

احراز بی درنگ هویت متقاضیان خدمات الکترونیک با استفاده از روش مبتنی بر یادگیری عمیق سه بُعدی با دو مسیر

ویدا اسماعیلی^۱، دانشجوی دکتری، محمود محصل فقهی^{۲*}، دانشیار

^۱ دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران - v.esmaeili@tabrizu.ac.ir

^۲ دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران - mohasselfeghi@tabrizu.ac.ir

چکیده: احراز هویت در فضای مجازی، حفظ امنیت ملی را به دنبال دارد و اگر با دقت بسیار بالا انجام شود، می تواند به عنوان پدافند غیرعامل برای تداوم خدمت رسانی در شرایط مختلف مورد توجه قرار گیرد. پژوهش حاضر با هدف پیشنهاد یک روش کاملاً کاربردی برای احراز هویت متقاضیان خدمات الکترونیک بصورت بی درنگ ارائه شده است. برای جلوگیری از حقه های احتمالی کاربران، در روش پیشنهادی از شناسایی حرکات ماهیچه های صورت و سنجهی بیومتریکی عنبیه استفاده شده است. عنبیه قابلیت اطمینان بیشتری را ایجاد می کند و قابل سرقت و جعل نیست؛ زیرا باید بصورت زنده در اختیار باشد. برای احراز هویت و به طور همزمان، شناسایی حرکات ماهیچه های صورت و تمییز فرد زنده از تصویر، یک روش مبتنی بر یادگیری عمیق سه بُعدی با دو مسیر پیشنهاد شده است. با توجه به ارزیابی ها مشخص شد که روش پیشنهادی اطمینان قابل توجهی را برای استفاده عموم فراهم می آورد و قابل اجرا در شرایط واقعی و عملی می باشد. با استفاده از روش پیشنهادی، دقت ۹۹/۹۹ درصد و میانگین صحت احراز هویت و شناسایی افراد در هر دو پایگاه داده CASME و CASME2 بیش از ۹۹/۵۰ درصد است.

واژه های کلیدی: احراز هویت، سنجه بیومتریکی عنبیه، شناسایی حرکات ماهیچه های صورت، یادگیری عمیق، 3D-DenseNet در دو مسیر.

Real-time Authentication for Electronic Service Applicants using a Method Based on Two-Stream 3D Deep Learning

Vida Esmacili ¹, Ph.D. candidate, Mahmood Mohassel Feghhi ^{2*}, Associate Professor

¹ Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran, v.esmaeili@tabrizu.ac.ir

² Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran, mohasselfeghhi@tabrizu.ac.ir

Abstract: Authentication in cyberspace aims to maintain national security, and if it is done with high accuracy, it can be considered as a passive defense for the continuity of service delivery under different conditions. The present research aims to propose a fully applicable method for authentication of the applicants for electronic services in real-time. In order to prevent the possible tricks of the users, in the proposed method, identification of facial muscle movements and iris biometric measurement have been used. The iris creates more reliability and cannot be stolen or faked; because it must be available live. A method based on two-stream 3D deep learning is proposed for authentication and simultaneously identifying facial muscle movements and distinguishing a living person from an image. According to the evaluations, it was found that the proposed method provides significant assurance for public use and is applicable in real and practical conditions. Using the proposed method, the accuracy of the proposed method is 99.99%, and the average precision of authentication and identifying people in both CASME and CASME2 databases is more than 99.5%.

Keywords: Authentication, iris biometrics, detection of facial muscle movements, deep learning, Two-Stream 3D-DenseNet

* Mahmood Mohassel Feghhi, mohasselfeghhi@tabrizu.ac.ir

۱. مقدمه

کنند و یا شخصی بخواهد با تصویر یا فیلم متعلق به شخص دیگر وارد سامانه شود.

احراز هویت، به معنای شناسایی و تأیید فرد مورد تقاضا، می‌تواند به روش‌های مختلفی انجام گیرد [۴]؛ مانند احراز هویت با شماره ملی یا کارت‌های معتبری که مختص افراد می‌باشد. اما این روش‌ها قابل اطمینان نیستند؛ زیرا می‌توانند در اختیار دیگری قرار گیرند. بنابراین، استفاده از سنج‌های بیومتریکی و زیستی منحصر بفرد که قابل سرقت و در اختیار دیگران قرار دادن نیست، مفید خواهد بود و می‌تواند خطاهای احتمالی را به حداقل برساند.

عنبیه، اثر انگشت، صدا و ... بخشی از خصیصه‌های فیزیکی و بیومتریکی هر فرد است که افراد مختلف را از یکدیگر متمایز می‌سازد. مشخصه و ویژگی‌های مربوط به اثر انگشت و عنبیه‌ی چشم در تمام طول عمر و در طول زمان، نامتغیر و ثابت هستند. دو چشم یک فرد کاملاً دارای بافت‌های مستقل از هم در عنبیه است و بافت عنبیه چشم‌های دوقلوها نیز با هم متفاوت است. تشخیص هویت مبتنی بر تصاویر عنبیه شامل تجزیه و تحلیل ویژگی‌هایی است که در بافت رنگی چشم محصور بین مردمک و عنبیه قرار دارند. بافت پیچیده عنبیه می‌تواند شامل ویژگی‌های مشخص و فراوانی از قبیل: شیارها و برآمدگی‌ها، بافت‌های زیگزاگی، حلقه‌ها و لکه‌ها باشد. بنابراین، عنبیه از اثر انگشت منحصر بفردتر است. اسکن عنبیه با استفاده از دوربین معمولی و بدون تماس فیزیکی انجام می‌گیرد. همچنین، تصویر برداری از فردی که عینک به چشم دارد و یا از لنز استفاده می‌کند نیز به راحتی صورت می‌گیرد. بافت عنبیه برای دوباره‌سازی به صورت مجازی، بسیار سخت و پیچیده است و شخصی که بخواهد از هویت عنبیه شخص دیگر به‌طور غیرقانونی استفاده کند؛ چاره‌ای جز اینکه بافت عنبیه آن شخص را به‌طور زنده در اختیار داشته باشد، ندارد. بنابراین، این روش از قابلیت اطمینان بالاتری

امروزه گرایش مردم به دریافت خدمات بصورت شبانه‌روزی و حتی در روزهای تعطیل از راه دور و در مکان‌های مختلف با استفاده از سرویس‌های برخط بیش از پیش است. با توجه به اهمیت خدماتی که سازمان‌هایی مانند نیروی انتظامی، قضایی، ثبت احوال و ... به افراد مختلف ارائه می‌نمایند، احراز هویت خصوصاً در سرویس‌های برخط امری ضروری است. زیرا احراز هویت با دقت بالا علاوه بر تأمین امنیت ملی و محافظت از اطلاعات مهمی که فقط به شخص مورد نظر داده خواهد شد، موجب استمرار در رساندن خدمات ضروری، تداوم اداره و مدیریت مردم و نیز حفاظت سایبری از کشور و شبکه‌های اطلاع‌رسانی می‌شود که به‌عنوان پدافند غیرعامل در مقابل تهدیدات دشمن و بیگانه می‌تواند مورد توجه قرار گیرد.

با این حال، هنوز کشور جمهوری اسلامی ایران در رابطه با امنیت سایبری با چالش‌هایی مواجه است [۱، ۲]. بنابراین، هویت فرد متقاضی خدمات باید به‌درستی احراز شود تا سازمان مربوطه مطمئن شود که خدمات به هویت واقعی ارائه می‌شود و امکان سوء استفاده برای دیگران وجود ندارد. به همین منظور، احراز هویت متقاضیان خدمات غیرحضوری الکترونیک بر مبنای سنج‌های بیومتریکی که امکان جعل در آن وجود ندارد؛ می‌تواند مسیر را برای توسعه خدمات هوشمند هموارتر سازد.

امنیت داده و اطلاعاتی که در ارائه خدمات الکترونیک غیرحضوری بین سرویس‌دهنده و متقاضی متبادل می‌شوند و نیز اجتناب از غافل‌گیری و حقه‌های احتمالی بسیار حائز اهمیت است. به‌طوری‌که، در اغلب موارد به اشتراک گذاشتن اطلاعات در اختیار دیگران به‌طور ناخواسته که ممکن است به‌دلیل اختلالات سامانه‌ای بروز کند، جبران‌ناپذیر است و می‌تواند تهدید امنیت ملی را نیز به‌همراه داشته باشد [۱، ۳]. حتی ممکن است دو فرد دوقلو بخواهند به‌جای یکدیگر خدمات دریافت

نهایت، دو نتیجه حاصل می‌شود: (۱) فرد، چه کسی است و (۲) تصاویر دریافتی متعلق به فرد متقاضی خدمات است یا خیر؛ و آیا در صورت و چهره‌ی وی، حرکاتی هرچند بسیار ریز مشاهده می‌شود یا اینکه تصویر شخص دیگر در مقابل دوربین ثبت تصویر قرار داده شده است. بلوک دیاگرام روش پیشنهادی در شکل ۱ نشان داده شده است.

نوآوری‌های پژوهش حاضر به شرح ذیل می‌باشند:

(۱) ارائه یک روش جدید و کاربردی شامل فیلتر گابور و الگوی باینری محلی یکنواخت و 3D-DenseNet در دو مسیر پیشنهادی برای احراز هویت متقاضیان خدمات غیر حضوری بصورت بی‌درنگ و بدون تأخیر.

(۲) شناسایی عنبیه و حرکات ماهیچه‌های صورت به‌طور همزمان.

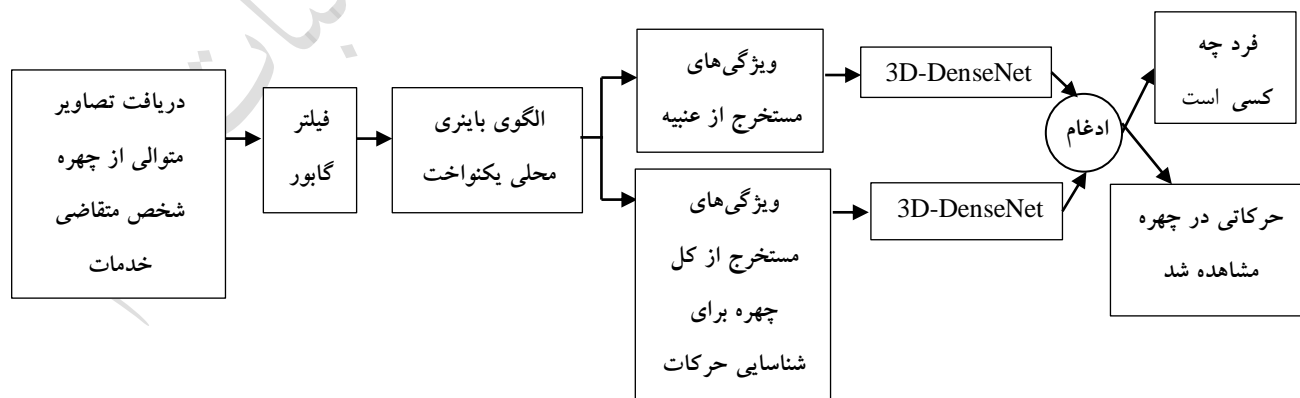
(۳) افزایش دقت و صحت احراز هویت متقاضیان دریافت خدمات الکترونیک.

در ادامه، پژوهش‌های پیشین در بخش ۲ مرور خواهند شد. در بخش ۳، شرح اجمالی روش پیشنهادی ارائه خواهد شد و نتایج به دست آمده و بحث بر روی آن‌ها در بخش ۴ آورده خواهند شد. در نهایت، نتیجه‌گیری مقاله در بخش ۵ خواهد بود.

برخوردار است. برای تشخیص زنده بودن عنبیه می‌توان از تغییرات و نوسان اندازه مردمک در برابر نور بهره گرفت.

امروزه استفاده از روش‌های یادگیری عمیق به دلیل سرعت پردازش و عملکرد مطلوب در بسیاری از حوزه‌ها از جمله پردازش تصویر مورد توجه قرار گرفته است [۵-۹]. در پژوهش حاضر، یک روش جدید مبتنی بر یادگیری عمیق و کاملاً کاربردی پیشنهاد شده است که می‌تواند بصورت بی‌درنگ برای احراز هویت متقاضیان خدمات غیرحضوری الکترونیک مورد استفاده قرار گیرد.

در روش پیشنهادی، ابتدا تصاویر متوالی از چهره شخص متقاضی خدمات دریافت می‌شوند. فیلتر گابور که در برابر ویژگی‌ها در چند مقیاس و چند جهت مقاوم است، و نیز الگوی باینری محلی یکنواخت که نسبت به چرخش تصویر نامتغیر است، بر روی این تصاویر اعمال می‌شوند. سپس، ویژگی‌های استخراج شده از تصاویر متوالی از کل چهره و ناحیه شامل عنبیه چشم در دو مسیر به‌طور همزمان به یک مدل پیشنهادی شبکه عصبی کانولوشنی سه بعدی (3D-DenseNet در دو مسیر) تغذیه می‌شوند. نگهداری ویژگی‌هایی با پیچیدگی کمتر، کارایی محاسباتی و گردش گرادیان قوی از مزایای این مدل هستند. در



شکل (۱): بلوک دیاگرام روش پیشنهادی.

۲. پژوهش‌های پیشین

در دهه‌های اخیر، استفاده از ویژگی‌های فیزیولوژیکی و بیومتریک افراد برای تشخیص هویت در سیستم حضور و غیاب کارکنان، سامانه‌ی ارائه خدمات هوشمند و ... مورد توجه است [۱۰-۱۱]. با این حال، اثر انگشت همانند امضا و دستخط ممکن است به صورت تقلبی توسط جاعل به دست آید [۱۱-۱۲]. تشخیص هویت به کمک امواج مغزی [۱۱] نیز نیاز به الکتروود و یک سری امکانات خاص دارد که برای احراز بی‌درنگ هویت و بدون تماس فیزیکی با فرد متقاضی خدمات غیر حضوری مناسب نیست. بنابراین، استفاده از عنبیه به عنوان یک معیار بیومتریک قابل اعتماد پیشنهاد می‌شود.

داگمن استفاده از توصیفگر گابور را برای استخراج ویژگی‌ها از تصاویر عنبیه، برای به دست آوردن بازدهی بالا پیشنهاد کرد [۱۳]. علاوه بر این، توصیفگرهای زیادی برای استخراج ویژگی پیشنهاد شده‌اند که از مزایای بافت پیچیده عنبیه استفاده می‌کنند. روش‌های مبتنی بر موجک هار [۱۴]، فیلتر بانک‌های فضایی [۱۵]، فیلتر بانک‌های جهتی [۱۶] و ویژگی‌های آماری باینری شده [۱۷] مورد بررسی قرار گرفته‌اند تا نرخ تشخیص را افزایش دهند.

با ورود تکنیک‌های یادگیری عمیق، رویکردهای جدیدی برای شناسایی عنبیه توسعه یافته‌اند. در سال ۲۰۱۶، DeepIris توسط لیو و همکاران ارائه شد [۱۸]. این کار به عنوان اولین تلاش برای حل مشکل تشخیص عنبیه با استفاده از شبکه عصبی کانولوشنی بود. این مدل از یک شبکه عصبی کانولوشنی و یک بانک فیلتر برای تأیید عنبیه تشکیل شده است. سپس، Gangwar و Joshi دو معماری عمیق‌تر به نام‌های DeepIrisNet A و DeepIrisNet B را پیشنهاد کردند [۱۹]. ژائو و کومار [۲۰] شبکه‌ای به نام UniNet را بر اساس شبکه‌های کاملاً کانولوشن ارائه کردند. آن‌ها یک تابع تلفات مربوط

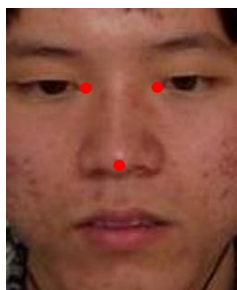
به تغییر تلفات سه گانه را معرفی کردند تا در مرحله تطبیق، بر روی عملیات انتقال بیت و ماسک ناحیه‌ی غیرعنبیه تمرکز کنند [۲۰]. وانگ و کومار [۲۱] مدلی را برای یادگیری شبکه بر اساس کرنل‌های کانولوشنی پیشنهاد کردند تا ویژگی‌های بیشتری را از تصاویر عنبیه به دست آورند. همچنین، استفاده از تکنیک‌های تنظیم دقیق (fine-tuning)، برای غلبه بر مشکل نیاز به مجموعه داده‌ی عنبیه زیاد در معماری‌های پیچیده مورد بررسی قرار گرفته است [۲۲]. مینایی و همکاران [۲۳] لایه‌های کانولوشنی یک VGG از قبل آموزش دیده را برای استخراج ویژگی عنبیه بررسی کردند و برای دستیابی به دقت بیشتر، تعداد مؤلفه‌های مختلف PCA را ارزیابی کردند. ژائو و همکاران روشی را بر اساس معماری شبکه کپسول معرفی کردند [۲۴]. آن‌ها چندین ساختار کانولوشن با عمق مختلف را با توجه به خروجی‌های مختلف شبکه‌های از پیش آموزش دیده برای اتصال به ساختار کپسول ساختند.

اخیراً، در [۲۵] یک روش تشخیص عنبیه بر اساس مدل‌های طبقه‌بندی تصویر از پیش آموزش دیده پیشنهاد شده است که در آن از فیلترهای لایه‌های اول شبکه‌های عصبی کانولوشنی به عنوان استخراج‌کننده ویژگی استفاده می‌شود و برای مجموعه داده‌های جدید، نیازی به fine-tuning ندارد. از آنجایی که ویژگی‌های انتخابی استخراج شده از لایه‌های کانولوشنی، سطح عنبیه را رمزگذاری می‌کنند؛ این مزیت را دارند که به موقعیت‌های فضایی خاص محدود نمی‌شوند. بنابراین، انجام فرآیند انتقال بیت در مرحله تطبیق ضروری نیست.

مرجع [۲۶] تشخیص عنبیه با مجموعه باز (بازشناسی در موقعیت‌های ناشناخته‌ای که مدل‌ها در طول آموزش یاد نگرفته‌اند) را بر اساس یادگیری عمیق پیشنهاد می‌کند. در این روش، با آموزش شبکه عمیق، ویژگی‌های عنبیه استخراج شده در نزدیکی مرکز ویژگی هر نوع تصویر عنبیه خوشه بندی می‌شوند.

۱.۳. پیش پردازش

برای جداسازی تصویر چهره از پس زمینه پیچیده، از فیلتر رنگ مدل مخلوط گاوسین استفاده می‌کنیم (برای مطالعه جزئیات بیشتر به [۲۷-۲۸] مراجعه شود). سپس، دو گوشه‌ی داخلی چشم و یک نقطه برجستگی نوک بینی در اولین فریم یافت شده و در سکانس‌های بعدی با استفاده از الگوریتم Kanade-Lucas-Tomasi [۲۹] ردیابی می‌شوند. بنابراین، چرخش سر و تغییرات اندازه صورت در سکانس‌ها بر اساس هماهنگ شدن با مختصات این سه نقطه تصحیح می‌شوند. سه نقطه مورد نظر در شکل ۲ نشان داده شده‌اند.



شکل (۲): سه نقطه مشخص شده در صورت.

برای هم‌ترازی چهره در فریم‌های متوالی (شکل ۳) از روش استفاده شده در [۳۰] استفاده می‌کنیم. فرض کنید که بردار $V = \{v_1, v_2, \dots, v_m\}$ نماد کل مجموعه‌ی برش‌های ویدئو و m نیز نشان‌دهنده‌ی تعداد نمونه‌ها باشد؛ در این صورت، b اُمین برش ویدئویی نمونه $v_b = (I_{b,1}, I_{b,2}, \dots, I_{b,nb})$ است، که در آن $I_{b,i}$ همان اُمین فریم و nb تعداد فریم برش ویدئویی v_b است. ابتدا از تصویر جلویی (اولین تصویر) صورت I_{mod} به عنوان مدل استفاده می‌کنیم. ۶۸ نقطه از مدل صورت I_{mod} با β استفاده از مدل شکل فعال^۱ شناسایی می‌شوند. برای نرمال‌سازی بیشتر تغییرات ناشی از نمونه‌های مختلف، از تابع تبدیل تصویر برای ایجاد ارتباط بین $I_{b,1}$ و مدل‌های چهره‌ی I_{mod} استفاده می‌

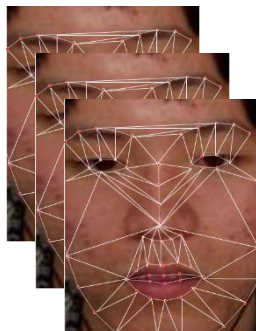
سپس، ویژگی‌های استخراج شده به یک فضای ویژگی جدید نگاشت شده و طبقه‌بندی می‌شوند. در نهایت، نمونه‌های کلاس ناشناخته با یک آستانه‌ی احتمال SoftMax تعیین می‌شوند. آزمایش‌ها نشان می‌دهند که این روش عملکرد نسبتاً خوبی در تشخیص عنبیه با مجموعه باز دارد و می‌تواند به طور مؤثر نمونه‌های عنبیه‌ی کلاس‌های ناشناخته را تشخیص دهد و تأثیر کمی بر توانایی تشخیص کلاس‌های شناخته شده نمونه‌های عنبیه دارد.

تاجایی که نویسندگان مقاله حاضر اطلاع دارند، روش‌های مورد استفاده در پژوهش‌های پیشین بر روی تنها یک تصویر عنبیه انجام شده‌اند و هیچ یک تصاویر متوالی عنبیه را برای تشخیص زنده بودن و شناسایی هویت فرد بصورت بی‌درنگ و بدون تأخیر مورد بررسی قرار نداده‌اند. همچنین، دقت تشخیص همچنان قابل ارتقاء است.

۳. روش پیشنهادی

در این بخش، روش پیشنهادی به تفصیل ارائه خواهد شد. برای احراز بی‌درنگ هویت متقاضیان خدمات غیرحضوری الکترونیکی، لازم است دوربین‌هایی در محل نصب شوند تا از چهره‌ی فرد فیلم‌برداری کنند. تعبیه چند دوربین در چند جهت مختلف برای تفکیک انسان واقعی از فیلم استفاده می‌شود. سپس، فیلم ضبط شده به تصاویر متوالی تبدیل شده و وارد مرحله پیش‌پردازش می‌شود. برای استخراج ویژگی و پس از آن طبقه‌بندی، فیلترهای گابور و الگوی باینری محلی یکنواخت به تصاویر (کل چهره و ناحیه چشم) اعمال می‌شوند و در دو مسیر به شبکه عصبی کانولوشنی سه بُعدی تغذیه می‌شوند. در نهایت، فرد بر اساس ویژگی‌های استخراج شده از عنبیه تشخیص داده می‌شود و وجود حرکات در چهره برای تشخیص زنده بودن فرد شناسایی می‌شود.

^۱ Active Shape Model



شکل (۳): هم‌ترازی چهره.

برای تشخیص ناحیه چشم و برش آن، ابتدا مقادیر $A1$ ، $A2$ و $A3$ با استفاده از رابطه (۳) مشخص می‌شوند [۲۸]:

$$A1 = (I_{g_x})^2, A2 = (I_{g_y})^2, A3 = (I_{g_x} I_{g_y})^2 \quad (3)$$

که در آن I_{g_x} و I_{g_y} به ترتیب گرادیان تصویر در جهت‌های x و y هستند. سپس، R از رابطه (۴) به دست می‌آید:

$$R = [(A1)(A2) - (A3)^2] - K(A1 + A2)^2 \quad (4)$$

نقاط با بالاترین مقادیر R در ناحیه مشخص شده به‌عنوان ناحیه چشم انتخاب می‌شوند. از آنجائی که یکی از عوامل مفید برای تعیین محل چشم، تغییرات نامنظم در سطوح روشنایی صلبیه (ناحیه سفید چشم) است [۲۸]. رگ‌های خونی در این ناحیه به آنتروپی در سطح صلبیه کمک می‌کنند. بنابراین، آنتروپی رنگ برای تعیین مردمک چشم محاسبه می‌شود. احتمال آنتروپی از رابطه (۵) به دست می‌آید [۲۸].

$$H(.) = - \sum_{q=1}^{num} pix(x_q) \log_2^{pix(x_q)} \quad (5)$$

کنیم و فریم‌های باقیمانده در یک سکانس می‌توانند در مدل چهره ثبت شوند. وزن میانه‌ی محلی $(W)^T$ برای محاسبه‌ی ماتریس انتقال برای ثبت چهره در تصویر نمونه بکار برده می‌شود. دلیل استفاده از ماتریس تبدیل این است که مدل شکل فعال ممکن است به اندازه کافی دقیق نباشد، بنابراین می‌تواند منجر به انحراف قابل توجهی از نقاط برای نقاط مشابه شود، حتی زمانی که صورت ثابت باقی می‌ماند. ماتریس انتقال T به صورت زیر (۱) است [۲۷]:

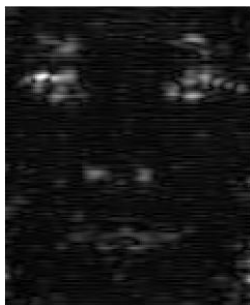
$$T_b = W(\beta(I_{mod}), \beta(I_{b,1})); \quad b=1, \dots, m$$

که $\beta(I_{b,1})$ هماهنگی ۶۸ نقطه از اولین فریم در کلیپ v_b است. سپس، فریم‌ها با استفاده از ماتریس T_b نرمال می‌شوند. تصویر نرمال شده I' طبق (۲) به‌عنوان تصویر انتقالی دو بعدی از تصویر اصلی، محاسبه می‌شود [۳۰]:

$$I'_{b,i} = T_b \times I_{b,i}; \quad i=1, \dots, n_b \quad (2)$$

که $I'_{b,i}$ آمین فریم نرمال شده کلیپ v_b است.

در آخرین مرحله، ناحیه‌ی چهره از تصویر نرمال شده از هر کلیپ و با استفاده از مستطیل تعریف شده در اولین فریم $I'_{b,i}$ برش داده می‌شود.



شکل (۴): تصویر اندازه‌ی گابور.

سپس، ویژگی‌های استخراج شده توسط فیلترهای گابور با استفاده از روش الگوی باینری محلی یکنواخت به صورت اعداد باینری رمزگذاری می‌شوند. در واقع، هر پیکسل یک تصویر اندازه‌ی گابور به وسیله آستانه‌گذاری پیکسل‌های f_p ($p=0,1,\dots,7$) همسایگی ۳ در ۳ با پیکسل مرکزی f_{center} و با در نظر گرفتن نتیجه به عنوان یک عدد باینری برچسب‌گذاری می‌شوند. در واقع، الگوی باینری محلی در تصویر اندازه‌ی گابور به شکل زیر (رابطه ۸) محاسبه می‌شود:

$$LBP(x_{center}, y_{center}) = \sum_{p=0}^7 Sgn(f_p - f_{center}) 2^p \quad (6)$$

$$Sgn(f_p - f_{center}) = 0; f_p < f_{center}$$

$$Sgn(f_p - f_{center}) = 1; f_{center} \leq f_p \quad (8)$$

به عنوان مثال، بخشی از یک تصویر با 3×3 پیکسل را در نظر بگیرید که در آن یک پیکسل در وسط (f_{center}) و هشت پیکسل در اطراف (f_p) قرار دارند. حال باید مقدار مربوط به هر پیکسل با مقدار پیکسل مرکزی مقایسه شود. به عبارت دیگر، تفاضل مقدار روشنایی هر پیکسل از مقدار روشنایی پیکسل مرکزی را محاسبه می‌کنیم. مقادیر پیکسل‌ها می‌تواند از ۰ تا ۲۵۵ باشد. سپس، برای تبدیل به الگوی باینری محلی (به صورت دودویی) مقادیر بزرگتر یا مساوی پیکسل مرکزی (مقادیر مثبت)،

که در آن $\text{pix}(x_q)$ احتمال پیکسل x_q را نشان می‌دهد و numb نشان دهنده تعداد پیکسل‌های هر بخش است. تابع احتمال هر پیکسل، فرکانس شدت رنگ پیکسل در ناحیه انتخاب شده تقسیم بر فرکانس کل پیکسل‌ها است. دو ناحیه با بالاترین آنتروپی در چشم به عنوان ناحیه شامل مردمک تعیین می‌شوند. مرکز این مناطق به عنوان مرکز مردمک در نظر گرفته می‌شود.

۲.۳. استخراج ویژگی

برای استخراج ویژگی، ابتدا، بانکی از فیلترهای گابور چندجهته و چندمقیاسه به نواحی برش داده شده (کل چهره و ناحیه شامل عنبیه چشم) برای دستیابی به تصاویر اندازه‌ی گابور^۳ در حوزه فرکانس اعمال می‌شود (شکل ۴). فیلتر گابور می‌تواند برای نمایش بهینه یک سیگنال در حوزه زمان و فرکانس به طور توأمان مورد استفاده قرار گیرد و به صورت زیر تعریف می‌شود [۳۱]:

$$G(X_0, Y_0, \alpha, \psi) = \frac{1}{2\pi\zeta^2} \exp\left[-\left(\frac{X_0^2 + Y_0^2}{2\zeta^2}\right)\right] \left[\exp(j\psi X_0) - \exp\left(-\frac{\psi^2 \zeta^2}{2}\right) \right]$$

که $X_0 = X \cos \alpha + Y \sin \alpha$, $Y_0 = -X \sin \alpha + Y \cos \alpha$ است، (X, Y) مختصات مکانی فیلتر گابور است، ψ و α معرف مقیاس و جهت فیلتر گابور هستند و $\zeta = \pi/\psi$ بیانگر انحراف معیار تابع گاوسی است.

بازنمایی گابور به وسیله کانولوشن با فیلتر گابور استخراج می‌شود [۳۱]:

$$G(X, Y, \alpha, \psi) = G(X_0, Y_0, \alpha, \psi) * I(X, Y) \quad (7)$$

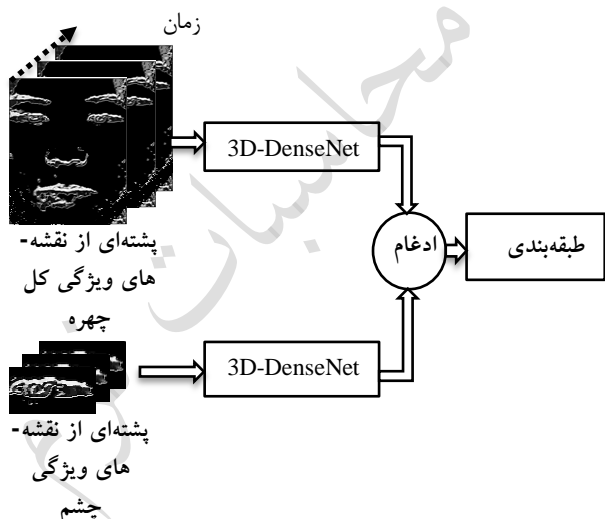
که * بیانگر عملگر کانولوشن است.

^۳ Gabor Magnitude

شکل (۶): نقشه‌ی ویژگی به دست آمده پس از اعمال الگوی باینری محلی یکنواخت (کل چهره (سمت راست) و ناحیه چشم شامل عنینه (سمت چپ)).

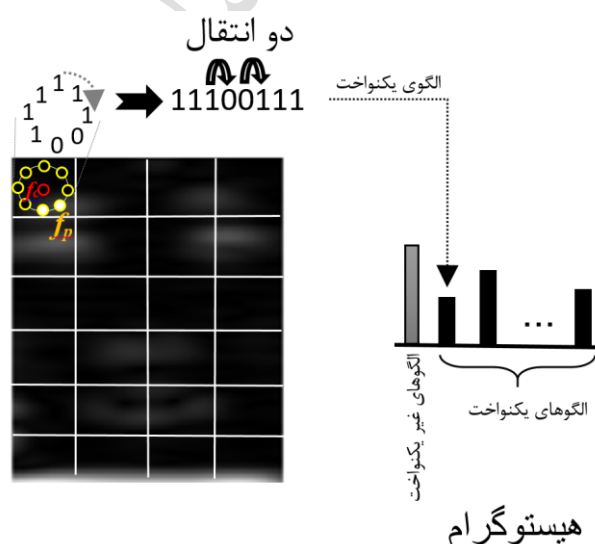
۳.۳. طبقه‌بندی

برای طبقه‌بندی ویژگی‌های استخراج‌شده و در نتیجه، احراز هویت و شناسایی زنده بودن افراد از مدل پیشنهادی شبکه‌ی عصبی کانولوشنی سه بعدی (3D-DenseNet) در دو مسیر استفاده می‌کنیم. تصاویر به دست آمده از مرحله‌ی قبل پس از تبدیل به فیلم، ورودی شبکه هستند. به طور واضح‌تر، پس از اعمال فیلترهای گابور و الگوی باینری محلی یکنواخت به تصاویر متوالی دریافت شده از دوربین، پشته‌ای از نقشه‌های ویژگی از کل ناحیه صورت و ناحیه چشم در امتداد بُعد زمان تولید می‌شود که به صورت فیلم در دو مسیر (در یک مسیر نقشه‌های ویژگی کل چهره و در مسیر دیگر، نقشه‌های ویژگی چشم) به طور همزمان به DenseNet سه بُعدی داده می‌شوند (شکل ۷).

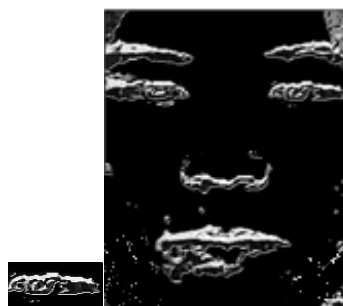


شکل (۷): پشته‌ای از نقشه‌های ویژگی از کل ناحیه صورت و ناحیه چشم در دو مسیر به طور همزمان به DenseNet سه بُعدی داده می‌شوند.

یک منطقی (۱) و مقادیر کوچکتر از آن (مقادیر منفی)، صفر منطقی (۰) در نظر گرفته می‌شوند. هشت عدد صفر یا یک کنار هم قرار می‌گیرند و اگر انتقال از صفر به یک یا بالعکس، کمتر یا مساوی دو باشد؛ آنگاه الگوی باینری محلی تولید شده، یکنواخت است و دارای بین^۴ جداگانه در هیستوگرام است. سایر الگوها در یک بین به نام الگوهای غیریکنواخت قرار می‌گیرند. برای درک بهتر، شکل ۵ را مشاهده کنید. نقشه‌ی ویژگی^۵ به دست آمده پس از اعمال الگوی باینری محلی یکنواخت در شکل ۶ قابل مشاهده است. این تصویر، ورودی مرحله بعد است.



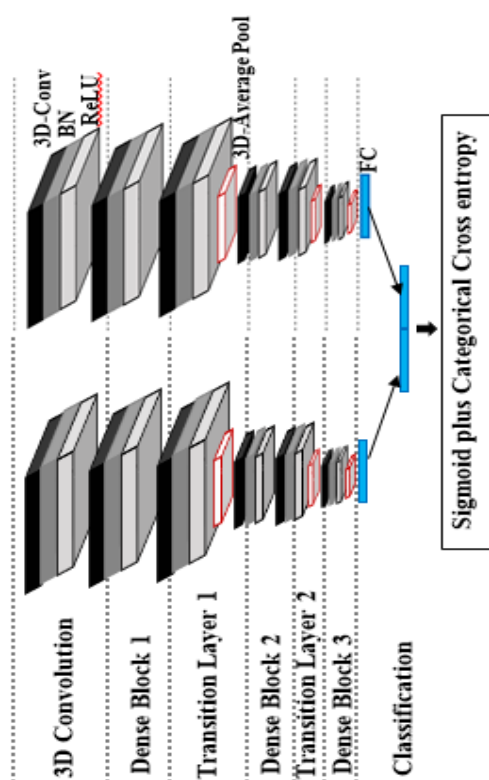
شکل (۵): محاسبه الگوی باینری یکنواخت بر روی تصویر اندازه گابور. پیکسل مرکزی f_{center} با رنگ قرمز و f_p ها با رنگ زرد مشخص شده‌اند.



⁴ Bin

⁵ Feature map

بودن تابع تصمیم و جلوگیری از اشباع است. در اتصال کامل دو مسیر به یکدیگر متصل می‌شوند. برای طبقه‌بندی، در لایه آخر از فعال‌ساز سیگموئید و **Categorical Cross entropy** استفاده می‌کنیم تا دو خروجی به‌طور همزمان فعال شود. یکی، فرد شناسایی شده را اعلام می‌کند و دیگری، وجود حرکات و تغییرات را در چهره وی مشخص می‌کند. معماری 3D-DenseNet در دو مسیر پیشنهادی در شکل ۸ نشان داده شده است.



شکل (۸): معماری 3D-DenseNet در دو مسیر پیشنهادی.

۴. آزمایش‌ها، نتایج و بحث

در این بخش، ابتدا پایگاه داده‌های مورد استفاده معرفی خواهند شد. سپس، نحوه‌ی اجرا و پیاده‌سازی روش پیشنهادی به تفصیل بیان خواهد شد. در نهایت، نتایج آزمایش‌ها و بحث بر روی نتایج ارائه خواهد شد. آزمایش‌ها بر روی یک رایانه شخصی RAM و Intel i7 3.5GHZ هشت گیگا بایت اجرا شده است.

در واقع، این شبکه می‌تواند حرکات ریز و درشت را در تصاویر متوالی یاد بگیرد و همچنین، فرد متقاضی خدمات را بر اساس ویژگی‌های منحصر بفرد وی شناسایی کند. برای ایجاد 3D-DenseNet، اطلاعات زمانی به تمام لایه‌های داخلی (یعنی: pooling سه بعدی، کانولوشن سه بعدی، ReLU، و کاملاً متصل (fully connected)) در DenseNet اضافه می‌شوند. لایه‌های سه بعدی با استفاده از مدل‌سازی اطلاعات زمانی، ویژگی‌های مکانی-زمانی را یاد می‌گیرند. برای کانولوشن سه بعدی، اطلاعات را در امتداد محور سوم (به عنوان مثال، بُعد زمانی) به کرنل، حجم ورودی و خروجی اضافه می‌کنیم. پشته‌های ورودی دارای سه بُعد (H: ارتفاع، Wi: عرض و L: طول) هستند. اگر نقشه‌های ورودی با اعمال افزودن صفر^۶، ۶×۶ باشند؛ آنگاه کرنل ۳×۳×۳ که به صورت یک مکعب است، بر روی آن‌ها اعمال می‌شود. در حقیقت، افزودن صفر برای اصلاح اندازه‌ی ویژگی است.

در pooling سه بعدی، ورودی $H \times W_i \times L \times C_M$ گرفته می‌شود. C_M کانال نقشه است. یک کرنل $H_k \times W_k \times L_k$ تعریف شده است. H_k ، W_k و L_k ارتفاع، عرض و طول کرنل هستند. گام، برابر با ۲ تعریف می‌شود. در نهایت، خروجی $H_o \times W_o \times L_o \times C_M o$ تولید می‌شود.

به طور کلی، ۳ بلوک متراکم^۷ در DenseNet سه بعدی با ۴۰ لایه وجود دارد. قبل از اولین بلوک متراکم، یک کانولوشن و کرنل سه بعدی روی نقشه‌ها استفاده می‌شود. دو لایه انتقال^۸ بین بلوک‌های متراکم وجود دارد که شامل کانولوشن و $TL \times 2 \times 2$ pooling است. در انتقال اول، TL برابر با ۱ است. در دیگری، برابر با ۲ می‌باشد. تابع فعال‌ساز ReLU برای افزایش غیرخطی

^۶ Zero padding

^۷ dense block

^۸ transition layer

۱.۴. پایگاه داده‌های مورد استفاده

روی تصاویر اندازه‌ی گابور محاسبه می‌شود. اندازه‌ی تصاویر به دست آمده را به 100×100 تغییر می‌دهیم. برای افزایش تعداد نمونه‌ها، هر ۱۰ تصویر نهایی به دست آمده‌ی متوالی را در پوشه‌ی جداگانه ذخیره می‌کنیم. فرض کنید یک نمونه در پایگاه داده CASME2 شامل ۸۰ تصویر متوالی است که عملیات اعمال فیلترهای گابور و الگوی باینری محلی یکنواخت بر روی آن‌ها انجام شده است. اکنون، تصاویر نهایی به دست آمده از این نمونه در ۸ پوشه ذخیره می‌شوند (تصاویر به دست آمده از فریم ۱ تا ۱۰ در پوشه اول، ۱۱ تا ۲۰ در پوشه دوم، و ...). همچنین، از تکنیک داده افزایی چرخش تصاویر استفاده می‌نماییم.

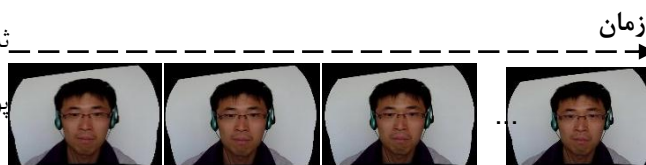
سپس، تصاویر هر پوشه را به ویدئو کلیپ تبدیل می‌کنیم که ورودی شبکه عصبی کانولوشنی سه بعدی پیشنهادی خواهد بود. کلیپ‌های مربوط به کل چهره را که در آن‌ها حرکت ماهیچه‌های صورت وجود دارد، با شماره در پوشه‌ی M ذخیره می‌کنیم. علاوه بر این، کلیپ‌های حاصل از تصاویر چشم را در پوشه‌های نامگذاری شده با نام شخص ذخیره می‌کنیم. همچنین، از تصاویر ثابت (بدون حرکت) چند کپی گرفته و کلیپ‌های حاصل را در پوشه DM ذخیره می‌کنیم.

حال، باید 3D-DenseNet در دو مسیر پیشنهادی را آموزش دهیم. پیاده‌سازی آن در پایتون با Tensor-flow انجام می‌شود. برای بهینه‌سازی از SGD^۹ استفاده می‌شود. اندازه دسته^{۱۰} برابر با ۱۰ انتخاب شده است. نرخ یادگیری اولیه $0.01/0.01$ در ۵۰ دوره^{۱۱} است. زمان آموزش $1/5$ روز طول کشید. نسبت داده‌های آزمایش به آموزش، $0.25/0.75$ است.

۳.۴. نتایج آزمایش‌ها و بحث

در این مقاله از پایگاه داده‌های CASME و CASME2 [۳۲-۳۳] که حرکات بسیار ریز و نامحسوس را در تصاویر متوالی چهره افراد نشان می‌دهند، استفاده شده است. قابل ذکر است که در برخی از نمونه‌ها، حرکات بیشتری مانند پلک زدن و باز و بسته شدن دهان نیز وجود دارد. همچنین، تصاویر هر دو پایگاه داده در شرایط واقعی ثبت شده‌اند.

۱۹۵ نمونه شامل تصاویر متوالی با رزولوشن 720×1280 پیکسل در CASME و ۲۴۷ نمونه دارای تصاویر متوالی با رزولوشن 240×280 پیکسل در CASME2 وجود دارد. نرخ فیلم‌برداری ۲۰۰ و ۶۰ فریم در ثانیه به ترتیب در CASME2 و CASME گزارش شده است [۳۲-۳۳]. به علاوه، در هر نمونه‌ی آن‌ها تصویر بدون حرکت چهره وجود دارد. نمونه‌ای از تصاویر متوالی یک نمونه در پایگاه داده CASME در شکل ۹ قابل مشاهده است.



شکل (۹): نمونه‌ای از تصاویر متوالی یک نمونه در CASME.

۲.۴. آزمایش‌ها و تنظیم پارامترها

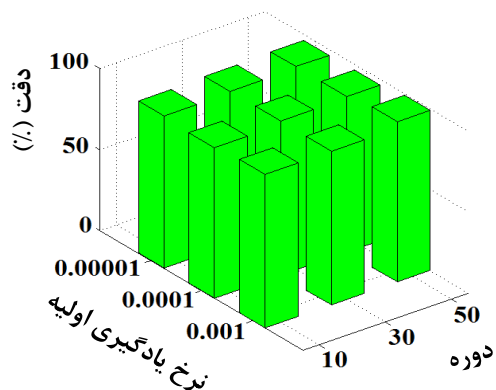
مختصات سه نقطه (نوک بینی و دو گوشه‌ی داخلی چشم) در اولین فریم تصاویر متوالی هر نمونه از پایگاه داده، مشخص شده و با الگوریتم Kanade-Lucas-Tomasi در سایر فریم‌ها ردیابی می‌شوند. سپس، هم‌ترازی چهره در فریم‌های متوالی انجام می‌شود و ناحیه‌ی چهره و چشم برش داده می‌شود.

فیلترهای گابور با ۳ مقیاس و ۲ جهت (۶ فیلتر) به تصاویر چهره و چشم اعمال می‌شوند و الگوی باینری محلی یکنواخت بر

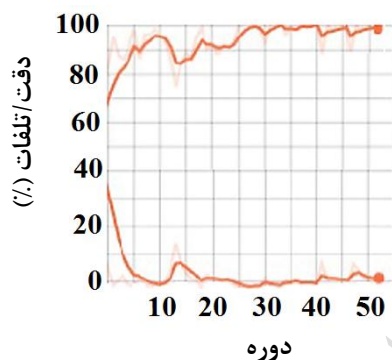
⁹ Stochastic Gradient Descent

¹⁰ batch size

¹¹ epoch



شکل (۱۱): دقت احراز هویت با نرخ یادگیری اولیه و اندازه دسته مختلف.



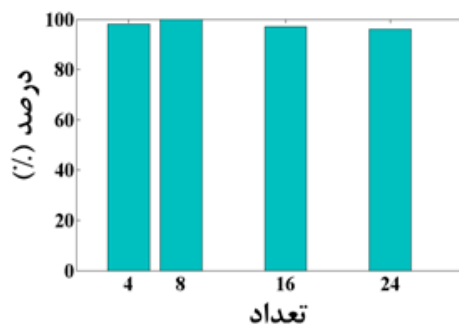
شکل (۱۲): نمودار دقت و تلفات در مرحله آموزش شبکه.

همانطور که مشاهده می‌شود، روش پیشنهادی نتایج مطلوبی را تولید کرده است. این نشان می‌دهد که اعمال فیلتر گابور و الگوی باینری محلی یکنواخت، ویژگی‌های مفید و مؤثری را در اختیار شبکه قرار می‌دهند. قابل ذکر است که پارامتر تعداد f_p بر نتیجه احراز هویت افراد تأثیرگذار است. در شکل ۱۰، مقادیر مختلف این پارامتر و درصد احراز هویت با استفاده از روش پیشنهادی آمده است. طبق نتایج به دست آمده، $f_p = 8$ بهترین نتیجه را تولید کرده است. همچنین، دقت احراز هویت با نرخ یادگیری اولیه در اندازه دسته مختلف، متغیر است که در شکل ۱۱ می‌توان مشاهده کرد. بهترین عملکرد با نرخ یادگیری اولیه 0.001 در 50 دوره است. نمودار دقت و تلفات در مرحله آموزش نیز در شکل ۱۲ نشان داده شده است.

درصد احراز هویت و تشخیص افراد با روش‌های مختلف استخراج‌کننده ویژگی در جدول ۱ آورده شده است.

جدول (۱): احراز هویت و تشخیص افراد با استفاده از روش‌های مختلف استخراج‌کننده ویژگی

روش	درصد احراز هویت (%)
نقشه لبه مبتنی بر گرادیان	۸۶/۴۹
الگوی باینری محلی	۹۰/۷۹
توزیع B	۹۲/۸۸
توزیع ویگنر-سوایل	۹۳/۱۵
فیلتر گابور	۹۴/۰۷
روش فازی	۹۶/۳۸
ترکیب الگوی باینری محلی و فیلتر گابور	۹۸/۳۴
روش پیشنهادی	۹۹/۹۹



شکل (۱۰): درصد احراز هویت افراد با مقادیر مختلف تعداد f_p .

طبق نتایج، میانگین صحت در هر دو پایگاه داده بیش از ۹۹/۵ درصد است. میانگین نرخ مثبت صحیح برای شناسایی صحیح افراد نیز حدود ۹۸٪ در هر دو پایگاه داده‌ی مورد استفاده است. بنابراین، احراز هویت متقاضیان خدمات غیرحضور در شرایط واقعی با استفاده از روش پیشنهادی با خطای بسیار کم انجام خواهد شد.

قابل ذکر است که دقت شناسایی و امنیت پایدار با رزولوشن تصاویر نسبت مستقیم دارد. شکل ۱۳ دقت شناسایی و احراز هویت شخص را با رزولوشن‌های مختلف تصاویر نشان می‌دهد. در نتیجه، استفاده از دوربین با رزولوشن بالا و نرخ فریم مناسب کارآمد خواهد بود. زیرا وضوح و رزولوشن پایین می‌تواند منجر به خطا و اشتباه در شناسایی شود. نرخ فریم نیز باید مناسب باشد تا در اثر حرکت، تصاویر مات تولید نشوند.

در جدول ۵، مقایسه‌ی دقت شناسایی افراد بر اساس عنبیه با استفاده از مدل‌های مختلف شبکه عصبی کانولوشنی مورد استفاده در پژوهش‌های پیشین و مدل پیشنهادی در پژوهش حاضر انجام شده است. این جدول، افزایش دقت احراز هویت را با مدل پیشنهادی تأیید می‌کند. همچنین، در جدول ۶، درصد شناسایی حرکات ریز در دو پایگاه داده مورد استفاده با سایر مقالات مرتبط مقایسه شده است که نشان دهنده عملکرد مطلوب روش پیشنهادی در مقابل سایر روش‌ها می‌باشد.

به علاوه، بهره‌وری روش پیشنهادی با استفاده از معیارهای شناخته شده اندازه‌گیری می‌شود. این معیارها حساسیت^{۱۲} (Sensitiv.)، امتیاز F1، ویژگی^{۱۳} (Specifi.) و صحت^{۱۴} (Preci.) هستند. برای محاسبه این معیارهای ارزیابی، از معادلات زیر استفاده می‌کنیم [۳۴]:

$$Sensitiv. = \frac{Tr.Po.}{Fal.Ne. + Tr.Po.} \quad (۹)$$

$$Preci. = \frac{Tr.Po.}{Fal.Po. + Tr.Po.} \quad (۱۰)$$

$$F1score = 2 \left(\frac{Preci. \times Sensitiv.}{Preci. + Sensitiv.} \right) \quad (۱۱)$$

$$Specifi. = \frac{Tr.Ne.}{Tr.Ne. + Fal.Po.} \quad (۱۲)$$

که در آن‌ها، منظور از Fal.Ne.، Fal.Po.، Tr.Ne.، Tr.Po. و به ترتیب مثبت صحیح، منفی صحیح، مثبت کاذب و منفی کاذب می‌باشد.

سپس، استراتژی اعتبارسنجی متقابل ۱۰ برابری به کار گرفته شده است. در این استراتژی، داده‌ها به طور تصادفی به ده دسته تقسیم می‌شوند که تقریباً دارای نسبت‌های مشابهی از برچسب‌های کلاس در هر دسته هستند. در ده آزمایش، نه مورد برای آموزش و یکی برای آزمایش استفاده می‌شود و در نهایت، میانگین نتایج گزارش می‌شود. نتایج حاصل از اعتبارسنجی متقابل با استفاده از روش پیشنهادی در پایگاه داده‌های CASME و CASME2 به ترتیب در جدول‌های ۳ و ۴ آورده شده است.

¹² Sensitivity

¹³ Specificity

¹⁴ Precision

جدول (۴): نتایج حاصل از اعتبارسنجی متقابل با استفاده از روش

پیشنهادی در پایگاه داده‌های CASME2.

حساسیت	ویژگی	صحت	امتیاز F1	
۰/۹۹	۱	۱	۰/۹۹	Fold1
۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹	Fold2
۰/۹۷	۰/۹۸	۰/۹۸	۰/۹۷	Fold3
۰/۹۸	۰/۹۹	۱	۰/۹۸	Fold4
۰/۹۵	۰/۹۹	۱	۰/۹۷	Fold5
۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹	Fold6
۰/۹۶	۰/۹۸	۱	۰/۹۷	Fold7
۰/۹۷	۰/۹۶	۱	۰/۹۸	Fold8
۰/۹۸	۰/۹۸	۰/۹۹	۰/۹۸	Fold9
۰/۹۹	۱	۱	۰/۹۹	Fold10
۰/۹۷۷	۰/۹۸۶	۰/۹۹۵	۰/۹۸	میانگین

جدول (۳): نتایج حاصل از اعتبارسنجی متقابل با استفاده از روش

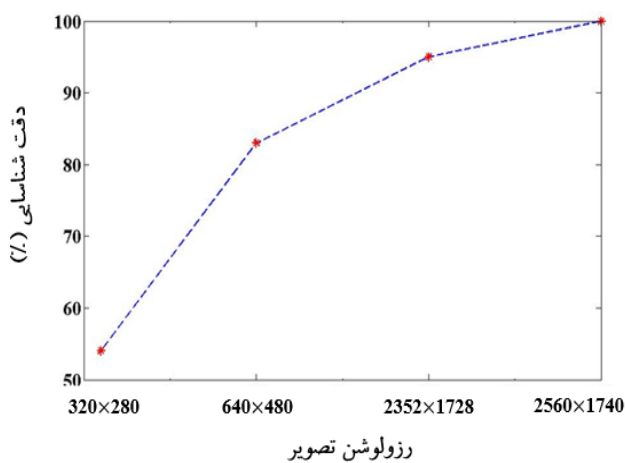
پیشنهادی در پایگاه داده‌های CASME.

حساسیت	ویژگی	صحت	امتیاز F1	
۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹	Fold1
۰/۹۸	۰/۹۸	۱	۰/۹۸	Fold2
۰/۹۹	۰/۹۹	۰/۹۹	۰/۹۹	Fold3
۰/۹۷	۰/۹۸	۱	۰/۹۸	Fold4
۰/۹۶	۰/۹۹	۱	۰/۹۷	Fold5
۰/۹۸	۰/۹۸	۱	۰/۹۸	Fold6
۰/۹۸	۰/۹۷	۰/۹۹	۰/۹۸	Fold7
۰/۹۹	۰/۹۶	۱	۰/۹۹	Fold8
۰/۹۹	۱	۱	۰/۹۹	Fold9
۰/۹۸	۰/۹۹	۱	۰/۹۸	Fold10
۰/۹۸	۰/۹۸	۰/۹۹۷	۰/۹۸	میانگین

جدول (۵): دقت شناسایی افراد بر اساس عنبیه با استفاده از مدل‌های

مختلف شبکه عصبی کانولوشنی.

دقت (%)	مدل
۹۳/۸	VGG-16
۹۹/۳	ResNet50
۹۹/۴	Xception
۹۹/۹	مدل پیشنهادی



شکل (۱۳): احراز هویت با تصاویر با رزولوشن متفاوت.

جدول (۶): درصد شناسایی حرکات ریز در دو پایگاه داده CASME و

CASME2

CASME2	CASME	مرجع
درصد شناسایی (%)	درصد شناسایی (%)	

- [۲] راستگو، محمد؛ جلالی، مهرداد، "تشخیص جرایم سایبری در ارتباطات بر خط با رویکرد داده کاوی"، نشریه علمی پدافند غیرعامل، سال یازدهم، شماره ۱، صص ۷۰-۶۳، ۱۳۹۹.
- [۳] نجف‌زاده کشتلی، سمیرا؛ آقاگل‌زاده، علی؛ کاظمی‌تبار، سید جواد، "بهبود امنیت پایگاه داده سیستم بیومتریک عنیبیه"، نشریه علمی علوم و فناوری‌های پدافند نوین، سال دوازدهم، شماره ۱، صص ۱۱۹-۱۰۷، ۱۴۰۰.

۸۸	۹۳	[35]
۹۵	۸۰	[36]
۸۶	۷۴	[37]
۹۸	۹۹	مدل پیشنهادی

۵. نتیجه‌گیری

- در این مقاله، به ارائه روشی پیشنهادی برای احراز هویت متقاضیان خدمات غیرحضوری الکترونیک بصورت بی‌درنگ بر مبنای سنج بیومترکی عنیبیه پرداخته شد. طبق نتایج، مشخص شد که استفاده از این روش، امنیت ملی پایدار را در فضای مجازی تا حد قابل قبولی افزایش می‌دهد. بنابراین، روش پیشنهادی می‌تواند به‌عنوان پدافند غیرعامل در فضای سایبری و هنگام استفاده از خدمات الکترونیک غیرحضوری مصونیت ایجاد کند. علت دقت بالا در احراز هویت اشخاص با روش پیشنهادی، ترکیب فیلترهای گابور و الگوی باینری محلی یکنواخت با شبکه عصبی کانولوشنی DenseNet سه بعدی در دو مسیر می‌باشد که علاوه بر احراز هویت، زنده بودن شخص را نیز تشخیص می‌دهد. کل زمان پردازش بین ۰٫۱ تا ۰٫۳ ثانیه بود. در آینده، این روش می‌تواند برای احراز هویت متقاضیان خدمات غیرحضوری الکترونیک در سایر ارگان‌ها مورد استفاده قرار گیرد.
- [۴] صنیعی، رعنا؛ قبله، عباس؛ ابراهیمی مقدم، محسن، "شناسایی هویت از روی نحوه راه رفتن با استفاده از مدل سلسله مراتبی فازی - اسپایکی"، مجله علوم رایانشی، شماره ۱۱، صص ۹۵-۸۰، ۱۳۹۷.
- [5] Esmaili, V., Mohassel Feghhi, M., Shahdi, S. O., "A comprehensive survey on facial micro-expression: approaches and databases", *Multimedia Tools and Applications*, pp. 1-46, 2022.
- [6] Esmaili, V., Mohassel Feghhi, M., "Diagnosis of Covid-19 Disease by Combining Hand-crafted and Deep-learning Methods on Ultrasound Data", *Journal of Machine Vision and Image Processing*, 2022.
- [7] Esmaili, V., Feghhi, M. M., Shahdi, S. O., "Early COVID-19 Diagnosis from Lung Ultrasound Images Combining RIULBP-TP and 3D-DenseNet", In *2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, pp. 1-5: IEEE, 2022.
- [۸] افتخاریان، محسن؛ نودهی، علی، "بهبود تشخیص و طبقه بندی توده های سرطان سینه بر اساس یادگیری عمیق و روش های پردازش تصویر"، مجله علمی محاسبات نرم، ۱۴۰۱.

مراجع

- [۱] کاویانی، حسن؛ میرسپاسی، ناصر؛ معمارزاده طهران، غلامرضا، "الگوی توسعه راهبردی منابع انسانی در حوزه امنیت سایبری نیروهای مسلح ج.ا.ایران"، فصلنامه علمی راهبرد دفاعی، سال هجدهم، شماره ۶۹، ۱۳۹۹.
- [۹] زارع مهرجردی، فاطمه؛ یزدیان دهکردی، مهدی؛ لطیف، علی - محمد، "ارزیابی روش‌های یادگیری کلاسیک و یادگیری عمیق در تجزیه و تحلیل احساسات داده‌های تلگرام فارسی"، مجله علمی محاسبات نرم، ۱۴۰۱.

- [18] Liu, N., Zhang, M., Li, H., Sun, Z., Tan, T., "Deepiris: Learning pairwise filter bank for heterogeneous iris verification", Pattern Recognit. Lett., vol. 82, no. 2, pp. 154-161, 2016.
- [19] Gangwar, A., Joshi, A., "DeepIrisNet: Deep iris representation with applications in iris recognition and cross sensor iris recognition", in Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 2301-2305, 2016.
- [20] Zhao, Z., Kumar, A., "Towards more accurate iris recognition using deeply learned spatially corresponding features", in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 3829-3838, 2017.
- [21] Wang, K., Kumar, A., "Toward more accurate iris recognition using dilated residual features", IEEE Trans. Inf. Forensics Security, vol. 14, no. 12, pp. 3233-3245, 2019.
- [22] Nguyen, K., Fookes, C., Ross, A., Sridharan, S., "Iris recognition with off-the-shelf CNN features: A deep learning perspective", IEEE Access, vol. 6, pp. 18848-18855, 2017.
- [23] Minaee, S., Abdolrashidiy, A., Wang, Y., "An experimental study of deep convolutional features for iris recognition", in Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB), pp. 1-6, 2016.
- [24] Zhao, T., Liu, Y., Huo, G., Zhu, X., "A deep learning iris recognition method based on capsule network architecture", IEEE Access, vol. 7, pp. 49691-49701, 2019.
- [25] Zambrano, J. E., Benalcazar, D. P., Perez, C. A., Bowyer, K. W., "Iris Recognition Using Low-Level CNN Layers Without Training and Single Matching", IEEE Access, vol. 10, pp. 41276-41286, 2022.
- [26] Sun, J., et al. "Open-set iris recognition based on deep learning", IET Image Processing, 2022.
- [27] Fini, M. R. R., Kashani, M. A. A., Rahmati, M., "Eye detection and tracking in image with complex background", In 2011 3rd International Conference on Electronics Computer Technology, pp. 57-61: IEEE, 2011.
- [28] Kashani, M. A. A., Arani, M. M., Fini, M. R. R., "Eye detection and tracking in images with using bag of pixels", In 2011 IEEE 3rd International Conference on Communication Software and Networks, pp. 64-68: IEEE, 2011.
- [۱۰] بزرگ‌خو، ناهید؛ هاشم‌زاده خوراسگانی، غلامرضا، "مؤلفه‌های تأثیرگذار بر فناوری رایانش ابری در ارائه خدمات هوشمند در سیستم حمل و نقل"، نشریه علمی فناوری اطلاعات و ارتباطات انتظامی، سال دوم، شماره دوم (پیاپی ۶)، صص ۴۲-۳۳، ۱۴۰۰.
- [۱۱] اعتمادی، زهرا؛ امینان مدرس، امیر مسعود، "تحلیل و مقایسه روش‌های نوین تشخیص هویت به کمک امواج مغزی در شرایط یکسان و در حالات ذهنی متفاوت"، نشریه علمی فناوری اطلاعات و ارتباطات انتظامی، دوره اول، شماره ۳، صص ۵۸-۴۷، ۱۳۹۹.
- [۱۲] فرهنگدپور، زینب؛ نیک‌مهر، هومان؛ منصوری‌زاده، محرم؛ طیب‌زاده قمصری، امید، "یک سیستم نوین هوشمند تشخیص هویت نویسنده فارسی زبان بر اساس سبک نوشتاری"، مجله علمی محاسبات نرم، دوره اول، شماره دوم، صص ۳۵-۲۶، ۱۳۹۱.
- [13] Daugman, J., "How iris recognition works", Essential Guide Image Process., vol. 14, no. 1, pp. 715_739, 2009.
- [14] Ng, T. W., Tay, T. L., Khor, S. W., "Iris recognition using rapid Haar wavelet decomposition", in Proc. 2nd Int. Conf. Signal Process. Syst., p. 820, 2010.
- [15] Ma, L., Wang, Y., Tan, T., "Iris recognition using circular symmetric filters", in Proc. 16th Int. Conf. Pattern Recognit., vol. 2, pp. 414-417, 2002.
- [16] Park, C.-H., Lee, J.-J., Smith, M. J., Park, K.-H., "Iris based personal authentication using a normalized directional energy feature", in Proc. Int. Conf. Audio Video-Based Biometric Person Authentication, pp. 224-232: Springer, 2003.
- [17] Czajka, A., Moreira, D., Bowyer, K. W., Flynn, P. J., "Domain-specific human-inspired binarized statistical image features for Iris recognition", in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), pp. 959-967, 2019.

- [29] Tomasi, C., Kanade, T., "*Detection and Tracking of Point Features*", School of Computer Science, Carnegie Mellon Univ. Pittsburgh, 1991.
- [30] Li, X., et al, "*Towards reading hidden emotions: a comparative study of spontaneous micro-expression spotting and recognition methods*", IEEE Trans Affect Comput, vol. 9, no. 4, pp. 563–577, 2018.
- [31] Hu, C., Jiang, D., Zou, H., Zuo, X., Shu, Y., "*Multi-task micro-expression recognition combining deep and handcrafted features*", In 2018 24th International Conference on Pattern Recognition (ICPR), pp. 946-951: IEEE, 2018.
- [32] Yan, W.J., Li, X., Wang, S.J., Zhao, G., Liu, Y.J., Chen, Y.H., Fu, X., "*CASME II: An improved spontaneous micro-expression database and the baseline evaluation*", PloS one, Vol. 9, No. 1, p. e86041, 2014.
- [33] Yan, W.J., Wu, Q., Liu, Y.J., Wang, S.J., Fu, X., "*CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces*", In 2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG), IEEE, pp. 1-7, 2013.
- [34] Nicholson, C.: *Evaluation metrics for machine learning Accuracy, precision, recall, and F1 defined*, <https://wiki.pathmind.com/accuracy-precision-recall-f1>, Accessed July 2022.
- [35] Esmaili, V., Mohassel Fegghi, M., Shahdi, S. O., "*Spotting micro-movements in image sequence by introducing intelligent cubic-LBP*", IET Image Processing, Vol. 16, No. 14, pp. 3814-3830, 2022.
- [36] Esmaili, V., Mohassel Fegghi, M., Shahdi, S. O., "*Automatic Micro-Expression Recognition using LBP-SIPI and FR-CNN*", AUT Journal of Modeling and Simulation, Vol. 54, No. 1, pp. 59-72, 2022.
- [37] Esmaili, V., Mohassel Fegghi, M., Shahdi, S. O., "*Micro-Expression Recognition based on the Multi-Color ULBP and Histogram of Gradient Direction from Six Intersection Planes*", Journal of Iranian Association of Electrical and Electronics Engineers, 2022.