سامانه، میزان نظارت انسان به حداقل می‌رسد و اجزای سامانه باید ضبط داده‌های زیست‌سنگی را بدون نظارت، کنترل کنند. در میان این سامانه‌ها، زیست‌سنگی چهره به دلیل استفاده گسترده در کاربردهای متنوع کنترل دسترسی فیزیکی و منطقی نقش برجسته‌ای دارد. علاوه بر این، بازشناسی چهره^۲ با مزایایی مانند ضبط اطلاعات غیرمستقیم و حسگرهای کم‌هزینه همراه است. یک مزیت اساسی که بازشناسی چهره نسبت به سایر زیست‌سنگی‌ها دارد: راحتی می‌باشد. این واقعیت که فناوری بازشناسی چهره می‌تواند به طور خودکار کاربر را از فاصله دور شناسایی کند، سبب برتری این سیستم از نظر راحتی و رضایتمندی بیشتر کاربر شده است و برای نمونه از زمان شیوع ویروس کرونا در جهان، کاربرد این سیستم به دلیل عدم تماس مستقیم با افراد فزونی یافته است و نیز احراز هویت الکترونیکی از راه دور به وسیله‌ی بازشناسی چهره، امری است که اکنون در بسیاری از خدمات بانکی گسترش یافته است. استفاده گسترده از سامانه‌های بازشناسی چهره، نگرانی‌های جدیدی را به ویژه در مورد آسیب‌پذیری سامانه‌ی ضبط داده ایجاد کرده است. با این حال، یکی از اصلی‌ترین موانع پیش روی سامانه‌های شناسایی زیست‌سنگی، هویت متقلبات است که از نظر مفهومی، حمله جعل^۳ نامیده می‌شود[۴]. حمله جعل، شرایطی است که در آن یک شخص یا برنامه، خود را به عنوان یکی دیگر معرفی کرده و با جعل داده‌ها و استفاده از یک مصنوع جعلی^۴ مانند تصویر و یا ویدیوی گرفته شده از شخص هدف و یا استفاده از ماسک سه‌بعدی، سعی در جعل هویت فرد، برای دسترسی به سامانه دارد. تشخیص زنده بودن^۵ یکی از حوزه‌های اصلی مورد توجه در زمینه‌ی زیست‌سنگی چهره است که شامل یک فرایند بررسی است که آیا زیست‌سنگی گرفته شده توسط سامانه‌ی تشخیص، واقعی است (یعنی زنده است) یا توسط مهاجمان تقلید شده است. این موارد، آسیب‌پذیری سامانه‌های بازشناسی چهره را در دنیای واقعی نشان می‌دهند. به دلیل اهمیت تشخیص زنده بودن چهره در بازار تجاری و تحقیقات رو به رشد این حوزه در سال‌های اخیر، سبب شده تا سازمان بین‌المللی استانداردسازی مختص با این حوزه استانداردی با عنوان تشخیص «حمله نمایش زیست‌سنگی»^۶ (PA) تعریف کند که به معنی نمایش (ارائه‌ی) زیست‌سنگی به یک زیرسامانه ضبط آن با هدف تداخل در عملکرد سامانه می‌باشد. تشخیص زنده بودن و تشخیص حمله جعل می‌توانند هم‌معنی یا زیرمجموعه «تشخیص حملات نمایش»^۷ (PAD) قرار بگیرند[۵].

در طی دهه‌های گذشته، پیشرفت‌های بی‌شماری در فناوری وجود داشته است که به ارائه امکانات جدید به مردم در قالب دستگاه‌ها و خدمات جدید کمک کرده است. به لطف این پیشرفت سریع در فناوری، به ویژه در علوم کامپیوتر و الکترونیک، امکان استقرار گسترده‌ی سامانه‌های زیست‌سنگی فراهم شده است. امروزه، این فناوری در فرایندهای زیادی مانند کنترل تردد، نظارت، احراز هویت در کاربردهای مختلف مانند تلفن‌های هوشمند، پزشکی قانونی و خدمات برخط مانند آموزش الکترونیکی و تجارت الکترونیکی وجود دارند. چهره، دومین زیست‌سنگی است که از نظر سه‌میه بازار بلافاصله پس از اثر انگشت قرار دارد و هر روزه تولید‌کنندگان بیشتری از بازشناسی چهره در محصولات خود مانند تلفن‌های همراه هوشمند استفاده می‌کنند. بازار جهانی شناسایی چهره نشان‌دهنده‌ی محبوبیت و قابل پذیرش بودن سامانه‌های تشخیص چهره برای کاربردهای مختلف هم در سازمان‌های دولتی و هم بخش خصوصی است. قابل به ذکر است که چهره در بیشتر اسناد شناسایی مانند گذرنامه‌ی سازگار با ایکائو یا کارت‌های

¹ Liveness detection

² Face Recognition

³Spoofing Attack

⁴ Spoof

⁵ Liveness Detection

⁶ Presentation Attack (PA)

⁷ Presentation Attack Detection (PAD)

شناسایی ملی پذیرفته شده است. از این رو تکامل سریع سامانه‌های بازشناسی چهره و تبدیل آن‌ها به برنامه‌های زمان واقعی، نگرانی‌های بسیاری را در مورد توانایی آن‌ها در مقاومت در برابر حملات نمایش، ایجاد کرده است.

اطلاعات و تصاویر ویدیویی از نحوه ایجاد مصنوعهای مختلف چهره در صفحات وب ارائه شده است. می‌توان تصویری از چهره فرد مورد نظر را با جستجو در شبکه‌های اجتماعی یا گرفتن تصویر چهره وی از مسافت طولانی، به راحتی دریافت کرد. این تصاویر مرجع بدون اطلاع قربانی هدف از حمله، کپی یا ضبط می‌شوند. سپس می‌توان از چنین تصاویری برای ایجاد مصنوعات چهره استفاده کرد تا سامانه بازشناسی چهره را فریب دهد. این عوامل محققان مختلف را بر آن داشته است تا به چالش‌های تشخیص حمله برای سامانه بازشناسی چهره بپردازند.

۱-۲-۳- ارزیابی سامانه‌های زیست‌سنجدی و تشخیص زنده بودن

یک سیستم زیست‌سنجدی ایده‌آل سیستمی است که با دریافت هر داده زیستی در مورد انطباق آن با داده‌های موجود در پایگاه داده، تصمیم درستی را اتخاذ کند. یک سیستم زیست‌سنجدی را می‌توان به عنوان یک سیستم تشخیص/بازشناسی الگو^۱ دید که ناگزیر امکان دارد تصمیمات نادرستی بگیرد. در ادامه انواع مختلف خطاهایی که ممکن است در یک سیستم زیست‌سنجدی رخ دهد را بررسی خواهیم کرد. برای دست پیدا کردن به یک دید جامع‌تر نسبت به رفتار خطاهای در سیستم‌های زیست‌سنجدی مطالعه ISO/IEC 19795 پیشنهاد می‌شود [۱].

در ادامه این بخش ابتدا به بررسی دلایل ایجاد خطا می‌پردازیم و در سپس به خطاهای ممکن در هریک از بخش‌های یک سیستم زیست‌سنجدی پرداخته می‌شود و معیارها و اصطلاحات مورد استفاده در حوزه‌ی زیست‌سنجدی را مورد بررسی قرار می‌دهیم.

۱-۲-۳-۱- انواع خطا و دلایل ایجاد آن

سه دلیل عمده در سیستم‌های زیست‌سنجدی می‌تواند باعث ایجاد خطا شود:

محدودیت اطلاعات: اطلاعات بدست آمده از برخی از ویژگی‌های زیستی ممکن است شامل پارامترهای کمی باشد. به عنوان مثال اسکن ضربان قلب پارامترهای زیادی در اختیار مان قرار نمی‌دهد. به همین دلیل این اسکن در بهترین حالت نیز در مقابل اثر انگشت اطلاعات زیادی برای شناسایی فرد در اختیار ما نمی‌گذارد. همچنین محدودیت اطلاعات می‌تواند به علت استفاده‌ی ناصحیح از حسگر نیز اتفاق بیافتد. حتی یک مقایسه کننده‌ی ایده‌آل نیز زمانی که اطلاعات دریافت شده با اطلاعات ثبت شده به هنگام ثبت‌نام همپوشانی نداشته باشد، نخواهد توانست دو الگو را طبق دهد.

محدودیت سازنده‌ی الگو: یک سازنده‌ی الگوی ایده‌آل به گونه‌ای است که تمامی داده‌های دریافت شده از سمت حسگر را به شکلی در الگو ذخیره کند. اما در عمل سازنده‌های الگو نمی‌توانند تمام داده‌های دریافت شده از حسگر را در الگو جای دهند و در طول این عمل ناگزیر بخشی از ویژگی‌ها حذف می‌شوند یا اطلاعاتی به اشتباه در الگو ذخیره می‌شوند. همین امر می‌تواند موجب به وجود آمدن خطا در سیستم شود.

محدودیت تغییرناپذیری: در نهایت قرار است الگوهای مربوط به یک فرد در هنگام مقایسه، مشابه یکدیگر؛ و الگوهای متعلق به افراد مختلف، متفاوت تشخیص داده شوند. یک مقایسه کننده‌ی ایده‌آل لازم است به درستی ارتباط بین دو الگویی که به یک نمونه تعلق دارند را (با وجود تفاوت شرایط اسکن) تشخیص دهد. این بار نیز در عمل یک مقایسه کننده نمی‌تواند ارتباط بین دو الگو را به طور کامل مدل کند (به عنوان مثال به علت کافی نبودن داده‌های آموزشی و یا به وجود آمدن داده‌های پیش‌بنیی نشده در هنگام تست) و در نتیجه الگوهای مختلف از یک نمونه ممکن است به درستی تطبیق داده نشوند.

¹ Pattern Recognition

۱-۲-۳-۲- مازول دریافت از حسگر، استخراج ویژگی

در یک سیستم خودکار زیست‌سنجدی ممکن است در هنگام جمع‌آوری اطلاعات، به دو نوع خطا برخورد کند: خطا عدم تشخیص^۱ و خطا عدم دریافت اطلاعات^۲. در ادامه اگر داده‌ی دریافت شده از حسگر دارای کیفیت پایینی باشد و امکان پردازش آن و استخراج ویژگی‌های کافی وجود نداشته باشد، در بخش استخراج ویژگی به خطا پردازش^۳ بر می‌خوریم.

۱-۲-۳-۳- مازول ساخت الگو

این مازول یک یا چند مجموعه از ویژگی‌های استخراج شده را دریافت می‌کند و یک الگو برای ویژگی مورد نظر برای فرد تولید می‌کند. در حین تولید الگو ممکن است به علت کم بودن تعداد ویژگی‌های دریافت شده، امکان ساخت الگو وجود نداشته باشد. در این صورت خطا در ثبت‌نام^۴ خواهیم داشت. معمولاً یک ارتباط معکوس بین این خطا و دقت سیستم برقرار است. با کاهش سختگیری برای تولید الگو، امکان ساخت الگوهایی با نویز بیشتر وجود خواهد داشت و موجب پایین آمدن دقت در بخش تطابق و مقایسه می‌شود.

۱-۲-۳-۴- مازول تطبیق

این مازول با تطابق دو الگو مقداری بین صفر تا یک به میزان شباهت آن‌ها اختصاص می‌دهد. سپس مازول تصمیم‌گیری با اعمال حداقل امتیاز لازم برای تطابق، در مورد مطابق بودن الگوها تصمیم‌گیری می‌کند. لازم به ذکر است که این مازول دقیقاً تطابق دو الگو را مورد بررسی قرار می‌دهد. و خطاها ممکن در این بخش عبارتند از: تطابق اشتباه^۵ برای حالتی که دو الگو که متعلق به دو فرد متفاوت هستند را مطابق تشخیص دهد؛ و عدم تطابق اشتباه^۶ حالتی که دو الگو که متعلق به یک فرد هستند را تطابق ندهند.

لازم به ذکر است خطاها یاد شده برای این مازول را با خطاها «پذیرش اشتباه^۷» و «عدم پذیرش اشتباه^۸» نباید یکی فرض کرد. پذیرش اشتباه و عدم پذیرش اشتباه به کاربردهایی مانند «تایید هویت» و «تعیین هویت» مرتبط هستند. اما با توجه به به کارگیری مازول تطبیق در این دو کاربرد، خطاها تطابق و عدم تطابق بر روی خطاها پذیرش و عدم پذیرش تاثیرگذار هستند. به عنوان مثال در کاربرد «تایید هویت» که ادعای کاربر مبنی بر داشتن هویتی مشخص مورد بررسی قرار می‌گیرد، با رخدادن خطا تطابق اشتباه بالا رفتن میزان خطا پذیرش اشتباه، و با رخدادن خطا عدم تطابق اشتباه بالا رفتن میزان خطا عدم پذیرش اشتباه را خواهیم داشت. علاوه بر این یک سیستم زیست‌سنجدی ممکن است برای پذیرش یا رد در کاربردهای یاد شده علاوه بر امتیاز تطابق از معیارهای دیگری نیز استفاده کند. به همین دلیل می‌توان گفت میزان خطا تطابق یا عدم تطابق در یک سیستم مستقل از کاربرد بوده و به همین علت معیار مناسب‌تری برای مقایسه به شمار می‌روند.

در کاربرد «تعیین هویت» مازول تطابق به صورت یک به چند عمل می‌کند که در ساده‌ترین فرم می‌توان آن را به صورت چند مقایسه‌ی یک به یک در نظر گرفت. در صورتی که دستگاه تنها افراد ثبت‌نام شده در سیستم قابل استفاده باشد، اصطلاحاً آن را «تعیین هویت مجموعه‌ی بسته^۹» می‌نامند که همواره تعدادی از افراد ثبت‌نام شده را به عنوان شبیه‌ترین الگوها باز می‌گرداند. این نوع خاص از تعیین هویت به ندرت در عمل مورد استفاده قرار می‌گیرند. اما در مقابل نوعی دیگر از تعیین هویت که در آن امکان

¹ Failure to Detect (FTD)

² Failure to Capture (FTC)

³ Failure to Process (FTP)

⁴ Failure to Enroll (FTE)

⁵ False Match

⁶ False Non-Match

⁷ False Accept

⁸ False Reject

⁹ Closed Set Identification

استفاده از سیستم توسط افرادی که در سیستم ثبت نام نکرده‌اند نیز وجود دارد، «تعیین هویت مجموعه‌ی باز^۱» گفته می‌شود. در این نوع تعیین هویت ممکن است کاربر با هیچ‌یک از نمونه‌های ثبت‌نام شده مطابقت نداشته باشد. از این پس در این مقاله منظور از «تعیین هویت» همان «تعیین هویت مجموعه‌ی باز» خواهد بود.

۱-۲-۵ نرخ خطای تایید هویت

در بخش گذشته خطاهای ممکن برای مازول تطابق را تعریف کردیم. این خطاهای برای مقایسه‌های یک به یک تعریف می‌شوند و از آنجایی که تایید هویت در سیستم‌های زیست‌سنگی نیز یک تطابق یک به یک است، خطای تطابق و عدم پذیرش را همان خطای تطابق در نظر می‌گیریم. در صورتی که الگوی ذخیره شده در پایگاه داده از یک فرد را T بنامیم و الگوی دریافت شده از حسگر را I بنامیم، یکی از دو فرضیه‌ی زیر را خواهیم داشت:

- H_0 : الگوهای T و I مربوط به دو هویت مجزا باشند.
- H_1 : الگوهای T و I مربوط به یک هویت واحد باشند.

و در نهایت تصمیم گیری سیستم تایید هویت، یکی از دو نتیجه‌ی زیر را تولید خواهد کرد.

- D_0 : عدم تطابق
- D_1 : تطابق

سیستم تایید هویت با استفاده از (I, s) (همان مقدار امتیازی که به تشابه بین I و T داده شده است) یکی از دو نتیجه‌ی فوق را باز می‌گرداند. اگر امتیاز داده شده از مقدار آستانه تعریف شده برای سیستم کمتر باشد، D_0 و در غیر این صورت D_1 به عنوان خروجی باز گردانده می‌شود. با توجه به معیارهای تعریف شده در بالا برای سیستم تایید هویت دو نوع خطای می‌توان انتظار داشت:

- نوع یک: تطابق اشتباہ (تصمیم گیری D_1 در حالی که H_0 برقرار باشد)
- نوع دو: عدم تطابق اشتباہ (تصمیم گیری D_0 در حالی که H_1 برقرار باشد)

معیار نرخ تطابق اشتباہ^۲ (FMR) احتمال وقوع خطای نوع یک و نرخ عدم تطابق اشتباہ^۳ (FNMR) احتمال وقوع خطای نوع دوم می‌باشد:

$$FMR = P(D_1|H_0)$$

$$FNMR = P(D_0|H_1)$$

برای اندازه‌گیری دقت تایید هویت یک سیستم زیست‌سنگی، به امتیازهای تعداد زیادی مقایسه نیاز داریم به شرطی که (الف) الگوی دریافتی و الگوی ثبت شده در پایگاه داده متعلق به یک هویت باشد (مقادیر $(s|H_1)$ p) (ب) الگوی دریافتی و الگوی ثبت شده در پایگاه داده متعلق به یک هویت نباشد (مقادیر $(s|H_0)$ p). در صنعت به توزیع امتیازها با شرط الف، توزیع اصل^۴ و به توزیع امتیازها با شرط ب، توزیع تقلب^۵ گفته می‌شود. در شکل ۱-۵ نمونه‌ای کلی از نمودار این دو توزیع ملاحظه می‌شود. در این شکل مقادیر FMR و FNMR برای آستانه‌ی t نیز نمایش داده شده‌است. مقادیر FMR و FNMR را می‌توان از روی این توزیع‌ها به کمک فرمول زیر محاسبه کرد (لازم به ذکر است فرض می‌شود که نمونه‌هایی با امتیازی بیشتر از t توسط الگوریتم پذیرفته می‌شوند و مقادیر کمتر از آن رد می‌شوند).

$$FNMR = \int_0^t P(s|H_1)ds$$

$$FMR = \int_t^1 p(s|H_0)ds$$

¹ Open Set Identification

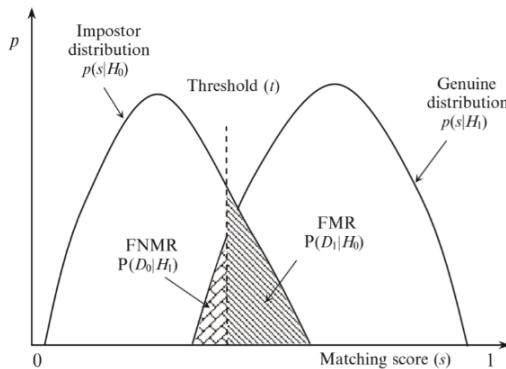
² False Match Rate (FMR)

³ False Non-Match Rate (FNMR)

⁴ Genuine Distribution

⁵ Impostor Distribution

تقریبا در همه سیستم‌های زیست‌سنگی یک تقابل بین مقادیر FNMR و FMR وجود دارد. از آنجایی که هردوی این مقادیر وابسته به مقدار t می‌باشند از این پس آن‌ها را به صورت $FMR(t)$ و $FNMR(t)$ نمایش می‌دهیم. با کاهش مقدار آستانه t در واقع مقدار حساسیت سیستم را کاهش دادیم و موجب می‌شود مقدار $FMR(t)$ افزایش پیدا کند. در مقابل برای بالا بردن ایمنی سیستم مقدار آستانه t را می‌توان بالا برد و با این کار مقدار $FNMR(t)$ افزایش پیدا خواهد کرد.



شکل ۱-۵ مقادیر FNMR و FMR برای مقدار آستانه t که بر روی نمودارهای توزیع اصل و توزیع تقلب نمایش داده شده است

طراح یک سیستم زیست‌سنگی که دستگاه را برای کاربردهای مختلف طراحی می‌کند، بهتر است که مقدار آستانه را قابل تنظیم قرار دهد و مقادیر نرخ خطای را برای مقادیر آستانه مختلط گزارش دهد. این گزارش عموما با استفاده از نمودارهای ROC¹ و DET² صورت می‌گیرد. نمودارهای ذکر شده این امکان را به کاربر می‌دهند که مستقل از مقدار آستانه، تقابل میان FMR و FNMR را برای یک سیستم زیست‌سنگی مورد بررسی قرار دهد. نمودار ROC از رسم $FMR(t)$ بر حسب $(1-FNMR(t))$ برای مقادیر مختلف آستانه t و همچنین نمودار DET به طور مشابه از رسم $FNMR(t)$ بر حسب $FMR(t)$ حاصل می‌شود و نتیجه‌هی واضح‌تری از مقایسه‌ی مستقیم دو خطای در اختیارمان قرار می‌دهد.

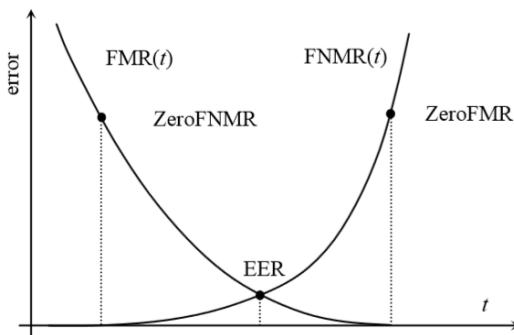
همچنین اصطلاحات دیگری نیز در سیستم‌های تایید هویت مورد استفاده قرار می‌گیرند که عبارتند از (شکل ۱-۶):

- مقدار EER³: مقدار خطایی که در آن مقدار FMR و FNMR برابر می‌شوند ($FMR(t)=FNMR(t)$) که t مقدار آستانه متناظر است. در عمل به علت گستره بودن نمودارهای توزیع (به علت محدود بودن نمونه‌های مورد بررسی و کوانتیزه کردن مقادیر امتیاز خروجی) نمی‌توان یک مقدار دقیق برای EER گزارش کرد. در این حالت به جای گزارش کردن یک عدد، بازه‌ای که در آن مقادیر FNMR و FMR برابر هستند گزارش می‌شود. گرچه EER معیاری مهم برای سیستم زیست‌سنگی به حساب می‌آید، اما به ندرت برای تنظیم مقدار آستانه به آن توجه می‌شود و عموماً آستانه سختگیرانه‌تر و با توجه به مقدار FMR مورد انتظار تنظیم می‌شود.
- مقدار ZeroFNMR : مقدار کمترین FMR که در آن هیچ خطای عدم تطابقی نخواهیم داشت.
- مقدار ZeroFMR : مقدار کمترین FNMR که در آن هیچ خطای تطابقی نخواهیم داشت.

¹ Receiver Operating Characteristic (ROC)

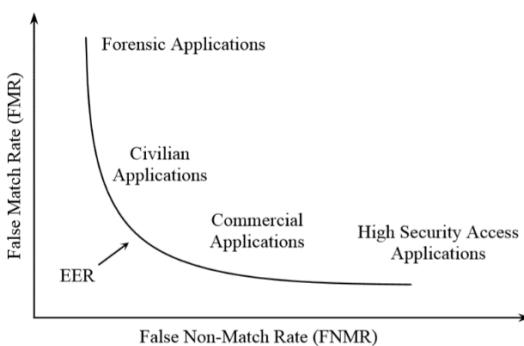
² Detection-Error Tradeoff (DET)

³ Equal-Error Rate



شکل ۱-۶ نمودارهای $FMR(t)$ و $FNMR(t)$ و مقادیر EER و $ZeroFNMR$ و $ZeroFMR$

دقتی که یک سیستم زیست‌سنگی انتظار می‌رود به شدت به کاربرد سیستم وابسته است (شکل ۱-۷). به عنوان مثال برای کاربردهای فضایی و تشخیص افراد مجرم، عدم وجود خطای FNMR اهمیت بسیار بیشتری نسبت به FMR دارد. این به آن معناست که ما ترجیح میدهیم که تا حد ممکن با خطای عدم تطابق شناسایی مجرمی را از دست ندهیم و در مقابل ریسک خطای تطابق‌هایی که سیستم تشخیص خواهد داد را بپذیریم و به صورت دستی آن‌ها را بررسی کنیم. در مقابل در کاربردی مانند سیستم‌های ورود به محل‌های بسیار حساس (مانند سیستم کنترل تردد بخش‌های حساس یک قرارگاه هسته‌ای) به حداقل رساندن مقدار FMR برایمان اهمیت پیدا می‌کند تا کسی به اشتباه اجازه‌ی ورود به آن بخش را پیدا نکند. در این حالت نیز طبیعی است اگر با ناراحتی کاربران مجازی که گاهی اجازه‌ی دسترسی پیدا نمی‌کنند مواجه شویم. تمامی کاربردهای سیستم‌های زیست‌سنگی را می‌توان مابین این دو مثال افراطی قرار داد. نمودار ROC بیانگر مقادیر خطای مورد انتظار برای کاربردهای مختلف را به صورت نسبی نمایش می‌دهد.



شکل ۱-۷ نقاط کار معمول برای کاربردهای مختلف سیستم‌های زیست‌سنگی بر روی نمودار ROC

۱-۳-۶ نرخ خطای تشخیص هویت

در مقایسه‌ی ویژگی‌های زیست‌سنگی به صورت «یک به چند» امکان به وجود آمدن «خطای عدم تشخیص هویت^۱» و «خطای تشخیص هویت نادرست^۲» وجود دارد که مشابه FMR و FNMR قابل محاسبه است. اما این خطاهای با در نظر گرفتن شرایط مسئله (به عنوان مثال این که از بین کل پایگاه داده تشخیص صورت می‌گیرد یا تنها یک بخش از پایگاه داده) از روی مقادیر FMR و FNMR نیز قابل تخمین می‌باشند. به همین علت در سیستم‌های زیست‌سنگی، معمولاً به گزارش خطاهای تایید هویت بسنده می‌شود و مقایسه‌ی سیستم‌ها نیز بر اساس همان FMR و FNMR صورت می‌گیرد.

¹ False Negative Identification-Error Rate (FNIR)

² False Positive Identification-Error Rate (FPIR)

۱-۲-۳-۷- مهم‌ترین معیارهای ارزیابی سامانه‌ی تشخیص زنده بودن

یک سامانه‌ی تشخیص زنده بودن علاوه بر معیارهای کلی یک سامانه زیست‌سنجداری دارای معیارهای منحصر به فرد خود می‌باشد. در این بخش، معیارهای ارزیابی PAD ارائه شده در سند ISO / IEC DIS 30107-3 سازمان بین‌المللی استانداردسازی را ارائه می‌دهیم. براساس چارچوب PAD همانطور که در ISO / IEC 30107-1: 2016^۱ تعریف شده است این معیارهای ارزیابی برای ایجاد آگاهی از یک روش ارزیابی و گزارش‌دهی یکنواخت برای کارهای آینده در این زمینه، تعریف شده‌اند. نهادهای دولتی درگیر در فرآیند استانداردسازی تمایل دارند تا این معیارها را در سامانه‌های عملیاتی خود اعمال کنند، علاوه بر این، بسیاری از مقالات دانشگاهی نیز این معیارها را اتخاذ کرده‌اند.

ISO / IEC DIS 30107-3 سه سطح ارزیابی PAD را معرفی می‌کند: (۱) ارزیابی زیر سامانه PAD: این سطح فقط یک سامانه PAD را ارزیابی می‌کند که ممکن است مبتنی بر سخت‌افزار یا نرم‌افزار باشد. (۲) ارزیابی زیرسامانه ضبط داده: یک زیر سامانه ضبط داده را ارزیابی می‌کند که ممکن است شامل الگوریتم‌های PAD باشد یا نباشد اما بیشتر روی خود حسگر زیست‌سنجدار متمرکز است. (۳) ارزیابی کامل سامانه: کل سامانه شامل زیرسامانه ضبط داده، زیرسامانه مقایسه و اعلام نتیجه را ارزیابی می‌کند.

زیر سامانه‌های PAD با استفاده از دو معیار مختلف ارزیابی می‌شوند: (۱) نرخ خطای دسته‌بندی حمله نمایش زیست‌سنجداری^۱ (APCER) که مقدار PAهایی که به اشتباه، واجد شرایط و واقعی دسته‌بندی شده‌اند را نشان می‌دهد. (۲) نرخ خطای دسته‌بندی نمایش‌های واجد شرایط^۲ (BPCER) که مقدار نمایش‌های واقعی (زنده) که به اشتباه PA دسته‌بندی شده‌اند را نشان می‌دهد [6].

¹ Attack Presentation Classification Error Rate (APCER)

² Bona fide Presentation Classification Error Rate (BPCER)

فصل ۲ بررسی سامانه‌های بازشناسی چهره و رویکردهای آن

شناسایی افراد با توجه به چهره عملی است که اکثراً ما انسان‌ها در زندگی روزمره نیز برای تشخیص هویت استفاده می‌کنیم. توانایی انسان برای انجام این کار قابل توجه بوده و تشخیص چهره‌ی افراد بسیاری که در طول عمر خود دیده‌ایم را حتی با وجود تغییراتی در چهره و یا پس از گذشت سال‌ها انجام می‌دهیم.

در بین زمینه‌های زیست‌سنگی نیز، احراز هویت به کمک چهره بسیار مورد توجه قرار گرفته است. مخصوصاً در سه دهه اخیر، موضوع بازشناسی چهره از یک موضوع تحقیقاتی علمی عبور کرده و پا به عرصه‌ی تکنولوژی و محصولات تجاری گذاشته است. کاربردهای این تکنولوژی از تشخیص هویت افراد در مرزهای بین‌المللی و جستجو به دنبال مجرمان تا نشانه‌گذاری^۱ صورت‌ها در شبکه‌های اجتماعی گسترده شده است [7].

اولین تلاش‌ها برای دسته‌بندی چهره در مقاله‌ی [8] در سال ۱۸۸۸ میلادی مورد بررسی قرار گرفت. روش پیشنهادی نویسنده در این مقاله بدین صورت است که خطوط نیم‌رخ چهره به صورت برداری ذخیره شود (شکل ۱-۲) و با محاسبه میانگین این بردارها و محاسبه فاصله‌ی هر بردار تا بردار میانگین، دسته‌بندی خطوط انجام شود [9].

¹ Tagging



شکل ۱-۲ نمایی از خطوط نیم‌رخ چهره که در روش [8] به کار رفته

در سال‌های اخیر نیز شناسایی و بازشناسی چهره در یک تصویر توسط کامپیوتر بسیار مورد توجه قرار گرفته است. علت این امر آن است که تشخیص هویت به کمک چهره مزیت‌هایی نسبت به سایر روش‌های زیست‌سننجی دارد که به صورت مختصر به برخی از آن‌ها اشاره می‌شود [10]:

بسیاری از دیگر روش‌های زیست‌سننجی نیازمند قرار گرفتن کاربر در حالتی خاص می‌باشد؛ به عنوان مثال برای ثبت اثر انگشت و یا هندسه‌ی دست نیاز است که کاربر دست خود را در محلی مشخص قرار دهد و همچنین برای اسکن عنبه و شبکه‌ی چشم نیاز است که فرد در موقعیت مشخصی نسبت به دوربین قرار گیرد. اما در بازشناسی چهره (مخصوصاً حالت دو بعدی) بدون نیاز به قرار گرفتن کاربر در حالتی خاص می‌توان با دوربین‌هایی از فاصله‌ی دور نیز چهره‌ی افراد را شناسایی کرد.

از آن جایی که تصویربرداری از چهره با وجود یک دوربین ثابت از فاصله‌ی دور امکان‌پذیر است. با وجود الگوریتم‌های مناسب برای بازشناسی چهره و پیش‌پردازش‌های مناسب، می‌توان مدلی ارائه داد که تا حدی نسبت به تغییرات زاویه‌ی دید، اندازه و روشنایی مقاوم باشد.

برای دریافت بسیاری از اطلاعات زیست‌سننجی نیاز است که همه‌ی افراد از یک دستگاه خاص استفاده کنند و در طول دریافت این اطلاعات ممکن است که بدن آن‌ها با دستگاه تماس پیدا کند. این امر می‌تواند موجب انتقال میکروب بین افراد شود اما بازشناسی چهره هیچ نیازی به برخورد فیزیکی با فرد مورد نظر نداشته و استفاده از آن هیچ خطری برای سلامتی انسان ندارد.

اما بازشناسی چهره دارای پیچیدگی‌هایی نیز می‌باشد که باعث می‌شود استفاده از آن سختی‌هایی را نیز به همراه داشته باشد. علت اصلی این امر شباهت فرم کلی چهره‌ی انسان‌هاست و این که ایجاد تمایز بین افراد در دسته‌های از چهره‌ها که شباهت زیادی با یکدیگر دارند دشوار است [11]. علاوه بر این، چهره‌ی انسان‌ها در طول زندگی حالت ثابتی ندارد. عوامل بسیاری می‌توانند باعث ایجاد تغییرات ظاهری چهره شوند؛ این عوامل را می‌توان به دو دسته تقسیم کرد: عوامل درونی و عوامل بیرونی [12]:

عوامل درونی به ماهیت فیزیکی چهره مرتبط هستند. عواملی مانند سن، حالت چهره، موهای چهره، عینک، آرایش و... که گاهی می‌توانند در مدت کوتاهی تغییرات زیادی در چهره‌ی فرد ایجاد کنند.

عوامل بیرونی نیز موجب می‌شوند که ظاهر چهره در مقابل نورهای متفاوت و یا با توجه به مکان ناظر تغییر پیدا کنند. از جمله‌ی تغییراتی که به این صورت ایجاد می‌شود می‌توان به تغییرات نور، ژست¹ صورت، اندازه، وضوح تصویر، تمرکز تصویر²، نویز و... می‌باشد.

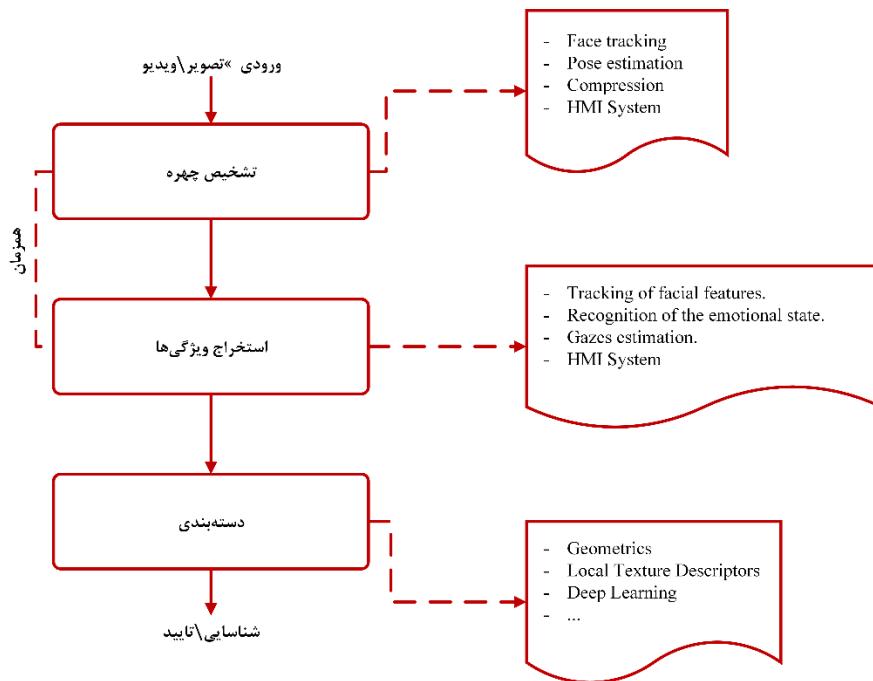
تحقیقات انجام شده نشان‌دهنده‌ی این موضوع است که سه عامل تغییراتی که به واسطه‌ی سن، تغییرات نور و تغییرات زاویه‌ی تصویربرداری ایجاد می‌شوند، مهم‌ترین مشکلاتی است که سیستم‌های بازشناسی چهره با آن مواجه هستند [13].

¹ Pose

² Image focus

۲-۱ مراحل اصلی در سیستم بازشناسی چهره

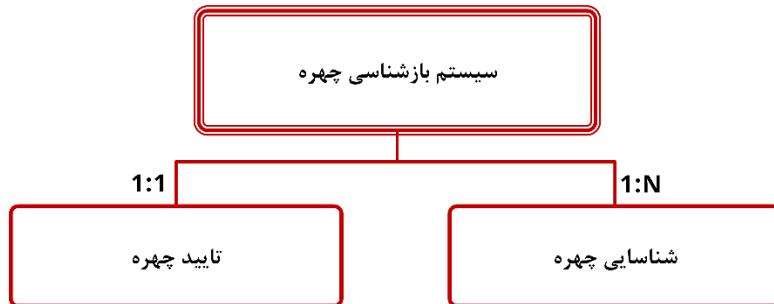
بازشناسی چهره به صورت خودکار شامل سه مرحله اساسی است [14]. همانطور که در شکل (شکل ۲-۲) مشاهده می‌کنید. ۱) تشخیص تقریبی چهره و نرمال‌سازی، ۲) استخراج ویژگی‌ها و نرمال‌سازی دقیق چهره و ۳) طبقه‌بندی یا دسته‌بندی (تأیید یا شناسایی).



شکل ۲-۲ طراحی استاندارد یک سیستم خودکار بازشناسی چهره

تشخیص چهره اولین قدم در سیستم خودکار بازشناسی چهره است. این مرحله تعیین می‌کند که آیا تصویر شامل چهره(ها) است یا خیر. با تشخیص چهره در تصویر وظیفه‌ی آن رهیابی مکان چهره(ها) در تصویر است [15]. مرحله استخراج ویژگی شامل استخراج یک بردار ویژگی از چهره شناسایی شده است که باید طوری تعیین گردد که برای نشان دادن چهره کافی باشد. طبقه‌بندی شامل تأیید و شناسایی است. برای تأیید، باید به یک هویت درخواست شده دسترسی داشته باشیم. با این حال، در شناسایی، چهره با چندین چهره دیگر مقایسه می‌شود. در بعضی موارد، بعضی از مراحل جدا نمی‌شوند. به عنوان مثال، از ویژگی‌های چهره (چشم، دهان و بینی) که برای استخراج ویژگی استفاده می‌شود، اغلب در هنگام تشخیص چهره نیز استفاده می‌شود. همانطور که در شکل ۲-۲ نشان داده شده است، می‌توان همزمان شناسایی و استخراج ویژگی‌ها را انجام داد. اگرچه سیستم‌های خودکار بازشناسی چهره باید سه مرحله ذکر شده در بالا را انجام دهند، اما هر مرحله به عنوان یک مسئله مهم تحقیقاتی در نظر گرفته می‌شود، به این دلیل که تکنیک‌های مورد استفاده برای هر مرحله نیاز به بهبود دارند و در کاربردهای مختلف استفاده می‌شوند. همانطور که در شکل ۲-۲ نشان داده شده است. به عنوان مثال، تشخیص چهره برای فعال‌سازی نظارت بر چهره ضروری است (رهیابی چهره) و استخراج ویژگی‌های چهره برای شناسایی حالت عاطفی فرد بسیار مهم است، که به نوبه خود در سیستم‌های تعامل انسان و ماشین^۱ (HMI) بسیار ضروری است. این گزارش عمدتاً بر استخراج ویژگی (و احتمالاً انتخاب ویژگی) و طبقه‌بندی متمرکز است. یک سیستم بازشناسی چهره با توجه به کاربرد می‌تواند برای تأیید چهره و یا شناسایی آن به کار رود (شکل ۳-۲).

^۱ Human–Machine Interaction (HMI)



شکل ۳-۲ طبقه‌بندی پروتکل‌های مختلف ارزیابی در بازشناسی چهره

می‌توان با صرف نظر از روش‌هایی که بر پایه‌ی دنباله‌ای از تصاویر کار می‌کنند، روش‌های بازشناسی چهره را با توجه به مدل چهره و روش جمع‌آوری داده‌ی آن می‌توان به دو دسته تقسیم کرد: ۱) روش‌هایی با محوریت عکس (دوبعدی) و ۲) روش‌های مبتنی بر ساختار سه بعدی چهره [16]. هر دوی این روش‌ها در وهله‌ی اول نیازمند دریافت تصویر فرد هستند. در ادامه‌ی این فصل پس از بررسی روش‌های دریافت تصویر، روند بازشناسی چهره به تفکیک دو بعدی و سه بعدی مورد بررسی قرار خواهد گرفت.

۲-۲ روش‌های دریافت چهره

با توجه به این که تمامی مراحل روند بازشناسی چهره بر پایه‌ی چهره‌ی دریافت شده پایه‌گذاری می‌شوند، عملیات دریافت چهره یک عنصر بسیار مهم در هر سیستم تشخیص چهره به حساب می‌آید. دریافت چهره را می‌توان از جنبه‌های مختلف تقسیم‌بندی کرد. دریافت می‌تواند با هدف بازشناسی چهره دو بعدی و یا سه بعدی انجام شود. همچنین حسگر ورودی می‌تواند به صورت دوربین معمولی (دریافت عکس دو بعدی) باشد و یا اطلاعات سه بعدی چهره را دریافت کند. از منظری دیگر نیز دریافت می‌تواند به صورت تک عکس و یا یک دنباله از عکس‌ها (یا فیلم) انجام شود و مبنای بازشناسی چهره قرار گیرد. در ادامه، این بخش را در دو دسته مورد بررسی قرار می‌دهیم؛ دریافت چهره با هدف تعیین چهره به صورت: ۱) دو بعدی و ۲) سه بعدی.

۲-۲-۱ تصاویر دوبعدی

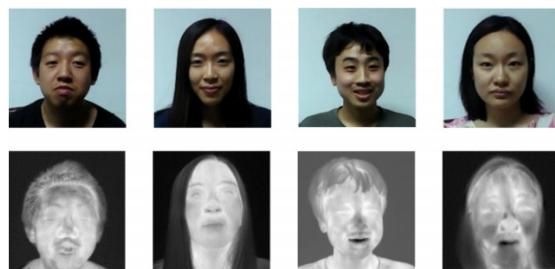
بسیاری از سیستم‌های زیست‌سنجی تنها با یک عکس دوبعدی از نمای رو به رو عملیات بازشناسی چهره را انجام می‌دهند. با وجود پیشرفت‌های بسیاری که در این زمینه در سال‌های اخیر انجام شده است، اما همچنان بازشناسی چهره تنها با استفاده از طیف مرئی در محیط‌های کنترل نشده با مشکلاتی همراه است. دلیل اصلی این امر این است که تغییرات تصاویر یک چهره ثابت در شرایط نوری مختلف و از زوایای مختلف بسیار زیاد است. عوامل تاثیرگذار دیگر مانند حالت‌های احساسی چهره و تغییرات ژست صورت نیز پیچیدگی‌های بیشتری را اضافه می‌کنند. همچنین دیگر چالش الگوریتم‌های بازشناسی چهره در تشخیص هویت افرادی است که بخشی از صورت آن‌ها پوشیده شده، مدل مو و ریش در آن‌ها به صورت اساسی تغییر کرده و یا آرایش کرده‌اند. همچنین تشخیص هویت افرادی که چهره‌ی خود را جراحی پلاستیک کرده‌اند تقریباً غیرممکن است [17].

بازشناسی چهره می‌تواند بر اساس تصاویر عادی (که به صورت روزمره با آن سر و کار داریم) انجام شود. اما در کاربرد بازشناسی چهره استفاده از تصاویر در شرایط عکس‌برداری مختلف دیگر نیز مورد توجه قرار گرفته است. امواج الکترومغناطیسی با طول موج کمتر از نور مرئی (مانند امواج X-Ray و فرابنفس) برای بدن انسان ضرر دارند بنابراین برای بازشناسی چهره نمی‌توانند به کار روند. اما امواج الکترومغناطیسی با طول موج بیشتر از نور مرئی و به طور خاص حسگرهای فروسرخ برای این منظور بسیار مورد استفاده قرار گرفته است [17].

در حالی که حسگرهای نور مرئی امواج با طور موج ۰.4 تا ۰.7 میکرومتر را دریافت می‌کنند، حسگرهای نور فروسرخ نسبت به تابش حرارتی در طول موج ۰.7 تا ۱۴ میکرومتر حساس هستند. این بازه طول موج خود شامل طیف «فروسرخ بازتاب شده» و

«فروسرخ حرارتی» می‌باشد. هر ماده‌ای با توجه به دما و جنسش، طیف حرارتی منتشر می‌کند. تابش حرارتی بدن انسان اکثرا در طول موج مربوط به فروسرخ طول موج بلند^۱ (LWIR) و مقداری نیز در طول موج مربوط به فروسرخ طول موج متوسط^۲ (MWIR) قرار می‌گیرد. طول موج این دو بخش به ترتیب ۳.۰ تا ۵.۰ میکرومتر و ۸.۰ تا ۱۴.۰ میکرومتر است. با دریافت طول موج‌های منتشر شده در این بازه‌ها، می‌توان یک تصویری حرارتی از محیط با دمای اتاق به دست آورد.

در عمل بازشناسی چهره نیز تصاویر حرارتی مورد استفاده قرار می‌گیرند. در شکل ۴-۲ نمونه‌ای از تصاویر معمولی چهره و تصاویر حرارتی آن‌ها ملاحظه می‌شود. مناطق روشن‌تر با درجه حرارت بالاتر (چشم) هستند و بدیهی است که با افزایش طول موج، سطح جزئیات تصویر چهره کاهش می‌یابد، یعنی بیشترین جزئیات در تصویر گرفته شده با دوربین معمولی و کمترین در امواج فروسرخ با طول موج زیاد است.



شکل ۴-۲ تصاویر چهره که توسط حسگرهای مختلف دریافت شده‌اند. ردیف بالا با حسگر نور مرئی و ردیف پایین نمایانگر تصاویر حرارتی از همان اشخاص است (در مقاله ذکر شده است که این حالت نمایش، بهترین نوع رنگ‌آمیزی است) [18]

با فدا کردن رنگ‌های مرئی می‌توان بازشناسی چهره را تنها بر اساس تصاویر حرارتی انجام داد. استفاده از این تصاویر در شرایطی که کنترلی بر شرایط نوری وجود ندارد بسیار مناسب است چون در محیط کاملاً تاریک نیز می‌توان از آن استفاده کرد. امواج الکترومغناطیس در بازه طیف حرارتی امواج بازتاب شده‌ی بسیار کمی را شامل می‌شوند؛ بنابراین امواج ساطع شده از پوست، مستقل از شرایط نوری، نمایش دهنده‌ی ویژگی‌های درونی چهره‌ی فرد خواهد بود. همانطور که در شکل ۵-۲ ملاحظه می‌شود تصاویر حرارتی در حالت‌های احساسی مختلف چهره نیز تغییرات ناچیزی می‌کند.



شکل ۵-۲ مقایسه‌ی تصویر نور مرئی با تصویر حرارتی در شرایط نوری و با حالت‌های احساسی مختلف

اما تصاویر حرارتی نیز در موقعیت‌هایی دارای محدودیت‌هایی هستند. به عنوان مثال برای شناسایی شخصی که عینک به چشم دارد یا در وسیله نقلیه متحرک قرار دارد. همان‌طور که در شکل ۶-۲ ملاحظه می‌شود، شبیه درصد اندکی از انرژی گرمایی را عبور می‌دهد و این موضوع باعث می‌شود که در تصاویر افرادی که در عینک به چشم دارند، اطلاعات ما از ناحیه‌ی اطراف چهره کم باشد. همچنین تغییر دمای بدن (به عنوان مثال پس از فعالیت بدنی) نیز تغییراتی در تصویر حرارتی ایجاد می‌کند.

¹ long-wave infrared

² mid-wave infrared



شکل ۶-۲ مشکل تصویر حرارتی در هنگام استفاده از عینک

روش‌های تشخیص هویت، برای کارکرد صحیح نیاز دارند که عکس ورودی آن‌ها یک وضوح حداقلی را دارا باشد. در صورت پایین بودن وضوح تصاویر، سیستم حداقل با دو مشکل «همترازی نادرست» و «کمبود ویژگی‌های تاثیرگذار» مواجه می‌شود [19]. در الگوریتم‌های بازشناسی چهره وضوح مشخصی برای تصاویر ورودی تعیین نشده است اما به عنوان مثال روش معروف مقادیر ویژه چهره^۱ (این الگوریتم در بخش ۲-۳-۱ بررسی خواهد شد) با تصاویر ورودی با ابعاد ۵۱۲ پیکسل کار می‌کند [20]. همچنین پایگاه داده FERET [21] که در این حوزه کاربرد بسیاری دارد، دارای تصاویری با ابعاد ۶۴۰ در ۴۸۰ یا ۴۹۰ پیکسل است. نمونه‌ی تصاویر این پایگاه داده در شکل ۷-۲ آورده شده است.

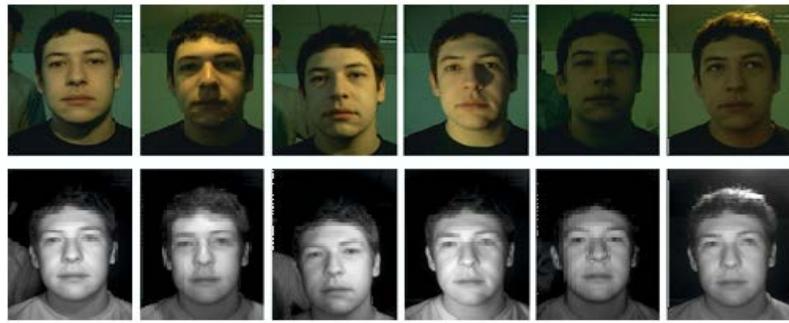


شکل ۷-۲ نمونه‌ی تصاویر پایگاه داده FERET [21]

در کاربردهای خاص که تصاویر وضوح کافی ندارند نیز روش‌های متفاوتی استفاده می‌شود و دقت آن‌ها به طبع پایین‌تر از روش‌های عادی است. در پژوهش [19] روش‌هایی برای بازشناسی چهره بر پایه‌ی تصاویر با وضوح کم آورده شده است. همانطور که پیش‌تر نیز به آن اشاره شد، در سامانه‌های بازشناسی چهره که از امواج فروسرخ حرارتی استفاده شده است، امکان بازشناسی چهره‌های پنهان و در شرایط نوری ضعیف وجود دارد. با این حال، به دلیل هزینه بالای حسگرهای حرارتی و ناپایداری آن در دماهای مختلف، استفاده از آن مطلوب نیست. در حالی که امواج فروسرخ نزدیک^۲ (NIR) به دلیل هزینه کم، توجه بیشتری را به خود جلب کرده است. امواج فروسرخ نزدیک (NIR) شامل تابش فرکانس‌های پایین در رنگ‌های قرمز در ناحیه مرئی است. شکل ۸-۲ تصاویری از یک نمونه چهره را نشان می‌دهد که هم تصاویر رنگی و هم NIR آورده شده است. در هر کدام از تصاویر در زوایای مختلفی به چهره نور تابیده شده است، همانطور که مشاهده می‌شود کیفیت تصاویر NIR چهره به سختی تحت تأثیر نور محیط قرار می‌گیرد. این امر برای بازشناسی چهره بسیار مفید خواهد بود.

¹ EigenFace

² near infrared (NIR)



شکل ۸-۲ ردیف بالا: تصویر رنگی از یک چهره در شرایط نوری مختلف. ردیف پایین: تصاویر مربوط به فیلتر NIR در همان شرایط نوری

مزیت اصلی زیست‌سنجی چهره با تصاویر حرارتی تغییرناپذیری در روشنایی است، زیرا می‌توان آن‌ها را در تاریکی و در شرایط دید کم ثبت کرد. اما، محدودیت فاصله ضبط وجود دارد زیرا میزان دقت با افزایش فاصله کاهش می‌یابد. از مشکلات دیگر تأثیر سلامت و وضعیت احساسی بر روی تصاویر حرارتی است علاوه بر موارد بیان شده، وجود یک سخت‌افزار خاص برای دریافت تصاویر سبب شده است تا هم‌چنان استفاده از تصاویر معمولی برای بازشناسی چهره بیشتر مورد استفاده قرار بگیرد.

۲-۱-۱ پایگاه داده‌های رایج برای بازشناسی چهره دوبعدی

مجموعه داده‌هایی وجود دارند که شامل تصاویر امواج فروسرخ نزدیک می‌باشند ازین جمله می‌توان به پایگاه داده ONVF و پایگاه داده INF و پایگاه داده CSIST اشاره کرد که در ادامه به طور مختصر به آن‌ها پرداخته می‌شود.

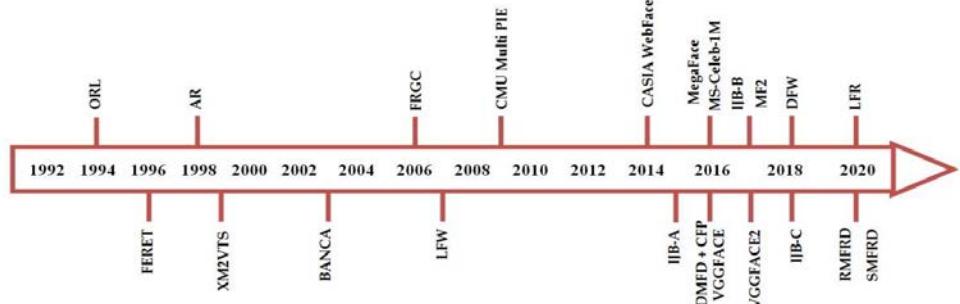
پایگاه داده ONVF : این پایگاه داده توسط تیم تحقیقاتی [12] جمع‌آوری شده است. تصاویر چهره از ۱۰۰۰ نفر، بدون محدودیت در روشنایی و موقعیت ضبط شده است. از هر سوژه حدود ۳۰ تصویر چهره NIR و ۳۰ تصویر معمولی ارائه شده است. در حین جمع‌آوری پایگاه داده، از دوربین‌های JAI با سنسور تصویر HM2131 1/2.7 استفاده شده است که به باند NIR حساس است. منبع نور فعال در طیف NIR بین ۷۸۰ نانومتر - ۱۱۰۰ نانومتر، بر روی دوربین نصب شده بود.

پایگاه داده INF شامل ۹۴ دانشجو از دانشگاه، از جمله ۵۷ مرد و ۳۷ زن است. در حین ضبط، از افراد خواسته شده است که جلوی دوربین بنشینند و تصاویر معمولی چهره آن‌ها جمع‌آوری شده است. فاصله دوربین تا چهره بین ۱۲۰-۸۰ سانتی متر بوده است که برای کاربر محدوده مناسبی است. ۵ تصویر چهره NIR برای هر سوژه با وضوح ۶۴۰ → ۴۸۰ پیکسل وجود دارد.

پایگاه داده CSIST : دو مجموعه تصویر در پایگاه داده [13] وجود دارد. Lab1 و Lab2 شامل ۵۰۰ تصویر NIR و ۵۰۰ تصویر قابل مشاهده است که از ۵۰ نفر گرفته شده است. در Lab2، ۱۰۰۰ تصویر NIR و ۱۰۰۰ تصویر قابل مشاهده از ۵۰ نفر تحت شرایط مختلف نوری جمع‌آوری شده است. اندازه تصویر برای پایگاه‌های داده Lab1 و Lab2 ۸۰ تا ۱۰۰ پیکسل است.

در ادامه مجموعه داده‌های رایج در سال‌های اخیر که برای رویکردهای بازشناسی چهره دوبعدی مناسب هستند، بررسی می‌شود.

شکل ۸-۲ مجموعه داده‌های ذکر شده به ترتیب زمانی پیدایش آن‌ها را خلاصه می‌کند.



شکل ۹-۲ تحوّلات مجموعه داده‌های دوبعدی بازشناسی چهره در طول زمان [22]

جدول ۱-۲ یک بررسی مقایسه‌ای از مجموعه داده‌های شناسایی چهره ذکر شده در بالا ارائه می‌دهد.

جدول ۱-۳ مجموعه داده‌های بازشناسی چهره دوبعدی در سال‌های مختلف به همراه تعداد تصاویر و تعداد افراد مختلف

مجموعه داده	سال پیدایش	تعداد تصاویر/ هر فرد	تعداد افراد	تعداد تصاویر
ORL [23]	۱۹۹۴	۱۰	۴۰	۴۰۰
FERET [24]	۱۹۹۶	-	۱۱۹۹	۱۴۱۲۶
AR [25]	۱۹۹۸	۲۶	۱۱۶	۳۰۱۶
XM2VTS [26]	۱۹۹۹	-	۲۹۵	-
BANCA [27]	۲۰۰۳	-	۲۰۸	-
FRGC [28]	۲۰۰۶	۷	-	۵۰۰۰۰
LFW [29]	۲۰۰۷	≈ ۲.۳	۵۷۴۹	۱۲۲۲۳
CMU Multi PIE [30]	۲۰۰۹	-	۳۳۷	>۷۵۰۰۰
CASIA WebFace [31]	۲۰۱۴	≈ ۴۶.۸	۱۰۵۷۵	۴۹۴۴۱۴
IJB-A [32]	۲۰۱۵	≈ ۱۱.۴	۵۰۰	۵۷۱۲
MegaFace [33]	۲۰۱۶	≈ ۱.۴	۶۹۰۵۷۲	۱۰۲۷۰۶۰
CFP [34]	۲۰۱۶	>۱۴	۵۰۰	۷۰۰۰
MS-Celeb-1M [35]	۲۰۱۶	۱۰۰	۱۰۰۰۰	۱۰ میلیون
DMFD [36]	۲۰۱۷	۶	۴۱۰	۲۴۶۰
VGGFACE [37]	۲۰۱۶	۱۰۰۰	۲۶۲۲	۲.۶ میلیون
VGGFACE [38]	۲۰۱۷	≈ ۳۶۲.۶	۹۱۳۱	۳.۳۱ میلیون
IJB-B [39]	۲۰۱۷	≈ ۳۶.۲	۱۸۴۵	۲۱۷۹۸
MF2 [40]	۲۰۱۸	≈ ۷	۶۷۲۰.۵۷	۴.۷ میلیون
DFW [41]	۲۰۲۰	≈ ۵.۲۶	۱۰۰۰	۱۱۱۵۷
IJB-C [42]	۲۰۲۰	≈ ۶	۳۵۳۱	۲۱۳۳۴
LFR [43]	۲۰۲۰	۱۰ - ۲۶۰	۵۴۲	۳۰۰۰۰
RMFRD [44]	۲۰۲۰	-	۵۲۵	۹۵۰۰۰
SMFRD [44]	۲۰۲۰	-	۱۰۰۰۰	۵۰۰۰۰۰

در اختیار داشتن مدل سه بعدی چهره و تشخیص هویت بر اساس آن موجب می شود که نسبت به تغییرات نور و زاویه در مقایسه با روش دو بعدی پایداری بیشتری داشته باشیم. بنابراین مدل های سه بعدی پتانسیل این را دارند که به کمک آن ها تشخیص هویت با دقت بیشتری انجام شود. روش های دریافت سه بعدی چهره دارای مشکلاتی نیز هستند؛ به عنوان مثال ساخت یک مش از چهره می تواند از نظر محاسباتی سنگین باشد و یا ممکن است به پاسخ بهینه همگرا نشود. علاوه این، مدل های سه بعدی به طور کامل نسبت به تغییرات نور مقاوم نیستند. این روش ها می توانند به وسیله ای منابع نور شدید و یا سطوح بازتاب کننده نور تحت تاثیر قرار بگیرد.

دو روش کلی برای تولید یک مدل سه بعدی چهره انسان وجود دارد: ۱) حسگر سه بعدی ۲) مدل morphable. این دو روش در ادامه در بخش های جداگانه بررسی می شوند.

۲-۲-۱ حسگر سه بعدی

در این روش یک حسگر سه بعدی به کمک تصاویر عمقی مختصات تقریبی مجموعه ای از نقاط چهره را در اختیارمان قرار می دهد. در این روش برای ساخت مدل سه بعدی چهره مراحل زیر طی می شود [45]:

- داده های عمق تصویر به سیستم مختصات دوربین برده می شوند به طوری که عمق تصویر (محور Z) در راستای محور کانونی دوربین قرار بگیرد.
 - داده های عمقی که از زوایای مختلف چهره جمع آوری شده اند با یکدیگر ترکیب می شوند.
 - مختصات سه بعدی به دست آمده از نقاط بهینه سازی می شوند تا به بهترین دقت خود برسند. در نهایت بر اساس ابر نقاط به دست آمده یک مش چند ضلعی از چهره فرد ایجاد می شود.
- با توجه به تکنولوژی های موجود در زمینه دریافت سه بعدی داده ها سه راه حل کلی برای دریافت تصویری سه بعدی مطرح شده است که در ادامه بررسی می شود [16]:

دوربین استریو: در این روش تعداد حداقل دو دوربین کالیبره شده^۱ از چهره به طور همزمان تصویربرداری می کنند. برای دست یابی به بازیابی دقیق دوربین ها باید با دقتی بالا کالیبره شده باشند. به وسیله ای این سیستم و یک مدل هندسی می توان مختصات فضایی هر نقطه مشترک در تصاویر را به طور دقیق محاسبه کرد.

تابش نور ساختار یافته: این روش با تصویربرداری از الگوهای نور ساختار یافته که بر روی چهره تابانده می شود کار می کند. اعوجاج ایجاد شده از الگوی تابیده شده بر روی صورت امکان استخراج اطلاعات عمق را در اختیارمان قرار می دهد. این روش بسیار ارزان قیمت بوده و به ما این امکان را می دهد تا با استفاده از تنها یک دوربین مدل سه بعدی چهره را استخراج کنیم. استخراج اطلاعات در این روش سریع بوده و برای برای دریافت اطلاعات بافت ظاهری نیز کافیست تصویری دیگر تحت نور عادی گرفته شود. توجه کنید که مدلی از چهره که به این طریق دریافت می شود تنها از یک نقطه دریافت شده و کاملا سه بعدی نیست. اصطلاحاً به آن تصویر دو و نیم بعدی (2.5D) گفته می شود. در شکل ۹-۲ نمونه ای از کاربرد این مشاهده می شود.

^۱ دوربین های کالیبره شده دوربین هایی هستند که مختصات و جهت گیری فضایی آن ها را نسبت به یکدیگر و همچنین نسبت به صحنه ای که از آن تصویر می گیرند به طور دقیق در اختیار داریم.



شکل ۱۰-۲ تصویری از نقاط نور فروسرخ که توسط تلفن همراه iPhone برای بازشناسی چهره بر روی چهره تابانده شده است^۱ در این سیستم بیش از ۳۰ هزار نقطه برای بازشناسی چهره تابانده می‌شود

حسگرهای لیزری: این روش‌ها نسبت به سایر روش‌ها بسیار دقیق‌تر هستند و در عوض کندتر و پر هزینه‌تر نیز هستند. به عنوان مثال پایگاه داده‌ی FRGC^۲ به کمک حسگر Minolta VIVID 910 دریافت شده است. این دستگاه که حاوی حسگری لیزری نیز می‌باشد در شکل ۱۰-۲ مشاهده می‌شود. پژوهش [46] نیز خلاصه‌ای از تعدادی حسگر صنعتی، دقق و ویژگی‌های آن‌ها آورده شده است.



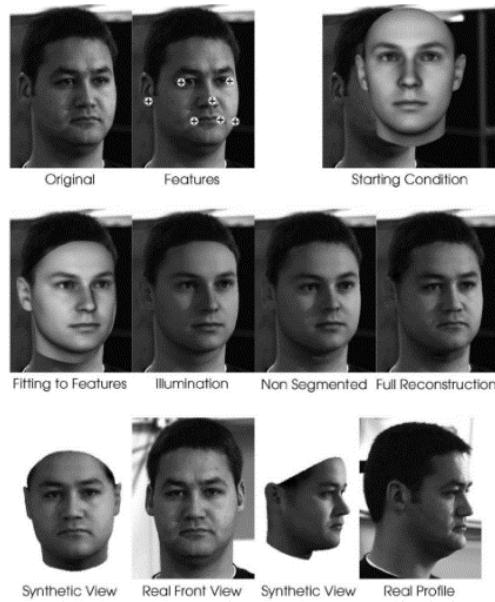
شکل ۱۱-۲ دستگاه VIVID 910 که پایگاه داده‌ی FCGD به کمک آن جمع‌آوری شده است

۲-۲-۲-۲ morphable مدل

روش دیگری که برای ساخت مدل سه‌بعدی چهره به کار می‌رود استفاده از مدل morphable است. ایده‌ی این روش این است که با تعریف تعداد زیادی پارامتر در چهره، مدل هر چهره‌ی دلخواه می‌توانند با تنظیم درست پارامترها توسط یک مدل سازنده‌ی چهره تولید شود. این روش نسبت به تغییرات اندازه، زاویه و چرخش مقاومت بیشتری دارد. اما از طرفی هزینه‌ی محاسباتی بالا و وابستگی مدل به تعداد و کیفیت پارامترها نیز از معایب مدل morphable به حساب می‌آید. شکل ۱۱-۲ نتایج حاصل از نمونه‌ی بارز این روش را نشان می‌دهد.

¹ <https://goshopen.blogspot.com/2017/11/this-is-how-iphone-x-face-id-work-look.html>

² Face Recognition Grand Challenge



شکل ۱۲-۲ ساخت مدل سه بعدی به وسیله‌ی یک تصویر بالا سمت چپ از چهره. به وسیله‌ی مدل ساخته شده می‌توان زوایای جدید از چهره را نیز تولید کرد. تصاویر تولید شده، در کنار تصاویر واقعی از همان زاویه در ردیف پایین ملاحظه می‌شود.[46]

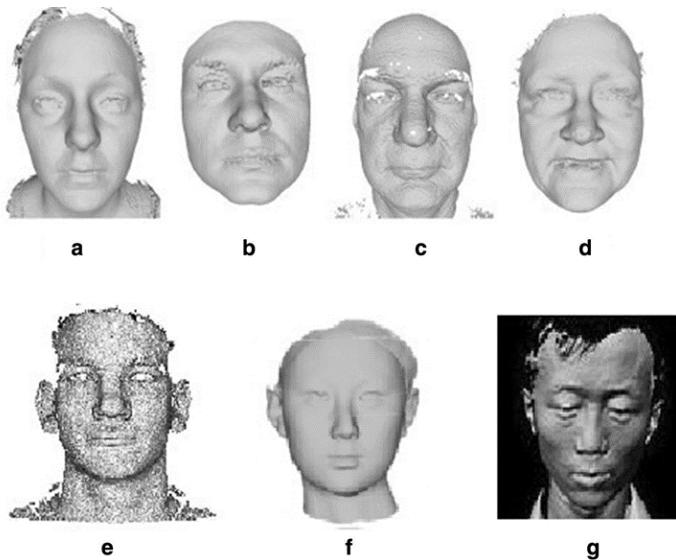
۲-۲-۳- پایگاه داده‌های رایج برای بازشناسی چهره سه بعدی

جدول ۲-۲ برخی از پایگاه داده‌های مهم برای بازشناسی چهره سه بعدی را با اطلاعات آن‌ها نشان می‌دهد.

جدول ۲-۲ خلاصه تطبیقی برخی از پایگاه‌های اطلاعاتی در دسترس برای شناسایی چهره سه بعدی

پایگاه داده	تاریخ ارائه	تعداد افراد	تعداد تصاویر	نوع داده
BU-3DFE	۲۰۰۶	۱۰۰	۲۵۰۰	مش
FRGC v1.0	۲۰۰۶	۲۷۳	۹۴۳	نگاشت عمق
FRGC v2.0	۲۰۰۶	۴۶۶	۴۰۰۷	نگاشت عمق
CASIA	۲۰۰۶	۱۲۳	۴۶۲۳	نگاشت عمق
ND2006	۲۰۰۷	۱۳۴۵۰	۸۸۸	نگاشت عمق
Bosphorus	۲۰۰۸	۱۰۵	۴۶۶۶	ابر نقطه
BJUT-3D	۲۰۰۹	۵۰۰	۱۲۰۰	مش
Texas 3DFRD	۲۰۱۰	۱۱۸	۱۱۴۰	نگاشت عمق
UMB-DB	۲۰۱۱	۱۴۳	۱۴۷۳	نگاشت عمق
BU-4DFE	۲۰۰۸	۱۰۱	۶۰۶۰۰ = ۶۰۶ توالی (فریم)	ویدیویی سه بعدی

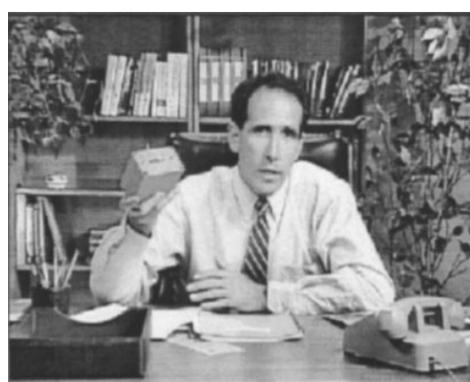
در شکل ۱۲-۲ یک مدل چهره سه بعدی از هفت پایگاه داده اصلی (a) CASIA (g) BJUT-3D (f) ND-2006 (e) 3DFRD (d) FRGC (c) Bosphorus (b) BU-3DFE (a) Texas را مشاهده می‌کنید.



شکل ۱۳-۲ یک مدل چهره سه بعدی از هفت پایگاه داده اصلی (a) Bosphorus (b) BU-3DFE (c) FRGC (d) ND-2006 (e) Texas 3DFRD (f) CASIA (g) BJUT-3D

۲-۳ بازشناسی چهره دو بعدی

عملیات مقایسه در فرآیند بازشناسی چهره با استفاده از یک عکس خاکستری مانند هر سیستم زیست‌سنجدی دیگر، مراحل مشابهی را طی می‌کند. به این صورت که ابتدا سیستم یک عکس حاوی چهره دریافت می‌کند، مکان چهره‌ی انسان را در عکس تشخیص می‌دهد، قسمت چهره از عکس بریده شده، نرمال می‌شود و ویژگی‌های آن استخراج می‌شود و بدین ترتیب الگوی تصویر صورت تشکیل می‌شود. در هنگام تشخیص هویت، این الگوی دریافت شده با الگوهای موجود در پایگاه داده مقایسه می‌شود. بدین ترتیب دو بخش اصلی این الگوریتم ۱) مکان‌یابی چهره و نرمال‌سازی و ۲) تشخیص هویت چهره خواهد بود. الگوریتم‌هایی که هر دو بخش را در بر می‌گیرند، الگوریتم‌های تشخیص چهره‌ی تمام اتوماتیک و الگوریتم‌هایی که تنها بخش دوم را شامل می‌شوند الگوریتم‌های نیمه اتوماتیک نامیده می‌شوند [۹]. در ادامه‌ی این بخش، جزئیات بیشتری در مورد هریک از این دو قسمت الگوریتم ارائه می‌شود.



شکل ۱۴-۲ نمونه عکسی که در کاربردی عملی برای بازشناسی چهره به کار می‌رود

امروزه عمل مکانیابی چهره در تصویر در بسیاری از کاربردهای تجاری به صورت بلاذرنگ بر روی تصویر انجام می‌شود. در ادامه‌ی این بخش الگوریتم‌های مکانیابی چهره در دو بخش «بر پایه‌ی تصویر» و «بر پایه‌ی ویژگی» مورد بررسی قرار خواهد گرفت و در هر بخش الگوریتم‌های مهم آن دسته بررسی خواهد شد.

۲-۱-۳- روش‌های بر پایه‌ی تصاویر

در روش‌هایی که بر پایه‌ی ویژگی‌های چهره بنا شده‌اند، در شرایط محیطی خاص و پیش‌بینی نشده شاهد کاهش شدید دقت الگوریتم‌ها بودیم. در دسته‌ی دیگری از روش‌ها نیز وجود دارند که شناسایی چهره را به عنوان یک مسئله‌ی تشخیص الگو فرموله می‌کنند و بدین ترتیب با استفاده از پایگاه داده‌ای از تصاویر چهره، مدل‌هایی را برای تشخیص چهره آموزش می‌دهند. در این‌گونه روش‌ها در صورتی که تنوع تصاویر پایگاه داده تا حد کافی باشد، حالت‌های پیش‌بینی نشده کمتر برای مدل اتفاق خواهد افتاد. روش‌های متنوعی در این زمینه مورد استفاده قرار گرفته است که از جمله‌ی آن‌ها می‌توان به روش‌های کاهش بعد، روش‌های آماری و همچنین شبکه‌های عصبی مصنوعی اشاره کرد.

یکی از تأثیرگذارترین الگوریتم‌های این زمینه الگوریتم ویولا-جونز^۱ است که در سال ۲۰۰۱ میلادی توسط ویولا و جونز در مقاله‌ی [47] ارائه شد. در ادامه‌ی این بخش به صورت اجمالی چگونگی کارکرد این الگوریتم بررسی می‌شود.

این روش سه ایده‌ی کلیدی دارد که آن را به یک مکانیاب چهره موفق بدل کرده است و امکان مکانیابی چهره به صورت بلاذرنگ را ممکن ساخته است: ۱) تصاویر انتگرالی ۲) استفاده از الگوریتم AdaBoost به عنوان طبقه‌بند^۲ و ۳) یافتن مکان‌های مهم تصویر به کمک دسته‌بندهای متوالی. هر یک از این سه بخش در ادامه اجمالاً مورد بررسی قرار می‌گیرد:

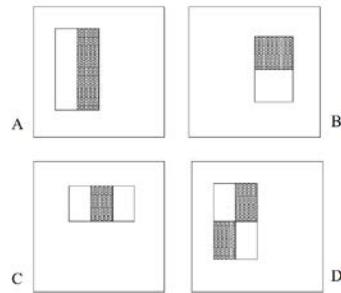
تصویر انتگرالی، در حقیقت ماتریسی است که از روی یک تصویر محاسبه می‌گردد و به ما اجازه می‌دهد که مجموع پیکسل‌های یک بخش مستطیلی از تصویر را تنها با چند عملیات تفریق ساده محاسبه کنیم. تصویر انتگرالی ii از روی تصویر i به این شکل محاسبه می‌گردد:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

که در آن (x, y) نمایانگر مقدار سطر x و ستون y از تصویر ii می‌باشد. تصویر انتگرالی این امکان را می‌دهد که ویژگی‌های هر-شکل تصویر با سرعت بالا محاسبه شوند. این ویژگی‌ها که برای اولین بار در همین پژوهش و به منظور مکانیابی چهره (و اشیاء) در تصویر معرفی شدند بر پایه‌ی موجک‌های هار نامگذاری شدند. نمونه‌هایی از این ویژگی‌ها در شکل ۱۵-۲ دیده می‌شود. برای محاسبه‌ی این ویژگی‌ها، لازم است مجموع مقادیر پیکسل‌هایی که در محوطه‌ی مستطیل خاکستری قرار می‌گیرند از مجموع مقادیر پیکسل‌هایی که در محوطه‌ی سفید قرار می‌گیرند کم شوند.

¹ Viola-Jones

² Classifier

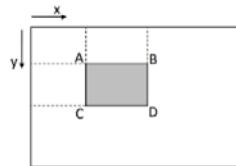


شکل ۱۵-۲ نمونه‌هایی از مستطیل‌هایی که برای استخراج ویژگی بر روی بخشی از تصویر قرار می‌گیرند.

به لطف تصویر انتگرالی محاسبه‌ی این ویژگی‌ها در هر اندازه‌ای در مرتبه‌ی زمانی ثابت قابل محاسبه خواهد بود. به عنوان مثال برای محاسبه‌ی مجموع مقادیر پیکسل‌ها در قسمت خاکستری شکل ۱۶-۲:

$$\sum_{(x,y) \in ABCD} i(x,y) = ii(D) + ii(A) - ii(B) - ii(C)$$

که بیانگر این موضوع است با کمک تصویر انتگرالی این محاسبات در مرتبه زمانی خطی قابل محاسبه بوده و این امر موجب می‌شود که سرعت اجرای الگوریتم بسیار بهبود یابد.

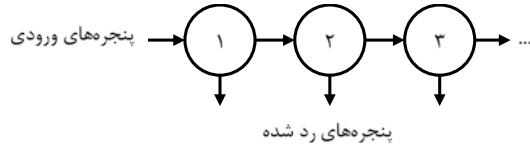


شکل ۱۶-۲ نحوه‌ی تعریف یک مستطیل در شکل

ایده‌ی کلیدی دوم این الگوریتم، استفاده از الگوریتم AdaBoost (Adaptive Boosting) برای تشخیص چهره در یک پنجره از تصویر است. به طور کلی انجام boosting در طبقه‌بندها به معنای ترکیب چند طبقه‌بند ضعیف برای دست‌یابی به یک طبقه‌بند قوی است.

تعداد بسیار زیادی ویژگی هار-شکل می‌تواند از یک پنجره از تصویر استخراج شود (تعداد آن‌ها از تعداد پیکسل‌های آن پنجره هم بیشتر است). اما تجربه نشان داده که در عمل به همه‌ی این ویژگی‌ها نیاز نداریم و با به‌کارگیری یک مجموعه کوچک از این ویژگی‌ها می‌توان به یک طبقه‌بند مؤثر دست یافت. اما انتخاب درست این مجموعه از ویژگی‌ها بسیار مهم خواهد بود. در این مقاله از الگوریتم AdaBoost به منظور «انتخاب مجموعه‌ی ویژگی‌ها» و «آموزش طبقه‌بندها» استفاده شده است.

ایده‌ی کلیدی سوم این مقاله آن است که در آن از تعدادی طبقه‌بند به صورت متوالی (آبشراری) استفاده شده است. یک پنجره از تصویر در صورتی که توسط اولین طبقه‌بند مورد قبول واقع شد، وارد طبقه‌بند دوم خواهد شد و به همین ترتیب در صورتی که پنجره وارد طبقه‌بند k -ام می‌شود که توسط همه‌ی طبقه‌بندهای قبلی مورد قبول واقع شده باشد (شکل ۱۷-۲). در این ساختار اولین طبقه‌بندها به گونه‌ای طراحی می‌شوند که بسیار سبک باشند (تعداد ویژگی‌های کمی در آن‌ها بررسی شود) و در عین حال تعداد بسیاری از پنجره‌های «غیر چهره» را رد کنند. اما پنجره‌هایی که احتمال ظاهر شدن چهره در آن‌ها هست، در این طبقه‌بندها قبول شده و وارد طبقه‌بندهای بعدی می‌شوند. طبقه‌بندهای بعدی با بررسی تعداد بیشتری از ویژگی‌ها، کندر و سخت‌گیرانه‌تر خواهند بود. با قبول شدن یک پنجره از تصویر در تمامی طبقه‌بندها آن پنجره به عنوان چهره تشخیص داده می‌شود.



شکل ۱۷-۲ روند آبشاری طبقه‌بندی اعمال شده بر روی پنجره‌های تصویر

در الگوریتم ارایه شده با وجود این که مجموع تعداد ویژگی‌های مورد بررسی در تمامی طبقه‌بندها) ۶۰۶۱ عدد است، اما در عمل با اجرای الگوریتم بر روی پایگاه داده‌ی MIT+CMU [48] به طور میانگین بر روی هر پنجره از تصویر تنها ۱۰ ویژگی بررسی شد. بدین ترتیب یک سری طبقه‌بند به صورت متوالی، به طرز چشم‌گیری سرعت عمل مکان‌یابی چهره افزایش پیدا می‌کند و الگوریتم به سمت نواحی از تصویر که احتمال وجود چهره در آن بیشتر است کشیده خواهد شد و در نهایت چهره‌های تصویر را مکان‌یابی خواهد کرد. نمونه‌ی خروجی این الگوریتم در شکل ۱۸ آمده است. همان‌طور که ملاحظه می‌شود، در این تصویر اکثر چهره‌ها به درستی تشخیص داده شده است.



شکل ۱۸-۲ نمونه‌ی از نتایج الگوریتم [48]

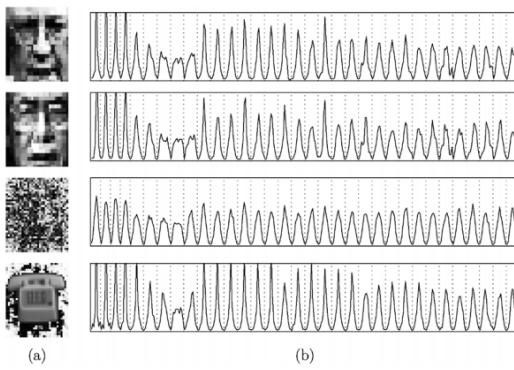
پس از این، مقالات بسیاری با الهام از ایده‌های کلیدی این پژوهش اقدام به بهبود الگوریتم‌های تشخیص چهره کردند از جمله [49] که در سال ۲۰۱۴ ارائه شد و با الهام‌گیری از ایده‌ی کلیدی سوم و دنباله‌ای ۲۲ تایی از طبقه‌بندها و ترکیب آن با ایده‌های جدیدتر، به بهترین دقیق در زمان خودش دست یافت.

در روشی دیگر که دسته‌ی «روش‌های بر پایه‌ی تصاویر» قرار می‌گیرد، از هیستوگرام‌های طیفی^۱ و ماشین بردار پشتیبانی (SVM)^۲ بهره گرفته شده است [50]. در این مقاله، با اعمال ۳۳ فیلتر مختلف بر روی تصویر و رسم هیستوگرام‌های آن‌ها، یک نمایش جدید از تصویر به دست آمده است. به عنوان مثال در شکل ۱۹-۲ هیستوگرام‌های خروجی چهار نمونه عکس آورده شده است. طبق این پژوهش، پس از نمایش تصاویر به این صورت، تصاویر چهره الگوی هیستوگرامی خاصی به خود می‌گیرند. بدین ترتیب با آموزش یک طبقه‌بند^۳ به کمک SVM، می‌توان یک پنجره از تصویر (که اندازه‌ی آن 21×21 در نظر گرفته شده) را دریافت کرد و چهره بودن یا نبودن آن را مشخص کرد. آموزش این مدل توسط ۴۵۰۰ تصویر چهره و ۸۰۰۰ تصویر غیر چهره انجام شده است.

¹ Spectral histograms

² Support vector machine

³ Classifier



شکل ۲-۱۹ هیستوگرام‌های طیفی چهار تصویر مختلف. (a) تصاویر ورودی (b) هیستوگرام‌های متناظر هریک از چهار عکس

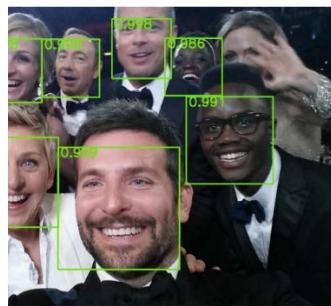
این الگوریتم در مرحله‌ی آزمایش، یک تصویر حاوی چهره را دریافت می‌کند و یک پنجره با اندازه‌ی ۲۱ پیکسل را بر روی آن می‌لغزاند و هیستوگرام‌های هر زیر-تصویر را ذخیره می‌کند. سپس ماتریس تشکیل شده از نتایج هیستوگرام‌ها به عنوان ورودی به مدل SVM داده می‌شود. این مدل نیز به عنوان خروجی، پنجره‌هایی بر روی چهره قرار گرفته‌اند را علامت می‌زند. این اعمال علاوه بر روی تصویر خام دریافت شده بر روی دو تصویر دیگر که به ترتیب نتیجه‌ی اعمال یک مرحله و دو مرحله اعمال فیلتر گوسی هستند نیز اعمال می‌شود تا به این ترتیب الگوریتم تا حدودی نسبت به تغییر اندازه‌ی چهره نیز مقاوم می‌شود. در شکل ۲۰-۲ نمونه‌ای از نتایج این الگوریتم آورده شده است. در این تصویر تمامی ۹ چهره‌ی موجود در تصویر به درستی تشخیص داده شده‌اند؛ اما یک بخش از تصویر نیز به اشتباه چهره تشخیص داده شده است. همان‌طور که از تصویر نتیجه پیداست، الگوریتم تا حدودی نسبت به تغییر اندازه‌ی صورت هم مقاوم است.



شکل ۲-۲۰ نمونه‌ای از نتایج الگوریتم ارائه شده توسط [51]

روش [51] نیز نمونه‌ی دیگری از الگوریتم مکان‌یابی چهره بر پایه‌ی تصویر است که در اخیرا بر پایه‌ی یادگیری عمیق پیاده‌سازی شده و در عین سادگی نسبت به سایر روش‌ها، بسیار موفق ظاهر شده است. شبکه‌ی عصبی مورد استفاده در این مقاله، بر پایه‌ی معماری معروف AlexNet [52] بنا شده و مانند نسخه‌ی استاندارد آن ورودی‌هایی با ابعاد 227×227 دریافت می‌کند. در این مقاله با استفاده از روش‌های افزایش داده، در نهایت مرحله‌ی آموزش با تعداد ۲۰۰ هزار

تصویر از چهره و ۲۰ میلیون تصویر غیر چهره انجام شده است. این روش موفق شد بدون نیاز به اطلاعاتی مانند نشانه‌گذاری صورت^۱ در دقت مکان‌بایی چهره گامی به سمت جلو برد [51]. در شکل ۲۱-۲ یک نمونه از خروجی الگوریتم ملاحظه می‌شود.



شکل ۲۱-۲ نمونه‌ای از خروجی الگوریتم ارائه شده در [51]

در ادامه‌ی این بخش، سایر ویژگی‌های کلاسیک که بیشتر در الگوریتم‌های دهه‌ی ۲۰۰۰ میلادی و سال‌های پیش از آن مورد استفاده قرار می‌گرفته است را بررسی خواهیم کرد.

۲-۳-۲ روش‌های بر پایه‌ی ویژگی

الگوریتم‌هایی که بر پایه‌ی ویژگی‌های تصویر تشخیص چهره را انجام می‌دهند را می‌توان بر اساس نوع ویژگی‌هایی که استفاده می‌کنند به چند دسته تقسیم کرد: ۱) تحلیل‌هایی در سطح پایین، ۲) تحلیل ویژگی‌ها و ۳) مدل‌های شکل فعال.^۲ در تحلیل‌های سطح پایین، الگوریتم‌ها عموماً عکس را بر پایه‌ی ویژگی‌های پیکسلی آن (نظیر روشنایی و رنگ آن) تقسیم‌بندی می‌کنند و طبیعت این ویژگی‌ها به گونه‌ای است که مبهم هستند. در تحلیل به کمک ویژگی‌ها، مدل با استفاده از حالت کلی هندسه‌ی صورت، به مفهومی کلی از چهره‌ی انسان دست پیدا می‌کند و ابهام آن‌ها به مراتب از ویژگی‌های سطح پایین کمتر است. در نهایت دسته‌ی مدل‌های شکل فعال است که با مدل مارها^۳ که در دهه‌ی ۱۹۸۰ آغاز شده و تا مدل‌های جدیدتری مانند PDM^۴ نیز ادامه داشته است که برای ردیابی لب و مردمک چشم نیز می‌تواند به کار رود [53]. در ادامه هریک از این سه دسته را به صورت مختصر مورد بررسی قرار خواهیم داد:

(۱) تحلیل‌های در سطح پایین خود می‌تواند به شکل‌های مختلف صورت بگیرد:

- لبه‌ها: به عنوان یکی از اولیه‌ترین ویژگی‌های پردازش تصویر، لبه‌ها برای اولین بار توسط Sakai و همکارانش در [54] به منظور مکان‌بایی چهره کار رفته است. روش‌های بسیار دیگری نیز در ادامه این راه را ادامه دادند. به عنوان مثال در مقاله‌ی [55] با استفاده از برچسب‌گذاری لبه‌ها به عنوان نیمه‌ی راست صورت، نیمه‌ی چپ صورت و خطوط مربوط به موها و در نظر گرفتن نسبت طلایی در چهره، چهره‌ها را در تصویر مکان‌بایی می‌کنند.
- روشنایی پیکسل‌ها: با در نظر گرفتن ساختار چهره‌ی انسان و این که معمولاً ابروها، ابروها و مردمک‌های چشم معمولاً در تصاویر تیره‌تر ظاهر می‌شوند، الگوریتم‌هایی وجود دارند (مانند [56]) که بر مبنای تفاوت نسبی روشنایی بخش‌های مختلف تصویر عمل می‌کنند.
- رنگ: بسیاری از روش‌های مکان‌بایی چهره در عکس بر مبنای نمایش RGB رنگ‌ها کار می‌کنند. این نمایش بر مبنای مقادیر سه رنگ قرمز و سبز و آبی کار می‌کند و معمول‌ترین روش برای نمایش رنگ‌های رنگ‌هاست. در این روش معمولاً برای بی‌اثر کردن شدت روشنایی، مقادیر رنگ‌ها را نرمال می‌کنند و نسبت آن‌ها را در نظر می‌گیرند. با این کار معمولاً رنگ پوست در یک

¹ Facial landmarks

² Active shape models

³ Snakes

⁴ Point Distributed Models

محدوده‌ی مشخص قرار خواهد گرفت [57]. در پژوهش‌های دیگر نیز روش‌هایی بر مبنای سایر نمايش‌های رنگ (مانند HSV [58]) نیز ارائه شده است.

۲) **تحلیل ویژگی‌ها:** همان‌طور که گفته شد تحلیل‌ها در سطح پایین ممکن است با ابهام همراه باشند به عنوان مثال در تحلیل ویژگی‌ها با استفاده از رنگ ممکن است بخش‌هایی از پس‌زمینه نیز که هم‌رنگ پوست هستند به اشتباه تشخیص داده شوند. با استفاده از تحلیل در سطح ویژگی‌ها می‌توان بخشی از کمبودهای تحلیل در سطح پایین را جبران کرد. تحلیل در سطح ویژگی‌ها را می‌توان به دو صورت انجام داد [53]:

- **جستجوی ویژگی:** در این نوع جستجو ابتدا به دنبال ویژگی‌های بارز چهره در تصویر می‌گردیم (به عنوان مثال می‌توان ویژگی‌های یک جفت چشم که در کنار یکدیگر قرار گرفته‌اند را مبنا قرار داد [58]). پس از مکان‌یابی این ویژگی‌ها، محدوده‌ای فرضی برای صورت در نظر گرفته می‌شود و با توجه به آن ویژگی‌های جزئی‌تر چهره را نیز در آن محدوده جستجو خواهیم کرد.

- **تطبیق اجزای صورت:** تکیه‌ی بسیاری از روش‌هایی که در بخش «جستجوی ویژگی» مورد بررسی قرار گرفتند، بر اطلاعاتی است که از چهره‌هایی در شرایط یکسان به دست آمده است. در این صورت این الگوریتم‌ها در شرایط محیطی متفاوت و با وجود پس‌زمینه‌های پیچیده به مشکل خواهند خورد. اما در دسته‌ای دیگر از پژوهش‌ها (مانند [60]) اجزای صورت به صورت جداگانه جستجو شده‌اند و ارتباط میان آن‌ها مورد بررسی قرار گرفته است. این روش‌ها برخی پژوهش‌ها دقت به دست آمده از روش جستجوی ویژگی را بهبود بخشیده‌اند.

۳) **مدل‌های شکل فعل:** برخلاف مدل‌هایی که تا به اینجا برای چهره ارائه شد، مدل‌های شکل فعل، بر اساس فیزیک واقعی صورت طراحی شده‌اند و سطح بالاتری از ویژگی‌های ظاهری را مبنا قرار می‌دهد [53]. این مدل‌ها عموماً به این صورت عمل می‌کنند که ابتدا خطوطی از مدل یک چهره به صورت تقریبی در اطراف محل قرارگیری چهره قرار می‌گیرند. پس از آن این خطوط به لبه‌هایی از تصویر که در همان محل استخراج شده‌اند قفل خواهند شد و شکل کلی صورت شکل خواهد گرفت. پس از آن با کمینه کردن مجموع خطاهای داخلی و خارجی سعی می‌شود بهترین حالت ممکن برای قرارگیری خطوط چهره محاسبه شود. خطاهای داخلی موجب می‌شود که شکل کلی خطوط از فرم صورت طبیعی یک انسان خارج نشود و خطاهای خارجی نیز بر اساس قرارگیری خطوط بر روی لبه‌ها در تصویر محاسبه می‌شوند و باعث جایگیری درست خطوط بر روی چهره در تصویر خواهند شد [61]. پژوهش‌های بسیاری سعی در بهبود این روش نیز داشته‌اند و به عنوان مثال با تغییر تابع خطا و در نظر گرفتن مکان چشم‌ها، الگوریتم را نسبت به تغییرات نور و حالت صورت مقاوم‌تر کرده‌اند.

در این مقاله روش‌های مکان‌یابی چهره در تصویر در دوسته‌ی «بر پایه‌ی تصویر» و «بر پایه‌ی ویژگی» مورد بررسی قرار گرفت [53]. اما دسته‌بندی‌های دیگری نیز برای این روش‌های در برخی منابع [62] ارائه شده که بر مبنای آن این روش‌ها به شکل‌های دیگر تقسیم‌بندی می‌شوند. به عنوان مثال در مقاله‌ی [63] به روش‌های مکان‌یابی چهره به چهار دسته‌ی زیر تقسیم‌بندی می‌شوند:

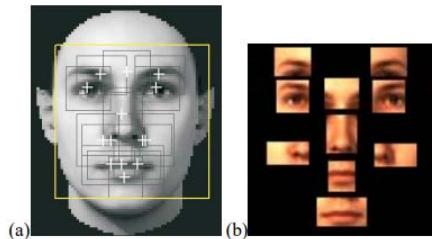
- **بر پایه‌ی دانش:** بر اساس قوانین از پیش تعریف شده، بر مبنای دانش انسان، چهره را در تصویر تشخیص می‌دهد.
- **بر پایه‌ی ویژگی‌های ثابت:** ساختاری از چهره را پیدا می‌کند که نسبت به تغییرات نور و زاویه‌ی دید مقاوم باشند.
- **انتتباق نمونه:** با مقایسه‌ی یک تصویر با نمونه‌های تصویر چهره از پیش ذخیره شده، در مورد چهره بودن یا نبودن تصویر جدید تصمیم می‌گیرد.
- **بر پایه‌ی ظاهر:** مدلی برای چهره بر پایه‌ی نمونه‌های تصویر دیده شده آموزش داده می‌شود و از این مدل برای مکان‌یابی در تصاویر جدید استفاده می‌شود.

۲-۳-۲ تشخیص هویت به کمک چهره

در یک سیستم تشخیص هویت به کمک چهره، پس از مکان‌یابی چهره در تصویر و پیش‌پردازش آن، وارد مرحله‌ی بعدی یعنی استخراج ویژگی از چهره و تشکیل الگوی چهره می‌شود. الگوریتم‌های بازناسی چهره را می‌توان در یک دسته‌بندی کلی به دو

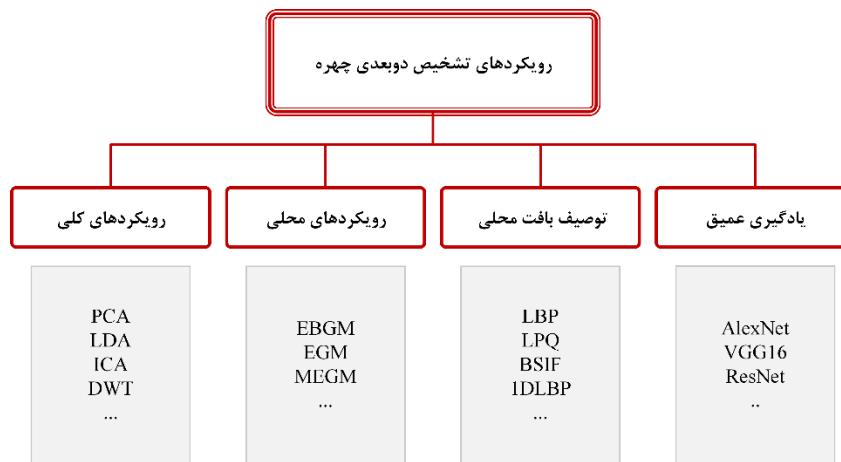
بخش تقسیم‌بندی کرد: ۱) روش‌های کلی ۲) روش‌های بر پایه‌ی اجزای صورت [9]. در روش‌های کلی، ویژگی‌های کل صورت در یک بردار ذخیره می‌شود. این بردار را می‌توان به عنوان ورودی به طبقه‌بند داد. اما در روش‌های بر پایه‌ی اجزاء صورت، هر یک از اجزاء به صورت جداگانه مکان‌یابی شده و ترکیب آن اجزاء با یکدیگر در تشخیص هویت چهره به کار می‌روند.

به عنوان یک روش بارز که «بر پایه‌ی اجزای صورت» پیاده‌سازی شده، می‌توان یکی از روش‌های ارائه شده در پژوهش [63] را بررسی کرد. مزیت این روش نسبت به روش‌های کلی این است که برای تغییر زاویه‌های جزئی در صورت، تغییراتی که در هر یک از اجزاء به تنها‌ی ایجاد می‌شود، به نسبت تغییرات کلی صورت بسیار کمتر است و بدین ترتیب سیستم نسبت به چرخش و تغییر حالت مقامات بیشتر نشان خواهد داد. شکل ۲۲-۲ نمایش دهنده اجزاء مورد استفاده در این الگوریتم تشخیص چهره است. این روش این اجزاء پس از تغییر اندازه با یکدیگر ترکیب شده و پس از آن با اعمال الگوریتم SVM به صورت «یکی در مقابل سایرین» مدلی برای تشخیص چهره از بین یک پایگاه داده آموزش داده شده است.



شکل ۲۲-۲ (a) تمامی ۱۴ جزء صورت که در الگوریتم [63] مکان‌یابی می‌شوند (b) ۱۰ جزء مورد استفاده در الگوریتم شناسایی هویت

روش‌های بازشناسی چهره را می‌توان از جنبه‌های مختلف طبقه‌بندی نمود به همین منظور در مقاله‌های مختلف طبقه‌بندی‌های مختلفی آورده شده که در ادامه به رایج‌ترین آن‌ها می‌پردازیم. طبق مقاله‌ی [22] رویکردهای دوبعدی را می‌توان به چهار دسته مختلف تقسیم کرد. ۱) روش‌های جامع (کلی)، ۲) روش‌های محلی (هندرسی)، ۳) روش‌های مبتنی بر توصیف بافت محلی و ۴) روش‌های مبتنی بر یادگیری عمیق که در شکل ۲۳-۲ نمونه‌هایی از آن‌ها مشاهده می‌شود.



شکل ۲۳-۲ یک طبقه‌بندی رویکردهای شناسایی چهره دوبعدی

در [9] خلاصه‌ای از روش‌ها و پژوهش‌های انجام شده در بازشناسی چهره آورده شده است. طبق این مقاله، الگوریتم‌های بازشناسی چهره را می‌توان در این چهار دسته قرار داد:

- **یادگیری کلی:** در این روش‌ها که بیشتر در دهه‌ی ۱۹۹۰ و اوایل دهه‌ی ۲۰۰۰ میلادی مورد توجه قرار گرفتند، تلاش بر این بود که به کمک یک پرآکنده‌گی فرضی، یک بازنمایی با تعداد ابعاد محدود برای هر چهره ارائه شود. اولین و بارزترین نمونه‌ی روش، «مقادیر ویژه چهره» [20] است که در بخش ۲-۳-۲ بررسی خواهد شد. این روش‌ها تحت شرایط محیطی مختلف معمولاً با مشکل مواجه می‌شوند. این روش که در ابتدای دهه‌ی ۱۹۹۱ میلادی ارائه شد، یکی از زمینه‌های رشد زمینه‌ی بازشناسی چهره به شمار می‌رود [64]. الگوریتم‌های بر پایه‌ی تطابق گراف‌ها^۱، مدل مخفی مارکف^۲، تطابق ویژگی هندسی^۳، تطابق نمونه‌ها^۴، نقشه‌ی خطوط لبه^۵ و همچنین SVM نیز از دیگر روش‌هایی هستند که در مسئله‌ی تشخیص هویت به کمک چهره به کار رفته‌اند.
- **ویژگی‌های محلی:** در دهه‌ی ۲۰۰۰ میلادی، روش‌هایی بر پایه‌ی ویژگی‌های محلی (مانند نتایج فیلترهای گابور) ارائه شد. این روش‌ها تا حدودی نسبت به شرایط محیطی مختلف مقاومت نشان می‌دادند اما فشردگی کافی را نداشتند و همچنین قابلیت ایجاد تمایز در آن‌ها کافی نبود. روش [65] که بر پایه‌ی فیلترهای گابور ارائه شد، به عنوان یک روش بارز در این بخش شناخته می‌شود.
- **یادگیری کم‌عمق:** در اوایل دهه‌ی ۲۰۱۰ میلادی روش‌هایی ارائه شدند که در آن‌ها توصیف‌گرهای محلی بر پایه‌ی یادگیری معرفی شدند. در واقع در این روش‌ها با توجه به پایگاه داده، فیلترهایی آموزش داده می‌شوند که بیشترین ایجاد تمایز را ایجاد می‌کنند. اما هنوز این روش‌ها مقاومت کافی در برابر تبدیل‌های غیر خطی و پیچیده‌ی چهره را نداشتند. پژوهش [66] نماینده این روش ارائه شده در این زمینه است.
- **یادگیری عمیق:** در سال ۲۰۱۴ میلادی با ارائه‌ی الگوریتم DeepFace [67] توسط تیم تحقیقاتی شرکت Facebook سری دیگری از روش‌های بازشناسی چهره بر پایه‌ی یادگیری عمیق کلید خورد. در این روش‌ها برخلاف روش‌های یادگیری کم‌عمق، تعداد لایه‌های زیادی به صورت متوالی به منظور استخراج ویژگی و تبدیل آن‌ها در نظر گرفته شده و بدین ترتیب در سطوح ویژگی‌های مختلفی با سطوح پیچیدگی مختلف شناسایی می‌شوند و این ویژگی‌ها نسبت به حالت چهره و شرایط محیطی نیز مقاوم هستند. لازم به ذکر است DeepFace برای اولین بار دقت الگوریتم‌های بازشناسی چهره را به دقت بازشناسی چهره توسط انسان (حدود ۹۷ درصد) رسانید. پس از ارائه‌ی DeepFace الگوریتم‌های دیگری نیز بر پایه‌ی یادگیری عمیق بازشناسی چهره کردند از جمله‌ی این روش‌ها می‌توان به FaceNet [70] و VGGFace2 [69] و FaceID [71] اشاره کرد.

۲-۳-۱-۱ الگوریتم مقادیر ویژه چهره

در سال‌های ۱۹۸۷ و ۱۹۹۰ میلادی در مقالات [72] و [73] با استفاده از تحلیل مولفه‌های اصلی (PCA)^۶، یک بازنمایی بهینه از تصویر چهره به کمک برداری از اعداد را ارائه شد. دی این مقالات نشان داده شد که تصویر هر چهره را می‌توان با همراه داشتن یک مجموعه تصویر استاندارد و یک بردار از ضرایب نمایش داد. پس از آن در سال ۱۹۹۱ میلادی در پژوهش [74] با الهام از پژوهش‌های یاد شده روشنی با عنوان «مقادیر ویژه چهره» (EigenFace) برای طبقه‌بندی تصاویر چهره ارائه شد.

در این روش ابتدا هریک از تصاویر موجود در پایگاه داده، از حالت آرایه‌ی دو بعدی $N \times N$ به برداری بزرگ با اندازه‌ی N^2 تبدیل می‌شود. بدین ترتیب هر تصویر در فضایی با تعداد ابعاد بسیار زیاد قابل نمایش خواهد بود و هر تصویر چهره بیانگر یک نقطه در این فضا خواهد بود. پس از این مرحله، با استفاده از الگوریتم PCA و با در نظر گرفتن واریانس هر مولفه از بردار، تعدادی بردار با اندازه‌ی N^2 (مقادیر ویژه چهره) محاسبه می‌شوند که بیانگر «عناصر تشکیل دهنده‌ی صورت استاندارد» هستند و با توجه به

¹ Graph matching

² Hidden Markov model

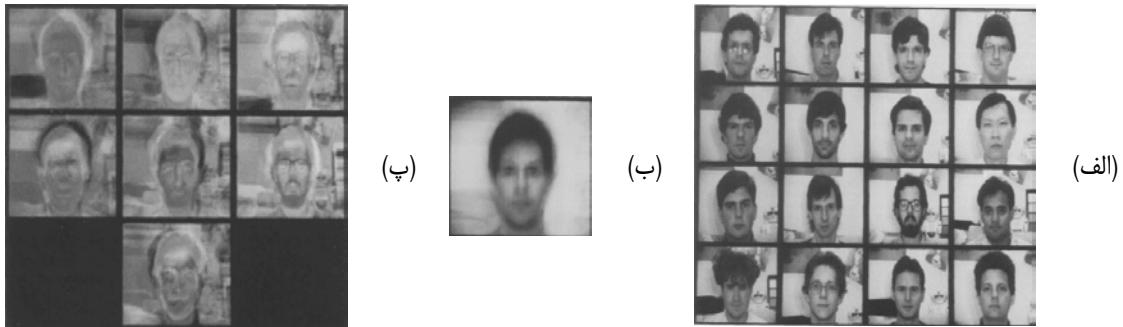
³ Geometrical feature matching

⁴ Template matching

⁵ Line edge map (LEM)

⁶ Principal component analysis

تمامی تصاویر موجود در پایگاه داده محاسبه شده‌اند. در شکل ۲۴-۲ یک نمونه پایگاه داده و مقادیر ویژه چهره مربوط به آن دیده می‌شود.



شکل ۲۴-۲ (الف) پایگاه داده‌ی نمونه (ب) تصویر میانگین این پایگاه داده (پ) هفت نمونه مقادیر ویژه چهره از همان پایگاه داده [75]

پس از محاسبه‌ی مقادیر ویژه چهره، تمامی تصاویر چهره در پایگاه داده را می‌توان با ترکیبی از آن‌ها نمایش داد. به عنوان مثال یکی از چهره‌های پایگاه داده ممکن است با تصویر میانگین به علاوه‌ی ۱۰ درصد از اولین مقدار ویژه چهره، به علاوه‌ی ۵۶ درصد از دومین مقدار ویژه چهره، منهای ۳ درصد از سومین مقدار ویژه چهره و الی آخر ساخته شود. بدین ترتیب هر چهره را می‌توان به صورت یکتا و با ترکیب خاصی از این ضرایب نمایش داد که نسبت به تصویر اصلی به صورت قابل توجهی حجم کمتری اشغال می‌کند. این تصویر را می‌توان با ترکیب خطی از تصاویر مقادیر ویژه چهره بازسازی کرد. در شکل ۲۵-۲ نمونه‌ای از یک تصویر و نمونه‌ی بازسازی شده‌ی آن توسط ضرایب ملاحظه می‌شود.



شکل ۲۵-۲ سمت چپ: یک اصلی یک چهره سمت راست: تصویر چهره بازسازی شده به کمک مقادیر ویژه چهره و ضرایب محاسبه شده برای آن چهره [74]

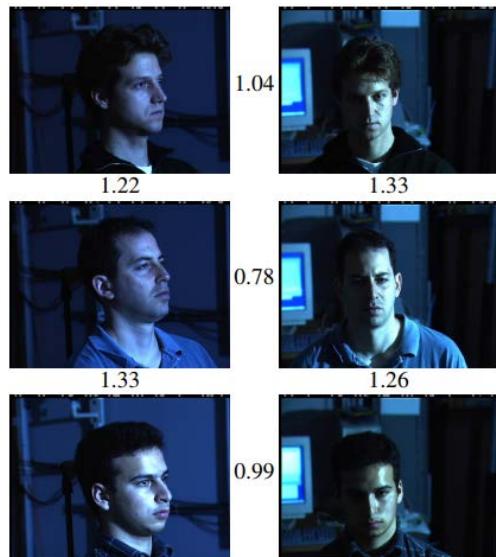
در سیستم زیست‌سنگی ارائه شده در مقاله [74]، ابتدا در فاز ثبت‌نام به ازای هر فرد تعداد چهار تصویر وجود دارد. بردار ضرایب مقادیر ویژه چهره به ازای تک تک این چهار تصویر محاسبه می‌شود و در نهایت با میانگین‌گیری از آن‌ها، یک بردار برای نمایش آن فرد در مدل به دست خواهد آمد. بدین ترتیب، در فضایی با ابعاد تعداد مقادیر ویژه چهره، هر فرد ثبت‌نام شده یک نقطه‌ی متناظر خواهد داشت. اما در فاز مقایسه، با آمدن یک تصویر چهره‌ی جدید، ابتدا بردار ضرایب مقادیر ویژه چهره آن محاسبه می‌شود. بدین ترتیب برای عکس جدید نیز یک نقطه در فضایی با ابعاد تعداد مقادیر ویژه چهره خواهیم داشت. فاصله‌ی اقلیدسی این نقطه تا نقاط متناظر هریک از افراد ثبت‌نام شده محاسبه می‌شود. تصویر جدید با نزدیک‌ترین نقطه در این فضا (در صورتی که فاصله‌ی آن از یک آستانه‌ی به خصوص بالاتر نباشد) تطابق داده خواهد شد [74].

در این مقاله دقت‌های ۹۶، ۸۵ و ۶۴ درصد به ترتیب برای شرایط تغییرات نور، جهت و اندازه گزارش شده است. پایگاه داده‌ی این مقاله شامل ۲۵۰۰ تصویر از ۱۶ فرد بوده است. وجود پس‌زمینه در این تصاویر، تاثیر بسیاری در دقت گزارش شده در این پژوهش گذاشته است [9]. پژوهش‌هایی دیگر پس از ارائه‌ی این روش، سعی در بهبود آن نیز داشته‌اند. به عنوان مثال [75] یکی از

این مقالات است که با الهام از این الگوریتم، الگوریتمی «بر پایه‌ی اجزای صورت» (رجوع شود به بخش ۲-۳-۲) ارائه داده است که در آن ویژگی‌هایی مانند مقادیر ویژه چهره را نه برای کل صورت بلکه برای اجزائی از صورت (به عنوان مثال چشم، بینی و دهان) به صورت جداگانه محاسبه کرده است.

۲-۳-۲ FaceNet الگوریتم

الگوریتم FaceNet [71] در سال ۲۰۱۵ توسط تیم تحقیقاتی شرکت Google ارائه شد. این روش از یادگیری عمیق برای بازشناسی چهره استفاده کرده. بر خلاف روش DeepFace [67] که یک مدل سه بعدی از چهره ساخته و برای همترازی و شناسایی از آن بهره می‌گیرد، FaceNet روش ساده‌تری برای بازشناسی چهره ارائه کرده و با افزایش تعداد پارامترها و لایه‌های شبکه، بار پردازشی بیشتری را بر روی آن قرار داده است. مقاومت الگوریتم در مقابل تغییرات نور و زاویه در شکل ۲۶-۲ دیده می‌شود. لازم به ذکر است که در این پژوهش در عکس‌های ورودی سیستم، از محدوده‌ی چهره گرفته شده‌اند (مانند تصاویر شکل ۲۶-۲) و فرض شده است که عمل مکانیابی پیش از آن انجام شده است.



شکل ۲۶-۲ اعداد بین هریک از دو تصویر، بیانگر فاصله‌ی صفر به معنای تصاویر همسان و فاصله‌ی ۴ به تصاویر دو هویت مجزا تعلق می‌گیرد. همان‌طور که ملاحظه می‌شود با در نظر گرفتن آستانه‌ی ۱/۱ می‌توان دسته‌بندی افراد در تصاویر را به درستی انجام داد.

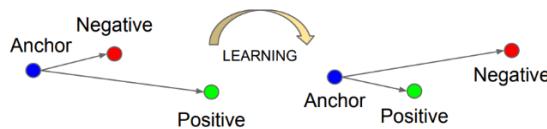
در این پژوهش با الهام از دو معماری شبکه‌های عصبی Inception [76] و Zeiler&Fergus [77] شبکه‌هایی با ساختار شکل ۲۷-۲ طراحی شده‌اند که پس از آموزش، مانند تابعی به فرم $f(x)$ عمل می‌کند که با دریافت یک تصویر از محدوده‌ی چهره مانند x یک بازنمایی از آن در فضای ویژگی \mathbb{R}^d تولید می‌کند به طوری که فاصله‌ی مجدد اقلیدسی^۱ برای تمام چهره‌ها (مستقل از شرایط تصویر) برای هویت‌های یکسان کم و برای هویت‌های متفاوت زیاد است. علاوه بر این، در این پژوهش فضای n -بعدی اقلیدسی (x) به یک ابرکره‌ی n -بعدی محدود شده است. به عنوان مثال می‌توان فرض کرد: $\|f(x)\|_2 = 1$.

^۱ Squared Euclidean distance



شکل ۲۷-۲ ساختار شبکه: ورودی داده‌ها به صورت دسته‌ای، شبکه‌ی عصبی پیچشی، نرم‌افزاری L2 در این مرحله بازنمایی چهره تولید می‌شود. در طول روند آموزش، بخش هزینه‌ی سه‌گانه پس از همه‌ی این بخش‌ها قرار می‌گیرد.

برای آموزش شبکه، در این پژوهش ازتابع هزینه‌ی سه‌گانه^۱ استفاده شده است. به کمک تابع هزینه‌ی سه‌گانه می‌توان شبکه را به گونه‌ای آموزش داد که یک تصویر مبنا (anchor) x_i^a متعلق به یک فرد خاص، به تمامی تصاویر مثبت (positive) x_i^p که متعلق به همان شخص است، نزدیک‌تر باشد تا هر تصویر منفی (negative) x_i^n که متعلق به هویتی دیگر است (شکل ۲۸-۲).



شکل ۲۸-۲ تابع هزینه‌ی سه‌گانه، فاصله‌ی بین تصاویر متعلق به یک هویت را کاهش داده و فاصله‌ی بین تصاویر متعلق به افراد مختلف را افزایش می‌دهد.

بدین ترتیب لازم است که:

$$\|x_i^a - x_i^p\|_2^2 + \alpha < \|x_i^a - x_i^n\|_2^2, \forall (x_i^a, x_i^p, x_i^n) \in \mathcal{T}$$

که در آن α حداقل فاصله‌ای است که می‌خواهیم بین مجموعه‌ی مثبت و منفی داشته باشیم، \mathcal{T} مجموعه‌ی همه‌ی سه‌تایی‌ها در داده‌های آموزش است. در این صورت تابع هزینه به این صورت تعریف می‌شود:

$$L = \sum_i^N \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+$$

بر اساس این پژوهش، تولید همه‌ی سه‌تایی‌های ممکن و استفاده از آن‌ها در مرحله‌ی آموزش، موجب کند شدن همگرایی می‌شود. پس به منظور همگرایی سریع‌تر، در مرحله‌ی آموزش تنها سه‌تایی‌هایی انتخاب می‌شوند که شرط یاد شده را نقض می‌کنند. اما همیشه انتخاب نزدیک‌ترین تصویر منفی و دورترین تصویر مثبت ممکن است موجب یادگیری نامناسب شود. به همین دلیل روشی که در نهایت در این پژوهش استفاده شده است به این صورت است: در هر دسته از داده‌ها که همزمان وارد شبکه می‌شوند، از تعدادی از افراد، هر فرد ۴۰ تصویر صورت قرار داده شده (به همراه چندین تصویر از افراد دیگر) و همه‌ی جفت تصویرهای مبنا-مثبت ممکن به همراه نزدیک‌ترین تصویر منفی به عنوان سه‌تایی برای آموزش آن دسته در نظر گرفته شده‌اند.

یکی از معماری‌هایی که در این پژوهش به عنوان شبکه‌ی عصبی پیچشی استفاده شده است، بر پایه‌ی معماری Zeiler&Fergus بوده و شامل ۲۲ لایه است. جزئیات این معماری که شامل ۱۴۰ میلیون پارامتر است در جدول ۳-۲ آورده شده است.

¹ Triplet loss function

جدول ۲-۳ لایه‌های یکی از شبکه عصبی‌های به کار رفته در مدل [71]

layer	size-in	size-out	kernel	param	FLPS
conv1	220×220×3	110×110×64	7×7×3, 2	9K	115M
pool1	110×110×64	55×55×64	3×3×64, 2	0	
rnorm1	55×55×64	55×55×64		0	
conv2a	55×55×64	55×55×64	1×1×64, 1	4K	13M
conv2	55×55×64	55×55×192	3×3×64, 1	111K	335M
rnorm2	55×55×192	55×55×192		0	
pool2	55×55×192	28×28×192	3×3×192, 2	0	
conv3a	28×28×192	28×28×192	1×1×192, 1	37K	29M
conv3	28×28×192	28×28×384	3×3×192, 1	664K	521M
pool3	28×28×384	14×14×384	3×3×384, 2	0	
conv4a	14×14×384	14×14×384	1×1×384, 1	148K	29M
conv4	14×14×384	14×14×256	3×3×384, 1	885K	173M
conv5a	14×14×256	14×14×256	1×1×256, 1	66K	13M
conv5	14×14×256	14×14×256	3×3×256, 1	590K	116M
conv6a	14×14×256	14×14×256	1×1×256, 1	66K	13M
conv6	14×14×256	14×14×256	3×3×256, 1	590K	116M
pool4	14×14×256	7×7×256	3×3×256, 2	0	
concat	7×7×256	7×7×256		0	
fc1	7×7×256	1×32×128	maxout p=2	103M	103M
fc2	1×32×128	1×32×128	maxout p=2	34M	34M
fc7128	1×32×128	1×1×128		524K	0.5M
L2	1×1×128	1×1×128		0	
total				140M	1.6B

الگوریتم ارائه شده بر روی پایگاه داده‌های LFW¹ و چهره‌های یوتیوب² آزمایش شده و بر روی آن‌ها به ترتیب به دقتهای 99.63% و 95.12% دست یافته است. شکل ۲-۹ اشتباهاتی که توسط الگوریتم ارائه شده در پایگاه داده LFW رخ داده است را نمایش می‌دهد.



شکل ۲-۹ تمامی جفت تصویرهای پایگاه داده LFW که به اشتباه طبقه‌بندی شدند. جفت تصاویر سه سطر اول، به اشتباه یکسان تشخیص داده شده‌اند و سایر جفت تصاویر به اشتباه متفاوت تشخیص داده شده‌اند.

در ادامه پیشرفته‌ترین روش‌های بازنگاری چهره به همراه توضیحات آن آورده شده است.

DeepFace

¹ Labeled Faces in the Wild

² YouTube Faces DB

از یک شبکه عصبی عمیق نه لایه با بیش از ۱۲۰ میلیون پارامتر برای بازشناسی چهره استفاده می‌کند. از خطای Softmax برای آموزش شبکه استفاده شده است و مجموعه داده‌های آموزش، یک مجموعه داده خصوصی با چهار میلیون تصویر چهره با بیش از ۴۰۰۰ هویت است. این سامانه همچنین روش پیش‌پردازشی موثری را که از یک مدل سه‌بعدی برای تراز کردن چهره‌ها در موقعیت استاندارد چهره استفاده می‌شود، پیاده‌سازی می‌کند. به طور خلاصه، موقفيت DeepFace به سه عامل اصلی مربوط می‌شود: (۱) مرحله پیش‌پردازش دقیق، (۲) معماری شبکه و (۳) داده‌های آموزش در مقیاس بزرگ. علاوه بر سامانه‌ی پیشنهادی، DeepFace همچنین یک سیستم تأیید چهره انتها به انتها^۱ را با استفاده از یک شبکه Siamese رائه می‌دهد. پس از آموزش، شبکه شامل یک لایه طبقه‌بندی است که برای تولید ویژگی برای دو تصویر به طور همزمان، تکرار می‌شود. بردارهای ویژگی تولید شده برای تصمیم‌گیری اینکه آیا دو شخص هستند با یکدیگر مقایسه می‌شوند.

VGGFace

VGGFace با الهام از VGGNet که نشان داد پیچیدگی‌های عمیق‌تر می‌توانند در تشخیص تصویر در مقیاس بزرگ موثرتر باشند، طراحی شده است، VGGFace همان مفهوم را برای بازشناسی چهره به کار می‌برد. نویسنده‌گان از نسخه اصلاح شده معماری ارائه شده در VGGNet استفاده کرده‌اند و روی مجموعه داده VGGFace آموزش داده‌اند. نویسنده‌گان دوتابع محاسبه خط، softmax triplet را ارزیابی کرده‌اند و نتیجه گرفته‌اند که از خطای سه‌گانه قطعاً عملکرد کلی بهتری را ارائه می‌دهد. با این وجود، گزارش شده که آموزش شبکه برای طبقه‌بندی با خطای softmax، آموزش را به میزان قابل توجهی آسان و سریع‌تر می‌کند.

انطباق الگو^۲

بعداً از سامانه‌ی VGGFace برای یادگیری انتقالی با تطبیق الگو استفاده شد. در این پیاده‌سازی، ویژگی‌های CNN عمیق حاصل از VGGNet از پیش آموزش دیده با SVM‌های خطی آموزش دیده، در زمان آزمون ترکیب می‌شود. گزارش شده است که های خطی SVM one-vs-rest، قدرت تفکیک فضای ویژگی را افزایش می‌دهند.

Baidu

در این مقاله شبکه‌ای را ارائه می‌دهند که شامل ۹ لایه‌ی پیچشی است که با محاسبه خطای سه‌گانه آموزش دیده‌اند. این سامانه تقریباً از یک دقت کامل بر روی مجموعه داده LFW خبر می‌دهد. نویسنده‌گان به این نتیجه رسیدند که خطای سه‌گانه، برای بررسی چهره مناسب‌تر است.

DLIB

كتابخانه‌ای است که به زبان C++ نوشته شده است و مولفه‌های نرم‌افزاری را با هدف قرار دادن تخصص‌هایی مانند داده‌کاوی، یادگیری ماشین، پردازش تصویر و جبر خطی ارائه می‌دهد. این كتابخانه شامل یک مولفه‌ی بازشناسی چهره است که از نسخه اصلاح شده ResNet-34 استفاده می‌کند. بردارهای ویژگی خروجی ۱۲۸ بعدی هستند و شبکه با استفاده از خطای سه‌گانه آموزش داده می‌شود. این شبکه بر روی مجموعه داده‌ای با ۳ میلیون تصویر آموزش دیده است.

مولفه‌ی بازشناسی چهره‌ی Dlib از آموزش انتقالی استفاده می‌کند تا انعطاف‌پذیری را به کاربران ارائه دهد. در طی مراحل ثبت‌نام، مدل از قبل آموزش دیده، بردارهایی را برای تفسیر تصاویر چهره تولید و ذخیره می‌کند. در طی فرایند شناسایی، فاصله اقلیدسی بین بردار ویژگی نمونه ورودی و هر یک از بردارهای ویژگی ذخیره شده محاسبه می‌شود. در طول طبقه‌بندی، اگر فاصله محاسبه شده زیر یک آستانه از پیش تعیین‌شده باشد، دو چهره دارای یک هویت هستند.

OpenFace

یک سامانه بازشناسی چهره تحت مجوز Apache 2.0 است. این سامانه با هدف از بین بردن فاصله بین سامانه‌های بازشناسی چهره در دسترس عموم و پیشرفت‌های سامانه‌های خصوصی با کیفیت بالا، توسعه یافته است. این سامانه مبتنی بر مفاهیمی است که در FaceNet و GoogleNet معرفی شده است. OpenFace از نسخه اصلاح شده شبکه nn4 از GoogleNet استفاده می‌کند که

¹ end-to-end

² Template adaptation

در FaceNet نیز مورد استفاده قرار گرفته است. DNN با استفاده از خطای سه‌گانه آموزش داده می‌شود. بردارهای ویژگی خروجی به دست آمده از این مدل آموزش دیده دارای ۱۲۸ بعد هستند. طبقه‌بندی چهره با استفاده از SVM خطی انجام می‌شود. با توجه به تصاویر چهره دارای برچسب از داده‌های آموزش، این سامانه برای هر چهره بردارهای مشخصه تولید می‌کند. سپس، بردارهای ویژگی به SVM داده می‌شوند که مدلی را بر اساس بردارهای ویژگی چهره ایجاد می‌کند. هنگامی که یک بردار ویژگی‌های چهره از یک تصویر چهره ناشناخته ارائه می‌شود، مدل SVM چهره ناشناخته را طبقه‌بندی می‌کند.

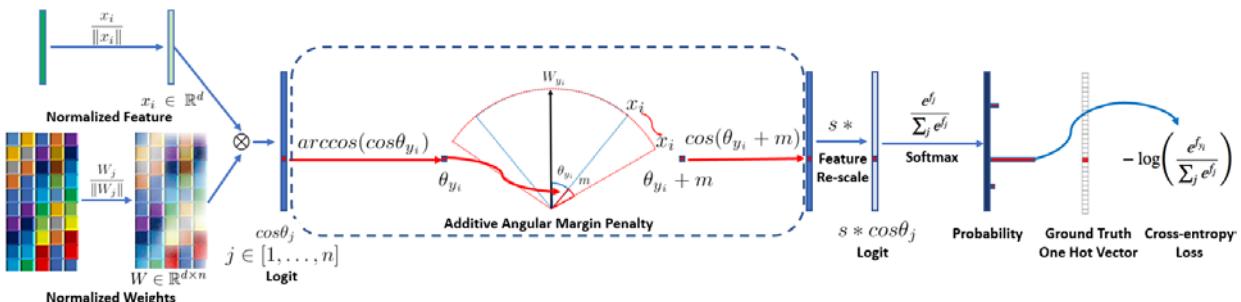
CosFace و SphereFace

دو سامانه‌ی بازشناسی چهره هستند که به ترتیب برای معرفی خطای CosFace و SphereFace استفاده شدن. هر دو سیستم از معماری ResNet-64 استفاده می‌کنند و با CASIA-WebFace آموزش داده شده‌اند. علاوه بر این، CosFace سیستم را با مجموعه داده خصوصی دیگری آموزش می‌دهد و عملکرد بالاتری را گزارش می‌دهد.

ArcFace

مجموعه‌ای از DNN‌ها (ResNet-100, ResNet-50 و ResNet-34) را همراه با خطای ArcFace پیاده‌سازی می‌کند. این سیستم یک بردار ویژگی ۵۱۲ بعدی را برای تصاویر چهره تولید می‌کند. DNN‌ها روی نسخه اصلاح شده مجموعه داده Ms Celeb dataset آموزش دیدند. در یک سری نتایج آزمایشی، نویسنده‌گان نشان می‌دهند که این پیاده‌سازی از اکثر نتایج پیشرفتی گزارش شده بهتر است.

یکی از چالش‌های اصلی در یادگیری ویژگی‌ها با استفاده از شبکه‌های عصبی پیچشی عمیق برای بازشناسی چهره در مقیاس بزرگ، طراحی توابع خطای مناسب است که قدرت تشخیص را افزایش می‌دهد. اخیراً یک روش تحقیق متداول این است که حاشیه‌ها را به منظور افزایش حداکثر قابلیت تفکیک چهره در توابع خطای ثابت به کار ببرند. در این مقاله، برای به دست آوردن ویژگی‌های متمایز‌کننده برای بازشناسی چهره، از یک خطای حاشیه‌ای زاویه‌ای افزایشی^۱ (ArcFace) استفاده می‌شود. نتایج نشان می‌دهد که ArcFace به طور مداوم از پیشرفت‌های ترین عملکرد برخوردار است و می‌تواند به راحتی با سربار محاسبه ناچیز پیاده‌سازی شود. شکل ۳۱-۲ فرایند محاسبه خطا و آموزش شبکه را نمایش می‌دهد. ضرب نقطه‌ای بین ویژگی DCNN و آخرین لایه کاملاً متصل که برابر با فاصله کسینوسی ویژگی و وزن بعد از نرمال‌سازی است، محاسبه می‌شود. از تابع معکوس کسینوس برای محاسبه زاویه بین ویژگی فعلی و وزن هدف استفاده می‌شود. پس از آن، یک زاویه‌ی حاشیه‌ای اضافی به زاویه هدف اضافه می‌شود و دوباره تابع کسینوس به دست آورده می‌شود (logit). سپس، همه logit‌ها را در مقیاس ویژگی s ضرب می‌کند و مراحل بعدی دقیقاً مشابه محاسبه خطای softmax است.



شکل ۳۰-۲ آموزش DCNN برای بازشناسی چهره تحت نظرات خطای [40]

¹ Additive Angular Margin Loss

جدول ۴-۲ نتایج مدل‌های مختلف را بر روی دو مجموعه داده‌ی LFW و YTF نشان می‌دهد.

جدول ۴-۲ عملکرد روش‌های مختلف بر روی دو پایگاه داده LFW و YTF

Method	LFW	YTF
DeepID	99.47	93.20
Deep Face	97.35	91.4
VGG Face	98.95	97.30
FaceNet	99.63	95.10
Baidu	99.13	-
Center Loss	99.28	94.9
Range Loss	99.52	93.70
Marginal Loss	99.48	95.98
SphereFace	99.42	95.0
SphereFace+	99.47	-
CosFace	99.73	97.6
MS1MV2, R100, ArcFace	99.83	98.02

شبکه عصبی تجمعی^۱ (NAN)

سامانه‌ای است که برای بازشناسی چهره ویدیویی طراحی شده است. این سامانه شامل یک شبکه عمیق و یک مولفه تجمعی است. شبکه عمیق بردارهای ویژگی را برای چهره‌ها در فریم‌های ویدیویی تولید می‌کند. مولفه‌ی تجمعی بردارهای ویژگی را تشکیل می‌دهد تا یک ویژگی واحد ایجاد کند. شبکه مورد استفاده در مقاله از معماری GoogLeNet با نرمال‌سازی دسته‌ای^۲ است.

۴-۲ بازشناسی چهره سه‌بعدی

با وجود پیشرفت قابل توجه الگوریتم‌های بازشناسی چهره‌ی دوبعدی، هنوز دقت این الگوریتم‌ها با تغییر شرایط نور و زاویه بسیار تحت تاثیر قرار می‌گیرد [78]. به همین دلیل تشخیص هویت انسان به کمک مدل سه‌بعدی چهره می‌تواند به عنوان یکی از گزینه‌های زیست‌سنگی قدرتمند به شمار رود. مقاومت این الگوریتم‌ها نسبت به تغییرات نور، زاویه‌ی چهره و حالت چهره بسیار بیشتر است، چون اساساً مدل سه‌بعدی چهره در شرایط نوری مختلف و از زوایای مختلف تغییری پیدا نمی‌کند [40]. مزیت اصلی این روش این است که شکل هندسی چهره‌ی انسان (و احتمالاً سر) در آن به طور دقیق استخراج می‌شود و با داشتن مدل دقیق‌تری از چهره، قدرت تفکیک‌پذیری مدل افزایش پیدا می‌کند [45]. اما به طول کلی می‌توان افزایش هزینه‌ها و پیچیده‌تر شدن مدل‌ها را می‌توان از معایب این روش نسبت به بازشناسی چهره‌ی دوبعدی در نظر گرفت.

در طول عملیات ثبت‌نام و مقایسه در یک سیستم بازشناسی چهره‌ی سه‌بعدی مانند هر سیستم زیست‌سنگی دیگر مراحل مختلفی طی می‌شود. در ادامه‌ی این بخش فعالیت‌های اخیر در حوزه‌های پیش‌پردازش، استخراج ویژگی و تطابق ویژگی‌ها مورد بررسی قرار می‌گیرد.

۴-۲-۱ پیش‌پردازش

داده‌های دریافت شده از حسگرها نمی‌توانند مستقیماً به منظور استخراج ویژگی به کار روند؛ زیرا داده‌های دریافت شده علاوه بر چهره شامل مواردی مانند مو، گوش، گردن، عینک و گردنبند نیز می‌تواند باشد که به تشخیص هویت فرد کمکی نمی‌کند. چون

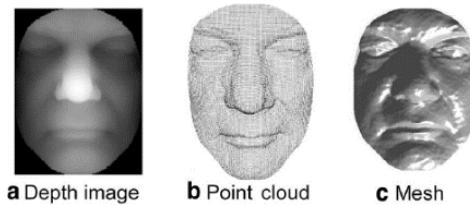
¹ Neural aggregation network (NAN)

² batch normalization

به عنوان مثال مو، گردن‌بند و عینک ممکن است تغییر کنند و مواردی مانند گردن و گوش نیز با توجه به نحوهی قرارگیری ممکن است شکل متفاوتی به خودشان بگیرند. به منظور دستیابی به دقت بالاتر در بازشناسی چهره چنین قسمت‌هایی از تصویر بهتر است حذف شود.

اولین مرحله در پیش‌پردازش داده‌ها تشخیص مکان و جهت‌گیری چهره است. پس از این مرحله با استفاده از یک تبدیل هندسی جهت صورت چرخیده می‌شود تا مستقیماً رو به روی محورهای دوربین قرار بگیرد. در ادامه مرحله‌ی پیش‌پردازش، از بخش‌هایی از چهره که به وضوح قابل شناسایی است (مانند بینی) استفاده می‌شود تا محدوده‌ی چهره از سایر بخش‌ها که قرار است حذف شود، جداسازی شود.

داده‌های پردازش شده‌ی چهره معمولاً به سه شکل تقسیر می‌شوند: تصویر عمقی^۱، ابر نقطه^۲ و بافت^۳. این سه نوع داده در شکل ۳۰-۲ ملاحظه می‌شود [78]. تصویر عمقی چهره که اصطلاحاً به آن تصویر دو و نیم بعدی (2.5D) نیز گفته می‌شود، در واقع یک بازنمایی دو بعدی از نقاط سه‌بعدی است. به عنوان مثال یک نمونه تصویر 2.5D می‌تواند تصویری خاکستری باشد که نقاط سیاه در تصویر بیانگر پس‌زمینه (با فاصله‌ی دور) در تصویر و نقاط سفید بیانگر نزدیک‌ترین سطوح به دوربین هستند. یک تصویر 2.5D از یک زاویه دید گرفته می‌شوند و به جای مدل‌سازی شکل سر و صورت، تنها اجازه‌ی مدل‌سازی سطح چهره را به ما می‌دهند [45].



شکل ۳۰-۲ سه فرمی که داده‌های سه‌بعدی چهره به صورت معمول پس از پیش‌پردازش به خود می‌گیرند [78]

۴-۲ رویکردهای سنتی بازشناسی چهره سه‌بعدی

بازشناسی چهره سه‌بعدی را می‌توان با استفاده از دو رویکرد کلی انجام داد. روش‌های سنتی یادگیری ماشین و روش‌های مبتنی بر یادگیری عمیق این دو رویکرد کلی هستند.

روش‌های سنتی به طور کلی به سه دسته تقسیم می‌شوند: رویکردهای کلی، محلی و ترکیبی. در رویکرد کلی، تمرکز بر شباهت چهره‌ها است. تمام چهره سه‌بعدی با تعریف مجموعه‌ای از ویژگی‌های سراسری توصیف می‌شود. تجزیه و تحلیل مولفه‌های اصلی و مدل‌سازی تغییر شکل‌ها از محبوب‌ترین روش‌های کلی هستند. رویکرد محلی ویژگی‌های هندسی چهره، عمدتاً چشم‌ها و بینی را بررسی می‌کند. راه حل ترکیبی ویژگی‌ها یا داده‌های کلی و همچنین محلی را ادغام می‌کند.

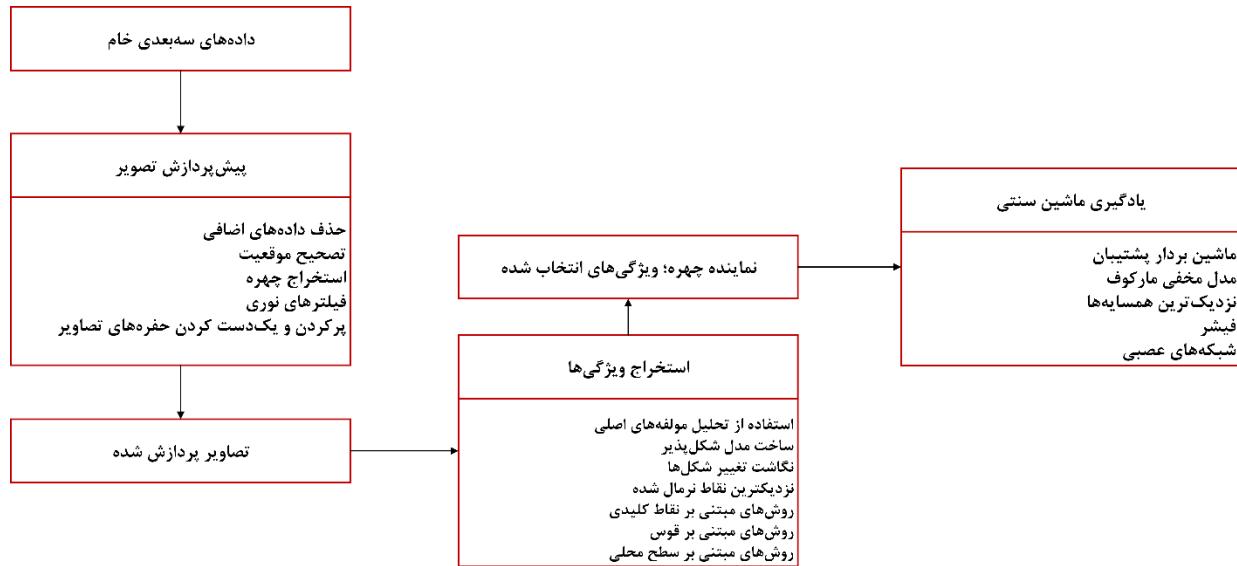
اگرچه مطالعات زیادی با استفاده از روش‌های کلی انجام شده است، اما به نظر می‌رسد که روش‌های محلی برای بازشناسی چهره به صورت سه‌بعدی مناسب‌ترند. در مقایسه با روش‌های کلی، روش‌های محلی از نظر جفت شدن قوی‌تر هستند و می‌توانند نتایج آزمایشی بهتری داشته باشند [79]. با این حال، اگر چهره رو به جلو باشد، و هیچ تفاوتی در حالت چهره وجود نداشته باشد، راه حل ترکیبی بسیار موثر است.

شکل ۳۱-۲ روند شناسایی چهره سه‌بعدی را با استفاده از روش‌های سنتی یادگیری ماشین نشان می‌دهد.

¹ depth image

² point cloud

³ mesh



شکل ۳۲-۲ خلاصه روند معمول در بازناسی سه‌بعدی [22]

در ادامه این روش‌ها به طور کلی در دو مرحله استخراج ویژگی و تطابق بررسی می‌شوند.

۲-۴-۱ استخراج ویژگی

برای استخراج، ذخیره و مقایسه ویژگی‌های چهره، سه روش وجود دارد: ۱) کلی: کل چهره فرد در یک بردار ویژگی تعریف می‌شود؛ ۲) محلی: هر یک از اجزاء چهره در بردارهای ویژگی مجزایی ذخیره می‌شود. در روش‌های نوع اول به هنگام مقایسه، چهره‌ی ورودی با چهره‌های موجود در پایگاه داده مقایسه می‌شود. اما در روش‌های نوع دوم برای هریک از اجزاء چهره ویژگی‌های مجزایی در پایگاه داده ذخیره شده است. هنگامی که یک چهره برای شناسایی به عنوان ورودی داده می‌شود، سیستم ابتدا ویژگی‌های هریک از اجزاء چهره‌ی دریافت شده را استخراج می‌کند و سپس در پایگاه داده به دنبال مجموعه‌های مشابه جستجو می‌کند. ۳) ترکیبی: روش‌هایی نیز دو نوع ویژگی نوع یک و دو را به صورت ترکیبی استفاده می‌کنند. این روش‌ها با صرف مقداری هزینه محاسباتی بیشتر معمولاً نتایج به بهتری دست پیدا می‌کنند [79]. در ادامه‌ی این بخش پژوهش‌های انجام شده با هریک از این ویژگی‌های کلی و محلی را مورد بررسی قرار می‌دهیم.

۲-۴-۲-۱ ویژگی‌های کلی

در این روش‌ها همانطور که گفته شد، کل چهره به عنوان یک بردار ویژگی در پایگاه داده ذخیره می‌شود و عملیات مقایسه نیز بین دو چهره‌ی کامل صورت می‌گیرد. چند نمونه از این روش‌ها که در یک دهه‌ی اخیر استفاده شده است در ادامه بررسی می‌شود:

- استفاده از تحلیل مولفه‌های اصلی (PCA): این روش مشابه روش «مقادیر ویژه چهره» است که در بخش دو بعدی بررسی شد اما این بار برای مدل‌سازی سه‌بعدی چهره.
- ساخت مدل شکل‌پذیر^۱: در این مقاله یک مدل شکل‌پذیر از سطح صورت تعریف شده است. برای هریک از افراد ثبت‌نام شده در سیستم، یک مدل شکل‌پذیر در پایگاه داده ذخیره می‌شود. بر پایه‌ی همین مدل، یک مازول مقایسه طراحی شده که به کمک آن می‌توان با دریافت یک تصاویر عمقی از چهره، آن را (حتی با وجود حالت‌های احساسی در چهره) با مدل سه‌بعدی ذخیره شده از چهره در حالت طبیعی (بدون احساس) مقایسه کرد.

^۱ Deformation modeling

- روش نگاشت تغییر شکل‌ها^۱ (SSDM) [83]: ایده این روش دسته‌بندی تغییرات را به دو دسته تقسیم می‌کند: ۱) تغییراتی که در برداشت‌های مختلف در چهره‌ی یک فرد ممکن است رخ دهد، و ۲) تغییراتی که مشخص می‌کند چهره‌ی جدید متعلق به یک هویت متفاوت با چهره‌ی اولیه است. در این روش برای مقایسه، ابتدا چهره‌ها در دو تصویر عمقی تراز می‌شوند و سپس یک نگاشت از تغییرات بین دو تصویر تعریف می‌شود. سپس بررسی می‌شود که این تغییرات می‌توانند در چهره‌ی یک فرد ایجاد شده باشد (مربوط به احساسات چهره باشد) یا این که دو تصویر متعلق به دو فرد مختلف است.
- روش نزدیکترین نقاط نرمال شده [84]: در این روش برای بالا بردن سرعت مقایسه بین تصاویر سه‌بعدی چهره، برای هر چهره ورودی، تعدادی از نزدیکترین چهره‌ها انتخاب می‌شوند که چهره‌ی ورودی تنها با آن‌ها مقایسه می‌شود و این عمل باعث می‌شود که امکان مقایسه‌ی با بار پردازشی بالاتر بین دو مدل سه‌بعدی چهره وجود داشته باشد.
- استفاده از یادگیری عمیق و مدل شکل‌پذیر [85]: در این مقاله مدلی با استفاده از یک پایگاهداده‌ی وسیع از افراد در حالت‌ها و وجهت‌های مختلف چهره، ارائه شده است. در بخش ابتدایی این مدل ابتدا با استفاده از تصاویر مصنوعی مدلی برای مکان‌یابی نقاط اصلی صورت آموزش داده شده است. از این نقاط برای استخراج ویژگی در مراحل بعدی و قرارگیری نقاط دقیق‌تر بر روش چهره استفاده شده و در نهایت مدلی مقاوم نسبت به تغییرات حالت چهره ارائه شده است.

۲-۴-۲-۱ ویژگی‌های محلی

این روش‌ها بر پایه‌ی ویژگی‌های مشترک محلی مشابه که از بخشی از صورت به دست آمده (مانند چشم‌ها و یا بینی) عمل می‌کنند. این روش‌ها از چند جنبه بر روش‌های کلی ارجاعیت دارند. به عنوان مثال در این روش‌ها نیازی به یک مدل کلی از چهره وجود ندارد پس بدین ترتیب تشخیص چهره‌هایی که بخشی از آن‌ها پوشیده شده، در این چنین مدل‌هایی ممکن می‌شود. همچنین روش‌های کلی نسبت به تغییرات حالت چهره بسیار حساس هستند در حالی که این مشکل توسط روش‌های مبتنی بر ویژگی‌های محلی تا حدی بهبود می‌یابد. در مجموع می‌توان گفت این مدل‌ها برای شناسایی و تایید هویت بهتر عمل می‌کنند در حالی که یافتن شباهت کلی بین افراد توسط روش‌هایی با ویژگی‌های کلی بهتر صورت می‌گیرد.

می‌توان اصلی‌ترین روش‌های مبتنی بر ویژگی‌های محلی را در سه دسته قرار داد [80]:

- روش‌های مبتنی بر نقاط کلیدی^۲: نقاط کلیدی به نقاطی از چهره گفته می‌شود که ویژگی‌های خاصی از نظر برجستگی داشته باشد. در این روش‌ها دو مرحله‌ی کلی وجود دارد: ۱) تشخیص نقاط کلیدی ۲) توصیف نقاط کلیدی. با توجه به این که تعداد زیادی از نقاط کلیدی با بردارهای ویژگی با ابعاد بالا هزینه‌ی پردازشی بالایی برای سیستم دارد؛ انتخاب نقاط کلیدی مؤثر از بین همه‌ی نقاط کلیدی ممکن در این روش‌ها بسیار مهم است.
- روش‌های مبتنی بر قوس^۳: در این روش‌ها از مجموعه‌ای از قوس‌ها از نقاط مختلف چهره به عنوان ویژگی برای نمایش شکل سه‌بعدی چهره استفاده می‌شود. روش‌های بر پایه‌ی این قوس را می‌توان در دو دسته قرار داد: ۱) بر پایه‌ی کانتور و ۲) بر پایه‌ی پروفایل. کانتور به قوس‌های بسته‌ای گفته می‌شود که هیچ خطی آن را قطع نمی‌کند. این خطوط محدوده‌های هم‌سطح در تصاویر عمقی را مشخص می‌کنند. نمونه‌ی این خطوط با دقت‌های مختلف در شکل ۳۲-۲ ملاحظه می‌شود. اما پروفایل به قوس‌هایی بازی گفته می‌شود که دارای نقطه‌ی شروع و پایان هستند. نمونه‌ی این خطوط در شکل ۳۳-۲ دیده می‌شود.
- روش‌های مبتنی بر سطح محلی^۴: این روش‌ها معمولاً اطلاعات هندسی محلی را از چندین بخش از سطح چهره و یا بخش‌هایی از چهره که در حالت‌های مختلف چهره ثابت باقی بماند، استخراج می‌کنند. از جمله‌ی این روش‌ها می‌توان به روش بر پایه‌ی بافت دودویی محلی^۵ (LBP) اشاره کرد. در مقاله [86] با الهام از ویژگی LBP برای بازناسی چهره‌ی دو بعدی روش 3DLBP بر پایه‌ی LBP برای بازناسی چهره‌ی سه‌بعدی مورد استفاده قرار گرفته است. این ویژگی بر روی تصویر عمقی محاسبه

¹ Signed shape difference map

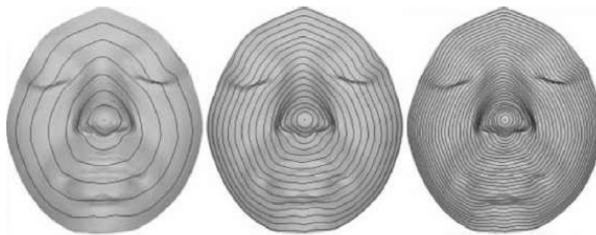
² keypoints-based

³ curve-based

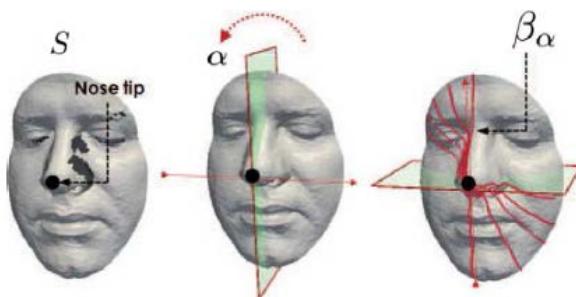
⁴ local surface-based

⁵ local binary pattern (LBP)

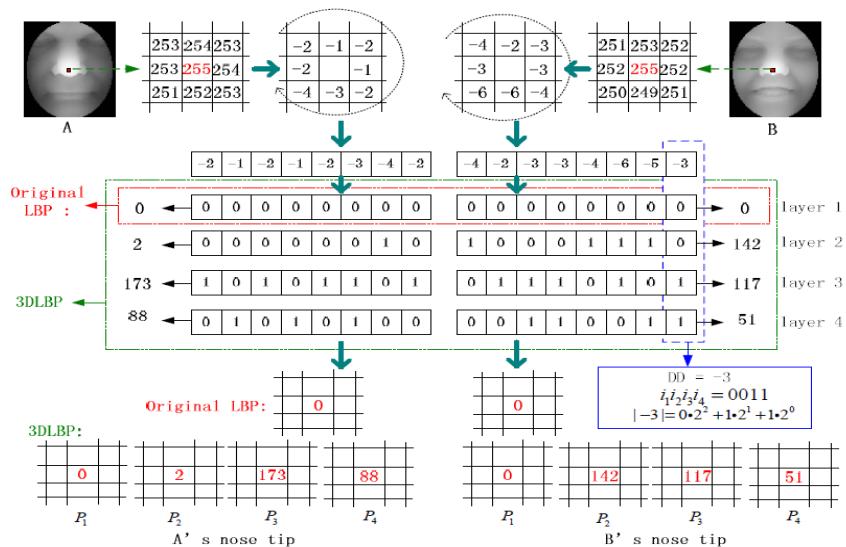
می‌شود. در شکل ۳۴-۲ یک نمونه از نحوه محاسبه این ویژگی و مقایسه‌ی آن با ویژگی LBP ملاحظه می‌شود. برای توصیف یک نقطه از تصویر عمقی مقدار عمق در آن با هشت پیکسل همسایه‌ی آن مقایسه می‌شود و تفاضل آن‌ها با نقطه‌ی مرجع ذخیره می‌شود (به عنوان مثال در شکل ۳۴-۲ این تفاضل در یک ظرف چهار بیتی ذخیره شده است) در نهایت مقدار تفاضل‌های ذخیره شده توصیف‌گر ویژگی در نقطه‌ی مورد نظر از تصویر است.



شکل ۳۳-۲ نمونه‌ای قوس‌های کانتور با دقت‌های مختلف بر روی چهره [87]



شکل ۳۴-۲ نمونه قوس‌های پروفایل بر روی چهره [88]



شکل ۳۵-۲ نمونه توصیف یک نقطه از دو تصویر از دو شخص متفاوت به کمک ویژگی 3DLBP ([86])

واحد انطباق در روش‌های بازشناسی چهره را می‌توان مستقل از ویژگی‌های استخراج شده مورد بررسی قرار داد. در تعدادی از پژوهش‌های مروی ([80], [89]) هسته‌ی الگوریتم‌های تطابق برای بسیاری از روش‌های مطرح آورده شده است. در ادامه‌ی این بخش چند روش تطابق که بیشتر مورد استفاده قرار گرفته‌اند بررسی می‌شوند و دقت آن‌ها بر روی پایگاه داده‌ی FRGC v2.0 [90] گزارش می‌شود. پژوهش‌های بسیاری که از این پایگاه داده به عنوان محک استفاده کرده‌اند، آن را به معیار مناسبی برای مقایسه‌ی کارایی الگوریتم‌ها تبدیل کرده است [80].

در برخی از روش‌هایی که در ادامه بررسی می‌شود یکتابع برای محاسبه‌ی فاصله (یا مشابهت) بین دو بردار ویژگی استفاده می‌شود. پرکاربردترین توابع تعریف شده در کاربرد بازشناسی چهره عبارت است از:

- فاصله‌ی اقلیدسی: تعریف این فاصله برای دو بردار $x = (x_1, x_2, x_3, \dots, x_p)$ و $y = (y_1, y_2, y_3, \dots, y_p)$ به صورت $dist(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_p - y_p)^2}$ است. به عنوان نمونه مقاله [91] از این معیار برای مقایسه‌ی دو بردار ویژگی استفاده کرده است.
- تشابه کسینوسی^۱: این معیار برای دو بردار غیر صفر از طریق ضرب داخلی دو بردار محاسبه می‌شود. به این صورت که $sim(x, y) = \frac{x \cdot y}{\|x\| \|y\|}$ که در آن $\|x\|$ همان اندازه‌ی بردار $(x_1, x_2, x_3, \dots, x_p)$ است که به صورت $\sqrt{x_1^2 + x_2^2 + x_3^2 + \dots + x_p^2}$ تعریف می‌شود. این معیار در واقع کسینوس زاویه‌ی بین بردار x و y را محاسبه می‌کند. در صورتی که مقدار تابع صفر باشد، دو بردار بر یکدیگر عمود بوده و هیچ شباهت ندارند؛ اما هرچه مقدار این تابع به یک نزدیک تر باشد، زاویه‌ی بین دو بردار کمتر و شباهت آن‌ها با یکدیگر بیشتر خواهد بود [92]. به عنوان مثال مقاله‌ی [93] از همین معیار به منظور مقایسه‌ی دو بردار ویژگی استفاده کرده است.
- فاصله‌ی ژئودزیک^۲: در صورتی که دو بردار ویژگی را در یک ابرفضا n بعدی از مرکز رسم کنیم و ابرکره n بعدی به مرکز مبدا مختصات را در دو نقطه‌ی a و b قطع کنند، فاصله‌ی ژئودزیک این دو بردار بیانگر طول کوتاهترین قوس بر روی ابرکره‌ی یاد شده است که دو نقطه‌ی a و b را به یکدیگر متصل می‌کند. به عنوان مثال مقاله‌ی [94] از این معیار به منظور مقایسه‌ی میزان تشابه بردارهای ویژگی دو چهره سه‌بعدی استفاده کرده است.
- فاصله‌ی Bhattacharyya: این معیار میزان شباهت بین دو توزیع احتمال را مشخص می‌کند. این معیار برای دو توزیع احتمال p و q که بر روی فضای X تعریف شده‌اند، به صورت $D_B(p, q) = -\ln \sum_{x \in X} \sqrt{p(x)q(x)}$ تعریف می‌شود. مقاله‌ی [95] نمونه‌ای است که استفاده‌ی این معیار در حوزه‌ی بازشناسی چهره‌ی سه‌بعدی.
- فاصله‌ی هاسدورف^۳: فضای هاسدورف در توپولوژی فضایی، به فضایی گفته می‌شود که در آن بتوان نقاط را با همسایگی از یکدیگر متمایز کرد. بین دو زیرمجموعه در این فضا فاصله‌ی هاسدورف تعریف می‌شود که از آن در برخی از مقالات بازشناسی چهره‌ی سه‌بعدی (از جمله [96]) به منظور تعیین میزان شباهت مجموعه ویژگی‌های دو چهره به کار رفته است.

حال در ادامه این بخش چند نمونه پرکاربرد از الگوریتم‌های انطباق چهره‌ی سه‌بعدی بررسی می‌شود.

۲-۴-۲-۱ رویکردهای مبتنی بر نزدیک‌ترین همسایه‌ها^۴

روش دسته‌بندی نزدیک‌ترین همسایه [97] یکی از پایه‌ای ترین روش‌ها در بازنمایی الگو^۵ است. در این روش با داشتن داده‌های آموزشی و تشکیل بردار ویژگی برای هر داده، آن را به عنوان یک نقطه در ابرفضا در نظر می‌گیریم. سپس یک معیار فاصله در این

¹ Cosine similarity

² Geodesic

³ Hausdorff distance

⁴ Nearest-neighbors

⁵ Pattern recognition

فضا تعریف می‌کنیم. برای هر داده‌ی جدید ابتدا بردار ویژگی را تشکیل می‌دهیم و آن نقطه‌ی جدید را نیز در فضا نمایش می‌دهیم و با توجه به معیار فاصله نزدیک‌ترین نقاط به آن را پیدا می‌کنیم. حال با در نظر گرفتن برحسب نزدیک‌ترین نقاط، دسته‌ی این داده‌ی جدید را مشخص می‌کنیم.

پژوهش [91] در سال ۲۰۰۷ میلادی با استفاده از همین روش یک الگوریتم تطابق برای مدل سه‌بعدی چهره ارائه کرده است که در آن معیار فاصله نیز «فاصله‌ی اقلیدسی^۱» در نظر گرفته شده است و به دقت ۹۶.۶۷٪ بر روی پایگاه داده‌ی FRGC v2.0 دست یافته است. مقاله‌ی [98] نیز از روش نزدیک‌ترین همسایه در سه استراتژی مختلف بر پایه‌ی هیستوگرام به منظور نمایش الگوی چهره استفاده کرده است و در بهترین حالت خود بر روی همان پایگاه داده به دقت ۹۴.۸۹٪ دست یافته است.

۲-۴-۲-۲ رویکردهای مبتنی بر ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبانی [99] یک روش برای دسته‌بندی داده‌هاست که در آن داده‌ها به فضایی با ابعاد بالاتر منتقل می‌شوند و این کار به گونه‌ای انجام می‌شود که پس از این تبدیل داده‌ها به صورت خطی قابل جداسازی باشند [92]. این مدل به منظور بازشناسی چهره‌ی سه‌بعدی نیز در سال‌های اخیر مورد استفاده قرار گرفته است. به عنوان مثال پژوهش [100] از مدل SVM به منظور تطابق چهره‌های سه‌بعدی استفاده کرده است و به دقت ۹۷.۶٪ بر روی پایگاه داده‌ی FRGC v2.0 دست یافته است. همچنین مقاله‌ی [101] نیز با استفاده از همین مدل در هسته‌ی ماژول تطابق خود به دقت ۹۷.۸٪ دست یافته است.

۲-۴-۲-۳ رویکردهای مبتنی بر پایه‌ی گراف

در برخی از مقالات بازشناسی چهره‌ی سه‌بعدی، با نمایش هر چهره به صورت یک گراف، مسئله‌ی مقایسه‌ی دو چهره را به مسئله‌ی مقایسه‌ی دو گراف معادل آن‌ها تبدیل کرده‌اند. به عنوان مثال پژوهش‌های [102] و [103] به همین ترتیب دو چهره‌ی سه‌بعدی را مقایسه کرده‌اند و به ترتیب در پایگاه داده‌ی FRGC v2.0 به دقت‌های ۹۷.۷٪ و ۹۹.۹٪ دست یافته‌اند.

۲-۴-۲-۴ رویکردهای مبتنی بر بر پایه‌ی ICP^۲

الگوریتم ICP یک الگوریتم بر پایه‌ی تکرار است که به کمک آن می‌توان تفاوت بین دو ابر نقاط را کمینه کرد. در بازشناسی چهره‌ی سه‌بعدی از این الگوریتم می‌توان به منظور ساخت مدل سه‌بعدی چهره نیز استفاده کرد. اما در برخی پژوهش‌ها از این الگوریتم به منظور تطابق دو مدل چهره استفاده شده است. به عنوان مثال مقاله‌های [104] و [105] هر دو از الگوریتم ICP در ماژول تطابق خود استفاده کرده‌اند و بر روی پایگاه داده‌ی FRGC v2.0 به دقت‌های ۹۷.۱٪ و ۹۷.۲٪ دست یافته‌اند.

۲-۴-۳ رویکردهای مبتنی بر یادگیری عمیق در بازشناسی چهره سه‌بعدی

از نظر تئوری، روش‌های مختلف مبتنی بر یادگیری عمیق کارآمد هستند و به تعریف منطقه مورد نظر^۳ (ROI) و استخراج و انتخاب ویژگی‌ها نیاز ندارند. در عمل، یادگیری عمیق نیاز به یک فرآیند یادگیری روی حجم زیادی از داده‌ها دارد با این حال، تعداد نگاشت‌های سه‌بعدی چهره بسیار محدود است که باعث می‌شود عملکرد تشخیص از نظر دقت (میزان تشخیص) بسیار بحرانی و غیرقابل اعتماد باشد.

برای روشن کردن این نکته، به عنوان مثال، مجموعه داده‌های FaceNet که در بازشناسی چهره دو بعدی با یادگیری عمیق استفاده می‌شود، از ۲۰۰ میلیون برای آموزش CNN عمیق است، در حالی که در بازشناسی چهره سه‌بعدی، بهترین مجموعه‌های داده شامل ۲ تا ۱۵ هزار داده است. به عنوان مثال، مجموعه داده معروف Bosphorus شامل حدود ۴ هزار تصویر و BU-3DFE حدود ۲.۵ هزار است. بنابراین واضح است که عملکرد بازشناسی چهره سه‌بعدی حتی اگر برخی از نویسنده‌گان عملکرد نسبتاً قابل

¹ Euclidean distances

² Iterative

³ definition of a region of interest (ROI)

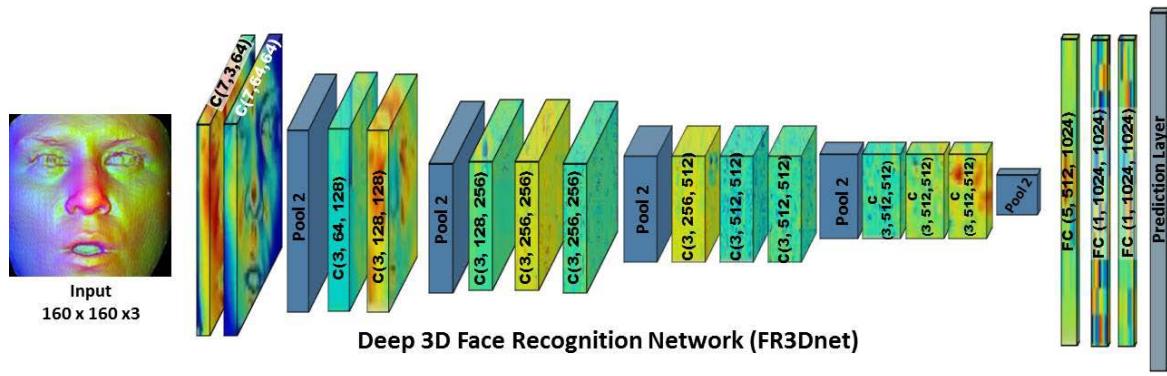
قابلی بدست آورند اما به دلیل هزینه‌های زیاد ناشی از پیچیدگی و محاسباتی، کارایی پایین‌تری از انتظارات ترویج‌کنندگان این روش‌ها دارد.

اخیراً رویکرد یادگیری عمیق نتایج امیدوارکننده‌ای در شناسایی چهره و برنامه‌های مربوطه به دست آورده است. با این حال، بازناسایی چهره سه‌بعدی مبتنی بر CNN به تازگی مطرح شده است. بیشتر سامانه‌های بازناسایی چهره سه‌بعدی بر روش‌های سنتی متتمرکز هستند که ویژگی‌های دستی را برای تأیید یا شناسایی استخراج می‌کنند. اما این روش‌ها به شدت به الگوریتم ترازبندی چهره و توصیف کننده‌های ویژگی متکی هستند که مقایسه پذیری را در مقایسه با رویکردهای یادگیری انتها به انتهای محدود می‌کنند. دو چالش برای بازناسایی چهره سه‌بعدی، چگونگی طراحی شبکه عمیق آگاه به ابر نقطه و عدم وجود پایگاه داده بزرگ چهره سه‌بعدی برای آموزش است. شبکه‌های موفق شناسایی چهره عمدتاً بر روی تصاویر کار می‌کنند. با این حال، رایج‌ترین قالب داده برای چهره سه‌بعدی، ابر نقطه و مش هستند که ساختار غیرشبکه‌ای آن‌ها به این حقیقت منجر شده که نتوانند به طور مستقیم در شبکه‌های دو بعدی کار کنند.

یکی از کارهای شناخته شده مقاله کیم و همکاران است [106]. آن‌ها نتایج را در سه مجموعه داده عمومی پس از تنظیم دقیق شبکه VGGFace [107] بر روی تصاویر با عمق سه‌بعدی گزارش کرده‌اند. برای تنظیم دقیق شبکه VGG-Face، از یک مجموعه داده متشکل از ۱۲۳,۳۲۵ تصویر عمق استفاده و سپس آن را به صورت جداگانه روی مجموعه داده‌های BU3DFE، Bosphorus و D-TEC³ (دو قلوها) آزمایش کرده‌اند. به جز مجموعه داده‌های Bosphorus، نتایج آن‌ها از روش‌های مرسوم پیشی نمی‌گیرد. علاوه بر این، آن‌ها نتایج مربوط به مجموعه داده‌های چالش برانگیز FRGCv2 را گزارش نکرده‌اند و مدل دقیق آن‌ها در دسترس عموم نیست.

FRGCv2 با ۴۴۶ هویت تا سال ۲۰۱۸ همچنان بزرگترین مجموعه داده معیار بازناسایی چهره سه‌بعدی بوده است. مطالعات نشان می‌دهد که اختلاف زیادی بین اندازه داده‌های چهره دو بعدی و سه‌بعدی وجود دارد. به دلیل عدم روش جایگزین برای جمع‌آوری چهره‌های سه‌بعدی واقعی برای آزمایش سامانه‌های بازناسایی چهره سه‌بعدی، پژوهش [108] یک پروتکل برای ادغام چالش برانگیزترین مجموعه‌های داده عمومی ارائه می‌دهد و این مجموعه داده را "LS3DFace" می‌نامد. یک مجموعه داده چهره سه‌بعدی از ۱۸۵۳ هویت با ۳۱۸۶۰ نمونه چهره ایجاد شده است. مجموعه داده تهیه شده هر سناریوی چالش برانگیز احتمالی را در بازناسایی چهره به ثبت می‌رساند و شامل تغییرات شدید در حالت، انسداد، داده‌های از دست رفته، نوع حسگر و شباهت‌های چهره به صورت دو قلوهای یکسان است.

این مقاله با الهام از موقوفیت شبکه‌های عمیق اخیر در شناسایی چهره دو بعدی، یک شبکه عصبی کانولوشن عمیق را پیشنهاد می‌دهد که مناسب داده‌های سه‌بعدی باشد. شبکه VGG برای تصاویر دو بعدی طراحی شده است که تغییرات قابل توجه بافت را روی ناحیه‌های کوچکی از تصویر نمایش می‌دهد. در مقابل، سطح چهره سه‌بعدی به طور کلی صاف است و از این رو فیلترهایی با اندازه هسته بزرگتر می‌توانند مجموعه‌ای بهتر از این نوع داده‌ها باشند. به عنوان مثال، تکه‌های یک سطح ۷*۷ دارای تنوع بیشتری نسبت به تکه‌های ۳*۳ هستند. این ادعا به طور تجربی با محاسبه میانگین واریانس و تعداد متوسط نقاط کلیدی بر روی هسته‌هایی با اندازه‌ی ۳، ۷ و ۹ در ۱۰,۰۰۰ تصویر سه‌بعدی که به طور تصادفی از داده‌های آموزش انتخاب شده است، اثبات می‌شود. واریانس هسته و تعداد متوسط نقاط کلیدی در هسته ۷*۷ به طور قابل توجهی بالاتر از اندازه ۳ است. ساختار FR3DNet ارائه شده در مقاله [109] را دنبال می‌کند اما با تغییر در لایه‌های پیچشی که جزئیات آن در شکل ۳-۵ آورده شده است. هدف مقاله این است که با یادگیری پارامترهای شبکه طراحی شده برای طبقه‌بندی $N = 100,005$ هویت، میانگین خطای پیش‌بینی پس از لایه softmax را به حداقل برساند. همچنین FR3DNet را با داده‌های ذخیره شده برای آزمون در مقایس بزرگ تنظیم کرده و آن را به عنوان FR3DNet_{FT} نشان می‌دهند.



شکل ۲-۳ معماری FR3DNet پیشنهادی [108]

جدول ۴-۲ جزئیات نتایج شناسایی LS3DFace را ارائه می‌دهد و آن‌ها را با رویکردهای عمیق و پیچشی پیشرفته مقایسه می‌کند. توجه داشته باشید که روش‌های متداول که تقریباً نتایج اشباع شده را در مجموعه داده‌های کوچک چهره سه‌بعدی گزارش می‌کنند، نمی‌توانند به دقت بالایی در مجموعه داده‌های بزرگ LS3DFace دست پیدا کنند و نشان می‌دهد که افزایش اندازه مجموعه تصاویر تأثیر معکوس زیادی بر عملکرد این الگوریتم‌ها دارد. FR3DNet بیش از ۱۴٪ از الگوریتم‌های مرسوم بازنگاری چهره سه‌بعدی و بهترین روش بازنگاری چهره دو بعدی با ۴.۷٪ عملکرد بهتری دارد. FR3DNetFT در مقایسه با VGG-Face هنگامی که نمونه‌ها چالش برانگیزتر هستند مانند مجموعه داده‌های 3D-TEC (دو قلوها) و UMBDB، ۸٪ بهتر عمل می‌کند.

جدول ۵-۲ جزئیات نتایج شناسایی LS3DFace

مدل/روش	Modality	مجموعه تصاویر										
		LS3DFace همین مقاله	FRGC [49]	BU3DFE [65]	BU4DFE [64]	Bosphorus [51]	CASIA [63]	GavabDB [41]	TexasFRD [23]	3D-TEC [61]	UMBDB [15]	ND-2006 [19]
GoogleNet [57]	RGB	53.97	21.51	50.76	65.41	63.44	85.91	-	53.08	79.95	65.78	24.14
Resnet152 [24]	RGB	15.05	13.53	8.04	9.64	7.05	52.85	-	20.94	72.66	34.08	10.92
VGG-Face [45]	RGB	90.85	87.92	97.68	96.51	96.39	94.18	-	99.73	83.30	81.54	82.86
GoogleNet [57]	3D	38.66	35.54	46.56	41.88	26.81	50.81	66.56	67.59	67.29	47.66	30.81
Resnet152 [24]	3D	12.49	14.40	5.80	10.13	3.84	25.34	44.26	16.25	60.98	22.20	12.08
VGG-Face [45]	3D	61.20	62.42	71.16	53.17	48.14	71.95	77.38	85.58	78.04	67.48	60.81
MMH [35]	3D + 2D	83.08	89.37	88.50	84.93	85.10	85.24	86.64	85.67	80.85	77.32	86.71
3D Keypoint [36]	3D	81.76	86.59	85.14	82.50	82.64	81.38	84.41	84.99	75.63	71.68	82.30
R3DM [9]	3D	82.89	87.50	87.13	83.21	86.06	84.51	85.60	85.47	78.27	77.11	84.84
K3DM [13]	3D	84.67	89.50	89.24	86.05	88.60	85.35	87.90	86.13	79.55	78.64	87.77
FR3DNet	3D	95.51	97.06	98.64	95.53	96.18	98.37	96.39	100.00	97.90	91.17	95.62
FR3DNet _{FT}	3D	98.75	99.88	99.96	98.04	100.00	99.74	99.70	100.00	99.12	97.20	99.13

مقایسه FR3DNet_{FT} در مجموعه داده‌های جدگانه در جدول ۵-۲ نشان می‌دهد که این شبکه با یک نمونه چهره دقیق برای هر شخص، دقیق‌تر از الگوریتم‌های مرسوم پیشرفته است. نتایج سایر روش‌ها از مقاله اصلی آن‌ها گزارش شده است.

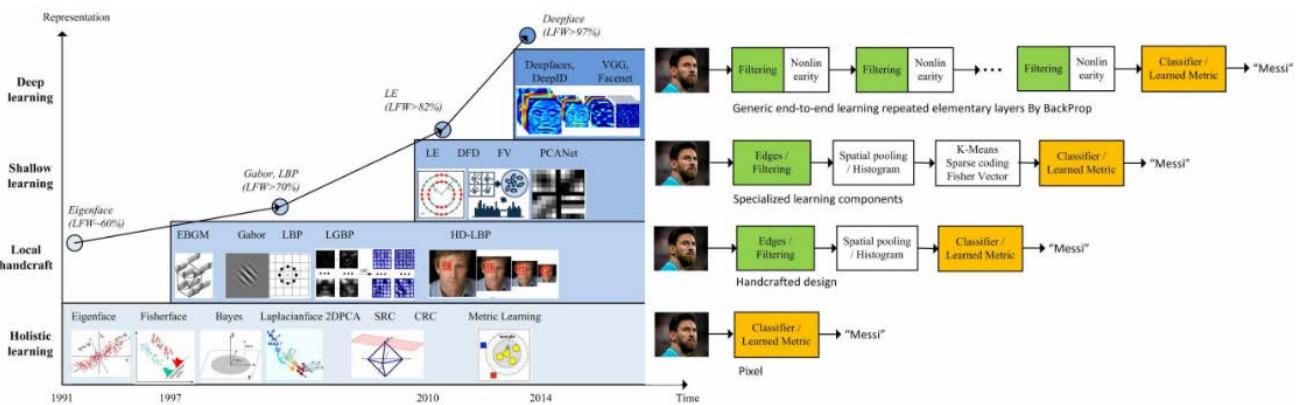
جدول ۲-۶ مقایسه ساده FR3DNetFT در مجموعه داده های جدگانه

	FRGC	BU3DFE	BU4DFE	Bosphorus	CASIA	GavabDB	TexasFRD	3D-TEC	UMBDB	ND-2006	روش/دیتاست
-	-	-	-	-	83.9	-	-	-	-	-	Xu <i>et al.</i> [63]
-	95.0	-	99.2	-	-	-	-	-	-	-	Kim <i>et al.</i> [29]
-	-	-	-	-	-	-	-	-	-	82.8	Faltemier <i>et al.</i> [19]
-	-	-	-	-	-	97.9	-	-	-	-	Gupta <i>et al.</i> [23]
97.8	-	-	-	-	-	-	97.2	-	-	-	Al-Osaimi <i>et al.</i> [4]
96.3	92.2	-	96.6	-	-	-	96.7	-	-	-	Li <i>et al.</i> [32,33]
96.3	94.0	-	-	-	-	96.3	-	-	73.1	-	Lei <i>et al.</i> [31]
96.2	95.9	94.2	96.4	82.5	95.4	98.0	95.9	69.3	95.3	-	Mian <i>et al.</i> [35]
98.5	96.2	96.0	98.6	85.4	96.5	98.1	92.6	78.6	96.8	-	Gilani <i>et al.</i> [13]
99.9	99.9	98.0	100.0	99.7	99.7	100.0	99.1	97.2	99.1	-	FR3DNet _{FT}

۵-۲ جمع‌بندی

در فصل دو، سامانه‌های زیست‌سنگی چهره (بازشناسی چهره) به طور کلی بیان گردید. سامانه‌های بازشناسی چهره دوبعدی و سه‌بعدی و رویکردهای این دو سامانه به طور مجزا بررسی و پایگاه داده‌های معروف آن‌ها آورده شد. در ادامه به جمع‌بندی رویکردها و نتایج مختلف روش‌های پیشرفت‌ههای بر روی پایگاه داده‌های معروف پرداخته می‌شود.

در پژوهش [64] نیز یک دسته‌بندی دقیق‌تر از الگوریتم‌های بازشناسی چهره و سیر پیشرفت آن‌ها ارائه شده است که خلاصه‌ی آن در شکل ۳۶-۲ ملاحظه می‌شود.



شکل ۳۷-۲ نمای کلی سیر تحول الگوریتم‌های بازشناسی چهره و دقت هریک (در محک LFW) از دهه ۱۹۹۰ تا به امروز [65]

یک معیار عمومی برای بازشناسی چهره است که به آن تطبیق جفت نیز می‌گویند. در جدول ۲-۶، عملکرد برخی از الگوریتم‌های معروف بر روی این مجموعه داده را مشاهده می‌کنید که در وبسایت آن آورده شده است^۱:

¹ <http://viswww.cs.umass.edu/lfw/>

جدول ۷-۲ نتایج رویکردهای مختلف بازشناسی چهره بر روی مجموعه داده LFW

الگوریتم	بازیابی چهره(%)	الگوریتم	بازیابی چهره(%)
Deep Face	97.35	FaceNet [87]	99.63
DeepFR	98.95	DeepID2+ [89]	99.47
Center Face	99.2	8 Baidu [91]	99.13
SphereFace	99.42	VGGFace [88]	99.13
Face++	99.50	FR+FCN [94]	96.45
DeepID	97.45	GaussianFace [96]	98.52
DeepID2	99.15	DeepID3 [97]	99.53
YouTu Lab, Tencent	99.80	PingAn AI Lab [98]	99.80
yunshitu	99.87	Deepmark [98]	99.23
Camvi	99.87	Innovative Technology [98]	99.88
Fisher vector faces	93.03	CMD+SLBP [100]	92.58
Simile classifiers	84.72	DFD [102]	84.02
LBP PLDA	87.33	LBP multishot [104]	85.17

با توجه به نتایج قابل قبول در رویکردهای مبتنی بر یادگیری عمیق در ادامه به صورت خلاصه چند نمونه از الگوریتم‌های محبوب در جدول ۷-۲ مشاهده می‌شود. این جدول خلاصه‌ی کارهای معروف را به ترتیب زمانی که بر روی مجموعه داده‌های LFW ارزیابی شده آورده است. این جدول شامل اطلاعات زمان انتشار، طراحی شبکه، تعداد شبکه‌ها، معیار آموزش، مجموعه آموزش و دقت است.

۲-۵-۱ تحلیل و مقایسه رویکردهای بازشناسی چهره

با توجه به جدول ۸-۲ و جدول ۹-۲ مشاهده می‌شود که رویکردهای مبتنی بر یادگیری عمیق به طور کلی نتایج بهتری را بر روی یک مجموعه داده به دست آورده‌اند. البته باید به این موضوع اشاره کرد که این نتایج برای رویکردهای دو بعدی می‌باشند و همانطور که بیان شد روش‌های بازشناسی چهره سه بعدی همچنان در حال پیشرفت می‌باشند.

نتایج مژو رجیدترین روش‌ها و پیشرفت‌های اخیر به ما نشان می‌دهد که افزایش چشمگیری در تحقیقات این حوزه طی پنج سال گذشته رخ داده است، به ویژه با ظهور رویکرد یادگیری عمیق که از محبوب‌ترین روش‌های بینایی ماشین به حساب می‌آید. علاوه بر این، پایگاه داده‌های متعدد چهره (دولتی و خصوصی) برای اهداف تحقیقاتی و تجاری در دسترس هستند و ویژگی‌های اصلی آن‌ها و پروتکل‌های ارزیابی ارائه شده است. تمرکز بر روی چهره‌های برچسب زده شده در پایگاه داده (LFW) از نظر روش، معماری، معیارها، دقت و پروتکل‌ها لازم است تا محققان بتوانند نتایج خود را با این پایگاه داده مرجع مقایسه کنند. درس‌های اصلی آموخته شده از این مطالعه مژوی این است که تشخیص چهره دو بعدی هنوز به پیشرفت‌های فنی آینده برای دستیابی به تجزیه و تحلیل تصاویر نیاز دارد. از سوی دیگر، توجه محققان به طور فزاینده‌ای با بازشناسی چهره سه بعدی جلب می‌شود. توسعه اخیر حسگرهای سه بعدی جهت جدیدی را برای بازشناسی چهره نشان می‌دهد که می‌تواند بر محدودیت‌های اصلی فناوری‌های دو بعدی غلبه کند، به عنوان مثال، تغییرات ظاهری، عامل پیری، حالت، تغییرات در شدت نور و به طور کلی در حالات چهره، داده‌های از

دست رفته، لوازم آرایشی و انسداد. اطلاعات هندسی ارائه شده توسط داده‌های چهره سهبعدی می‌تواند دقت تشخیص چهره را در شرایط نامساعد اکتسابی بهبود بخشد. با این حال، فقدان پایگاه داده بازشناسی چهره سهبعدی مانع بهره‌برداری از روش‌های مبتنی بر یادگیری عمیق می‌شود. همچنین، تفسیر حالت چهره سهبعدی، شناسایی تغییرات در سن و یادگیری انتقالی سه چالش دیگر این روش است که هنوز در آغاز کار خود هستند و نیاز به تحقیقات بیشتری دارد. به طور طبیعی، این پیشرفت‌های جدید در بازشناسی چهره باید چهار هدف را برآورده کند: سریع بودن (پاسخ فوری از دید کاربر)، دقت نزدیک به ۱۰۰، امنیت مطلوب، تجهیزات مینیاتوری و قابل حمل.

جدول ۲-۸ مقایسه‌ی روش‌های مختلف یادگیری ماشین بر روی داده‌های LFW

دقت	مجموعه داده	معیار آموزش	تعداد شبکه	شبکه	سال	روش
97.35 _ 0.25	Facebook (4.4 M, 4 K)	Softmax	3	CNN-9	2014	DeepFace
99.53 _ 0.10	WDRef + CelebFaces + (290 k, 12 k)	Contrastive Softmax + JB ¹	25	VGGNet	2015	DeepID3
99.63 _ 0.09	Google (200 M, 8 M)	Triplet Loss	1	GoogleNet	2015	FaceNet
99.77	Private Database (1.2 M, 18 K)	Triplet Loss	10	CNN-9	2015	BAIDU
98.95	VGGFace (2.6 M, 2.6 K)	Triplet Loss	1	VGGNet	2015	VGGFace
99.86	MS-Celeb 1M (3 M, 80 k)	COCO Loss	1	ResNet-128	2017	COCO Loss
99.42	CASIA WebFace (494 k, 10 k)	A-Softmax	1	ResNet-64	2018	SphereFace
99.83	MS-Celeb-1M (5.8 M, 85 k)	ArcFace	1	ResNet-100	2019	ArcFace
99.2 _ 0.04	CASIA WebFace (494 k, 10 k)	Softmax with center loss	1	GoogleNet	2020	Ben Fredj work
99.83	DeepGlint-MS1M (3.9 M, 86 K)	ArcFace Loss, LMC loss	1	ResNet-100	2020	ACNN LMC
98.13, 99.03, 99.07	CASIA WebFace (494 k, 10 k)	SDLMC loss, DLMC loss	1	ResNet32	2020	SDLMC DLMC

¹ JB: Joint Bayesian

فصل ۳ بررسی سامانه تشخیص زنده بودن و رویکردهای آن

سامانه‌های تشخیص چهره به طور فزاینده‌ای در انواع فرایندها و برنامه‌ها به کار گرفته می‌شوند. به دلیل این کاربرد گسترده، آن‌ها مجبورند در برابر حملات بسیار متنوع، مقاومت کنند. در میان همه این تهدیدهای حملات نمایش زیست‌سنجدی بسیار زیاد است. در این پژوهش، مروری بر نقاط قوت و آسیب‌پذیری چهره به عنوان یک ویژگی زیست‌سنجدی، ارائه می‌شود. حملات اصلی نمایش زیست‌سنجدی، تفاوت بین روش‌های مختلف، اقدامات مقابله‌ای PAD مربوطه و پایگاه داده‌های عمومی را که می‌توان برای ارزیابی روش‌های جدید حفاظت استفاده کرد، توصیف می‌شود. نقاط ضعف اقدامات متقابل مورد بررسی قرار می‌گیرد.

بانک‌های اطلاعاتی موجود منابع مفیدی برای مطالعه حملات نمایش هستند، اما ممکن است روش‌های PAD با استفاده از آن‌ها در همه سناریوهای احتمالی حمله قوی نباشد و همچنین با گسترش فناوری و کاهش هزینه‌های تولید ماسک‌های سه‌بعدی و ارتقاء کیفیت نمایشگرها، مهاجمان می‌توانند به راحتی اقدام به حمله به سامانه کنند. بنابراین، جمع‌آوری پایگاه‌های داده جدید با سناریوهای جدید به منظور توسعه روش‌های موثرتر PAD، مهم است. به طور معمول، روش‌های PAD برای مبارزه با یک نوع حمله مشخص (به عنوان مثال، تصاویر چاپ شده)، بازیابی شده از یک مجموعه داده خاص، توسعه یافته است. با این حال، هنگام آزمایش این روش‌ها در برابر سایر مصنوعات جعلی (به عنوان مثال، حملات پخش مجدد ویدئو)، معمولاً سامانه قادر به تشخیص آن‌ها نیست. یک درس مهم باید از این واقعیت آموخته شود: هیچ راهکار برتر PAD وجود ندارد که در همه شرایط بهتر از بقیه باشد. بنابراین دانستن اینکه از کدام روش باید در برابر هر نوع حمله استفاده کرد، یک عنصر اصلی است. در این پژوهش با استفاده از روش‌های مختلفی که ثابت شده در برابر انواع خاصی از مصنوعات مقاوم است، به منظور ایجاد طرح‌های همجوشی استفاده خواهد شد. از طرف دیگر، با پیشرفت مداوم فناوری، دستگاه‌های سخت‌افزاری و روش‌های نرم‌افزاری جدید همچنان ظاهر می‌شوند. پیگیری این پیشرفت‌ها در طی پژوهش از اهمیت زیادی برخوردار است زیرا برخی از پیشرفت‌ها می‌تواند کلید توسعه روش‌های جدید و کارآمد برای تشخیص زنده بودن باشد که از این جمله می‌توان به تمرکز بر تحقیق در مورد ماهیت بیولوژیکی ویژگی‌های زیست‌سنجدی به عنوان مثال، جریان خون اشاره کرد [110].

برای اولین بار در سال ۲۰۰۱ تشخیص زنده بودن، در یک مقاله در رابطه با امنیت اطلاعات به کار برده شد^[111]. یکی از مقالات ابتدایی در مورد زنده بودن توسط استفانی شوکرز در سال ۲۰۰۲ منتشر شد. در این مقاله در مورد "تشخیص زنده بودن براساس شناخت اطلاعات فیزیولوژیکی به عنوان نشانه‌های زنده بودن" صحبت شده است^[112].

در سال‌های اخیر، میزان دسترسی و علاقه به استفاده از حسگرهای زیست‌سنجدی برای احراز هویت کاربران افزایش یافته است ازین رو احتمال حمله به سامانه از طریق حسگر زیست‌سنجدی نیز فعالان این حوزه را با محدودیت‌هایی مواجه کرده است و در همین راستا در سال ۲۰۱۶ سازمان استاندارد جهانی استاندارد ISO / IEC 30107-1 از طریق آن می‌توان واقایع حمله نمایش را تعیین آوردن پایه‌ای برای PAD از طریق تعریف اصطلاحات و ایجاد چارچوبی است که از طریق آن می‌توان واقایع حمله نمایش را تعیین و شناسایی کرد تا بتوان آن‌ها را برای تصمیم‌گیری بعدی و فعالیت‌های ارزیابی عملکرد، طبقه‌بندی و ابلاغ نمود.

در ادامه به تعریف اصطلاحات مختلف بر اساس استاندارد ISO / IEC 30107-1 در این حوزه، پرداخته می‌شود^[113].

حمله نمایش: نمایش (ارائه‌ی) یک مصنوع یا مشخصه‌ی انسانی به زیرسامانه ضبط زیست‌سنجدی به روشی که می‌تواند در سیاست مورد نظر سامانه، تداخل ایجاد کند.

ابزار حمله نمایش^۱ (PAI): به مشخصه‌های زیست‌سنجدی یا شی مورد استفاده در حمله نمایش، PAI گفته می‌شود. مجموعه PAI شامل مصنوع‌ها می‌باشد اما نکته حائز اهمیت این است که در بردارنده مشخصه‌های زیست‌سنجدی بی‌روح (به عنوان مثال اجساد مرده) یا مشخصه‌های تغییریافته (به عنوان مثال اثر انگشت تغییریافته، عمل جراحی چهره) که در حمله به کار می‌روند نیز می‌باشد.

تشخیص حمله نمایش: به تشخیص خودکار حمله‌ی نمایش به یک سامانه ضبط مشخصه‌های زیست‌سنجدی گفته می‌شود.

زنده‌بودن: کیفیت یا حالت زنده بودن که توسط خصوصیات آناتومیکی (به عنوان مثال جذب نور توسط پوست یا خون)، واکنش‌های غیرارادی یا عملکردهای فیزیولوژیکی (به عنوان مثال واکنش عنبیه به نور، نبض) و واکنش‌های ارادی (مانند فشار دادن انگشتان در تشخیص هندسه دست) مشهود است.

تشخیص زنده بودن: تشخیص خصوصیات آناتومیکی یا واکنش‌های غیرارادی یا داوطلبانه، به منظور تعیین اینکه آیا نمونه زیست‌سنجدی ضبط شده از یک موجود زنده گرفته می‌شود. روش‌های تشخیص زنده بودن به عنوان زیر مجموعه‌ای از PAD تعریف شده است.

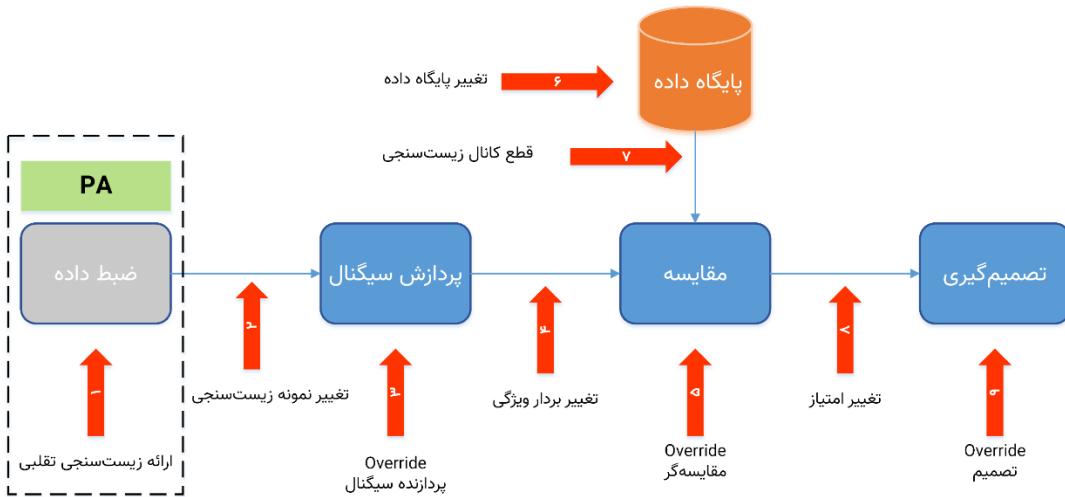
جعل: ایجاد تداخل در یک سامانه زیست‌سنجدی با ارائه یک مصنوع.

مصنوع: شی مصنوعی یا نمایش جعلی از مشخصه‌ی زیست‌سنجدی.

۱-۳ ساختار کلی سامانه تشخیص زنده بودن

یک سامانه زیست‌سنجدی می‌تواند تحت حملات مختلفی قرار بگیرد (شکل ۱-۳). PA حمله به حسگر ضبط داده‌های زیست‌سنجدی در یک سامانه زیست‌سنجدی است که عملکرد طبیعی آن را مختل می‌کند^[113].

^۱ Presentation attack instrument (PAI)



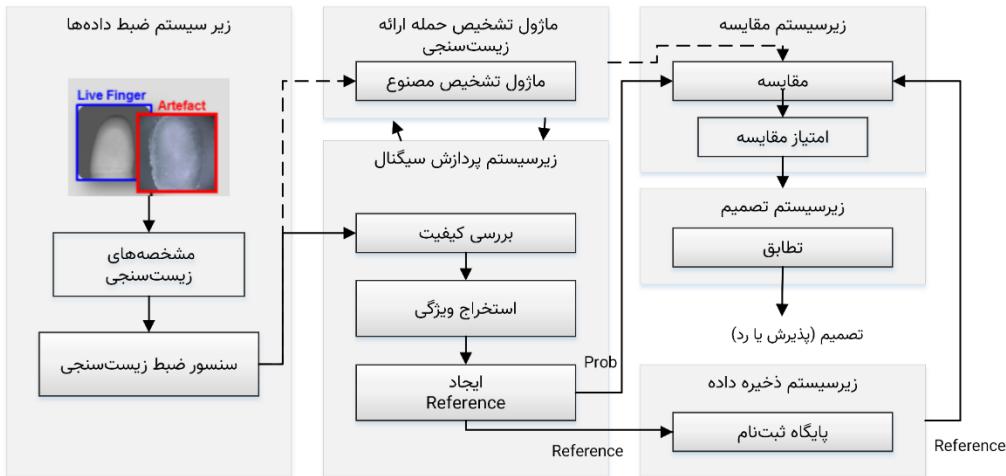
شکل ۳-۱ انواع حملات به یک سامانه زیست‌سنجی [113]

در حال حاضر ما بر روی PA، یعنی حملات علیه حسگر یک سامانه تشخیص زیست‌سنجی متوجه شده‌ایم. با این‌سانسازی نقاط خاصی از سامانه تشخیص، از جمله کانال‌های ارتباطی، تجهیزات و زیرساخت‌های درگیر، می‌توان از حملات غیرمستقیم (نقاط ۹-۲ در شکل ۳-۱) جلوگیری کرد. روش‌های مورد نیاز برای بهبود این مأموریت‌ها بیشتر مربوط به امنیت سایبری است تا زیست‌سنجی، بنابراین در این بحث پوشش داده نمی‌شوند. از طرف دیگر، حملات نمایش فقط یک آسیب‌پذیری زیست‌سنجی است که با سایر راه حل‌های امنیتی فناوری اطلاعات مشترک نیست و نیاز به اقدامات متقابل خاصی دارد.

یک سامانه تشخیص زیست‌سنجی از زیرسامانه‌های مختلف ضبط مشخصه‌های زیست‌سنجی، پردازش سیگنال و استخراج ویژگی، مقایسه، تصمیم‌گیری و زیرسامانه ثبت در پایگاه داده تشکیل شده است. در صورت اضافه کردن زیرسامانه‌ی تشخیص حمله‌ی نمایش به این سامانه، این زیرسامانه می‌تواند در محل‌های مختلفی واقع شود:

- پس از زیرسامانه ضبط داده
- در زیرسامانه ضبط داده
- پس از زیرسامانه پردازش سیگنال
- پس از زیرسامانه مقایسه یا تصمیم‌گیری

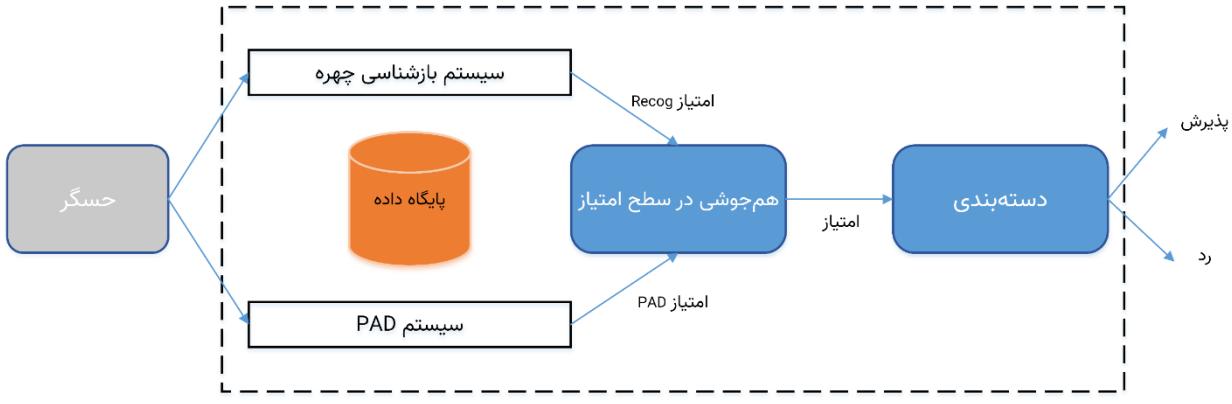
شکل ۳-۲، تصویری از یک چارچوب کلی از سامانه زیست‌سنجی با تشخیص حمله نمایش را نشان می‌دهد [113].



شکل ۲-۳ چارچوب کلی سامانه زیست‌سنگی با تشخیص حمله نمایش [113].

۲-۳ تشخیص زنده بودن در سامانه بازشناسی چهره

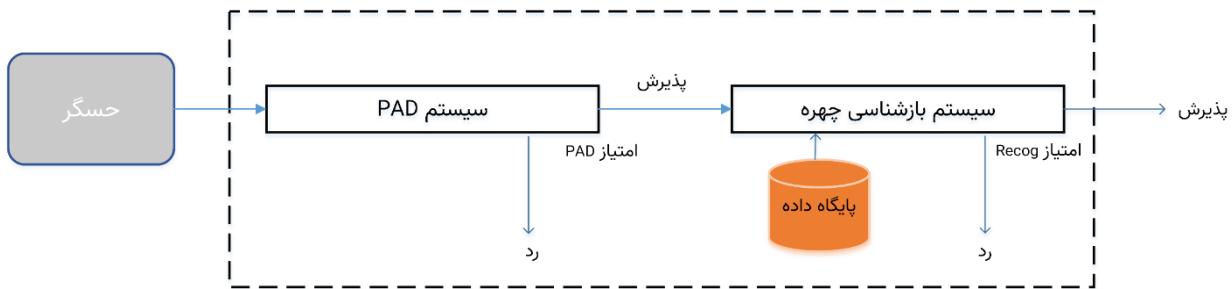
چهره متداول‌ترین سنجه زیست‌سنگی است که در برنامه‌های شخصی و تجاری عمومی استفاده شده است، مانند دسترسی به تلفن همراه، رایانه‌های شخصی، بانکداری مجازی، فروندگاهها و کنترل مرزها. زیست‌سنگی چهره به دلیل قابلیت دسترسی به آن در مقایسه با سایر زیست‌سنگی‌ها مانند اثر انگشت و عنبیه مشهور است. با این حال، این مزیت همچنانی باعث ضعف در موقعیت‌های کینه‌توزانه می‌شود و مهاجمان، چهره را به راحتی تقلید می‌کنند [114]. آسیب‌پذیری سامانه شامل سه نوع حمله از جمله در سطح حسگر زیست‌سنگی، سطح ویژگی و حملات سطح امتیاز است [114-117]. حمله عکس چاپی و حملات پخش مجدد فیلم هر دو به دلیل سادگی و هزینه پایین بسیار رایج می‌باشند. تشخیص حمل چهره به عنوان یک روش زیست‌سنگی برای اولین بار در سال ۱۹۹۲ توسط هوگستدن پیشنهاد شد [118]. پس از آن، در سال ۲۰۰۲ مطالعه مفصلی در مورد تقلب و تشخیص حمل چهره صورت گرفت [114]. کالریدر اولین روش زنده بودن را در سال ۲۰۰۵ براساس تنفس ساختاری تصاویر چهره پیشنهاد کرد [115]. به منظور ایجاد یک سامانه بازشناسی چهره‌ی مقاوم در برابر حملات نمایش زیست‌سنگی، باید راهکار مناسب PAD انتخاب شود. پس از آن، ادغام اقدامات متقابل PAD با سامانه بازشناسی چهره می‌تواند در سطوح مختلف، یعنی در سطح امتیاز یا تصمیم‌گیری انجام شود [119]. اولین احتمال شامل استفاده از همجوشی سطح امتیاز است که در شکل ۳-۳ نشان داده شده است. به دلیل سادگی و نتایج خوبی که این ادغام در تلفیق سامانه‌های زیست‌سنگی چندوجهی به دست آورده است، یک روش محظوظ به حساب می‌آید. در این حالت، داده‌های زیست‌سنگی همزمان به سامانه بازشناسی چهره و هم به سامانه PAD وارد می‌شوند و هر کدام امتیازات خود را محاسبه می‌کنند. سپس، امتیاز هر سامانه در یک امتیاز نهایی جدید ترکیب می‌شود و برای تعیین اینکه نمونه از یک کاربر واقعی است یا نه، استفاده می‌شود.



شکل ۳-۳ ادغام سامانه PAD با سامانه تشخیص چهره در سطح امتیاز[119].

مزیت اصلی این روش سرعت آن است، زیرا هر دو پیمانه، به عنوان مثال، PAD و پیمانه‌ی تشخیص چهره، عملیات خود را همزمان انجام می‌دهند. این واقعیت را می‌توان در سامانه‌هایی با مشخصات محاسبات موازی خوب مانند سامانه‌های دارای پردازنده چند هسته‌ای/چند نخی مورد بهره‌برداری قرار داد. روش متداول دیگر برای ترکیب سامانه‌های PAD و تشخیص چهره، یک طرح سریال است، همانطور که در شکل ۴-۳ نشان داده شده است، در آن سامانه PAD ابتدا تصمیم خود را می‌گیرد و فقط اگر نمونه‌ها از یک فرد زنده تشخیص داده شوند، سپس هویت مربوط به نمونه زیست‌سنگی را جستجو می‌کند. از طرف دیگر، در طرح سریال به دلیل تأخیرهای پی در پی PAD و پیمانه‌های تشخیص چهره، میانگین زمان برای دسترسی بیشتر خواهد بود. با این حال، این روش از کار اضافی سامانه شناسایی چهره در صورت حمله PAD جلوگیری می‌کند، زیرا محاسبه در مراحل اولیه پایان می‌یابد[120].

.122]



شکل ۴-۳ ادغام سامانه PAD با سامانه تشخیص چهره به صورت موازی[112].

۳-۳ پایگاه داده‌های رایج تشخیص زنده بودن برای چهره

NUAA PI DB: بانک اطلاعات جعل NUAA [158] اولین پایگاه داده PA است که به صورت عمومی منتشر شد. کل پایگاه داده شامل ۱۵ سوژه و ۱۲۶۱۴ نمونه است که تصاویر آن‌ها با استفاده از وب‌کم ضبط شده است. مصنوع چاپ شده نیز با گرفتن عکس چهره با کیفیت بالا برای هر سوژه با استفاده از دوربین DSLR تولید می‌شود، سپس با استفاده از چاپگر HP رنگی روی کاغذ عکاسی ۷۰ گرم A4 چاپ می‌شود.

PRINT-ATTACK DB: پایگاه PRINT-ATTACK DB [159] از ۵۰ سوژه و ۴۰۰ نمونه تشکیل شده است که نمونه‌های چهره واقعی آن‌ها با استفاده از لپ‌تاپ مکبوک ۱۳ اینچی اپل گرفته شده است. نمونه‌های حمله با ثبت تصویر چهره با کیفیت بالا و با

شرایط کنترل شده و نامطلوب با استفاده از دوربین ۱۲.۱ مگاپیکسلی Canon PowerShot SX150 IS تولید می‌شود. سپس، این تصاویر چهره با کیفیت بالا با استفاده از چاپگر لیزری رنگی Triumph-Adler DCC 2520 روی کاغذ A4 ساده چاپ و سپس با تنظیم دستی و ثابت به دوربین ارائه می‌شود.

REPLAY-ATTACK DB: بانک اطلاعات پخش مجدد [160] با پخش مجدد فیلم ضبط شده واقعی و با استفاده از iPhone و iPad ضبط می‌شوند. این پایگاه داده بستری را برای توسعه الگوریتم‌های PAD چهره با هدف حملات پخش مجدد ویدئو فراهم می‌کند و شامل ۴۰۰ نمونه می‌باشد.

CASIA FAS DB: بانک اطلاعات CASIA FAS [161] مشابه پایگاه داده حمله ویدیویی پخش مجدد است با این تفاوت که نمونه‌های حمله با استفاده از سه وضوح مختلف (وضوح پایین، متوسط و زیاد) جمع‌آوری می‌شوند و شامل ۶۰۰ نمونه است پایگاه داده حمله 3D Mask [162]: اولین پایگاه داده ماسک سه‌بعدی در دسترس عموم است و شامل ۱۷ سوژه و ۲۵۵ نمونه است که ماسک‌های سه‌بعدی آن‌ها توسط thatsmyface.com تهیه شده است. ضبط تصاویر با استفاده از دستگاه Kinect انجام شده است تا هم به تصاویر عمق و هم رنگ بدهد.

CelebA-Spoof DB: اگرچه پیشرفت‌های امیدوارکننده‌ای در روش‌های تشخیص زنده بودن، حاصل شده است، اما کارهای موجود هنوز در مدیریت حملات پیچیده و تعیین دادن به سناریوهای واقعی دارای ضعف می‌باشند. دلیل اصلی این است که مجموعه داده‌های فعلی از نظر کمی و تنوع محدود هستند. برای غلبه بر این موانع، یک مجموعه داده بزرگ CelebA-Spoof، در مقاله مورد بررسی [157] با همین عنوان (CelebA-Spoof) ارائه شده است: تعداد CelebA-Spoof شامل ۶۲۵.۵۳۷ تصویر از ۱۰،۱۷۷ هدف است، به طور کلی بزرگتر از مجموعه داده‌های موجود است. تنوع: تصاویر جعلی از ۸ صحنه (۲ محیط * ۴ شرایط روشنایی) با بیش از ۱۰ حسگر گرفته شده است. اطلاعات مازاد (حاشیه نویسی^۱): CelebA-Spoof شامل ۱۰ نوع جعل و همچنین ۴۰ ویژگی مازاد است که از مجموعه داده اصلی CelebA گرفته شده است.

YALE-RECAPT DB: پایگاه داده بازپخش شده YALE [156] نتیجه گرفتن دوباره عکس از پایگاه داده گستردۀ Yale Face است که از نورپردازی‌های مختلف استفاده می‌کند. در این پایگاه داده مصنوعات چهره با استفاده از دو دوربین متفاوت با وضوح بالا با استفاده از نمایشگرهایی که نمونه‌های چهره را نمایش می‌دهند گرفته می‌شوند و شامل ۲۵۶۰ نمونه است. در جدول ۷-۳ اطلاعات مربوط به پایگاه داده‌های ذکر شده، آورده شده است. اطلاعات مربوط به نوع داده‌ها، تعداد افراد و نمونه‌های واقعی و مصنوعی جعلی را در این جدول مشاهده می‌کنید.

جدول ۱-۳ ویژگی پایگاه داده‌های رایج

پایگاه داده	نوع	نوع جعل	تعداد نمونه‌ها	افراد
	داده		(واقعی / جعلی)	
Celeba Spoof [157]	تصویر	تصویر چاپ شده و ویدیو بازپخش شده، ماسک سه‌بعدی	۵۳۷/۶۲۵	۱۰۱۷۷
NUAA PI DB [158]	تصویر	تصویر چاپ شده	۷۵۰۹/۵۱۰	۱۵
PRINT-ATTACK DB [159]	ویدیو	تصویر چاپ شده و ماسک	۲۰۰/۲۰۰	۵۰
REPLAY-ATTACK DB [160]	ویدیو	ویدیو بازپخش شده	۱۰۰۰/۲۰۰	۵۰

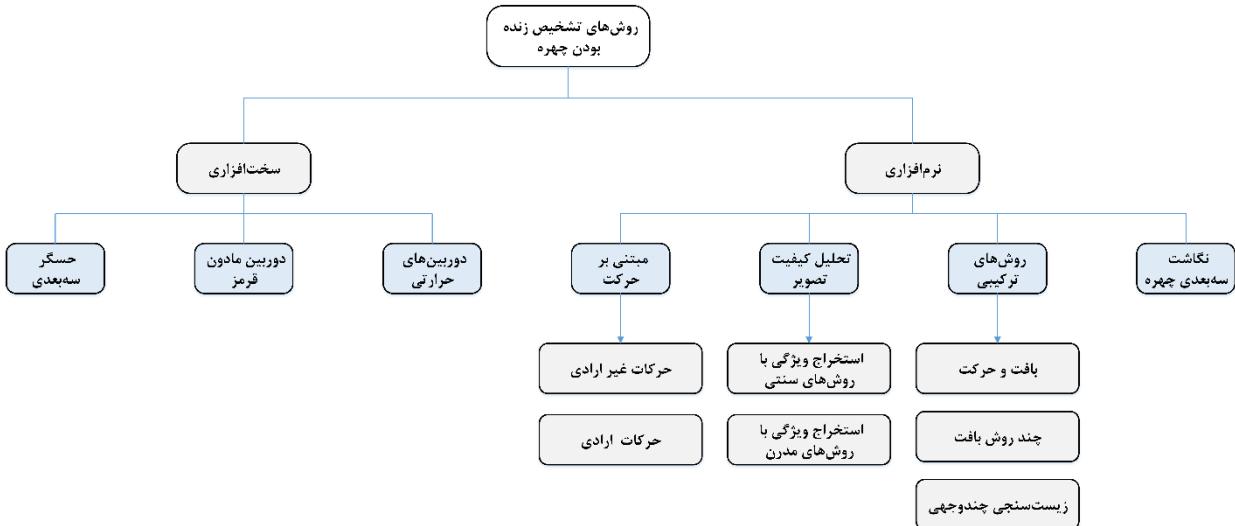
^۱ Annotation

پایگاه داده	نوع	نوع جعل	تعداد نمونه‌ها	افراد
(واقعی / جعلی)				
CASIA FAS DB [161]	ویدیو	تصویر چاپ شده و ویدیو بازپخش شده	۴۵۰/۱۵۰	۵۰
MASK-ATTACK DB [162]	ویدیو	تصویر چاپ شده و ویدیو بازپخش شده، ماسک سه بعدی	۸۵/۱۷۰	۱۷
YALE-RECAPT DB [156]	تصویر	تصویر چاپ شده	۱۹۲۰/۶۴۰	۱۰

۳-۴ رویکردهای تشخیص زنده بودن در سامانه‌ی بازشناسی چهره

هیچ طبقه‌بندی پذیرفته شده جهانی برای رویکردهای مختلف تشخیص زنده بودن وجود ندارد. تحمیل یک طبقه‌بندی شسته و رفته بر رویکردهای تشخیص زنده بودن چهره‌ی موجود، ساده نیست. در راستای تعیین روش مناسب برای این پژوهش، بهتر است اجزای مختلف یک سامانه تشخیص زنده بودن بررسی شود. روش‌ها و گونه‌شناسی‌های موجود، مورد مطالعه قرار بگیرد. در ادامه به چند گونه‌شناسی رایج در مقاله‌های معتبر این حوزه پرداخته می‌شود. در پژوهش [128] چهار روش مختلف پیشنهاد داده شده است: ۱) روش‌هایی که از حسگرهای موجود برای تشخیص هر سیگنال از ویژگی‌های زنده استفاده می‌کنند. ۲) استفاده از سخت‌افزار اختصاصی برای شناسایی شواهدی از زنده بودن که همیشه امکان استقرار آن وجود ندارد. ۳) با یک روش پاسخ به چالش که در آن از کاربر خواسته می‌شود تا با سامانه تعامل داشته باشد. ۴) با استفاده از الگوریتم‌های تشخیص ذاتی که در سامانه پیاده‌سازی شده است.

در [129] سه طبقه‌بندی ارائه شده است: ۱) روش‌های مبتنی بر بافت ۲) روش‌های مبتنی بر حرکت ۳) روش‌های یادگیری عمیق. پژوهش [130] روش‌های موجود در برابر حملات نمایش دو بعدی را به پنج دسته طبقه‌بندی می‌کند: مبتنی بر بافت، مبتنی بر کیفیت تصویر، رویکردهای پویا، ویژگی‌های آموخته شده و روش‌های ترکیبی. طرح‌های مبتنی بر بافت عمدهاً تفاوت الگوهای ریز بافتی چهره‌ها و مصنوعات واقعی را با کمک توصیف‌گرهای مختلف کشف می‌کنند. رویکردهای پویا، از اطلاعات زمانی برای کشف الگوهای حرکتی در فریم‌های ویدیویی بهره‌برداری می‌کنند. یادگیری عمیق برای استخراج ویژگی‌های سازگار استفاده و رویکرد دیگر، توسعه روش‌های PAD مبتنی بر روش‌های ترکیبی است که با ترکیب ویژگی‌های مختلف از نقاط قوت هر زمینه بهره‌مند می‌شوند. نویسنده‌گان در [131] رویکردهای تشخیص زنده بودن را به روش‌های مبتنی بر نرم افزار و روش‌های مبتنی بر سخت‌افزار طبقه‌بندی نموده‌اند(شکل ۳-۵). روش‌های نرم افزاری را نیز در ۴ دسته قرار داده‌اند. (۱) روش مبتنی بر حرکت: پلک زدن، حرکت دهان و چرخش سر (۲) روش مبتنی بر بافت: این روش بر ویژگی‌های بافت کاغذ تصاویر چاپ شده متمرکز است و زنده بودن را براساس طیف فرکانس تشخیص می‌دهد. (۳) روش زیست‌سنگی چندوجهی، دو یا چند نوع شناسایی زیست‌سنگی، به عنوان مثال، عنبیه و چهره و همچنین شناسایی صورت و اثر انگشت، در برابر ساختگی استفاده می‌کند. (۴) روش‌های یادگیری عمیق. با توجه به مطالعات بیان شده یکی از طبقه‌بندی‌هایی که می‌توان بیان کرد به است.



شکل ۳-۵ طبقه‌بندی رویکردهای تشخیص زنده بودن به دو طبقه‌ی نرم‌افزاری و سخت‌افزاری [131]

با وجود پیشرفت عملکرد سامانه‌های بازشناسی چهره، هنوز در برابر حملات مختلفی به ویژه چاپگرهای سه‌بعدی، آسیب‌پذیر می‌باشند. بنابراین، محققان چندین روش را برای محافظت از سامانه‌های بازشناسی چهره در برابر این آسیب‌پذیری‌ها پیشنهاد و تجزیه و تحلیل کرده‌اند. بر اساس روش‌های پیشنهادی، روش‌های تشخیص زنده بودن چهره به دو دسته اصلی تقسیم می‌شوند: روش‌های مبتنی بر سخت‌افزار و روش‌های مبتنی بر نرم‌افزار. روش مبتنی بر سخت‌افزار به یک دستگاه اضافی برای تشخیص ویژگی زیست‌سنجه‌ی خاص مانند تعییر انگشت، فشار خون و ترمومتر چهره نیاز دارد[123]. این دستگاه حسگر در سامانه احراز هویت زیست‌سنجه‌ی گنجانده می‌شود. برخی از دستگاه‌های کمکی، مانند تجهیزات مادون قرمز، در مقایسه با دستگاه‌های ساده، از دقت بالاتری برخوردار می‌باشند. با این حال، دستگاه‌های کمکی گران هستند و اجرایی شدن آن‌ها دشوار است[124]. روش مبتنی بر نرم‌افزار ویژگی‌های زیست‌سنجه‌ی را که از طریق یک حسگر استاندارد دریافت شده‌اند، استخراج می‌کند تا ویژگی‌های واقعی را از ویژگی‌های جعلی تشخیص دهد. استخراج ویژگی پس از دستیابی ویژگی‌های زیست‌سنجه‌ی توسط حسگر، مانند ویژگی‌های بافت در تصویر آغاز می‌شود[125].

این گزارش روش‌های مبتنی بر نرم‌افزار را پوشش می‌دهد. روش‌های موجود در این رده‌بندی را می‌توان به سه نوع اصلی تقسیم کرد (شکل ۳-۶): (۱) روش‌های تعاملی، (۲) روش‌های غیرتعاملی(منفعل) و (۳) ترکیب روش‌ها

تشخیص زنده بودن تعاملی

فناوری‌های تشخیص زنده بودن "فعال" کاربران را ملزم به انجام یک کار ساده مانند پلک زدن، چرخاندن سر یا حرکت تلفن خود به جلو و عقب می‌کنند. این امر منجر به سه مسئله می‌شود: اول، کلاهبرداران می‌توانند عکسی با برش‌های محدوده‌ی چشم را از دستگاه استفاده کنند یا ویدیویی را برای فریب سامانه نشان دهند. دوم، راهکارهای پاسخ به چالش، مهاجمان را در حالت آماده باش قرار می‌دهد که در حال بررسی هستند و در آخر اینکه، روش‌های فعل، اصطکاک ایجاد می‌کنند به گونه‌ای که روند احراز هویت را کند و نرخ از قلم انداختن^۱ را افزایش می‌دهد یعنی ممکن است به هر دلیلی پاسخ کاربر مطابق استانداردهای سیستم نباشد (کاربر مراحل چالش تعیین شده را به درستی انجام ندهد) و هر حرکت این پتانسیل را دارد تا دوباره از کاربر درخواست شود تا آن را تکرار کند و همین امر تجربه‌ی کاربر را دچار مشکل می‌کند.

¹ DropOut

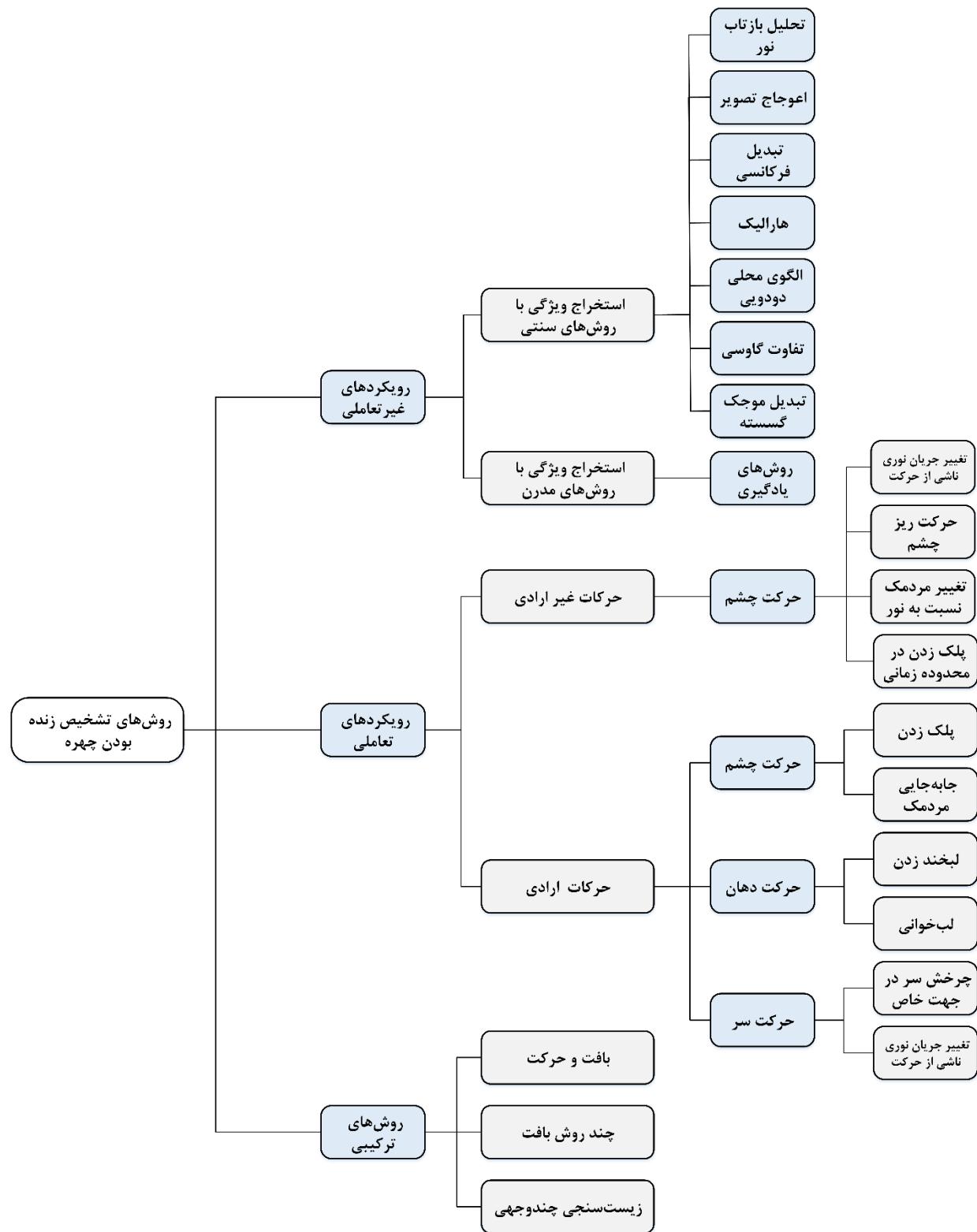
تشخیص زنده بودن غیرتعامی

یک روش نوین و ابتکاری برای تشخیص زنده بودن که به عنوان تشخیص غیرفعال شناخته می‌شود و مبتنی بر هوش مصنوعی می‌تواند، چهره‌ای را که به سامانه تشخیص چهره ارائه می‌شود را بررسی کند که آیا واقعی است (زنده است) یا جعلی. روش غیرفعال هیچ نشانه‌ای به کاربران در حال آزمایش نمی‌دهد و کاربران نیازی به انجام هیچ گونه حرکت اضافی ندارند. تشخیص زنده بودن غیرفعال در پس زمینه اتفاق می‌افتد و از این جهت این فرآیند به هیچ وجه به کاربر هشدار نمی‌دهد و به همین دلیل کشف چگونگی دور زدن این فناوری برای متقلبان سخت‌تر است [126,127].

تشخیص زنده بودن غیرفعال به طور کلی برتر از فعال است:

- از نظر امنیتی در سامانه‌های زیست‌سنگی چهره، قوی‌تر است.
- فرایند روان‌تر و آسان‌تری دارد.
- سریع‌تر است.
- نرخ از قلم انداختن را به میزان قابل توجهی کاهش می‌دهد.

لازم به ذکر است با توجه به اینکه استفاده از سامانه‌های بازشناسی چهره امروزه گسترش یافته است، امنیت و قابلیت اطمینان این سامانه‌ها به طور قابل توجهی به یک موضوع مهم تبدیل شده است. از بین آن‌ها تشخیص زنده بودن که هدف آن شناسایی زنده یا تقلبی بودن چهره‌ی ارائه شده است، بسیار مورد توجه قرار گرفته است. اگرچه پیشرفت امیدوارکننده‌ای حاصل شده است، اما کارهای موجود هنوز در مدیریت حملات پیچیده و تعییم دادن به سناریوهای واقعی دارای ضعف می‌باشند. به همین منظور استفاده از روش‌های ترکیبی که شامل هر دو روش فعال و غیرفعال باشد تا حدودی کارایی سامانه را در دنیای واقعی بالا می‌برد.



شکل ۳-۶ گونه‌شناسی روش‌های تشخیص زنده بودن با استفاده از چهره

در ادامه به بررسی و توضیح الگوریتم‌های رایج در پژوهش‌های معتبر پرداخته می‌شود.

۳-۴-۱- ۱- الگوی دودویی محلی^۱ (LBP)

رویکردهای مبتنی بر بافت، مبتنی بر تجزیه و تحلیل الگوهای ریزبافت در نمونه تصویر چهره است. این نوع روش در تشخیص مصنوعات تصویر بسیار موفقیت‌آمیز است، زیرا این روش می‌تواند بین ویژگی‌های مصنوع مانند وجود رنگدانه‌ها (به دلیل نقص چاپ)، بازتاب آینه‌وار و سایه (به دلیل استفاده از نمایشگر) به طور ممتازی تمایز قائل شود. مشهورترین و پرکاربردترین رویکرد، مبتنی بر الگوهای دودویی محلی است. روش LBP برای اولین بار در پژوهش مات و همکاران [132] بررسی شد. آن‌ها LBP‌های چند مقیاس را برای تجزیه و تحلیل بافت تصاویر چهره به کار گرفتند. سپس، از طبقه‌بندی ماشین بردار پشتیبان^۲ (SVM) برای طبقه‌بندی داده‌های جمع‌آوری شده و تولید نتایج استفاده کردند.

روش پیشنهادی را می‌توان به چهار مرحله تقسیم کرد: ابتدا با برش دادن و نرمال کردن، تصویر چهره به یک تصویر ۶۴ * ۶۴ پیکسل تبدیل می‌شود. در مرحله بعد، یک عملگر LBP به تصویر چهره نرمال شده اعمال می‌شود (شکل ۵-۳) و تصویر چهره LBP حاصل به قسمت‌های ۳ * ۳ که همپوشانی دارند، تقسیم می‌شوند. هیستوگرام‌های هر منطقه محاسبه و در یک هیستوگرام دودویی ۵۳۱ جمع‌آوری می‌شود. دو هیستوگرام دیگر از کل تصویر چهره با استفاده از عملگرهای LBP مختلف محاسبه می‌شود و هیستوگرام‌های محاسبه شده را به هیستوگرامی که قبلاً محاسبه شده بود، اضافه نمودند. سرانجام، از یک طبقه‌بندی SVM غیرخطی با هسته عملکرد شعاعی برای تعیین اینکه آیا تصویر ورودی یک چهره زنده است، استفاده شده است. روش مبتنی بر ریزبافت، از نظر محاسباتی سریع است و نیازی به تعامل کاربر ندارد. فراتر از آن، از ویژگی‌های بافتی که برای تشخیص زنده بودن استفاده می‌شود، می‌توان برای تشخیص چهره نیز استفاده کرد و یک فضای ویژگی منحصر به فرد برای ترکیب بازناسایی چهره و تشخیص زنده بودن را فراهم می‌کند. بعلاوه، این روش همچنین می‌تواند برای شناسایی حملات تقلب با استفاده از ماسک یا مدل‌های سه‌بعدی چهره استفاده شود، زیرا پوست دارای ساختار منحصر به فردی مانند منافذ پوست است در حالی که چهره‌های جعلی به ندرت چنین جزئیاتی را دارند و اگر که از ماسک سه‌بعدی برای فریب سامانه استفاده شود. از آنجا که مواد استفاده شده در ماسک سه‌بعدی در مقایسه با پوست واقعی الگوهای بافتی متفاوتی را نشان می‌دهد، ویژگی‌های LBP در تشخیص این تغییرات کاملاً موفق هستند.

در این پژوهش برای بررسی نتایج از شاخص سطح زیر نمودار^۳ (AUC) و نرخ خطای تساوی^۴ (EER) استفاده می‌شود، که به ترتیب AUC برابر با ۰.۹۶ و EER معادل با ۲.۹ به دست می‌آید.



شکل ۳-۷ نمونه تصویر پردازش شده در پژوهش [132]

¹ Local Binary Pattern (LBP)

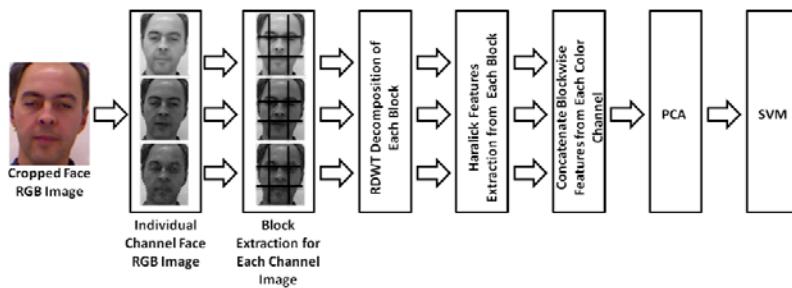
² Support Vector Machine (SVM)

³ Area Under the Curve (AUC)

⁴ Equal error rate (EER)

۳-۴-۱-۲ ویژگی‌های هارالیک^۱

آگاروال و همکاران [133] استفاده از ویژگی‌های هارالیک بافت چهره را پیشنهاد داده‌اند. ناحیه چهره بریده شده به کanal‌های RGB منفرد تقسیم و سپس به بلوک‌های غیرهمپوشان تبدیل می‌شود. برای هر بلوک، تبدیل موجک گستته^۲ (RDWT) بر روی چهار باند فرعی شامل: تقریب H_H، عمودی H_V، مورب H_D اجرا می‌شود. ویژگی‌های هارالیک همه بلوک‌ها برای چهار زیرباند گفته شده و همینطور تصویر اصلی بریده شده، استخراج می‌شود. بردار ویژگی نهایی از الحق ویژگی‌های اصلی هارالیک و ویژگی‌های هارالیک برای چهار باند فرعی از همه بلوک‌ها تشکیل و در نهایت توسط SVM طبقه‌بندی می‌شوند (شکل ۳-۶).



شکل ۳-۸ نمودار روش پیاده‌سازی شده مبتنی بر ویژگی‌های هارالیک [133]

در این پژوهش با توجه به مختصات چشم‌ها، برای فریم‌هایی که از ویدیوها به دست آمده است، ناحیه چهره مشخص می‌گردد. ویژگی‌های که پیش‌تر بیان شد استخراج می‌شود و نتایج برای هر دو طبقه‌بندی فریم و ویدیو به دست می‌آید. هر فریم از ویدیو به عنوان تصویر واقعی (زنده) و یا جعلی دسته‌بندی می‌شود و نتایج هم از نظر صحت طبقه‌بندی درست^۳ و هم از نظر نرخ نصف خطای کل^۴ (HTER) گزارش شده است.

جدول ۱-۳ نتایج طبقه‌بندی فریم و ویدیو را با استفاده از الگوریتم پیشنهادی بدون طبقه‌بندی بلوک و با طبقه‌بندی بلوک به طور خلاصه بیان می‌کند. الگوریتم پیشنهادی شامل ویژگی‌های الحق از کanal‌های جدگانه R، G و B است. بنابراین عملکرد کanal‌های جدگانه، را نیز ارزیابی کرده‌اند. در مجموعه آزمایش‌های صورت گرفته، الگوریتم پیشنهادی، بدون تقسیم به بلوک‌ها، HTER برابر با ۴.۱٪ دارد. با استفاده از نسخه مبتنی بر بلوک، الگوریتم پیشنهادی مقدار ۳ درصد را ارائه می‌دهد. جالب است بدانید که فقط با استفاده از کanal قرمز با استخراج ویژگی بلوک‌ها، HTER برابر با صفر در هر دو مجموعه اعتبارسنجی و آزمایش ارائه می‌شود. نویسنده‌گان معتقدند که این بدان دلیل است که از بین سه کanal، کanal R تغییرات را بهتر از دو کanal دیگر کد می‌کند و از این رو نتایج بهتری بدست می‌آید. نتایج مبتنی بر ویدئو با استفاده و بدون استفاده از تجزیه و تحلیل مولفه‌های اصلی^۵ (PCA) و استخراج ویژگی بلوک بررسی می‌شود. کاهش ابعاد نه تنها به کاهش ابعاد ویژگی بلکه به تشخیص حمله جعل مبتنی بر ویدئو نیز کمک می‌کند. به جای ویژگی‌های الحق شده از همه فریم‌ها به یک بردار ویژگی نهایی، فریم‌ها جدگانه طبقه‌بندی می‌شوند و میانگین امتیاز همه فریم‌ها به عنوان امتیاز نهایی فیلم محاسبه می‌شود. این رویکرد همچوشی سطح امتیاز، HTER برابر با ۰.۵۹ درصد را دارد.

¹ Haralik

² Redundant Discrete Wavelet Transform (RDWT)

³ Classification Accuracy

⁴ Half Total Error Rate (HTER)

⁵ Principal Component Analysis (PCA)

جدول ۲-۳ نتایج طبقه‌بندی فریم و ویدیو را با استفاده از الگوریتم پیشنهادی (هارالیک) [۲۶]

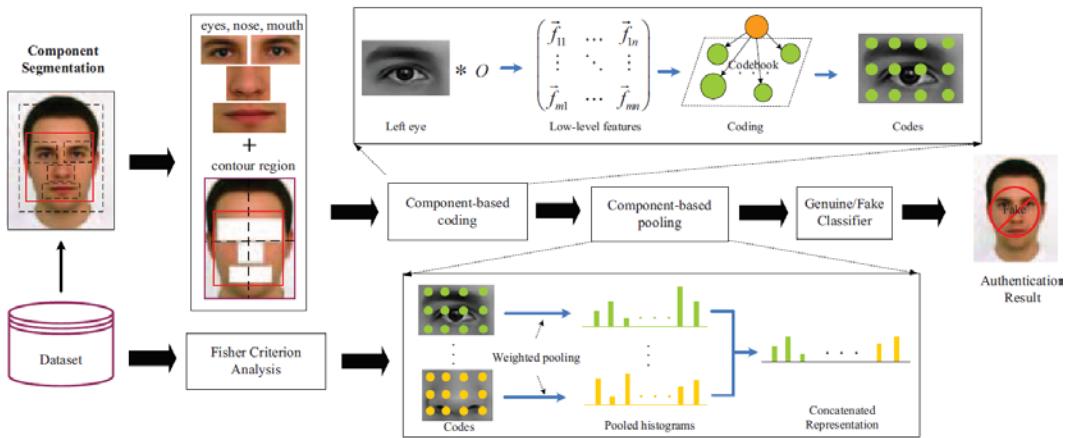
روش	کانال‌های رنگی	دقت طبقه‌بندی		HTER	
		اعتبارسنجدی	تست	اعتبارسنجدی	تست
فریم با بلک	RGB	۹۵.۹	۹۷.۸	۶.۰	۳.۰
	R	۱۰۰	۱۰۰	۰.۰	۰.۰
	G	۹۶.۰	۹۸.۰	۵.۰	۲.۰
	B	۹۶.۰	۹۷.۶	۵.۰	۳.۰
فریم بدون بلک	RGB	۹۰.۲	۹۵.۸	۱۰.۱	۴.۱
	R	۹۰.۹	۹۴.۴	۷.۱	۴.۳
	G	۷۷.۷	۸۲.۵	۲۷.۹	۲۰.۰
	B	۸۲.۵	۸۶.۳	۱۸.۴	۱۴.۶
ویدیو با بلک	RGB	۱۰۰	۱۰۰	۰.۰	۰.۰
	R	۹۹.۹	۱۰۰	۰.۰	۰.۰
	G	۹۶.۱	۹۸.۰	۵.۰	۲.۰
	B	۹۸.۸	۹۹.۲	۱.۰	۰.۰
ویدیو بدون بلک	RGB	۹۷.۵	۹۷.۶	۲.۰	۲.۰
	R	۷۹.۲	۸۵.۹	۱۶.۰	۱۰.۰
	G	۶۸.۹	۷۷.۲	۴۱.۰	۲۸.۰
	B	۷۸.۰	۸۳.۹	۲۵	۱۸.۰

۳-۴-۱-۳ استفاده از هیستوگرام شبکه‌ای جهت‌دار^۱ (HoG) و تفاوت گاوی^۲ (DoG)

در پژوهش یانگ و همکاران [۱۳۴] همانطور که در شکل ۲-۳ نشان داده شده است، ابتدا کل چهره را که آن را چهره جامع (H-Face) نامیده است، استخراج می‌کند. پس از آن، H-Face را به شش جز (بخش)، شامل منطقه کانتور، ناحیه چهره، ناحیه چشم چپ، ناحیه چشم راست، ناحیه دهان و ناحیه بینی تقسیم می‌کند. علاوه بر این، منطقه کانتور و چهره را به شبکه‌ای ۲*۲ تقسیم می‌کند (۸ حالت). برای همه دوازده ترکیب، ویژگی‌های سطح پایین (به عنوان مثال، LBP، HOG) استخراج می‌شود. با توجه به ویژگی‌های محلی استخراج شده، کدگذاری مبتنی بر مولفه‌ها، برای به دست آوردن کدهای محلی انجام می‌شود. سپس کدها را با یک توصیف‌کننده سطح بالا با وزن حاصل از تجزیه و تحلیل معیار فیشر با هم الحقق می‌کنند. سرانجام، مشخصات را برای طبقه‌بندی به یک SVM می‌دهند (شکل ۲-۳).

¹Histogram of Oriented Gradients (HoG)

²Difference of Gaussian (DoG)



شکل ۹-۳ نمودار روش پیاده‌سازی شده HoG و DoG در پژوهش [۳۴]

در این مقاله عملکرد تشخیص را با روش DoG [135] و روش پیشرفتی مبتنی بر MsLBP [136] مقایسه می‌کند. نتایج پژوهش در جدول ۲-۳ آورده شده است. به ترتیب با "۱" و "۲" مشخص می‌شوند. به طور مشابه، "H-Face" و "Face, MsLBP" به ترتیب با ۳ و ۴ نشان داده می‌شوند. روش پیشنهادی نیز با ۵ نشان داده می‌شود.

جدول ۳-۳ نتایج طبقه‌بندی بر روی داده‌های پایگاه داده NUAA با استفاده از الگوریتم پیشنهادی HoG و DoG [۳۴]

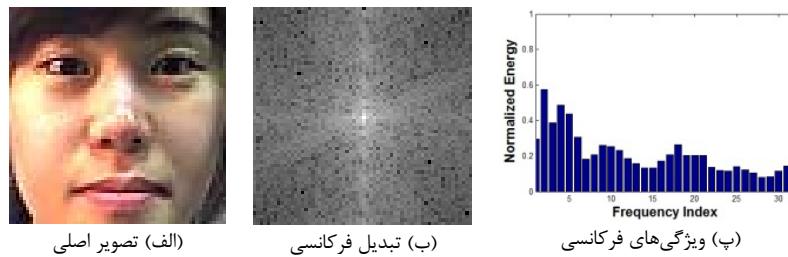
۵	۴	۳	۲	۱	معیار
۰.۹۷۷	۰.۹۲۷	۰.۷۴۹	۰.۸۱۸	۰.۷۴۶	صحبت دسته‌بندی
۰.۹۹۸	۰.۹۹۰	۰.۸۷۳	۰.۸۳۰	۰.۷۱۷	AUC
۰.۰۱۹	۰.۰۴۸	۰.۲۳۹	۰.۲۳۳	۰.۳۵۹	EER

۴-۱-۴-۳ تجزیه و تحلیل فرکانس

در مقاله‌ی کیم و همکاران [137]، یک روش تشخیص زنده بودن مبتنی بر تجزیه و تحلیل فرکانس و بافت برای تشخیص ماسک‌های کاغذی دو بعدی از چهره‌های زنده ارائه شده است. برای تجزیه و تحلیل فرکانس، روش مبتنی بر طیف توان [138] را انجام داده‌اند که نه تنها از اطلاعات فرکانس پایین بلکه از اطلاعات در ناحیه‌ها با فرکانس بالا نیز بهره‌برداری می‌کند. علاوه بر این، روش توصیف مبتنی بر الگوی دودویی محلی [139] برای تجزیه و تحلیل بافت‌های موجود در تصاویر چهره استفاده شده است. استخراج اطلاعات فرکانسی از تصاویر چهره داده شده به شرح زیر است:

ابتدا تصویر چهره داده شده با استفاده از تبدیل فوریه گسسته دو بعدی به دامنه فرکانس تبدیل می‌شود. تصویر اصلی چهره در شکل ۸-۳ (الف) و اندازه تصویر ورودی تبدیل شده توسط فوریه در شکل ۸-۳ (ب) نشان داده شده است. توجه داشته باشید که تبدیل فوریه تصویر، طوری تغییر یافته است که مولفه فرکانس صفر در مرکز طیف قرار دارد و نتیجه تبدیل شده به چند گروه از حلقه‌های متحدم‌المرکز تقسیم می‌شود. اختلاف شعاع بین هر جفت حلقه همسایه ۱ می‌باشد. سپس یک مجموعه با ۳۲ حلقه متحدم‌المرکز از یک تصویر با اندازه ۶۴*۶۴ (عرض) تولید می‌شود. هر حلقه نشان‌دهنده یک منطقه مربوطه در باند فرکانس است، یعنی یک حلقه با شعاع کوچک حاوی اطلاعات فرکانس پایین تصویر داده شده است. سرانجام، بردار ویژگی را می‌توان با بهم پیوستن مقادیر متوسط انرژی تمام حلقه‌های متحدم‌المرکز بدست آورد. از آنجا که مقادیر متوسط برای ناحیه‌های مختلف اجزای

فرکانسی (حلقه‌های متحدمالمرکز) در مقدار زیادی متفاوت است، از نرمال‌سازی استفاده شده است. ویژگی فرکانس حاصل در شکل ۸-۳ (پ) نشان داده شده است.



شکل ۱۰-۳ تصویر اصلی و تبدیل به حوزه فرکانس [۳۷]

روش‌ها بر روی دو پایگاه داده که جمع‌آوری شده بود، مورد آزمایش قرار گرفت، عملکرد پایگاه داده تصاویر وب کم (یکی از مجموعه داده‌ها) در جدول ۳-۳ نشان داده شده است.

جدول ۴-۳ نتایج طبقه‌بندی با استفاده از الگوریتم تحلیل فرکانس پیشنهادی [۱۳۷]

ادغام	LBP	فرکانس	
۹۱.۵۷	۸۸.۴۲	۸۸.۱۳	چهره زنده
۹۳.۹۱	۸۸.۸۹	۹۰.۸۲	عکس
۹۰.۳۱	۸۸.۰۵	۸۶.۱۱	ماسک کاغذی چاپ شده
۸.۴۳	۱۱.۵۸	۱۱.۸۷	EER

۳-۴-۵-۱ استخراج ویژگی‌های حالت‌های رنگ تصاویر در کانال‌های مختلف فضاهای رنگی

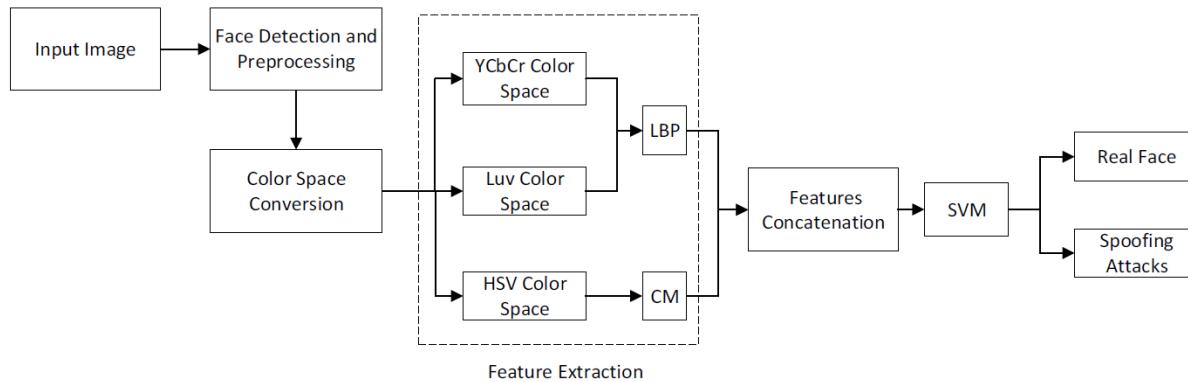
در مقاله‌هایی و لو [۱۴۰] یک روش تشخیص زنده بودن چهره مبتنی بر ترکیب مدل‌های مختلف فضاهای رنگی ارائه شده است به این صورت که تصویر چهره را به یک مدل فضای رنگی متفاوت تغییر دهید تا تفاوت در ویژگی بافت را کشف کنید. در مرحله اول، تصویر RGB ورودی به ۴ مدل فضای رنگی مختلف YCbCr، HSV و Lab و Luv تبدیل می‌شود و سپس ویژگی‌های LBP و SVM رنگ^۱ (CM) برای مدل‌های فضای تک رنگ و ترکیبات آن‌ها استخراج می‌شود. سرانجام، ویژگی‌ها به یک طبقه‌بندی EER داده می‌شوند. از طریق آزمایش‌ها، مشخص شد که استخراج ویژگی‌های LBP در فضای رنگی HSV و YCbCr همراه با استخراج ویژگی‌های CM در فضای رنگی HSV، می‌تواند به طور دقیق بین چهره‌های اصلی و حملات جعل تمایز ایجاد کند. آزمایش‌ها در دو مجموعه داده عمومی CASIA-FASD و Replay-Attack اجرا شده و در مقایسه با الگوریتم‌های دیگر، نتایج نشان می‌دهد که روش پیشنهادی عملکرد تشخیص زنده بودن چهره را تا حد زیادی بهبود داده است.

چهره واقعی دارای ویژگی‌های سه‌بعدی است که به دلیل بازتاب نور و سایه‌ها و سایر عوامل باعث ایجاد تفاوت در جزئیات بافت در مقایسه با تصویر جعلی می‌شود. بنابراین، اگر تصویر RGB به سایر فضاهای رنگی تبدیل شود، با در نظر گرفتن اجزای رنگی مربوطه در فضاهای مختلف رنگ، برخی از ویژگی‌های متفاوت بین تصویر چهره واقعی و حمله جعل را می‌توان یافت.

نمودار روش تشخیص جعل چهره که بر اساس ترکیب مدل‌های مختلف فضای رنگی ارائه شده در شکل ۹-۳ نشان داده شده است. فرایند به این شرح می‌باشد: ۱) تشخیص ناحیه چهره و پیش‌پردازش روی تصویر ورودی انجام می‌شود، و تصویر چهره توسط

^۱ color moment (CM)

الگوریتم تشخیص چهره Viola-Jones استخراج می‌شود و تصویر صورت استخراج شده برش داده و اندازه آن 64×64 پیکسل نرمال می‌شود. ۲) تصویر RGB پردازش شده را به ترتیب به فضای رنگی YCbCr و Luv و HSV تبدیل کرده و ویژگی‌های LBP را در فضای رنگی CM و ویژگی‌های Luv YCbCr و ویژگی‌های CM را در فضای رنگی HSV استخراج می‌شود^(۳) ویژگی‌های استخراج شده در مرحله ۲ به بردارهای ویژگی سراسری یک بعدی تبدیل می‌شود.^(۴) از SVM برای طبقه‌بندی و خروجی نتایج استفاده می‌شود.



شکل ۳-۱۱ نمودار روش پیاده‌سازی شده مبتنی بر حالت‌های رنگی در پژوهش [۱۴۰]

همانطور که در جدول ۴-۳ مشاهده می‌شود، نتیجه‌ی آزمایش‌ها بر روی داده‌های REPLAYATTACK نشان می‌دهد که اختلاف بین فضاهای رنگی زیاد می‌باشد. در این آزمایش، با مقایسه عملکرد فضای تک رنگ و فضای رنگی ترکیبی، به این نتیجه رسیده‌اند که تلفیق چندین ویژگی استخراج فضای رنگی می‌تواند ویژگی‌های مفیدی را در فضاهای رنگی مختلف ضبط کند، که می‌تواند به طور موثر عملکرد تشخیص الگوریتم را بهبود بخشد. نتایج نشان داد که استخراج ویژگی‌های LBP در فضای رنگی HSV می‌تواند به طور موثری عملکرد تشخیص زنده بودن چهره را بهبود بخشد. ویژگی‌های CM در فضای رنگی Luv و YCbCr و ویژگی‌های CM در فضای رنگی HSV می‌تواند به طور موثری عملکرد تشخیص زنده بودن چهره را بهبود بخشد.

جدول ۳-۵ نتایج طبقه‌بندی با استفاده از الگوریتم مبتنی بر حالت‌های رنگی پیشنهادی [۱۴۰]

CM		LBP		فضای رنگی
(%) HTER	(%) EER	(%) HTER	(%) EER	
۱۷.۳۹	۱۸.۵	۸.۷۱	۹.۰۱	RGB
۵.۹۸	۶.۸۳	۶.۷۶	۵.۰۵	YCbCr
۵.۲۳	۶.۶۹	۹.۵۳	۱۱.۷۸	HSV
۸.۹۹	۱۰.۰۶	۱۶.۹۹	۱۸.۹۵	Lab
۹.۵۱	۱۰.۶۷	۸.۸۲	۹.۳۲	Luv

۳-۴-۲ رویکردهای تعاملی

این روش به حرکات چهره و سر به صورت خاص توجه می‌کند، هدف آن بررسی ویژگی‌های زیست‌سنگی با آزمایش حرکات مورد انتظار کاربران است که از این جمله پلک زدن، حرکت دهان (مانند لب‌خوانی و تشخیص لبخند) و چرخش سر (حرکت بالا و پایین، چپ و راست و یا یک الگوی خاص) می‌باشد. زیرمجموعه‌های از این روش به عنوان روش‌های چالش و پاسخ نیز دسته‌بندی می‌شوند، سامانه برای تشخیص به مشاهده حرکت خاصی از کاربر نیاز دارد که خود آن را تعریف کرده است (می‌توان این حرکت‌ها را به عنوان همان چالش‌ها در نظر گرفت). برخی از رویکردها از حرکات خود به خودی (غیرارادی) صورت برای تشخیص استفاده

می‌کنند (تغییر مردمک چشم نسبت به نور). روش‌های مبتنی بر حرکت از اطلاعات حرکتی مانند پلک زدن و حرکت لب برای تشخیص چهره واقعی از جعلی استفاده می‌کنند. به عنوان مثال، [141] زمینه تصادفی مشروط را برای مدل‌سازی مراحل مختلف پلک زدن بررسی می‌کند. فرض اصلی این است که چهره‌های واقعی الگوهای حرکتی مختلفی را در مقایسه با جعلی نشان می‌دهند. علاوه بر این، پلک زدن یکی از نشانه‌هایی است که در [142] برای تشخیص حملات جعل مانند حمله با کاغذ چاپ شده پیشنهاد شده است. در [143]، کالریدر و همکاران از حرکت لب برای تشخیص زنده بودن چهره استفاده می‌کنند. روش‌های پیشنهادی در [144] نشانه‌های صوتی و تصویری را برای تأیید زنده بودن چهره ترکیب می‌کنند. بائو و همکاران [145] با استفاده از زمینه‌های جریان نوری یک اقدام متقابل ارائه می‌دهد که تفاوت بین حملات عکس دو بعدی و چهره‌های واقعی را تخمین می‌زنند. این روش شامل تجزیه و تحلیل الگوهای حرکتی برای تشخیص حمله جعل است اما با حملات ماسک، حرکت‌های سر انسان را می‌توان به راحتی تکرار کرد. این روش‌های مبتنی بر حرکت در صورت حمله با عکس خم شده، حمله با عکس برش داده شده و حمله ماسک ممکن است، شکست بخورند.

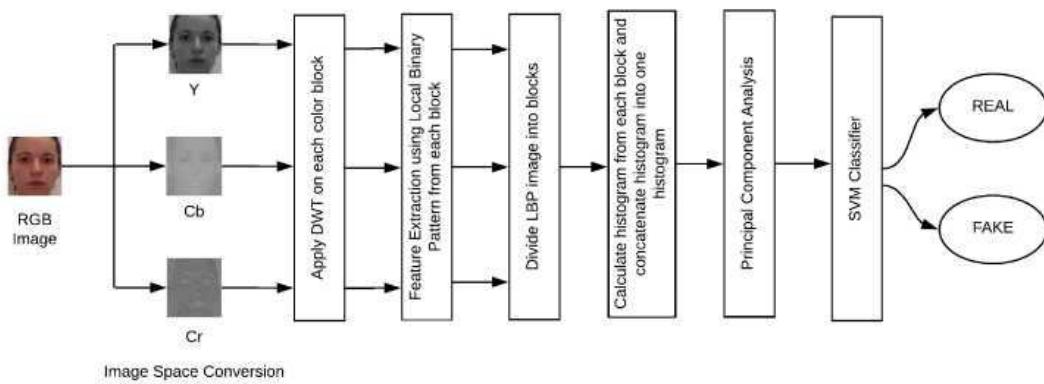
۴-۳-۴ رویکردهای مبتنی بر ترکیب

روش‌های مبتنی بر ترکیب در اکثر موقعیت‌ها از نظر کارایی و نتایج در دنیای واقعی بهتر از روش‌های دیگر عمل می‌کنند. رویکردهای ترکیبی می‌توانند در سه دسته ترکیب چند روش مبتنی بر بافت، ترکیب روش مبتنی بر بافت و حرکتی و روش‌های مبتنی بر ترکیب چند زیست‌سنجدی قرار بگیرند.

۴-۳-۱ ترکیب چند روش مبتنی بر بافت

روش ترکیبی، ویژگی‌های مختلف را در سطح ویژگی یا سطح امتیاز با هم پیوند می‌زند تا عملکرد تشخیص، بهبود یابد. ترکیب ویژگی‌های مختلف بافت، یکی از راه‌های مستقیم برای هم‌جوشی ویژگی‌ها است. روش‌های ترکیبی بافت و اندازه‌گیری کیفیت تصویر نیز عملکرد امبدوارکننده‌ای با پیچیدگی محاسباتی کم را نشان می‌دهد. این روش‌ها محدودیت‌های روش‌های مبتنی بر حرکت را بهبود می‌بخشد، اما هزینه محاسبه بالاتری دارند. تلفیق روش‌های مختلف و مطالعه چگونگی ترکیب ویژگی‌های مختلف برای ساختن چارچوب‌های PAD موثرتر و مقاوم، در سال‌های اخیر به طور فزاینده‌ای مورد توجه قرار گرفته است. پژوهش مهاره و تریپاتی [146] نمونه‌ای از روش ترکیبی است. نمودار روش پیشنهادی در شکل ۳-۱۰ نشان داده شده است. روش پیشنهادی براساس ویژگی‌های مبتنی بر بافت است. این رویکرد، ترکیب تبدیل موجک گستته^۱ (DWT) و الگوی دودویی محلی است. در حقیقت، رویکرد ترکیبی DWT و LBP با موفقیت در سامانه بازشناسی چهره برای تأیید کاربران مختلف اعمال شده است اما در این پژوهش، برای حملات جعل که توسط ماسک‌های صورت سه‌بعدی رخ می‌دهد، استفاده می‌شود. LBP معمولاً روش تحلیل بافتی است اما هدف استفاده از DWT، تجزیه سیگنال اصلی به فرکانس پایین و فرکانس بالا در دامنه فرکانسی است. ماسک‌ها فاقد مولفه‌های فرکانس بالا هستند. در اینجا به جای تصاویر خاکستری از تصاویر رنگی استفاده می‌شود زیرا قسمت رنگ‌نمایی در تصاویر خاکستری وجود ندارد و به جای RGB روی فضای رنگی YCbCr آزمایش شده است زیرا در YCbCr می‌توان قسمت روشناهی و رنگ‌نمایی را جدا کرد که در RGB امکان‌پذیر نیست. فریم‌های به دست آمده از DWT سپس توسط LBP پردازش می‌شوند. تصویر به دست آمده به بلوک‌هایی با اندازه مساوی و غیرهمپوشان تقسیم می‌شود. بردار ویژگی جداگانه برای هر بلوک به هم متصل تا یک بردار ویژگی واحد برای یک تصویر تشکیل دهد. بردار ویژگی حاصل به شکل هیستوگرام برای طبقه‌بندی استفاده می‌شود تا چهره واقعی و ماسک از هم تمیز داده شود.

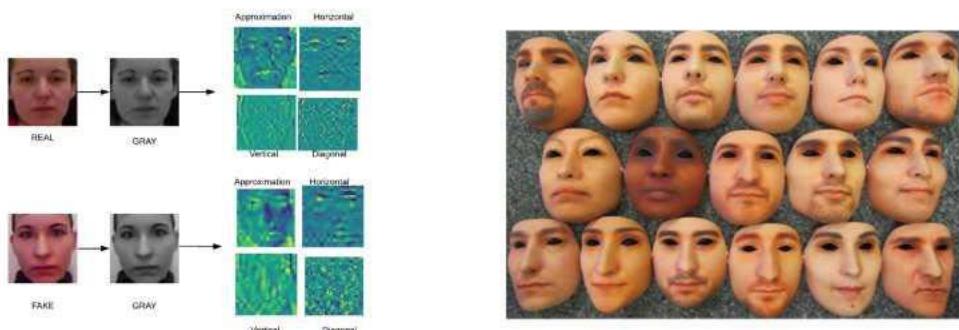
^۱ Discrete Wavelet Transform (DWT)



شکل ۱۲-۳ نمودار روش پیاده‌سازی شده الگوریتم ترکیب چند روش مبتنی بر بافت در پژوهش [146]

مجموعه داده و تنظیمات مربوط به آزمایش‌های صورت گرفته در این پژوهش در ادامه آورده شده است.

۱) مجموعه داده: پایگاه داده 3D Mask Attack 3D در این پژوهش مورد آزمایش قرار گرفته است. ماسک‌های رزینی از پایگاه داده 3DMAD در شکل ۱۱-۳ نشان داده شده است. هر فریم شامل (الف) یک تصویر عمق (ب) تصویر رنگی مربوطه (ج) موقعیت چشم می‌باشد. ۲) تنظیمات آزمایش: برای تحلیل و بررسی، هر ۱۰ فریم در نظر گرفته شده است. نتایج از نظر میزان دقیقت و HTER گزارش شده است. آزمایش به این گونه است که در هر بار، ۱۷ هدف از ۱۸ هدف موجود در مجموعه داده برای آموزش و یک هدف برای تست قرار داده می‌شد و در نهایت نتایج ۱۷ نفر به طور متوسط با یک روش خاص ترکیب می‌شود. قبل از تست، فیلم‌ها به فریم تبدیل می‌شوند و چهره‌ها از فریم بریده می‌شوند. تمام چهره‌ها به اندازه 64×64 نرمال می‌شوند. از LBP برای استخراج ویژگی استفاده می‌شود که به دو پارامتر نیاز دارد، تعداد نقاط همسایه P و شعاع دایره R و عملکرد آزمایش‌ها برای مقادیر مختلف P و R نشان داده شده است. از آنجایی که می‌تواند ترکیب‌های مختلفی از P و R وجود داشته باشد، آزمایش‌ها را برای سه مقدار P و R محدود کردند: (۸,۱)، (۱۶,۲) و (۲۴,۳).



شکل ۱۳-۳ نمونه‌ی تصاویر موجود در پایگاه داده مورد استفاده و ماسک‌هایی که به عنوان ابزار استفاده شده است [146]

آزمایش با روش پایه LBP آغاز می‌شود. سپس آزمایش‌هایی را بر روی بلوک‌ها با اندازه‌های مختلف و LBP چندمقیاس بر روی روش پیشنهادی اعمال کردند. دقیقت و HTER در روش پیشنهادی در جدول ۵-۳ آورده شده است. اثربخشی LBP در مقیاس چندگانه: LBP(8,1) و LBP(24,3) از LBP(16,2) بهتر عمل کرده اند. دقیقت (LBP(16,2)) و (LBP(24,3)) تقریباً برابر است. ماسک

چهره واقعی در منحنی و لبه‌های تصویر تفاوت دارند. ماسک‌ها در مقایسه با چهره واقعی لبه‌های تیزتری دارند. (8,1) LBP قادر به تشخیص ویژگی‌های غالب بافت نیست زیرا (8,1) LBP ناحیه‌ی بسیار کوچکی را در مقایسه با دو مورد دیگر در نظر می‌گیرد که منجر به از دست دادن اطلاعات می‌شود.^(۳) تجزیه و تحلیل بر اساس تصویر و بلوک: با اجرای آزمایش‌ها بر روی کل تصویر، یعنی ۶۴*۶۴، حداکثر دقت ۹۷.۱٪ را با استفاده از (24,3) LBP به دست آمد. همانطور که تعداد بلوک‌ها افزایش پیدا کرد، دقت از ۹۷.۱٪ بهبود پیدا کرد. اما افزایش بلوک‌ها باعث افزایش زمان محاسبه نیز می‌شوند.^(۴) نتایج رویکرد پیشنهادی: روش پیشنهادی (DWT + LBP(24,3)) بدون استفاده از بلوک‌ها، دقت ۹۷.۱٪ و HTER برابر با ۲.۵٪ را به همراه دارد. بهترین نتیجه برای (DWT + LBP(24,3)) با اندازه بلوک ۱۶*۱۶ با دقت ۹۹.۹۶٪ و HTER برابر با ۰.۰۱٪ بدست آمد.^(۵) محدودیت رویکرد پیشنهادی: نقطه ضعف روش پیشنهادی این است که تعداد ویژگی‌ها بسیار زیاد است. طول بردار ویژگی روش پیشنهادی برای LBP(24,3) ۴۹۹۲ است.

جدول ۳-۶ نتایج طبقه‌بندی با استفاده از الگوریتم ترکیب چند روش مبتنی بر بافت پیشنهادی^[۱۴۶]

	DWT+LBP(32*32)	DWT+LBP(16*16)	DWT+LBP(16*16)+PCA
(8,1)	۹۵.۸۹	۹۹.۹۲	۳۶.۳۸
(16,2)	۹۹	۹۹.۹۵	۹۶.۱۱
(24,3)	۹۹.۶۵	۹۹.۹۶	۹۹.۲۸

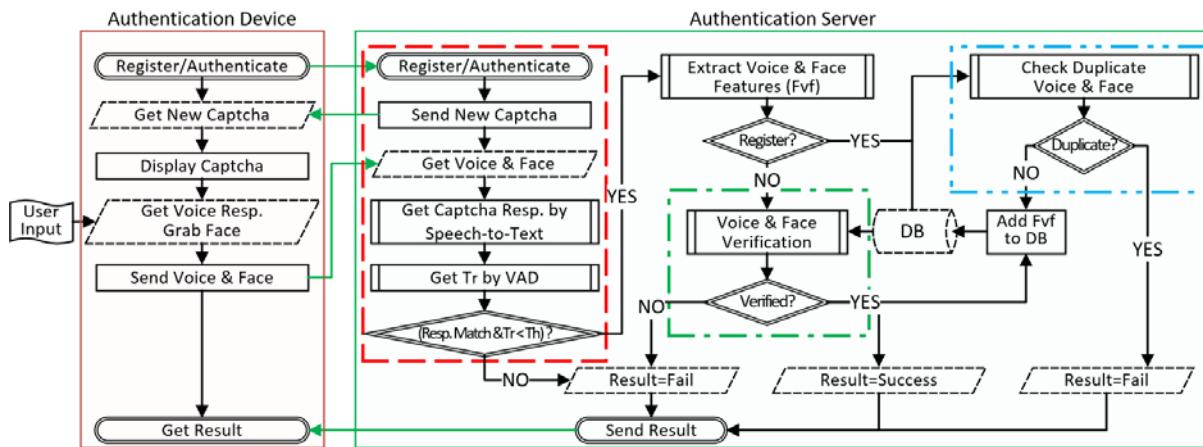
۴-۳-۲-۶ ترکیب چند روش زیست‌سنگی برای تشخیص زنده بودن

احراز هویت مبتنی بر چهره و صدا امروزه، به طور فزاینده‌ای محبوب شده است که یکی از دلایل آن استفاده آسان می‌باشد. کاربران اکنون می‌توانند با استفاده از تلفن همراه، خود را در خدمات آنلاین، احراز هویت کنند. بسیاری از خدمات تشخیص چهره/صدا که در دسترس عموم است حتی در ابتدایی ترین حملات نیز آسیب‌پذیر هستند. همه این موارد را می‌توان با سوء استفاده از دوربین و میکروفون تلفن قربانی یا از طریق حساب رسانه‌های اجتماعی او بدست آورد. در این پژوهش^[۱۴۷]، سامانه Real TimeCaptcha (rtCaptcha) ارائه شده است. این سامانه چالش لازم برای تشخیص زنده بودن را مشخص می‌کند که به عنوان کپچا^۱ رمزگذاری شده است. هنگامی که کاربر با استفاده از کپچا احراز هویت می‌کند، از او خواسته می‌شود هنگام اعلام پاسخ به کپچا، یک فیلم "سلفی" بگیرد. برای نشان دادن قابلیت استفاده و امنیت rtCaptcha، در این پژوهش یک مطالعه برای اندازه‌گیری زمان پاسخ انسان در رایج‌ترین طرح‌های کپچا انجام شده است. آزمایش‌ها نشان می‌دهد که به لطف سرعت انسان در حل کپچا، دشمنان برای فریب این سامانه‌ها و شکست دادن، باید در کمتر از ۲ ثانیه کپچا را حل کنند.

گردش کار rtCaptcha در شکل ۱۲-۳ خلاصه شده است. در این کار سرور، یک چالش را برای دستگاه احراز هویت (تلفن همراه) تولید و ارسال می‌کند و زمان پاسخگویی به دستگاه احراز هویت را اندازه‌گیری می‌کند. در دستگاه احراز هویت، چالش برای کاربر نشان داده می‌شود و پاسخ صوتی کاربر را ضبط می‌کند. برنامه‌ای که در دستگاه اجرا می‌شود، با استفاده از دوربین جلوی تلفن، به طور تصادفی چندین عکس فوری از کاربر می‌گیرد. نمونه‌های چهره گرفته شده با استفاده از کانال امن ایجاد شده به سرور ارسال می‌شوند. سرور با ضبط پاسخ صوتی دریافت شده با استفاده از یک الگوریتم استاندارد گفتار به متن، بررسی اولیه پاسخ را انجام می‌دهد تا تعیین کند آیا پاسخ با کپچا مطابقت دارد. در صورتی که زمان پاسخ به چالش از زمان تعیین‌شده‌ای که سامانه

^۱ Captcha

مشخص می‌کند، بیشتر باشد. آزمایش زنده بودن را بدون نتیجه و درخواست احراز هویت/ثبت‌نام را رد می‌کند. اگر پاسخ بدست آمده از بررسی‌های اولیه بگذرد، تجزیه و تحلیل قوی‌تری انجام تا اعتبار نمونه‌های صوتی و چهره موجود را بررسی کند.



شکل ۳-۱۴ نمودار روش پیاده‌سازی شده چند روش زیست‌سنجی در پژوهش [147]

۴-۴-۳ رویکرد شبکه عصبی عمیق در تشخیص زنده بودن

به دنبال موفقیت رویکردهای مبتنی بر یادگیری عمیق برای تشخیص چهره، روش‌های مبتنی بر CNN برای PAD چهره افزایش یافته است. یکی از دلایلی که محققان به دنبال استفاده از شبکه‌های عمیق در این زمینه هستند این است که با افزایش PAI‌ها، طراحی ویژگی‌های صریح دستی که بتواند جعلی را از واقعی تمایز کند، دشوارتر می‌شود.

در یکی از اولین کارها در این زمینه، یانگ و همکاران [148] یک CNN مشابه معماری ImageNet پیشنهاد کردند که لایه خروجی فقط برای دو خروجی پیکربندی شده است. در این کار نویسنندگان داده‌های آموزشی را با مقیاس‌های مختلف افزایش می‌دهند. برای استخراج بردار ویژگی برای هر تصویر ورودی، از آخرین لایه کاملاً متصل CNN آموزش دیده، استفاده می‌شود. سپس بردار ویژگی با استفاده از SVM، طبقه‌بندی می‌شود. کارهای اخیر بر روی CNN بر معماری‌های جدید CNN متمرکز شده است. لوسنا و همکاران [149]، یک شبکه عمیق FASNet4 را برای جعل چهره پیشنهاد داده‌اند. آن‌ها از VGGNet16 استفاده می‌کنند و فقط با حذف یک لایه کاملاً متصل^۱ (FC) و تغییر اندازه دو لایه FC بعدی به ۲۵۶ واحد و ۱ واحد، فقط قسمت کاملاً بالایی متصل به شبکه را تغییر می‌دهند. FASNet بهبود کمی نسبت به SpoofNet [150] در دو مجموعه داده 3DMAD و REPLAY-ATTACK که در هر دو کار استفاده می‌شود، نشان می‌دهد. ناگل و دوبی [151] عملکرد سه معماری مختلف CNN را با هم مقایسه می‌کنند: Inception-v3 [152] و دو نسخه ResNet [153] که شامل یک ۵۰ لایه‌ای و ۱۵۲ ResNet ۱۵۲ لایه‌ای می‌باشد. برای هر معماری، آن‌ها شش آزمایش را با آموزش شبکه‌هایی با تنظیمات پارامترهای مختلف انجام داده‌اند. مطالعه آن‌ها بر اساس مجموعه داده‌های MSU-MSFD که یک مجموعه داده کوچک است، می‌باشد. آن‌ها از flip فریم‌ها برای افزایش داده‌ها استفاده کردند. بهترین نتیجه به دست آمده در این کار، دقت ۹۷.۵٪، تولید شده توسط ResNet152 با وزن گرفته شده از ImageNet است و در آنجا فقط لایه‌های متراکم نهایی با استفاده از داده‌های MSU-MSFD دوباره آموزش دیده‌اند. آزمایش‌های آن‌ها همچنین نشان می‌دهد که استفاده از نرخ یادگیری پایین‌تر، ممکن است منجر به ایجاد تمایز بیشتری در آزمایش‌ها بر روی Face-PAD شود. لی و همکاران از CNN ترکیبی [154] استفاده کردند. ناحیه صورت به زیرمجموعه‌های مستطیلی تقسیم می‌شود

^۱ Fully Connected (FC)

و یک شبکه VGG-Face جداگانه برای هر ناحیه، آموزش داده می‌شود. با اتصال بردارهای خروجی از آخرین لایه کاملاً متصل هر CNN، یک بردار ویژگی ساخته می‌شود. سپس این بردار ویژگی با استفاده از SVM طبقه‌بندی می‌شود.

خو و همکاران [155] شبکه‌ای از LSTM و CNN را برای استخراج ویژگی‌ها رمزگذاری می‌کنند که هم اطلاعات زمانی و هم مکانی را شامل می‌شود. ورودی شبکه LSTM-CNN به جای فریم‌های منفرد، یک فیلم کوتاه است. LSTM در بالای CNN متصل شده است تا اطلاعات زمانی موجود در ویدئو را مدل کند. نویسنده‌گان نشان می‌دهند که این شبکه می‌تواند از CNN ساده و همچنین ویژگی‌های مختلف ساخته شده دستی بهتر عمل کند. لیو و همکاران [110] یک CNN و یک شبکه LSTM را ترکیب می‌کنند. در این معماری، CNN بر روی فریم‌های ویدئویی منفرد (تصاویر) آموزش داده می‌شود تا نگاشت‌های ویژگی‌های تصویر و همچنین نگاشت‌های عمق ناحیه چهره را استخراج کند. شبکه LSTM نگاشت ویژگی تولید شده توسط CNN را می‌گیرد و برای استخراج سیگنال در عکسبرداری از راه دور^۱ (rPPG) از فیلم، آموزش می‌بیند. آن‌ها نتایج مربوط را بر روی مجموعه داده‌های OULU-NPU ارائه می‌دهند.

به طور کلی، مجموعه داده‌های فعلی برای آموزش CNN‌ها کوچک است. از ابتدا بیشتر کارهایی که شامل CNN هستند با استفاده از یادگیری انتقالی سازگار شده‌اند. یکی از جدیدترین پژوهش‌ها که بر روی یک پایگاه داده‌ی بزرگ صورت گرفته است، روش ارائه شده در مقاله CelebA-Spoof [157] می‌باشد که علاوه بر ارائه یک پایگاه داده بزرگ در این حوزه یک روش بر پایه شبکه عصبی نیز پیشنهاد داده است. در این حوزه روش مورد استفاده در یک چارچوب چند منظوره منحصر به فرد، تحت عنوان شبکه تعبیه اطلاعات کمکی^۲ (AENet) اجرا می‌شود. در کنار دسته‌بندی دودویی معمول (جعلی و واقعی)، اطلاعات معنایی غنی به عنوان کارهای کمکی می‌تواند عملکرد و تعمیم‌پذیری تشخیص زنده بودن چهره را در طیف گسترده‌ای از حملات جعل افزایش دهد.

با در دست داشتن CelebA-Spoof، یک شبکه ساده و در عین حال قدرتمند به نام AENet در پژوهش طراحی شده است. در روش به کار گرفته شده مشاهدات مختلفی در نظر گرفته شده است: ۱) تحلیل تاثیر اطلاعات هندسی بر روی نوع جعل‌های مختلف و آشکار ساختن حساسیت اطلاعات هندسی به شرایط نوری مختلف. اطلاعات هندسی شامل نگاشت عمق^۳ و نگاشت بازتاب^۴ است. ۲) اطلاعات معنایی کمکی، از جمله ویژگی‌های چهره و نوع جعل، نقش مهمی در بهبود عملکرد دسته‌بندی دارد. ۳) ایجاد ۳ محکزنیCelebA-Spoof براساس اطلاعات کمکی. پژوهشگران ادعا می‌کنند که با ویژگی‌های مقایس بزرگ و تنوع بیشتر که در وجود دارد، می‌توانند شکاف بین مجموعه داده‌های ضد جعل چهره و صحنه‌های واقعی را از بین برند. همانطور که در شکل ۱۳-۳ نشان داده شده است، یک شبکه ساده و در عین حال فعال طراحی شده است که علاوه بر شاخه طبقه‌بندی اصلی دودویی (به رنگ سبز)، ۱) شاخه معنایی را ترکیب می‌کند (به رنگ نارنجی) تا از ظرفیت کمکی ویژگی‌های معنایی در مجموعه داده استفاده کند و ۲) اطلاعات کمکی هندسی موجود را در این چارچوب چند منظوره منحصر به فرد ارزیابی می‌کند.

در این پژوهش تأثیر خصوصیات معنایی منحصر به فرد بر AENet بررسی شده است. از آنجا که APCER توانایی طبقه‌بندی تصاویر جعلی را نشان می‌دهد، نتایج APCER نشان می‌دهد که در مقایسه با سایر ویژگی‌های معنایی، نوع جعل (ابزار مورد استفاده برای حمله به سامانه) به معنای واقعی، عملکرد طبقه‌بندی تصاویر جعلی AENet را تحت تاثیر قرار می‌دهد. از آنجا که BPCER توانایی طبقه‌بندی تصاویر زنده را بازتاب می‌دهد، نشان می‌دهد که ویژگی‌های معنایی در نظر گرفته شده کارایی سامانه را بهبود می‌بخشد. همچنین نگاشت‌های عمق می‌تواند عملکرد پایه را افزایش دهد به ویژه بهترین عملکرد را در نوع جعل 'replay' دارد و

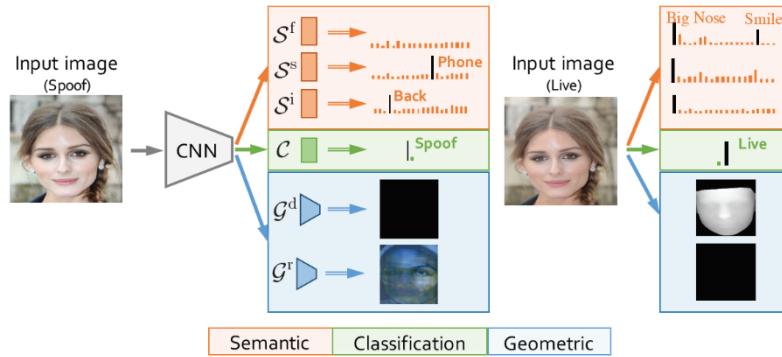
¹ Remote photo-plethysmography (rPPG)

²Auxiliary Information Embedding Network (AENet)

³Depth Maps

⁴Reflection Maps

بدترین عملکرد را در نوع ‘print’ دارد. نگاشت بازتاب نیز در ‘replay’ عملکرد را بهبود می‌دهد و در ‘A4’ بسیار بدتر از پوستر عمل می‌کند، اگرچه هر دو از نوع جعل print هستند. دلیل اصلی تفاوت زیاد در عملکرد در بین این سه نوع جعل برای AENet است که یادگیری نگاشت به روشنایی و تاریکی حساس است.



شکل ۳-۱۵ نمودار روش پیاده‌سازی شده CelebA-Spoof در پژوهش [157]

۳-۳ جمع‌بندی

در جدول ۳-۶ اطلاعات روش‌های بیان شده در گزارش، به طور خلاصه آورده شده است.

جدول ۳-۷ خلاصه اطلاعات روش‌های بیان شده برای تشخیص زنده بودن

نتایج	نوع حمله	پایگاه داده	روش	نویسنده و سال
سطح زیر نمودار (AUC) برابر با ۰.۹۹ و نرخ خطای تساوی (EER) معادل با ۲.۹	تصویر چاپ شده و نمایش از روی صفحه	NUAA	مبتنی بر الگوهای دودویی محلی (LBP)، طبقه‌بندی با ماشین بردار SVM پشتیبان	Matta و همکاران [132] ۲۰۱۱
بدون تقسیم به بلوک‌ها، HTER برابر با ۴.۱٪	نمایشگر		استفاده از ویژگی‌های هارالیک بافت چهره، جداسازی کانال‌های رنگی	
مبتنی بر بلوک‌ها، HTER برابر با ۷/۳	ماسک و بازپخش ویدیو	3DMAD, CASIAFASD, MSU-MFSD	تصویر، استفاده از تبدیل موجک گسسته، بلوک‌بندی تصویر، طبقه‌بندی با SVM	Agarwal و همکاران [133] ۲۰۱۶
استخراج از کانال قرمز با استخراج ویژگی بلوک‌ها، HTER برابر با صفر				
بهترین نتیجه حاصل از ترکیب روش‌ها و تقسیم صورت به ناجیه‌های مختلف می‌باشد.	تصویر چاپ شده و نمایش از روی صفحه	NUAA, CASIA, PRINT-ATTACK	استخراج ویژگی با هیستوگرام شبیه‌های جهتدار (HoG) و تفاوت LBP	Yang و همکاران [134] ۲۰۱۳
NUAA Dataset:	نمایشگر			

نويسنده و سال	روش	پايگاه داده	نوع حمله	نتایج
Ojala و همکاران [137] ۲۰۰۲	SVM	دو پايگاه داده شخصی	ناحیه‌بندی صورت، طبقه‌بندی با روش مبتنی بر طیف توان، روش توصیف مبتنی بر الگوی دودویی	، AUC: ۰.۹۷۷ Accuracy: ۰.۹۹۸
He و همکاران [140] ۲۰۱۹	ترکیب مدل‌های مختلف فضاهای Replay-Attack، CASIA-FASD	دو پايگاه داده شخصی دو بعدی	ماسک‌های کاغذی بازپخش ویدیو	EER در الگوی دودویی محلی: ۱۱.۵۸ EER در فرکانسی: ۱۱.۸۷ EER در ادغام: ۸.۴۳ بهبود عملکرد تشخیص با استخراج ویژگی‌های LBP در فضای رنگی YCbCr و ویژگی‌های لحظه‌ای HSV در فضای رنگی
Mahore و همکاران [146] ۲۰۱۸	روش ترکیبی مبتنی بر ویژگی‌های بافت، ترکیب تبدیل موجک گسسته و الگوی دودویی محلی در فضای رنگی YCbCr	3D Mask Attack	ماسک سه‌بعدی	بدون استفاده از بلوک‌ها، دقت ۹۷.۱٪ HTER ۲.۵٪ بلوک‌بندی با اندازه بلوک ۱۶*۱۶ با دقت HTER ۰.۰۱ و ۹۹.۹۶٪ از نمونه‌های اصلی به درستی ۹۷٪
Uzun و همکاران [147] ۲۰۱۸	ترکیب زیست‌سنگی صدا و گفتار، استفاده از چالش و پاسخ (کپچا)، تبدیل گفتار به متن	CASIA Face Anti-Spoofing	تصویر چاپ شده، ماسک و بازپخش ویدیو	شناسایی شده در سامانه تشخیص گفتار دقت درستی پاسخ به کپچا توسط شرکت‌کنندگان ۸۹.۲ درصد در زمان ۰.۹۸ ثانیه عملکرد بهبود سامانه با ویژگی‌های معنایی به ویژه نوع جعل، تأثیر اطلاعات هندسی در تشخیص در بازپخش ویدیو
Zhang و همکاران [157] ۲۰۲۰	استخراج ویژگی با شبکه عصبی، استفاده از ویژگی معنایی و هندسی تصویر برای طبقه‌بندی	Celeba spoof	تصویر چاپ شده، ماسک دو بعدی و بازپخش ویدیو	عملکرد بهبود سامانه با ویژگی‌های معنایی به ویژه نوع جعل، تأثیر اطلاعات هندسی در تشخیص در بازپخش ویدیو

با استفاده گسترده از هوش مصنوعی در زندگی واقعی، بازناسی چهره به وسیله‌ای مهم برای تحقق امنیت تبدیل شده است. به منظور جلوگیری از حملات مخرب، مواجهه با جعل و تقلب مسئله‌ی مهمی را ایجاد کرده است. از ابتدای روش استخراج ویژگی‌ها به شیوه‌ی دستی بر اساس بافت تصویر، اطلاعات زنده بودن، کیفیت تصویر و اطلاعات عمق و سپس استفاده از یادگیری عمیق برای استخراج خودکار ویژگی، همراه با به روز رسانی شبکه، یادگیری انتقالی، ادغام ویژگی‌ها و دامنه به طور کلی، مطالعه تشخیص زنده بودن به طور مداوم به روز شده و بهبود یافته است و کارآیی و دقت تشخیص اکنون به وضعیت قابل توجهی رسیده است. با این حال، تأثیر اندازه پایگاه داده بر دقت تشخیص بسیار اهمیت دارد. چگونه می‌توان از تشخیص دقیق زنده بودن اطمینان حاصل کرد، در حالی که توانایی تعمیم حمله‌ها و جعل‌های ناشناخته را به دلیل کمبود داده نداریم. در عین حال، هم به روزرسانی شبکه

و هم آموزش انتقالی ایده‌های خوبی برای بهبود عملکرد الگوریتم‌های تشخیص زنده بودن است که شایسته تحقیقات بیشتر توسعه محققان بیشتر است.

۳-۵-۱ تحلیل و مقایسه رویکردهای تشخیص زنده بودن

از نتایج ارزیابی ارائه شده، می‌توان دریافت که تشخیص زنده بودن چهره هنوز یک مشکل بسیار چالش برانگیز است. به طور خاص، عملکرد روش‌های فعلی هنوز پایین‌تر از الزامات اکثر برنامه‌های کاربردی در دنیا واقعی (به ویژه از نظر قابلیت تعمیم) هستند.

با این حال، همه ویژگی‌هایی که به صورت دستی استخراج می‌شوند توانایی تعمیم محدودی را دارند، زیرا به اندازه کافی قدرتمند نیستند که بتوانند همه تغییرات احتمالی را در شرایط مختلف ضبط چهره به دست آورند. ویژگی‌های آموزش دیده که توسط شبکه‌های عصبی عمیق استخراج می‌شوند در مقایسه با حجم محدود داده‌های آموزشی، دارای ابعاد بسیار بالایی هستند و از بیش‌برازش و در نتیجه تعمیم‌پذیری ضعیف رنج می‌برند. بنابراین، ویژگی‌های یادگیری که قادر به تمایز بین یک چهره واقعی و هر نوع PA باشند، احتمالاً با توجه به شرایط، بسیار متفاوت هستند.

همانطور که قبلاً گفته شد، ویژگی‌های یادگیری که بتوانند به حد کافی بین چهره‌های واقعی و PA‌های مختلف، تمایز قائل شوند هنوز یک چالش بزرگ است. البته، این نوع مسائل (مربوط به قابلیت‌های تعمیم مدل‌های داده محور) در زمینه بینایی ماشین، بسیار فراتر از تشخیص زنده بودن چهره است.

برای مقابله با تمام حملاتی که قبلاً دیده شده است، روند مناسب، ترکیب چندین روش است. با این حال، با توجه به چالش‌های ذکر شده در ایجاد مجموعه داده و همچنین پیشرفت‌های تکنولوژیکی که کاربران کلاهبردار می‌توانند به منظور توسعه حملات پیچیده به آن دسترسی پیدا کنند، روش‌های تشخیص زنده بودن ممکن است مجبور به شناسایی حملاتی شوند که در مجموعه داده آموزشی آن گنجانده نشده است.

در فصل سه تأثیرگذارترین روش‌های تشخیص زنده بودن بیان گردید که می‌توانند در سناریوهایی که کاربر فقط به دوربین‌های RGB دستگاه‌های عمومی دسترسی دارد، به کار گرفته شود. روش‌های مبتنی بر بافت که بیشترین استفاده را در تشخیص زنده بودن دارند و به ویژه روش‌های مبتنی بر بافت پویا، قادر به تشخیص تقریباً همه انواع حملات هستند. علاوه بر این، روش‌های مبتنی بر ویژگی‌های بافتی که با استفاده از یادگیری عمیق آموزش دیده‌اند، در مقایسه با روش‌های مبتنی بر ویژگی‌های بافت دستی، به طور چشمگیری عملکردهای تشخیص زنده بودن چهره را بهبود بخشیده است. از سوی دیگر، روش‌های تشخیص زنده بودن مبتنی بر روش‌های هندسی سه‌بعدی می‌توانند به قابلیت‌های نسبتاً بهتری برسند، حتی اگر هنوز در برابر حملات پخش مجدد ویدئو یا شرایط پیچیده روشناهی آسیب‌پذیر باشند. با این وجود، پیچیدگی محاسباتی این روش‌ها موضوعی است که باید برای برنامه‌های زمان واقعی در نظر گرفته شود. تا حدی به دلیل پیچیدگی مسئله تشخیص زنده بودن چهره، تنوع زیاد در حملات احتمالی و عدم وجود مجموعه داده که شامل نمونه‌های کافی با تنوع کافی باشد، همه رویکردهای فعلی هنوز از نظر تعمیم محدود هستند.

فصل ۴ تحلیل نیازهای پروژه

تغییرات و پیشرفت در دنیای فناوری و دیجیتال، سازمان‌ها را بر آن داشته تا قبل از اینکه زیر امواج سهمگین این تحولات غرق شوند، سازمان خود را متناسب با تغییرات عصر حاضر متحول کنند. در فضای سازمانی، تحول را یک فرآیند فراگیر می‌دانند که منجر به جهت‌گیری سازمان در مسیری نوین شده و آن را به سطحی بسیار بالاتر از اثربخشی ارتقا دهد. تحول دیجیتال عبارتست از اتخاذ استراتژیک در رابطه با فناوری‌های دیجیتال به منظور بهبود فرایندها، افزایش بهره‌وری، مدیریت ریسک‌های کسب و کاری و بهبود خدمات ارائه شده به مشتریان. استراتژی‌های مدرن تحول دیجیتال، به شکل روزافزون راههای پیچیده و سنتی را با فناوری‌های دیجیتالی ساده‌تر جایگزین می‌کنند.

تعدادی فناوری وجود دارد که می‌توانند به فرایند تحول کمک کنند. کلان داده‌ها^۱، تجزیه و تحلیل و گزارش‌دهی فوری^۲، رایانش ابری^۳ و تلفن همراه، اینترنت اشیا^۴، هوش مصنوعی^۵ و یادگیری ماشین^۶، برخی از گزینه‌های تکنولوژی هستند که سازمان‌ها در فرآیند تحول دیجیتال خود به کار می‌برند.

¹ Big data

² Realtime analysis and reporting

³ Cloud computing

⁴ Internet of Things

⁵ Artificial intelligence

⁶ machine learning

دولت نیز در راستای حل مشکلات کشور، دست به سوی دنیای فناوری دراز کرده و سند تحول تهیه کرده است. در این سند آمده که با بهره‌گیری از تحول دیجیتال می‌توان چالش‌های کلان کشور را حل و به بهبود شرایط زیست شهروندان، کسب‌وکارها و دولت کمک کرد. همچنین وزارت کشور نیز مکلف است تا با همکاری قرارگاه پدافند سایبری و چند نهاد دولتی دیگر پلیس هوشمند مبتنی بر فناوری دیجیتال را در ایران طراحی کند و توسعه دهد.

با رشد جهانی فناوری‌های جدید نظیر اینترنت اشیا و هوشمندشدن شهرها و زیرساخت‌های آن، نیروهای نظامی و انتظامی که همواره جز صنایع پیشرو هستند، در سراسر جهان با این روند جهانی همراه شده و در حال کاربردی ساختن تکنولوژی‌های جدید در ساختار خود هستند. اگرچه که عبارت هوشمندسازی کلی بوده و شامل راهکارهای مختلف و متنوع نرمافزاری و سختافزاری می‌شود. در ایران نیز هوشمندسازی خدمات پلیس جز برنامه‌های این ارگان است و در آینده‌ای نزدیک شاهد تحولات اساسی در خدمات الکترونیک پلیس با بهره‌مندی از این ابزارهای هوشمند به ویژه فناوری‌های هوش مصنوعی خواهیم بود. بر این اساس چندین تیم تخصصی درون و برونو سازمانی و بخش‌های مختلف از پلیس‌های تخصصی در قالب کارگروه مشترکی روی پروژه‌های متعدد در حال بررسی، اقدام و عمل هستند.

در حوزه خدمات مرتبط با پلیس، در سال‌های اخیر شاهد تغییراتی اساسی در نحوه خدمت‌رسانی به مردم، از ایجاد سامانه‌های هوشمند و استفاده از تجهیزات نوین با بهره‌مندی از هوش مصنوعی در حوزه‌های مختلف از جمله گذرنامه، گواهینامه، خدمت سربازی گرفته تا خدمات الکترونیکی، دوربین‌های کنترل ترافیک و تکمیل باندهای اطلاعاتی و آگاهی هستیم.

در حال حاضر ارائه خدمات الکترونیک نیروی انتظامی به خوبی انجام می‌شود. اصلاح فرآیند هوشمندسازی جزء اولویت‌های پلیس است و در حال حاضر برای دریافت گذرنامه دیگر نیازی به مراجعه به دفاتر «پلیس⁺¹⁰» نیست البته به شرطی که احراز هویت برای پلیس در سایت انجام شده باشد و بعد از آن مردم می‌توانند گذرنامه خود را درب منازل تحويل بگیرند. حرکت به سمت هوشمندسازی، اجتناب‌نایدیزیر است و به عبارتی دیگر باید میز پلیس را در منازل برد و مردم از طریق سامانه با نیروی انتظامی ارتباط داشته باشند و خدمات را با سرعت بیشتری به مردم ارائه دهند که این از اهداف پلیس هوشمند و زیرساخت‌های آن در حال آماده‌سازی است.

سامانه‌های هوشمند پلیس در بستر اپلیکیشن «پلیس همراه من» و سایت «پلیس راهور» و برخی سامانه‌های دیگر بارگذاری شده است. مردم با مراجعه به این سامانه‌ها می‌توانند به سهولت از خدمات راهنمایی و رانندگی ناجا بهره‌مند شوند. برای مقایسه‌ی خدمات ارائه شده در دفاتر پلیس⁺¹⁰ و سامانه‌های هوشمند در ادامه به بررسی خدمات این دفاتر و سامانه‌های هوشمند پرداخته می‌شود.

۴-۱ دفاتر پلیس⁺¹⁰

دفاتر خدمات الکترونیک انتظامی زیرمجموعه‌ای از خدمات الکترونیکی ارائه شده توسط پلیس در ایران است که در دفاتری به نام دفاتر خدمات الکترونیک انتظامی (یا دفاتر پلیس⁺¹⁰) ارائه می‌شود. اعم خدمات پلیس⁺¹⁰ شامل خرید دفترچه راهنمای وظیفه عمومی، خدمات اینترنتی وظیفه عمومی، تعویض کارت پایان خدمت و معافیت، صدور المثنی کارت پایان خدمت و معافیت، تعویض و تمدید گواهینامه، پیگیری کارت سوت، صدور المثنی کارت سوت، صدور و تمدید گذرنامه، صدور وضعیت خلافی خودرو، طرح اطلاعات اقتصادی خانوارها، صدور گواهی پروانه کسب، صدور گواهی عدم سوء پیشینه و بیمه حوادث مسافران عازم به خارج از کشور می‌شود. طرح پلیس⁺¹⁰ با سه خدمت عمومی گذرنامه، گواهینامه، اجرائیات (تخلفات راهنمایی و رانندگی) در دهه فجر انقلاب اسلامی سال ۱۳۸۲ آغاز شد و به تدریج تعداد دفاتر در سطح کشور افزایش پیدا کرد و بر اساس آیین‌نامه و کانون آن‌ها به

بخش خصوصی واگذار گردید. هدف‌گذاری اولیه، ارائه ۱۰ خدمت از خدماتی بود که پیشتر در سازمان نیروی انتظامی (ناجا) ارائه می‌شد.

۴-۱-۱- فهرست خدمات قابل انجام در دفاتر پلیس + ۱۰

۱. خدمات گواهینامه شامل تمدید و تعویض انواع گواهینامه رانندگی و صدور المتنی
۲. خدمات صدور گذرنامه شامل اخذ مدارک متقاضی گذرنامه و ثبت اطلاعات در سیستم
۳. خدمات اجراییات شامل صدور صورت وضعیت خلافی خودرو
۴. خدمات صدور و تمدید پروانه کسب (اماکن)
۵. استعلام تشخیص هویت (سوء پیشینه کیفری)
۶. ثبت درخواست‌های مشمولین (نظام وظیفه)
۷. رسیدگی غیرحضوری به شکایات صورت وضعیت
۸. صدور و المتنی کارت هوشمند سوخت
۹. تغییر آدرس مالکان خودرو

۴-۱-۲- اپلیکیشن پلیس + ۱۰

از این اپلیکیشن می‌توان برای خدمات زیر استفاده کرد:

۱. خرید سریال دفترچه وظیفه عمومی
۲. ثبت درخواست ارسال مجدد کارت‌های عودتی (کارت‌های معافیت یا پایان خدمت برگشت‌خورده از پست)
۳. ثبت درخواست تمدید و تعجیل اعزام به خدمت
۴. استعلام‌های خلافی خودرو، موتور
۵. آخرین وضعیت پایان خدمت
۶. آخرین وضعیت گواهینامه
۷. نمره منفی گواهینامه
۸. وضعیت کارت سوخت
۹. وضعیت تمدید یا تعجیل
۱۰. وضعیت ارسال مجدد کارت پایان خدمت.
۱۱. اپلیکیشن پلیس یار همراه

پلیس یار همراه یکی از دو اپلیکیشن نیروی انتظامی است که نسبت به دیگری امکانات بیشتری دارد. اپلیکیشن به سه قسمت راهور، وظیفه عمومی و پلیس + ۱۰ تقسیم می‌شود. رابط کاربری آن ساده است و خدمات تقریباً در تبعهای مختلف دسته‌بندی شده‌اند.

بخش اطلاع‌رسانی راهور علاوه بر وضعیت آب‌وهوای و مهمنه‌ترین اخبار، محدودیت‌های ترافیکی را به شما نشان می‌دهد، البته تمام این‌ها در قالب لینک به سایت اصلی است و از روش متفاوت و خوبی برای نمایش اطلاعات استفاده نشده است. همچنین امکان دارد خبرهایی را ببینید که مربوط به سال ۱۳۹۲ است و همچنان جزو اخبار اخیر دسته‌بندی شده است؛ بنابراین مخصوصاً در بخش آب‌وهوای انتظار زیادی نداشته باشید و سعی کنید از سایر اپلیکیشن‌های آب‌وهوای استفاده کنید که به‌روزتر هستند و رابط کاربری بهتری هم دارند. با این حال «فهرست تخلفات رانندگی» که در این بخش قرار گرفته، می‌تواند کاملاً کاربردی باشد که هزینه و نمره منفی هر تخلف را در گرافیک مناسبی نمایش داده است. در بخش خدمات راهور هم امکان مشاهده لیست خلافی خودرو که یکی از رایج‌ترین علل مراجعت افراد به پلیس + ۱۰ است، فراهم شده. البته برای استفاده از این قابلیت باید در اپلیکیشن ثبت‌نام کنید و بعد از آن وارد حساب کاربری خود شوید. برای مشاهده لیست خلافی کافی است شماره سریال کارت وسیله نقلیه را داشته باشید. همچنین می‌توانید بارکد کارت را هم با استفاده از اسکنر داخلی اپلیکیشن بخوانید.

۴-۱-۲- وظیفه عمومی

در بخش وظیفه عمومی علاوه بر اینکه به راحتی می‌توانید به فایل قانون وظیفه عمومی دسترسی داشته باشید، آخرین اخبار هم به ترتیب زمان برای شما لیست شده است که خوشبختانه از مهرماه امسال به خوبی به‌روز شده و می‌توان برای پیگیری اخبار سربازی روی آن حساب کرد. خدمات مربوط به وظیفه عمومی هم به خرید اینترنتی دفترچه و محاسبه جریمه (برای افراد مشمول خرید سربازی) محدود می‌شود. بخش محاسبه جریمه قابلیت ساده‌ای است که به چند ضرب و تقسیم محاسبه می‌شود اما بودن آن در این اپلیکیشن احتمالاً به درد خیلی‌ها بخورد.

۴-۱-۳- پلیس +

بخش پلیس + ۱۰ این اپلیکیشن در حقیقت کپی قسمتی از اپ دیگر نیروی انتظامی است. در بخش اطلاع‌رسانی مدارکی را که برای کارهای مختلف نیاز دارید، مشاهده می‌کنید و در یک قابلیت کاملاً کاربردی می‌توانید دفاتر خدمات الکترونیکی پلیس را روی نقشه به صورت دقیق مکان‌یابی کنید و مسیر رسیدن به این دفاتر را از جایی که هستید، بیابید. رابط کاربری این بخش به خاطر اینکه با ابزارهای گوگل طراحی شده بسیار سریع و زیباست و می‌توان آن را یکی از بهترین قسمت‌های اپ پلیس به شمار آورده.

۴-۲- اپلیکیشن پلیس من

اپلیکیشن «پلیس من» محصولی جدید با امکانات متنوع از شرکت پژوهش و توسعه ناجی است که با همکاری و هماهنگی فاوا ناجا تولید شده است. این سامانه با هدف سهولت در دسترسی مردم به خدمات الکترونیک انتظامی تهیه و تولید شده است. امکان احراز کاربر جهت ارائه خدمات الکترونیکی متعدد در حوزه انتظامی و پلیسی، مسیریابی، امکان نمایش دفاتر پلیس + ۱۰، کلانتری‌ها، ادارات گذرنامه و سایر مکان‌های مرتبط با پلیس روی نقشه و امکان مسیریابی و ناویگری کاربر به هر یک از مکان‌های انتخاب شده از روی نقشه از کارکردهای این سامانه می‌باشد. نمایش اطلاعات وضعیت پلاک‌های فعال و غیرفعال کاربر، استعلام خلافی خودرو، نمایش میزان بدھی (خلافی) خودرو و امکان پرداخت قبوض جریمه از طریق سامانه و رویت تصویر دوربین ثبت تخلفات (برای اولین بار) از دیگر کارکردهای پلیس من است. نمایش اطلاعات مربوط به وضعیت صدور کارت و سند خودرو و نمایش شماره پیگیری مرسوله و استعلام نمره منفی گواهینامه با دریافت شماره گواهینامه نیز از دیگر کارکردهای این سامانه است. نمایش آخرین

وضعیت گواهینامه، استعلام گذرنامه، استعلام وضعیت خروج از کشور افراد دارای گذرنامه (برای اولین بار)، امکان ثبت درخواست صدور و تعویض اینترنتی گذرنامه (به صورت آزمایشی در چند شهر کشور)، امکان انجام نقل و انتقال اینترنتی خودرو (به صورت آزمایشی در یک مرکز تعویض پلاک تهران) از دیگر کارکردهای این اپلیکیشن است.

یکی از مشکلاتی که برخی افراد با آن رو به رو هستند، بی اطلاعی افراد دارای گذرنامه از وضعیت خروج از کشور بود، ممکن است برخی افراد به دلایلی نظیر بدھی مالیاتی، دستور قضائی امکان خروج از کشور را نداشته باشند، اما چون از آن بی خبر بودند اقدام به تهیه بلیت هواییما و رزرو هتل کرده اما حین خروج با مشکل مواجه می شوند. در همین راستا و با هدف ارتقاء خدمات الکترونیک پلیس، در نرم افزار پلیس من امکان استعلام وضعیت خروج از کشور افراد دارای گذرنامه، اعتبار گذرنامه، پرداخت عوارض خروج از کشور فراهم شده است.

در گذشته برای گرفتن خلافی خودرو، استعلام ممنوع الخروجی، ارائه درخواست تمدید گذرنامه و یا حتی مشاهده نمرات منفی یک گواهینامه باید به مراکز پلیس + ۱۰ مراجعه می شد. حالا اما نیروی انتظامی با توسعه برنامه پلیس من برای پلتفرم تلفن همراه اندروید و iOS نیازهای کاربران از مراجعه حضوری را کاهش داده است تا در بازه زمانی کرونا از ورود به صفحه های شلوغ پلیس + ۱۰ خودداری شود. پلیس من را می توان به عنوان برنامه پلیس + ۱۰ نیز بدانیم که خدمات خود را به صورت آنلاین ارائه می دهد و قطعا هر شهروند ایرانی، نیاز به استفاده از این اپلیکیشن را خواهد داشت. بخش اپلیکیشن پلیس من شامل خدمات پلیس راهور و گذرنامه می شود.

۴-۲-۱ امکانات بخش راهور در اپلیکیشن پلیس من

این بخش شامل خدمات زیر می باشد(شکل ۱-۴):

۱. استعلام تخلفات رانندگی خودرو و موتورسیکلت
۲. نوبت دھی اینترنتی شماره گذاری (تعویض پلاک)
۳. آخرین وضعیت گواهینامه
۴. گزارش تخلف حمل و نقل عمومی
۵. نمره منفی
۶. وضعیت پلاک ها
۷. استعلام کارت و سند خودرو
۸. شماره گذاری اینترنتی



شکل ۱-۴ بخش خدمات راهور در اپلیکیشن پلیس من

۲-۲-۴ امکانات بخش گذرنامه در اپلیکیشن پلیس من

این بخش شامل خدمات زیر می‌باشد (شکل ۲-۴):

۱. وضعیت خروج از کشور افراد دارای گذرنامه
۲. پرداخت عوارض خروج از کشور
۳. ثبت درخواست صدور و تعویض اینترنتی گذرنامه



شکل ۴-۲ بخش خدمات گذرنامه در اپلیکیشن پلیس من

۱۲۰ - ۳ راهور

سایت rahvar120.ir که مربوط به خدمات راهنمایی و رانندگی است پورتالی دارد که از طریق آن می‌توان در کنار اخبار و اطلاع‌رسانی‌های خود به خدمات مختلفی دسترسی داشت. در بخش کارت سوخت هم می‌توانید در کنار پیگیری صدور کارت سوخت، مدارک موردنیاز برای صدور المثنی را پیدا کنید. اگر صاحب خودرویی هستید می‌توانید در بخش اصلاح اطلاعات مالکین خودرو سایت راهور، اطلاعات خود را ویرایش کنید یا تخلفات را مشاهده کنید و جریمه‌هایتان را بپردازید. برای اطلاع از تخلف‌ها هم کافی است شماره بارکد پستی کارت ماشین خود را وارد کنید. البته محتویات این بخش از سایت آفلاین است و تقریباً هر یک هفته یک بار به روزرسانی می‌شود و به هیچ‌وجه به بانک اطلاعاتی پلیس راهور متصل نیست.

۴-۴ ضرورت احراز هویت الکترونیکی

نیروی انتظامی با وجود محدودیت‌های بودجه و تجهیزات، بخش‌های قابل توجهی از خدمات مهم پلیس را که مردم بیشتر با آن درگیر هستند، هوشمند کرده است اما تا رسیدن به نقطه نهایی هنوز فاصله زیادی وجود دارد. با توجه به اهداف هوشمندسازی و سامانه‌های هوشمند بیان شده و خدماتی که آن‌ها ارائه می‌دهند، همچنان مشاهده می‌شود که افراد برای بسیاری از کارهای ضروری خود نیاز به مراجعه حضوری به دفاتر را دارند. به طور مثال، افرادی که برای تعویض کارت قدیمی معافیت و پایان خدمت اقدام نکرده‌اند، می‌توانند برای تعویض کارت معافیت و پایان خدمت به هوشمند از طریق دفاتر خدمات الکترونیک انتظامی (پلیس+۱۰) اقدام کنند و همچنین خدماتی مانند استعلام تشخیص هویت به دلیل فرایند احراز هویت، همگی نیاز به حضور در دفاتر را دارند. همچنین بخش‌های دیگری از سایت راهور ۱۲۰ مثل گواهینامه، گذرنامه و پیشگیری انتظامی تنها شامل اطلاعات و مدارک لازم می‌شود و به جز صدور دفترچه کفالت برای اتباع خارجی، خبری از خدمت دیگری نیست.

در برنامه توسعه به ازای هزار نفر، پنج پلیس باید باشد و این تعداد در بعضی از کشورها هشت نفر است ولی در حال حاضر در ایران حدود نصف این تعداد پلیس مشغول خدمات رسانی هستند که موجب فشار بر روی نیروی انتظامی می‌شود. موضوع

هوشمندسازی، بخشی از خلاً نیروی انتظامی را می‌پوشاند و هرچه این هوشمندسازی در جهت درست و کامل پیش رود علاوه بر احتیاج به نیروی انسانی کمتر به راحتی کاربر و هم‌چنین کاهش هزینه‌های سازمان کمک می‌کند که قدم اول در این راستا، پیاده‌سازی احراز هویت برخط (آنلاین) می‌باشد.

قابل به ذکر است که نیروی انتظامی جمهوری اسلامی ایران یک نهاد عظیم با رسته‌های مختلف است. شمار نیروی انسانی آن در سراسر ایران به بیش از چندصد هزار نفر می‌رسد. از این رو، ارائه خدمات به نیروهای این نهاد به دلیل گستردگی و پراکندگی، خود یکی دیگر از چالش‌های موجود است. کنترل این افراد و نظارت بر آن‌ها به خصوص به دلایل امنیتی نیازمند احراز هویت می‌باشد که به دلایل بیان شده، این نظارت تنها در صورت الکترونیکی بودن (از راه دور) دارای توجیه اقتصادی، امنیتی و مدیریتی می‌باشد.

هم‌چنین در بعضی از شهرستان‌ها به دلیل کمبود دفاتر پلیس + ۱۰ و یا ایام خاص سال (شکل ۳-۴)، مراجعه به این دفاتر افزایش پیدا می‌کند. اما در صورت کاهش لزوم مراجعه به این دفاتر به دلیل انتقال خدمات آن‌ها به سامانه‌های هوشمند از طریق احراز هویت الکترونیکی، می‌توان آسایش کاربران به خصوص برای شهرستان‌ها و مناطق روستایی را تأمین نمود.



(ب) علت ازدحام، تعداد کم دفاتر در شهرستان



(الف) علت ازدحام، دریافت گذرنامه برای سفر در ایام اربعین

شکل ۳-۴ ازدحام مردم در مقابل دفاتر پلیس + ۱۰

باید به این نکته اشاره کرد که با تحول دیجیتالی و هوشمندسازی، همواره نیازهای جدیدی به وجود می‌آید که راه حل‌های مناسب آن‌ها نیز در همین حوزه ارائه می‌شود. وجود بستری مانند احراز هویت الکترونیکی (از راه دور) می‌تواند انعطاف کافی برای اجرای چنین راه حل‌هایی را در زمینه امنیتی و خدمات پلیس ایجاد کرده و توانایی کنترل یکپارچه بر روی کاربران و مدیریت اطلاعات آن‌ها را در اختیار نهادها قرار دهد.

با توجه به موارد بیان شده، ضرورت ارائه خدمات غیرحضوری با رشد روزافزون خدمات برخط و افزایش تقاضای مردم برای آن، به ویژه در شرایطی مانند بحران بیماری کرونا، موضوعی بدیهی است که همه سازمان‌ها و نهادهای ارائه‌دهنده خدمات را به سمت بهره‌گیری از آن سوق داده است و مورد تاکید نهادهای بالادستی کشور شامل قانون‌گذاران و سیاست‌گذاران است. احراز هویت، پیش‌نیاز ارائه هرگونه خدمات غیرحضوری توسط پلیس است و لازم است افراد قبل از دریافت خدمات (به ویژه خدمات مهم و حساس)، احراز هویت شوند. بنابراین، همه سرویس‌های ارائه شده به مردم، قبل از دریافت توسط افراد، با فرآخوی سرویس احراز هویت، فرد گیرنده خدمات را شناسایی می‌کنند و این کار باید به صورت غیرحضوری و از راه دور باشد. بنابراین، خلاصه ضرورت‌های انجام طرح احراز هویت غیرحضوری در پلیس عبارتند از:

- تقاضای روزافزون ارائه خدمات الکترونیکی و غیرحضوری از سمت مردم و لزوم بهبود تجربه مشتری^۱ به دلیل سادگی و سرعت کار
 - تاکید نهادهای قانون‌گذار بر ارائه خدمات الکترونیکی و غیرحضوری به مردم به ویژه با تشدید موضوع در شرایط بحران کرونا
 - نیاز به کاهش مراجعات حضوری افراد (از نظر سلامتی، ترافیک، ...)
 - ضرورت افزایش امنیت و اشراف اطلاعاتی پلیس با تکمیل پایگاه‌های داده افراد به ویژه در تکمیل اطلاعات زیست‌سنگی، تهیه زیرساخت‌های استفاده از آنها، یکپارچه‌سازی و پیگیری (شفافیت)
 - لزوم کاهش خطاهای انسانی و سواستفاده افراد از اطلاعات و اسناد
- همان‌طور که بیان گردید، احراز هویت برای ارائه بسیاری از خدمات الکترونیکی انتظامی در کاربردهای مختلفی مانند حوزه گواهینامه، گذرنامه، وظیفه عمومی، کارت سوخت، اعلام سرقت، دریافت سابقه عدم سو پیشینه و ... ضروری است، به گونه‌ای که از تعداد زیاد خدمات الکترونیکی قبل ارائه در دفاتر ارائه خدمات الکترونیکی انتظامی، در حال حاضر تعداد بسیار محدودی از آن‌ها (بر اساس اطلاعات بیان شده) به صورت غیرحضوری و برخط ارائه می‌شود چراکه ارائه این خدمات به صورت غیرحضوری نیاز به احراز هویت افراد دارد. از این‌رو، با راهاندازی این سرویس، ارائه خدمات غیرحضوری در بسیاری از کاربردهای الکترونیکی انتظامی ممکن خواهد بود.

کاربران این سرویس دو گروه هستند: ارائه دهنده‌های خدمات الکترونیکی (پلیس) و گیرنده‌گان خدمات الکترونیکی (عامه مردم). در سمت ارائه دهنده، هر کدام از سرویس‌های الکترونیکی انتظامی در هر کدام از حوزه‌های فعلی (گذرنامه، گواهینامه، ...) می‌توانند با فراخوانی این سرویس و در صورت تایید هویت افراد توسط آن، به ارائه خدمت مرتبط به افراد بپردازن. از طرف دیگر، همه افراد متقاضی استفاده از خدمات غیرحضوری، می‌توانند با انجام دستورالعمل احراز هویت در بستر فراهم شده برای دریافت خدمات (مانند اپلیکشن موبایل یا نسخه تحت وب)، ابتدا کار احراز هویت خود را انجام داده و پس از آن، خدمت مربوطه را دریافت کنند.

۴- ۵ بیان مزايا و فواید عملیاتی احراز هویت غیرحضوری

اجرای طرح مزايا و دستاوردهای عملیاتی مختلفی را برای گروه‌های مختلف دخیل در آن به همراه خواهد داشت. در ادامه، اين مزايا به تفکیک دو گروه ارائه دهنده خدمات (پلیس) و دریافت خدمات (مردم) ذکر می‌شود.

- ### ۴- ۱ منافع برای ارائه دهنده خدمات (پلیس)
- کمک به تحقق پلیس هوشمند و ارائه خدمات غیرحضوری توسط نیروی انتظامی
 - همراستایی با اهداف بالادستی کشور و در راستای تحقق دولت الکترونیک
 - تصویرسازی نوآورانه و بهروز بودن مبتنی بر فناوری از پلیس
 - کمک به تحقق مسئولیت‌های پلیس در ارائه ساده و آسان خدمات با رعایت سلامتی شهروندان به ویژه در بحران‌هایی مانند شیوع کرونا

¹ User Experience

- کمک به اشراف اطلاعاتی پلیس و فراهم شدن امکان اعمال کنترل‌های امنیتی قوی‌تر و دقیق‌تر بر اساس روش‌های فناورانه مبتنی بر بیومتریک و در نتیجه ارتقاء امنیت
- کاهش مراجعات حضوری به دفاتر و کم کردن مشکلات ناشی از آن
- کاهش احتمال خطای انسانی و یا کم‌توجهی نیروهای انسانی
- فراهم کردن دسترسی ۷*۲۴ و حتی در روزهای تعطیل به خدمات
- فراهم کردن امکان نگاشت اطلاعات مختلف افراد به همدیگر در پایگاه داده‌های مختلف (کد ملی، اطلاعات گذرنامه، اطلاعات گواهینامه و ...)
- تسريع در طبیق‌پذیری: با تغییر مقررات، سیستم‌های کنترل دسترسی باید به طور متناوب تغییر کنند. فرایندهای احراز هویت در مواردی که نیاز به تغییر سریع دارد، می‌تواند به سادگی در سامانه به‌روز می‌شود و خیلی سریع با شرایط جدید سازگار شود.
- یکپارچه‌سازی: eKYC در بیشتر موارد، با استفاده از API‌ها، قابلیت احراز هویت را به آسانی به سایر سامانه‌ها اضافه می‌کند. همچنین، داده‌های مشتری، اسناد و اطلاعات به طور ایمن در سوابق الکترونیکی او ذخیره می‌شوند و در صورت لزوم در سایر سامانه‌ها قابل استفاده هستند.
- پیگیری/گزارش: داده‌های دیجیتالی جمع‌آوری شده در فرایند احراز هویت قابل انتقال به سیستم‌های تحلیل، ممیزی، پیگیری و گزارش‌دهی هستند و فرصت‌هایی را برای بهینه‌سازی و تحلیل استراتژیک ایجاد می‌کنند.

۴-۵-۲ منافع برای دریافت کننده خدمات (مردم)

- افزایش سرعت دریافت خدمات
- کاهش مراجعات حضوری به دفاتر و دستیابی به مزایای ناشی از آن (ترافیک، زمان، سلامتی و ...)
- فراهم کردن دسترسی شبانه‌روزی و حتی در روزهای تعطیل به خدمات انتظامی
- امکان دریافت خدمات به صورت ساده و آسان
- صرفه جویی در زمان افراد با حذف مراجعه حضوری و منتظر ماندن در دفاتر
- صرفه جویی در هزینه با توجه به کاهش تردد
- کمک به سلامتی و جلوگیری از شیوع در مواردی مانند بحران کرونا
- ایجاد اختیارات سلف‌سرمایش و تسهیل فرایندها
- بهبود تجربه مشتری در دریافت خدمات

۴-۵-۳ تحلیل راهبردی طرح احراز هویت غیرحضوری

طرح احراز هویت غیرحضوری از نظر راهبردی دارای اهمیت زیادی است چرا که اجرای آن منجر به تحقق بسیاری از اهداف کلان سازمانی نیروی انتظامی و حتی ملی می‌گردد. این سرویس پیش نیاز ارائه خدمات مختلف غیرحضوری است و تجربه موفق آن در پلیس راهگشای فعالیت‌های مشابه در سطح ملی خواهد بود. برخی از راهبردهای ملی و سازمانی که با اجرای طرح eKYC محقق و یا تقویت می‌گردند، عبارتند از:

توسعه پلیس هوشمند و فناور

- فراهم کردن بستری برای توجه به سلامت شهروندان در شرایط ویروس کرونا
- ارتقاء منزلت شهروندی و رعایت حداکثری حقوق آنها
- ارتقاء امنیت ملی و افزایش اشرافیت اطلاعاتی پلیس
- توسعه خدمات الکترونیک انتظامی
- کاهش اصطکاک و روپرتویی ملموس پلیس با مردم
- ارتقاء اعتبار و جایگاه پلیس

۴-۵-۴ نگرانی‌های رایج در احراز هویت الکترونیکی

در حالی که احراز هویت الکترونیک (به هر شیوه‌ای که در سازمان و یا شرکت‌ها پیاده‌سازی شود) سبب ارتقای عملکرد سازمانی و کاهش هزینه‌ها و رضایت کاربران می‌شود و به طور قابل ملاحظه‌ای کارآمدی را افزایش می‌دهد، همچنین می‌تواند سازمان‌ها را در معرض خطرات سطح بالایی از تقلب، ورود غیر مجاز و عدم رعایت قوانین قرار دهد. با این حال، یک فرآیند داخلی جامع، همراه با فناوری پیچیده امنیتی که امروزه مطالعات زیادی بر روی آن‌ها انجام شده است، می‌تواند خطرات را به میزان قابل توجهی کاهش داده و قابلیت اجرای این فناوری را افزایش دهد.

به طور کلی، داده‌های دیجیتالی نسبت به فرم‌های مرسوم کاغذی از امنیت بیشتری برخوردارند، زیرا ردبایی آن‌ها آسان‌تر است، به ویژه اگر از طریق یک نرم‌افزار خاص دریافت شود. محصولات نرم‌افزاری یا پلتفرم‌های ارائه دهنده خدمات احراز هویت الکترونیکی نیز از فناوری رمزگذاری و رمزگشایی در کنار سایر زیرساخت‌های امنیتی استفاده می‌کنند تا جلوی ورود غیر مجاز افراد را بگیرند.

با وجود این، هنوز تعدادی از خطرات وجود دارد که سازمان‌ها هنگام اجرای احراز هویت الکترونیکی باید در نظر بگیرند:

خطر تقلب و قابلیت اطمینان: در حالی که استفاده از یک نرم افزار اختصاصی نرم‌افزار می‌تواند به حل هرگونه مشکل احراز هویت در فرآیند کمک کند، هنوز خطر جعل و تقلب کلاهبرداران برای سازمان‌ها وجود دارد زیرا فناوری می‌تواند به خطر بیفتد یا هک شود. یکی از راه‌های مقابله با این امر، اتخاذ برخی اقدامات اضافی ضدکلاهبرداری مانند دریافت تأیید هویت از طریق یک تماس تلفنی است.

خطر افشاء اطلاعات کاربران: علاوه بر قوانین ملی و بین‌المللی در حوزه احراز هویت، سازمان‌ها همچنین باید قوانین و مقررات ارائه اسناد، افشاگری‌ها و سایر اطلاعات را در مراحل مختلف فرایند رعایت کنند. در صورت عدم رعایت تعهدات سازمان‌ها، حریم خصوصی کاربران به خطر می‌افتد. اگر اطلاعات محروم‌انه در مورد کاربران به دلیل ضعف‌های امنیتی، مدیریتی و یا سوءاستفاده کلاهبرداران افشا شود، این درز اطلاعات ممکن است به خسارات مالی به کسب و کار، پیگرد قانونی یا حتی ورشکستگی منجر شود. حفاظت از اطلاعات محروم‌انه یک نیاز تجاری و در بسیاری از موارد نیز نیاز اخلاقی و قانونی است. این مشکلات تنها مختص احراز هویت الکترونیک نیست و بخشی از مشکلات گستره‌ای است که در سراسر جهان گریبانگیر سامانه‌های اطلاعاتی و رایانه‌ای هستند. هر ساله سازمان‌های بسیاری هدف جرائم مرتبط با امنیت اطلاعات، از حملات ویروسی گرفته تا کلاهبرداری‌های تجاری از قبیل سرقت اطلاعات حساس تجاری و اطلاعات محروم‌انه کارت‌های اعتباری، قرار می‌گیرند.

با توجه به رویکرد پلیس در سال‌های اخیر، مشاهده می‌شود که هوشمندسازی پلیس از راهبردهای اساسی و اولویت‌دار است که به صورت جدی دنبال می‌شود. نیروی انتظامی یکی از دستگاه‌هایی است که علاوه بر برقراری نظم و امنیت از مرزها تا درون شهرها به مردم خدمات رسانی می‌کنند. در مجموعه عظیم نیروی انتظامی با وجود مأموریت‌های بسیار، می‌طلبد که در ارائه خدمت به مردم تحول ایجاد و خدمت‌رسانی را با سرعت و دقت بیشتر و بهره‌برداری از فناوری‌های روز ارائه کند. یکی از اقدامات نیروی انتظامی برای کاهش حضور مردم در مراکز پلیس راهاندازی اپلیکیشن «پلیس من» است.

رمز موافقیت پلیس در جوامع کنونی اشراف اطلاعاتی، هوشمندسازی تجهیزات، تربیت نیروهای متعدد و متخصص بوده و نیروی انتظامی نیز برای کارآمدی بیشتر، در پی رسیدن به هوشمندسازی است. از مزایای اصلی هوشمندسازی پلیس رویکرد جامعه‌محوری و کاهش برخورد فیزیکی با مردم است. ایجاد سامانه‌های هوشمند و استفاده از تجهیزات نوین با بهره‌مندی از هوش مصنوعی در حوزه‌های مختلف از جمله گذرنامه، گواهینامه، خدمت سربازی و خدمات الکترونیکی و دوربین‌های کنترل ترافیک و تجمعی بانک‌های اطلاعاتی و آگاهی منجر به تحولاتی ژرف و بنیادی در دنیای دیجیتال و خدمات الکترونیکی پلیس به منظور تسريع و تسهیل در ارائه خدمت به مردم خواهد بود.

در همین جهت با بررسی خدمات پلیس و سامانه‌های مختلفی که این نهاد برای خدمت‌رسانی به مردم ارائه کرده است، مشاهده می‌کنیم که بسیاری از خدمات پلیس تنها به صورت حضوری و با مراجعه‌ی افراد به دفاتر امکان‌پذیر خواهد بود. احرار هویت الکترونیکی را می‌توان به عنوان یک راهکار کارآمد جهت رفع شکاف بین وضعیت موجود و مطلوب دانست.

محدودیت بودجه و تجهیزات برای پاسخگویی حضوری به تعداد زیادی از مراجعه‌کنندگان، محدودیت نیروهای پلیس و تمرکز آن‌ها بر روی موضوعات مهم، ارائه خدمات به نیروهای انسانی خود پلیس و کنترل آن‌ها در رسته‌های مختلف این نهاد، محدودیت تعداد دفاتر خدماتی در شهرستان‌ها و عدم دسترسی روستاها به این دفاتر، ازدحام مراجعه به دفاتر در ایام خاصی از سال، پیدایش نیازهای جدید و بیشترین سازگاری برای رفع این نیازها از نظر اقتصادی و تجهیزاتی، سبب شده است تا نیاز به احرار هویت الکترونیکی بسیار مورد توجه قرار بگیرد.

همچنین احرار هویت الکترونیک در خدمات نیروی انتظامی موجب می‌شود پلیس با توجه به استخراج اطلاعات لازم، بتواند در اجرای مأموریت‌هاییش آگاهانه عمل کند و از طرفی، چون پلیس با سازمان‌های مختلف و مردم در ارتباط است، باید نسبت به مأموریت‌ها و خدماتش و هر آنچه پیرامون خود می‌افتد، اشراف همه جانبه داشته باشد.

فصل ۵ راهکار پیشنهادی پروژه

احراز هویت غیرحضوری (eKYC) روندی است که در آن برای شناسایی یا تأیید هویت مشتریان، از یک سیستم پرس و جوی شناسایی دیجیتالی (و معمولاً ملی) استفاده می‌کنند و در برخی موارد، اطلاعات اصلی آن‌ها را بازیابی می‌کنند. سیستم‌های eKYC می‌توانند با کاهش یا از بین بردن رویه‌های کاغذی و ثبت سوابق، روند کار را بهبود ببخشند و این باعث کاهش هزینه و زمان صرف شده در تأیید می‌شود و ارائه خدمات به کاربران کم درآمد را سودآورتر می‌کند. اگر این روش مبتنی بر زیست‌سنجد باشد، eKYC ممکن است به ضبط سنجه‌های اطلاعاتی مشتری نیاز داشته باشد که ممکن است شامل اثر انگشت، اسکن چهره، صدا و عنبیه باشد که از میان آنها چهره به دلیل دسترسی آسان و کم هزینه مردم به اخذ تصویر چهره و همچنین دقت مطلوب آن، در احراز هویت غیرحضوری کاربرد بیشتری بافته است.

فرایند زیست‌سنجد و تشخیص چهره در این روند باید از به روزترین تکنولوژی‌های هوشمند بهره‌مند باشد. در این فناوری باید راههای جلوگیری از ورود و یا دسترسی افراد غیر مجاز و تقلب نیز در نظر گرفته شده باشد. یکی از مهم‌ترین ویژگی‌های سیستم‌های احراز هویت افراد از راه دور با استفاده از فناوری چهره، تشخیص زنده بودن فرد است که اصالت و غیرتقلی بودن تصویر ارسالی را بررسی می‌کند.

مفهوم احراز هویت غیرحضوری می‌تواند گستره‌دار نظر گرفته شود و شامل رویکردهای مختلف و استفاده از زیست‌سنجدی‌های مختلف یا روش‌های متنوع تشخیص زنده بودن باشد با توجه به قلمرو این طرح، فقط شامل احراز هویت با چهره و تشخیص زنده بودن از روی ویدئو است. به منظور قرار گرفتن این دو سرویس در اختیار پلیس و به کارگیری آن در سایر خدمات پلیس، بعد از برگزاری جلسات مشورتی با ذینفعان و صاحبانظران در ناجا، دو راه حل ارائه می‌شود که در ادامه این دو

راه حل و مزایا و معایب آن بیان خواهد شد و انتظار می‌رود یکی از آن‌ها برای شروع بهره‌برداری از سرویس‌های تطبیق چهره و تشخیص زنده بودن انتخاب و بستر عملیاتی کردن آن فراهم شود.

۱-۵ راه حل شماره یک: استفاده از رابطه‌های برنامه نویسی و API و SDK

در این روش، سرویس‌های پایه سامانه‌های احراز هویت شامل تطبیق چهره و تشخیص زنده بودن به صورت API و SDK در اختیار سایر سامانه‌ها قرار می‌گیرد. در حال حاضر، سامانه‌های ناجا برای احراز هویت از سامانه SSO استفاده می‌نمایند که در اپلیکیشن پلیس من کاربران را با روش‌های معمول مبتنی بر رمز عبور احراز هویت می‌کند. نحوه اتصال این سامانه به سایر سامانه‌ها به صورت شکل ۱-۵ است. به این ترتیب، هر سامانه برای احراز هویت کاربران، نام کاربری و رمز عبور کاربر را دریافت و برای سامانه SSO ارسال کرده و این سامانه معتبر بودن آن را مشخص می‌نماید.



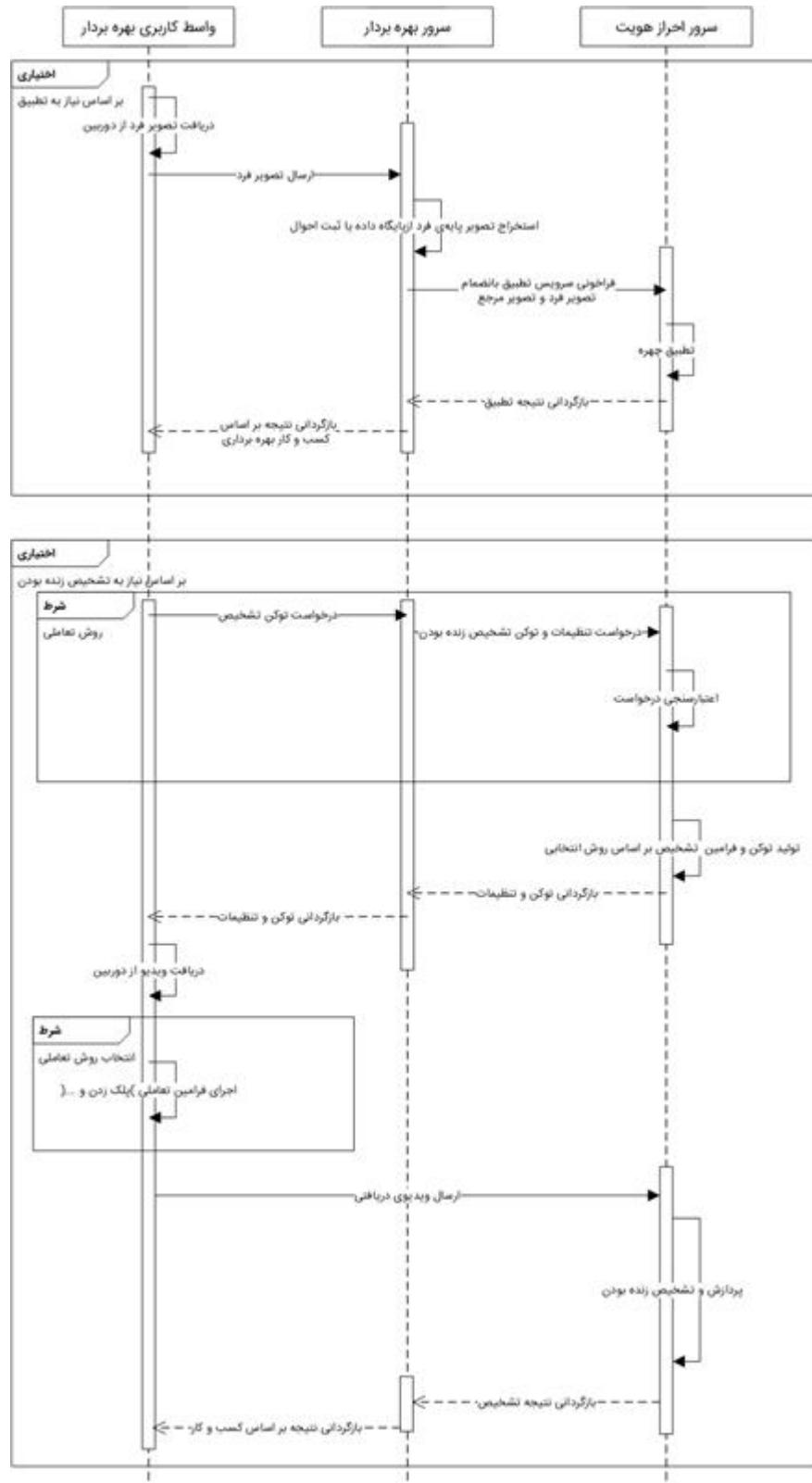
شکل ۱-۵ زیرساخت احراز هویت رایج مبتنی بر SSO ناجا و ارتباط با سامانه‌ها و سازمان‌های آن

با توجه به مدل موجود، یکی از راه حل‌ها، قرار دادن احراز هویت با تطبیق چهره و تشخیص زنده بودن داخل خود SSO می‌باشد. به این ترتیب برای احراز هویت کاربران توسط سامانه SSO از ترکیبی از سرویس‌های احراز هویت مبتنی زیست‌سنگی نیز استفاده شده و در حالت پیشرفته‌تری برای ثبت‌نام اولیه کاربران سامانه‌ها نیز می‌توان از این سرویس‌ها استفاده کرد. اتصال این سرویس‌ها به سامانه SSO همانند شکل ۲-۵ است.



شکل ۲-۵ اتصال سرویس‌های تطبیق چهره و تشخیص زنده بودن به SSO ناجا

در شکل زیر نمودار ترتیب (sequence diagram) برای راه حل پیشنهادی اول آورده شده است که در آن ارتباط بین سرور احراز هویت، سرور بearer بردار و برنامه کاربری سمت کاربر (وایط کاربری bearer بردار) آورده شده است. توجه شود که در این راهکار، به ازای هر سرویس دیگری نیز نیاز به طی چنین مراحلی هست و لازم است یکپارچه‌سازی سرویس‌های احراز هویت و bearer بردار به صورت مشابهی صورت بگیرد.



شكل ۳-۵ نمودار ترتیب راه حل پیشنهادی مبتنی بر SDK و API

۵-۲ راه حل شماره دو: استفاده از درگاه احراز هویت

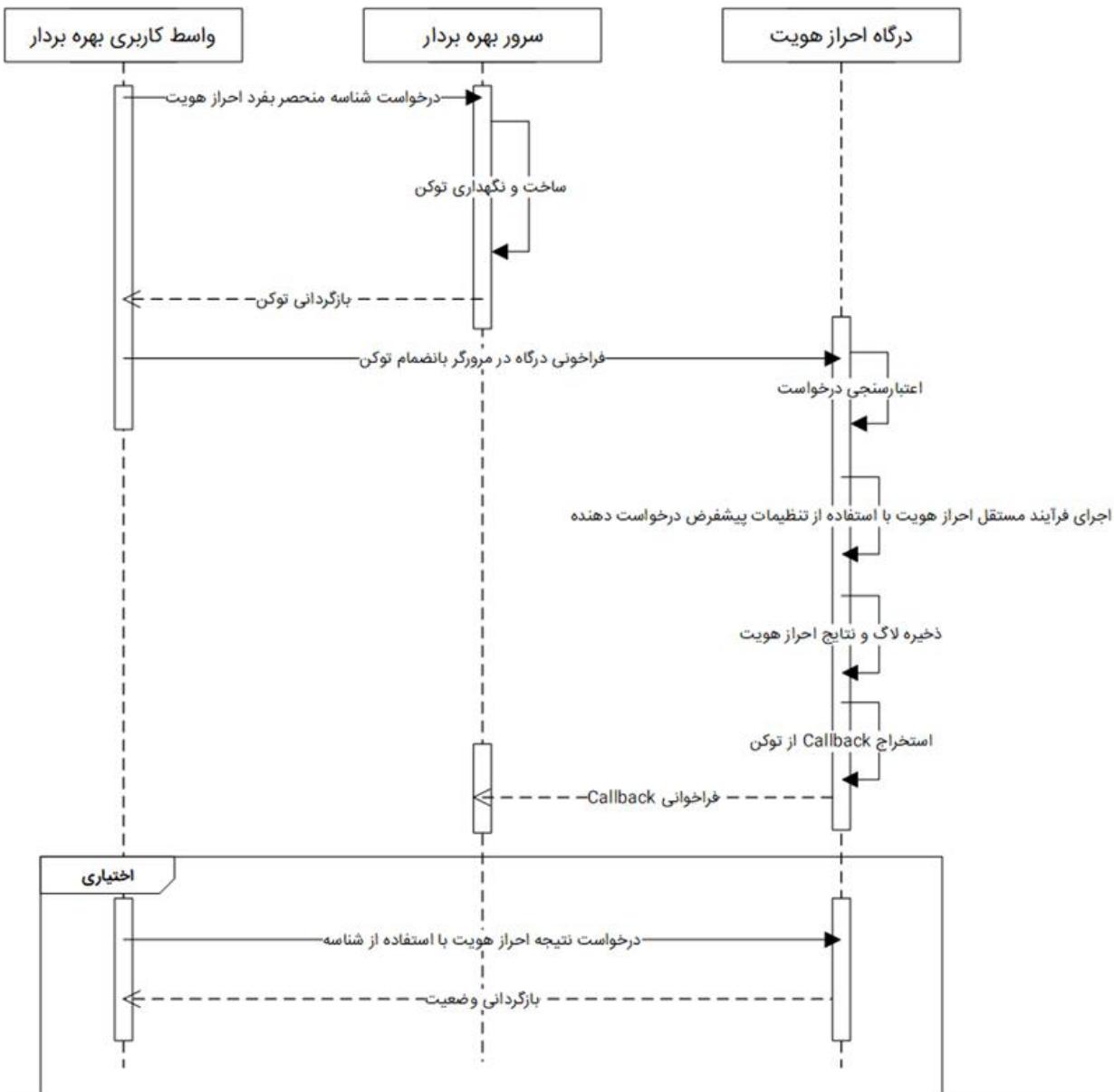
راه حل دوم، رابط احراز هویت دیگری به نام درگاه^۱ دارد که در آن مانند درگاه‌های بانکی، کاربران به سمت یک صفحه ثالث هدایت می‌شوند و روند احراز هویت در آن صفحه (سایت) صورت می‌گیرد و نتیجه موفق و یا ناموفق بودن آن برای سرویس گیرنده بازگردانده می‌شود. این راهکار مستقل از سامانه SSO بوده و تمام پیاده‌سازی‌های آن سمت تیم مجری می‌باشد. برای استفاده از این روش، کاربران در هنگام مراجعه به نرم‌افزاری مانند پلیس من و یا سامانه‌های داخلی، در زمان احراز هویت وارد وبسایت درگاه شده و اطلاعات هویتی نظری تصویر چهره و سایر اطلاعات مورد نیاز را وارد کرده و سامانه بعد از اعتبارسنجی داده‌های ورودی، تایید و یا عدم تایید را به سرویس گیرنده (مثلاً اپلیکیشن پلیس من) اطلاع می‌دهد (شکل ۳-۵).



شکل ۴-۵ ارائه خدمات احراز هویت از راه دور توسط درگاه احراز هویت

در شکل زیر نمودار ترتیب مربوط استفاده از سرویس احراز هویت مبتنی درگاه در سرویس‌ها و برنامه‌های کاربردی ناجا ارائه شده است. همان‌طور که مشخص است، در این حالت عمدۀ مسائل توسط خود درگاه مدیریت می‌شود و کارهای یکپارچه‌سازی و پیاده‌سازی خاص منظوره نیاز نخواهد بود.

¹ Gateway



شکل ۵-۵ نمودار ترتیب راه حل پیشنهادی مبتنی بر درگاه

۳-۵ جمع‌بندی

با توجه به اطلاعات دریافتی از کارشناسان ناجا، دو راهکار ۱- مبتنی بر یکپارچه‌سازی با API و ۲- استفاده از درگاه مستقل پیشنهاد شده است که بر اساس اولویت‌بندی مورد نظر می‌توان هر کدام از آنها را ارائه کرد. برای کمک بیشتر به تصمیم‌گیری، در جدول ۱-۵ مزايا و معایب هر کدام از دو روش پیشنهادی آورده شده است.

جدول ۱-۵ مزایا و معایب راهحلهای ارائه شده مبتنی بر API و مبتنی بر درگاه

راه حل	مزایا	معایب
اتصال از طریق API	یکپارچگی بیشتر در نرم افزارهای ناجا ایجاد کرده و باعث ایجاد انعطاف‌پذیری بیشتر در این نرم افزارها شده و می‌توان با توجه به نیاز، تغییرات خاصی را اعمال کرد	نیاز به پیاده‌سازی بیشتر در نرم افزارها و سرویس‌های دریافت کننده خدمات برای مدیریت روال‌های احراز هویت (مانند نگهداری اطلاعات وارد شده توسط کاربران، وضعیت احراز هویت آن‌ها، مراحل مورد نیاز برای احراز هویت و ...)
اتصال از طریق درگاه	نرم افزارهای استفاده کننده از سرویس احراز هویت غیرحضوری نسبت به روال‌های احراز هویت مستقل و بی‌خبر بوده و استفاده از آن مستقل از پیجیدگی‌های روال احراز هویت است (پیاده‌سازی آسان).	یکپارچگی کمتر با سامانه‌های استفاده کننده از سرویس‌های احراز هویت و انعطاف پایین‌تر

مراجع

- [1] Maltoni, D., Maio, D., Jain, A. K., and Prabhakar, S., *Handbook of Fingerprint Recognition*. Springer London, 2009
- [2] tutorialspoint website, “biometric_modality_selection.”
https://www.tutorialspoint.com/biometrics/biometric_modality_selection.htm
- [3] Yingzi, D. E., *Biometrics: From Fiction to Practice*. Pan Stanford Publishing, 2013.
- [4] M. S. Nixon, *Handbook of Biometric Anti-Spoofing*, Verlag London: Springer, 2014.
- [5] ISO/IEC JTC1 SC37 Biometrics: ISO/IEC 30107-2. Information technology - biometric presentation attack detection – Part 2: data formats. International organization for standardization. 2017
- [6] Newton, E., “Overview of the ISO / IEC 30107 Project Authentication Use Case Comparison,” pp. 1–13.
- [7] Givens, G. H., Beveridge, J. R., Phillips, P. J., Draper, B., Lui, Y. M., and Bolme, D., “Introduction to face recognition and evaluation of algorithm performance,” *Comput. Stat. Data Anal.*, vol. 67, pp. 236–247, 2013.
- [8] Galton, F., “Personal Identification and Description 2,” *Nature*, vol. 38, no. 973, pp. 173–177, 1888.
- [9] Tolba, A. S., El-Baz, A. H., and El-Harby, A. A., “Facial Recognition: A literature Review,” *Int. J. Signal Process.*, vol. 2, no. 2, pp. 88–103, 2006.

- [10] Jafri, R. and Arabnia, H. R., “A Survey of Face Recognition Techniques,” *J. Inf. Process. Syst.*, vol. 5, no. 2, pp. 41–68, 2009.
- [11] Nastar, C. and Mitschke, M., “Real-time face recognition using feature combination,” in *Proceedings - 3rd IEEE International Conference on Automatic Face and Gesture Recognition, FG 1998*, 1998, pp. 312–317.
- [12] Gong, S., McKenna, S. J., and Psarrou, A., “From Images to Face Recognition,” *Image Process. Imp. Coll. Press*, 1999.
- [13] Ding, X. and Fang, C., “Discussions on Some Problems in Face Recognition,” in *Lncs*, vol. 3338, Springer, 2004, pp. 47–56.
- [14] Chihoui, M.; Elkefi, A.; Bellil,W.; Ben Amar, C. A Survey of 2D Face Recognition Techniques. *Computers* 2016, 5, 21.
- [15] Benzaoui, A.; Bourouba, H.; Boukrouche, A. System for automatic faces detection. In *Proceedings of the 2012 3rd International Conference on Image Processing, Theory, Tools and Applications (IPTA)*, Istanbul, Turkey, 15–18 October 2012; pp. 354–358.
- [16] Petrovska-Delacrétaz, D., Chollet, G., and Dorizzi, B., *Guide to biometric reference systems and performance evaluation*. Springer, 2009.
- [17] Kong, S. G., Heo, J., Abidi, B. R., Paik, J., and Abidi, M. A., “Recent advances in visual and infrared face recognition - A review,” *Comput. Vis. Image Underst.*, vol. 97, no. 1, pp. 103–135, 2005.
- [18] Wang, P. and Bai, X., “Regional parallel structure based CNN for thermal infrared face identification,” *Integr. Comput. Aided. Eng.*, no. Preprint, pp. 1–14, 2018.
- [19] Wang, Z., Miao, Z., Jonathan Wu, Q. M., Wan, Y., and Tang, Z., “Low-resolution face recognition: A review,” *Vis. Comput.*, vol. 30, no. 4, pp. 359–386, Apr. 2014.
- [20] Turk, M. A. and Pentland, A. P., “Face recognition using eigenfaces,” in *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586–591.
- [21] Chen, S., Mau, S., Harandi, M. T., Sanderson, C., Bigdeli, A., and Lovell, B. C., “Face recognition from still images to video sequences: A local-feature-based framework,” *Eurasip J. Image Video Process.*, vol. 2011, p. 11, 2011.
- [22] Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Taleb-Ahmed, A. Past, Present, and Future of Face Recognition: A Review. *Electronics* 2020, 9, 1188.
- [23] Samaria, F.S.; Harter, A.C. Parameterization of a Stochastic Model for Human Face Identification. In *Proceedings of the 1994 IEEE Workshop on Applications of Computer Vision*, Sarasota, FL, USA, 5–7 December 1994; pp. 138–142.
- [24] Phillips, P.J.; Wechsler, H.; Huang, J.; Rauss, P. The FERET database and evaluation procedure for face recognition algorithms. *Image Vis. Comput.* 1998, 16, 295–306.
- [25] Martinez, A.M.; Benavente, R. The AR face database. *CVC Tech. Rep.* 1998, 24, 1–10.

- [26] Messer, K.; Matas, J.; Kittler, J.; Jonsson, K. Xm2vt sdb: The extended m2vts database. In Proceedings of the 1999 2nd International Conference on Audio and Video-based Biometric Person Authentication (AVBPA), Washington, DC, USA, 22–24 March 1999; pp. 72–77.
- [27] Bailliére, E.A.; Bengio, S.; Bimbot, F.; Hamouz, M.; Kittler, J.; Mariéthoz, J.; Matas, J.; Messer, K.; Popovici, V.; Porée, F.; et al. The BANCA Database and Evaluation Protocol. In Proceedings of the 2003 International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), Guildford, UK, 9–11 June 2003; pp. 625–638.
- [28] Phillips, P.J.; Flynn, P.J.; Scruggs, T.; Bowyer, K.W.; Chang, J.; Ho
man, K.; Marques, J.; Min, J.; Worek, W. Overview of the face recognition grand challenge. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), San Diego, CA, USA, 20–26 June 2005; pp. 947–954.
- [29] Huang, G.B.; Mattar, M.; Berg, T.; Learned-Miller, E. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments; Technical Report; University of Massachusetts: Amherst, MA, USA, 2007; pp. 7–49.
- [30] Gross, R.; Matthews, L.; Cohn, J.; Kanade, T.; Baker, S. Multi-PIE. *Image Vis. Comput.* **2010**, *28*, 807–813.
- [31] CASIA Web Face. Available online: <http://www.cbsr.ia.ac.cn/english/CASIA-WebFace-Database.html> (accessed on 21 July 2019).
- [32] Klare, B.F.; Klein, B.; Taborsky, E.; Blanton, A.; Cheney, J.; Allen, K.; Grother, P.; Mah, A.; Burge, M.; Jain, A.K. Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1931–1939.
- [33] Shlizerman, I.K.; Seitz, S.M.; Miller, D.; Brossard, E. The MegaFace benchmark: 1 million faces for recognition at scale. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 4873–4882
- [34] Sengupta, S.; Cheng, J.; Castillo, C.; Patel, V.M.; Chellappa, R.; Jacobs, D.W. Frontal to Profile Face Verification in theWild. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 7–10 March 2016; pp. 1–9.
- [35] Guo, Y.; Zhang, L.; Hu, Y.; He, X.; Gao, J. Ms-Celeb-1m: A dataset and benchmark for large-scale face recognition. In Proceedings of the 14th European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 8–16 October 2016.
- [36] Wang, T.Y.; Kumar, A. Recognizing Human Faces under Disguise and Makeup. In Proceedings of the 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), Sendai, Japan, 29 February–2 March 2016; pp. 1–7.
- [37] Parkhi, O.M.; Vedaldi, A.; Zisserman, A. Deep Face Recognition. In Proceedings of the 2015 British Machine Vision Conference, Swansea, UK, 7–10 September 2015; pp. 41.1–41.12.
- [38] Cao, Q.; Shen, L.; Xie, W.; Parkhi, O.M.; Zisserman, A. VGGFace2: A dataset for recognizing faces across pose and age. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG), Xi'an, China, 15–19 May 2018; pp. 67–74.

- [39] Whitelam, C.; Taborsky, E.; Blanton, A.; Maze, B.; Adams, J.; Miller, T.; Kalka, N.; Jain, A.K.; Duncan, J.A.; Allen, K. IARPA Janus Benchmark-B face dataset. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 592–600.
- [40] Nech, A.; Shlizerman, I.K. Level playing field for million scale face recognition. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 3406–3415.
- [41] Kushwaha, V.; Singh, M.; Singh, R.; Vatsa, M. Disguised Faces in the Wild. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018; pp. 1–18.
- [42] Maze, B.; Adams, J.; Duncan, J.A.; Kalka, N.; Miller, T.; Otto, C.; Jain, A.K.; Niggel, W.T.; Anderson, J.; Cheney, J.; et al. IARPA Janus benchmark-C: Face dataset and protocol. In Proceedings of the 2018 International Conference on Biometrics (ICB), Gold Coast, QLD, Australia, 20–23 February 2018; pp. 158–165.
- [43] Elharrouss, O.; Almaadeed, N.; Al-Maadeed, S. LFR face dataset: Left-Front-Right dataset for pose-invariant face recognition in the wild. In Proceedings of the 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, 2–5 February 2020; pp. 124–130.
- [44] Wang, Z.; Wang, G.; Huang, B.; Xiong, Z.; Hong, Q.; Wu, H.; Yi, P.; Jiang, K.; Wang, N.; Pei, Y.; et al. Masked Face Recognition Dataset and Application. arXiv 2020, arXiv:2003.09093v2.
- [45] Abate, A. F., Nappi, M., Riccio, D., and Sabatino, G., “2D and 3D face recognition: A survey,” *Pattern Recognit. Lett.*, vol. 28, no. 14, pp. 1885–1906, 2007.
- [46] Blanz, V. and Vetter, T., “Face Recognition Based on Fitting a 3D Morphable Model,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1063–1074, 2003.
- [47] Viola, P. and Jones, M., “Rapid object detection using a boosted cascade of simple features,” in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001, vol. 1, pp. I-511-I-518.
- [48] Rowley, H. A., Baluja, S., and Kanade, T., “Neural Network-Based Face Detection,” *{IEEE} Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 1, pp. 23–38, 1998.
- [49] Mathias, M., Benenson, R., Pedersoli, M., and Van Gool, L., “Face detection without bells and whistles with supplementary material,” in Eccv, 2014, pp. 720–735.
- [50] Waring, C. A. and Liu, X., “Face detection using spectral histograms and SVMs,” *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 35, no. 3, pp. 467–476, 2005.
- [51] Farfade, S. S., Saberian, M., and Li, L.-J., “Multi-view Face Detection Using Deep Convolutional Neural Networks,” in Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, 2015, pp. 643–650.
- [52] Krizhevsky, A. and Hinton, G. E., “ImageNet Classification with Deep Convolutional Neural Networks,” in Neural Information Processing Systems, 2012, pp. 1–9.

- [53] Hjelmås, E. and Low, B. K., "Face detection: A survey," *Comput. Vis. Image Underst.*, vol. 83, no. 3, pp. 236–274, 2001.
- [54] Sakai, T., Nagao, M., and Kanade, T., Computer analysis and classification of photographs of human faces. Kyoto University, 1972.
- [55] Govindaraju, V., "Locating human faces in photographs," *Int. J. Comput. Vis.*, vol. 19, no. 2, pp. 129–146, 1996.
- [56] Lam, K. M. and Yan, H., "Facial feature location and extraction for computerized human face recognition," in National Conference Publication - Institution of Engineers, Australia, 1994, vol. 1, pp. 167–171.
- [57] Jie Yang and Waibel, A., "A real-time face tracker," in Proceedings Third IEEE Workshop on Applications of Computer Vision. WACV'96, 1996, pp. 142–147.
- [58] Yang, L. L. and Robertson, M. a., "Multiple-face tracking system for general region-of-interest video\ncoding," in Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101), 2000, vol. 1, pp. 347–350.
- [59] Yang, G. and Huang, T. S., "Pattern Recognition Human face detection in a complex background," *Pattern Recognit.*, vol. 27, no. 1, pp. 10–11, 2015.
- [60] Sumi, Y. and Ta, Y. O., "Detection of face orientation and facial components using distributed appreance modeling," in In Proceeding of 1st International Workshop on A utomatic Face and Gesture Recognition, 1995, pp. 254–259.
- [61] Huang, C., Chen, C., and Chu, H., "Human Facial Feature Extraction for Face," *Pattern Recognit.*, vol. 25, no. 12, pp. 1435–1444, 1992.
- [62] Zhang, C. and Zhang, Z., "A Survey of Recent Advances in Face Detection," Tech. Report, Microsoft Res., no. June, p. 17, 2010.
- [63] Heisele, B., Ho, P., and Poggio, T., "Face recognition with support vector machines: Global versus component-based approach," in Proceedings of the IEEE International Conference on Computer Vision, 2001, vol. 2, pp. 688–694.
- [64] Sharif, M., Naz, F., Yasmin, M., Shahid, M. A., and Rehman, A., "Face recognition: A survey," *J. Eng. Sci. Technol. Rev.*, vol. 10, no. 2, pp. 166–177, 2017.
- [65] Liu, C. and Wechsler, H., "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, 2002.
- [66] Cao, Z., Yin, Q., Tang, X., and Sun, J., "Face recognition with learning-based descriptor," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, vol. 91, no. 6, pp. 2707–2714.
- [67] Yaniv Taigman Ming Yang Marc'Aurelio Ranzato Lior Wolf and Tel, "DeepFace - Closing the Gap to Human-Level Performance in Face Verification," in A Multi-Center, Randomized, Controlled Evaluation of the Safety and Efficacy of LASIK With Cross-linking Performed With the KXL System and Photrex ZD[TM] (Riboflavin Ophthalmic Solution) Compared to LASIK Alone for Hyperopia and Hyperopic Astigmatism, 2014, pp. 1701–1708.

- [68] Sun, Y., Wang, X., and Tang, X., “Deep Learning Face Representation by Joint Identification-Verification,” in Advances in neural information processing systems, 2014, pp. 1988–1996.
- [69] Parkhi, O. M., Vedaldi, A., and Zisserman, A., “Deep Face Recognition,” in Proceedings of the British Machine Vision Conference 2015, 2015, vol. 1, no. 3, pp. 41.1-41.12.
- [70] Cao, Q., Shen, L., Xie, W., Parkhi, O. M., and Zisserman, A., “VGGFace2: A dataset for recognising faces across pose and age. (arXiv:1710.08092v2 [cs.CV] UPDATED),” in ArXiv e-prints, 2017, vol. 1710, pp. 67–74.
- [71] Bombardelli, F., “FaceNet: A Unified Embedding for Face Recognition and Clustering Felipe Bombardelli FaceNet: A Unified Embedding for Face Recognition and Clusteri Introduction Algorithm Results References,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 815–823.
- [72] Sirovich, L. and Kirby, M., “Low-dimensional procedure for the characterization of human faces,” J. Opt. Soc. Am. A, vol. 4, no. 3, p. 519, 1987.
- [73] Sirovich, L. and Kirby, M., “Application of the KL Procedure for the Characterization of Human Faces,” J. Opt. Soc. Am. A, vol. 4, no. 1, pp. 591–524, 1987.
- [74] M., T. and A., P., “Eigenfaces for recognition,” J. Cogn. Neurosci., vol. 3, no. 1, pp. 71–86, 1991.
- [75] Pentland, Moghaddam, and Starner, “View-based and modular eigenspaces for face recognition,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. CVPR-94, pp. 84–91, 1994.
- [76] Zeiler, M. D. and Fergus, R., “Visualizing and understanding convolutional networks - features learnt by convolutional networks,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2014, vol. 8689 LNCS, no. PART 1, pp. 818–833.
- [77] Szegedy, C. et al., “Inception-v1: Going deeper with convolutions,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, vol. 07-12-June, pp. 1–9.
- [78] Zhou, S. and Xiao, S., “3D face recognition: a survey,” Human-centric Comput. Inf. Sci., vol. 8, no. 1, p. 35, 2018.
- [79] Soltanpour, S.; Boufama, B.; Wu, Q.M.J. A survey of local feature methods for 3D face recognition. *Pattern Recognit.* 2017, 72, 391–406
- [80] Soltanpour, S., Boufama, B., and Wu, Q. M. J., “A survey of local feature methods for 3D face recognition,” *Pattern Recognit.*, vol. 72, pp. 391–406, 2017.
- [81] Russ, T., Boehnen, C., and Peters, T., “3D face recognition using 3D alignment for PCA,” in Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, 2006, vol. 2, pp. 1391–1398.
- [82] Lu, X. and Jain, A., “Deformation modeling for robust 3D face matching,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 8, pp. 1346–1357, 2008.
- [83] Wang, Y., Liu, J., and Tang, X., “Robust 3D face recognition by local shape difference boosting,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1858–1870, 2010.

- [84] Mohammadzade, H. and Hatzinakos, D., “Iterative closest normal point for 3D face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 2, pp. 381–397, 2013.
- [85] Gilani, S. Z., Mian, A., and Eastwood, P., “Deep, dense and accurate 3D face correspondence for generating population specific deformable models,” *Pattern Recognit.*, vol. 69, pp. 238–250, 2017.
- [86] Huang, Y., Wang, Y., and Tan, T., “Combining Statistics of Geometrical and Correlative Features for 3D Face Recognition,” in *BMVC*, 2006, pp. 879–888.
- [87] Samir, C., Srivastava, A., Daoudi, M., and Klassen, E., “An intrinsic framework for analysis of facial surfaces,” *Int. J. Comput. Vis.*, vol. 82, no. 1, pp. 80–95, 2009.
- [88] Drira, H., Amor, B. Ben, Srivastava, A., Daoudi, M., and Slama, R., “3D face recognition under expressions, occlusions, and pose variations,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2270–2283, 2013.
- [89] Bowyer, K. W., Chang, K., and Flynn, P., “A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition,” *Comput. Vis. Image Underst.*, vol. 101, no. 1, pp. 1–15, 2006.
- [90] Phillips, P. J. et al., “Overview of the face recognition grand challenge,” in *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on*, 2005, vol. 1, pp. 947–954.
- [91] Li, X. and Zhang, H., “Adapting geometric attributes for expression-invariant 3D face recognition,” in *Shape Modeling and Applications, 2007. SMI’07. IEEE International Conference on*, 2007, pp. 21–32.
- [92] Han, J., Kamber, M., and Pei, J., *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.
- [93] Inan, T. and Halici, U., “3-D face recognition with local shape descriptors,” *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 2, pp. 577–587, 2012.
- [94] Gupta, S., Markey, M. K., and Bovik, A. C., “Anthropometric 3D face recognition,” *Int. J. Comput. Vis.*, vol. 90, no. 3, pp. 331–349, 2010.
- [95] Berretti, S., Werghi, N., Del Bimbo, A., and Pala, P., “Selecting stable keypoints and local descriptors for person identification using 3D face scans,” *Vis. Comput.*, vol. 30, no. 11, pp. 1275–1292, 2014.
- [96] Koudelka, M. L., Koch, M. W., and Russ, T. D., “A prescreener for 3D face recognition using radial symmetry and the Hausdorff fraction,” in *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, 2005, p. 168.
- [97] Cover, T. and Hart, P., “Nearest neighbor pattern classification,” *IEEE Trans. Inf. theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [98] Tang, H., Yin, B., Sun, Y., and Hu, Y., “3D face recognition using local binary patterns,” *Signal Processing*, vol. 93, no. 8, pp. 2190–2198, 2013.
- [99] Boser, B. E., Guyon, I. M., and Vapnik, V. N., “A training algorithm for optimal margin classifiers,” in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144–152.

- [100] Lei, Y., Bennamoun, M., and El-Sallam, A. A., "An efficient 3D face recognition approach based on the fusion of novel local low-level features," *Pattern Recognit.*, vol. 46, no. 1, pp. 24–37, 2013.
- [101] Lei, Y., Bennamoun, M., Hayat, M., and Guo, Y., "An efficient 3D face recognition approach using local geometrical signatures," *Pattern Recognit.*, vol. 47, no. 2, pp. 509–524, 2014.
- [102] Berretti, S., Del Bimbo, A., and Pala, P., "3D face recognition using isogeodesic stripes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2162–2177, 2010.
- [103] Mian, A. S., Bennamoun, M., and Owens, R., "Keypoint detection and local feature matching for textured 3D face recognition," *Int. J. Comput. Vis.*, vol. 79, no. 1, pp. 1–12, 2008.
- [104] Chang, K. I., Bowyer, K. W., and Flynn, P. J., "Multiple nose region matching for 3D face recognition under varying facial expression," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1695–1700, 2006.
- [105] Faltemier, T. C., Bowyer, K. W., and Flynn, P. J., "A region ensemble for 3-D face recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 3, no. 1, pp. 62–73, 2008.
- [106] D. Kim, M. Hernandez, J. Choi, and G. Medioni. Deep 3D face identification. arXiv preprint arXiv:1703.10714, 2017 .
- [107] O. M. Parkhi, A. Vedaldi, A. Zisserman, et al. Deep face recognition. In BMVC, page 6, 2015 .
- [108] S. Zulqarnain Gilani and A. Mian, "Learning from Millions of 3D Scans for Large-Scale 3D Face Recognition," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 1896-1905, doi: 10.1109/CVPR.2018.00203.
- [109] O. M. Parkhi, A. Vedaldi, A. Zisserman, et al. Deep face recognition. In BMVC, page 6, 2015 .
- [110] Liu Y, Jourabloo A, Liu X, Learning deep models for face anti-spoofing: binary or auxiliary supervision. In: Proceeding of IEEE computer vision and pattern recognition, Salt Lake City, USA, 2018 .
- [111] Liveness Website, Liveness Detection, <https://www.liveness.com/>, 2020 .
- [112] C Schuckers, S. A. and Measures Stephanie C Schuckers, A.-S. A., "NOT FOR GENERAL DISTRIBUTION: TO BE PUBLISHED Spoofing and Anti-Spoofing Measures," *Inf. Secur. Tech. Rep.*, vol. 7, no. 4, pp. 56–62, 2002
- [113] Newton, E., "Overview of the ISO / IEC 30107 Project Authentication Use Case Comparison," pp. 1–13
- [114] Schuckers, Stephanie AC. "Spoofing and anti-spoofing measures." *Information Security Technical Report*, vol. 7, no. 4, pp. 56-62, 2002
- [115] Kollreider, Klaus, Hartwig Fronthaler, and Josef Bigun. "Evaluating liveness by face images and the structure tensor." In Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05), pp. 75-80, 2005
- [116] Galbally, Javier, Sébastien Marcel, and Julian Fierrez. "Biometric antispoofing methods: A survey in face recognition." *IEEE Access*, vol. 2, pp. 1530-1552, 2014
- [117] Bagga, Manpreet, and Baljit Singh. "Spoofing detection in face recognition: A review." In 3rd International Conference on Computing for Sustainable Global Development (INDIACOM), pp. 2037-2042, 2016.

- [118] Hoogsteden, C., and P. CROSS. ""Public access to GPS: government duty, economic rationality or international philanthropy?"" *CISM Journal*, vol. 46, no. 1, pp. 41-53, 1992 .
- [119] Chingovska I, AnjosA,Marcel S. Anti-spoofing in action: joint operation with a verification system. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 98–104, 2013.
- [120] de Freitas Pereira T, Anjos A, De Martino JM, Marcel S. Can face anti-spoofing countermeasures work in a real world scenario? In: International conference on biometrics (ICB), pp 1–8, 2013.
- [121] Fierrez J, Morales A, Vera-Rodriguez R, Camacho D. Multiple classifiers in biometrics. Part 1: Fundamentals and review. *Inf Fusion* 44:57–64, 2018.
- [122] Ross AA, Nandakumar K, Jain AK. *Handbook of multibiometrics*, Springer, 2006.
- [123] Galbally, J., S. Marcel, and J. Fierrez, Biometric Antispoofing Methods: A Survey in Face Recognition. Access, IEEE, 2: p. 1530-1552, 2014.
- [124] Socolinsky, D.A. and A. Selinger. A comparative analysis of face recognition performance with visible and thermal infrared imagery. in *Pattern Recognition, Proceedings. 16th International Conference on 2002*, 2002.
- [125] Pietikäinen, M. and A. Hadid, Texture features in facial image analysis, in *Advances in Biometric Person Authentication* , Springer. p. 1-8, 2005.
- [126] Active and passive liveness detection. <https://www.biometricupdate.com/202002/active-and-passive-liveness-detection-for-biometric-face-authentication-explored-in-id-rd-whitepaper>.
- [127] Active and passive liveness detection. <https://www.idrnd.ai/passive-vs-active-facial-liveness-detection-which-is-best/>
- [128] Hernandez-Ortega, J., Fierrez, J., Morales, A., and Galbally, J., "Introduction to face presentation attack detection," *Adv. Comput. Vis. Pattern Recognit.*, no. April, pp. 187–206, 2019.
- [129] TZ-CHIA TSENG1, TENG-FU SHIH, C.-S. F., "Anti-Spoofing of Live Face Authentication on Smartphone," 2020.
- [130] Jia, S., Guo, G., Xu, Z., and Wang, Q., "Face presentation attack detection in mobile scenarios: A comprehensive evaluation," *Image Vis. Comput.*, vol. 93, no. xxxx, 2020.
- [131] Xin, Y. et al., "A survey of liveness detection methods for face biometric systems," *Sens. Rev.*, vol. 37, no. 3, pp. 346–356, 2017.
- [132] Maatta, J., Hadid, A. and Pietikainen, M. , "Face spoofing detection from single images using micro-texture analysis", *IEEE international joint conference on Biometrics (IJCB)*, pp. 1-7, 2011.
- [133] A. Agarwal, R. Singh and M. Vatsa, Face Anti-spoofing using Haralick Features, In *Proceedings of IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2016.
- [134] Yang, J., Lei, Z., Liao, S., and Li, S. Z. , Face liveness detection with component dependent descriptor. *International Conference on Biometrics (ICB)*, pp. 1–6, IEEE. DOI: 10.1109/icb.2013.6612955. 1, 6, 2013.

- [135] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li. A face antispooing database with diverse attacks. In A. K. Jain, A. Ross, S. Prabhakar, and J. Kim, editors, ICB, pages 26–31. IEEE, 2012.
- [136] J. Maatta, A. Hadid, and M. Pietikainen. Face spoofing detection from single images using micro-texture analysis. In Biometrics (IJCB), 2011 International Joint Conference on, pages 1 –7, oct. 2011.
- [137] Face Liveness Detection Based on Texture and Frequency Analyses_2012.
- [138] H. S. Choi, R. C. Kang, K.T. Choi, A. T. B. Jin, and J.H. Kim. Fake-Fingerprint Detection using Multiple Static Features. Optical Engineering, 48(4), 2009.
- [139] T. Ojala, and M. Pietikainen. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7): 971-987, 2002.
- [140] He, J. and Luo, J., “Face Spoofing Detection Based on Combining Different Color Space Models,” 2019 IEEE 4th Int. Conf. Image, Vis. Comput. ICIVC 2019, pp. 523–528, 2019.
- [141] G. Pan, et al., “Eyeblink-based anti-spoofing in face recognition from a generic webcam,” 2007 IEEE 11th International Conference on Computer Vision, 2007.
- [142] Common Objects in Context, <http://cocodataset.org/#home>, 2019.
- [143] K. Kollreider, et al., “Evaluating liveness by face images and the structure tensor,” In IEEE Workshop on Automatic Identification Advanced Technologies, pp. 75-80, 2005.
- [144] A. Anjos, et al., “Face anti-spoofing: Visual approach” In Handbook of Biometric Anti-Spoofing, pp. 65-82, Springer, 2014.
- [145] W. Bao, et al., “A liveness detection method for face recognition based on optical flow field,” In Image Analysis and Signal Processing, 2009, IASP 2009 International Conference, pp. 233-236, IEEE, 2009.
- [146] Mahore, A. and Tripathi, M., “Detection of 3D Mask in 2D face recognition system using DWT and LBP,” 2018 IEEE 3rd Int. Conf. Commun. Inf. Syst. ICCIS 2018, pp. 18 22, 2019.
- [147] Uzun, E., Chung, S. P. H., Essa, I., and Lee, W., “rtCaptcha: A Real-Time CAPTCHA Based Liveness Detection System,” pp. 1–15, 2018.
- [148] Yang J, Lei Z, Li SZ, Learn convolutional neural network for face anti spoofing. CoRR. <http://arxiv.org/abs/1408.5601>, arXiv:1408.5601, 2014.
- [149] Lucena O, Junior A, Hugo GMV, Souza R, Valle E, De Alencar Lotufo R. Transfer learning using convolutional neural networks for face anti-spoofing. In: Karray F, Campilho A, Cheriet F (eds.) Proceedings of international conference on image analysis and recognition (ICCIAR), Springer International Publishing, Cham, pp 27–34, 2017.
- [150] Menotti D, Chiachia G, Pinto A, Schwartz WR, Pedrini H, Falco A, Rocha A., Deep representations for iris, face, and fingerprint spoofing detection. IEEE Trans Inf Forensics Secur 10(4):864–879. <https://doi.org/10.1109/TIFS.2015.2398817>, 2015.

- [151] "Nagpal C, Dubey SR. A performance evaluation of convolutional neural networks for face anti spoofing. CoRR. <https://arxiv.org/abs/1805.04176>, arXiv:1805.04176, 2018.
- [152] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, Rethinking the inception architecture for computer vision. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 2818–2826. <https://doi.org/10.1109/CVPR.2016.308>, 2016.
- [153] He K, Zhang X, Ren S, Sun J, Deep residual learning for image recognition. In: Proceedings of IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778. <https://doi.org/10.1109/CVPR.2016.90>, 2016.
- [154] Li L, Xia Z, Li L, Jiang X, Feng X, Roli F, Face anti-spoofing via hybrid convolutional neural network. In: Proceedings of international conference on the frontiers and advances in data science (FADS), pp 120–124. <https://doi.org/10.1109/FADS.2017.8253209>, 2017 .
- [155] Xu Z, Li S, Deng W, Learning temporal features using LSTM-CNN architecture for face anti-spoofing. In: Proceedings of 3rd IAPR asian conference on pattern recognition (ACPR), pp 141–145. <https://doi.org/10.1109/ACPR.2015.7486482>, 2015.
- [156] Peixoto B, Michelassi C, Rocha A, Face liveness detection under bad illumination conditions. In: International conference on image processing (ICIP), pp 3557–3560, 2011
- [157] Zhang, Y. et al., “CelebA-Spoof: Large-Scale Face Anti-Spoofing Dataset with Rich Annotations,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 12357 LNCS, pp. 70–85, 2020 .
- [158] Tan X, Li Y, Liu J, Jiang L, Face liveness detection from a single image with sparse low rank bilinear discriminative model. Comput Vis–ECCV 504–517, 2010 .
- [159] Anjos A, Marcel S, Counter-measures to photo attacks in face recognition: a public database and a baseline. In: International joint conference on biometrics (IJCB), pp 1–7, 2011 .
- [160] Chingovska I, Anjos A, Marcel S, On the effectiveness of local binary patterns in face anti-spoofing. In: IEEE BIOSIG, 2012 .
- [161] Zhang Z, Yan J, Liu S, Lei Z, Yi D, Li SZ, A face antispoofing database with diverse attacks. In: International conference on biometrics (ICB), pp 26–31, 2012.
- [162] Erdogmus N, Marcel S, Spoofing face recognition with 3D masks. IEEE Trans Inf Forensics Secur 9(7):1084–1097, 2014.