بسم الله الرحمن الرحيم

دانشکدهٔ کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

پایان‌نامهٔ کارشناسی ارشد

طراحی و پیاده سازی سیستم احراز هویت تحت وب از طریق تشخیص چهره

دانشجو:

ایمان فیروزیان

استاد راهنما:

دکتر مرتضی زاهدی

استاد مشاور:

دکتر علیرضا احمدی فرد

مهندس مرضیه رحیمی

پایان‌نامه جهت اخذ درجهٔ کارشناسی ارشد

تابستان ۱۳۹۲

دانشگاه صنعتی شاهرود

دانشکده: کامپیوتر و فناوری اطلاعات

گروه: هوش مصنوعی

پایان‌نامهٔ کارشناسی ارشد آقای ایمان فیروزیان

تحت عنوان: طراحی و پیاده سازی سیستم احراز هویت تحت وب از طریق تشخیص چهره

در تاریخ ..... توسط کمیته تخصّصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با درجهٔ ..... مورد پذیرش قرار گرفت.

استاد راهنما	امضاء	استاد مشاور	امضاء
نام و نام خانوادگی : دکتر مرتضی زاهدی		نام و نام خانوادگی : دکتر علیرضا احمدی فرد مهندس مرضیه رحیمی	

اساتید داور	امضاء	نمایندهٔ تحصیلات تکمیلی	امضاء
نام و نام خانوادگی : دکتر حمید حسن پور		نام و نام خانوادگی :	
نام و نام خانوادگی : دکتر امیدرضا معروضی			
نام و نام خانوادگی :			
نام و نام خانوادگی :			

سرآغاز

به نام خدای که نام او راحت روح است و پیغام او مفتاح فتوح است و سلام او وقت صباح مؤمنان را

صبوح است و ذکر او مرهم دل مجروح است و مهر او بلانشینان را کشتی نوح است.

ای کریمی که بخشندهٔ عطایی و ای حکیمی که پوشندهٔ خطایی و ای صمدی که از ادراک خلق جدایی

و ای احدی که در ذات و صفات، بی همتایی و ای خالقی که راهنمایی و ای قادری که خدایی را سزایی، جان ما را

صفای خود ده و دل ما را هوای خود ده و چشم ما را ضیای خود ده

و ما را آن ده که آن به و مگذار ما را به که و مه.

الهی، عذر ما را بپذیر؛ بر عیب های ما مگیر. الهی در دل های ما جز تخم محبّت خود مکار و بر تن و جان های ما جز

الطاف و مرحمت خود مگار و بر کشته های ما جز باران رحمت خود مبار.

(یکی از مناجات های خواجه عبدالله انصاری)

تقدیم به:

پدر و مادر عزیزتر از جانم

آنانی که با راهنمایی ها و زحمات بی دریغشان مسیر پیشرفت و خوشبختی را آشکار ساخته،

دعای خالصانه شان بدرقه راه و وجودشان دلگرمی وجودم است.

آنانی که کوشیدند تا بدانم هدف غایی وجود این است

که هستی را پروردگاریست بی مانند.

قدردانی

حال که به لطف و رحمت لایتناهی حضرت حق، مراحل این پایان نامه رو به اتمام نهاده، بر خود لازم دانسته تا از همه دوستانی که در پیشبرد اهداف این پایان نامه اینجانب رامساعدت و یاری نموده اند، سپاس و قدردانی به عمل آورم.

ابتدا بر خود لازم می دانم تا از زحمات و پشتیبانی شبانه روزی، بی دریغ و بی شائبه استاد محترم، جناب آقای دکتر مرتضی زاهدی که راهنمائی این تحقیق را بر عهده داشته و از هرگونه راهنمایی و مساعدت مضایقه نکردند، تشکر و قدردانی ویژه ای داشته باشم. بی شک بدون حمایت و پشتیبانی ایشان انجام این تحقیق مقدور نمی بود.

هم چنین از جناب آقای دکتر علیرضا احمدی فرد و خانم مهندس مرضیه رحیمی که به عنوان اساتید مشاور، نقشی ارزنده و مکمل در انجام این پایان نامه داشتند سپاس گزاری می نمایم.

از جناب آقای دکتر علی اکبر پویان و همچنین جناب آقای دکتر حمید حسن پور، ریاست محترم دانشکده کامپیوتر و فناوری اطلاعات دانشگاه صنعتی شاهرود به جهت مساعدت هایشان در طی دوره کارشناسی ارشد قدردانی می نمایم.

در انتها سپاس گزار کلیه دوستان و عزیزانی هستم که در مراحل مختلف تحصیلی مرا یاری نمودند؛ هرچند مجالی برای بیان نام تک تک این دوستان نیست، لیکن همیشه خود را مرهون لطف و مهربانی آن ها دانسته و امید دارم به لطف الهی روزی توان جبران محبت هایشان را داشته باشم.

تعهد نامه

اینجانب ایمان فیروزیان دانشجوی دوره کارشناسی ارشد رشته کامپیوتر گرایش هوش مصنوعی دانشکده کامپیوتر و فناوری اطلاعات دانشگاه صنعتی شاهرود نویسنده پایان‌نامه طراحی و پیاده سازی سیستم احراز هویت تحت وب از طریق تشخیص چهره تحت راهنمائی دکتر مرتضی زاهدی متعهد می‌شوم.

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

.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

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

چکیده

در سال‌های اخیر، برخی سیستم‌های نرم‌افزاری تحت وب نیازمند افزودن یک لایه امنیتی برای احراز هویت کاربران خود می‌باشد. توسعه دهندگان وب و متقاضیان مربوطه دیگر، نیاز به روش‌هایی برای احراز هویت تحت وب کاربرانی همانند کاربران ارشد سیستم‌ها، کاربران آموزش الکترونیک و موارد مشابه دارند. در این پایان‌نامه، احراز هویت تحت وب از طریق تشخیص چهره با استفاده از دوربین وب‌کم[1] انجام شده است. این پایان‌نامه یک سیستم عملیاتی تشخیص چهره تحت وب را ارائه می‌دهد. این سیستم تشخیص چهره به سه بخش اپلت[2] تشخیص کاربر، اپلت افزودن کاربر و سرور جاوای تشخیص چهره تقسیم بندی می‌شود. عملیات اخذ تصویر از وب‌کم و عملیات چهره یابی به همراه پردازش‌های مورد نیاز در اپلت جاوای سمتِ کلاینت صورت می‌گیرد. عملیات‌های اصلی آموزش و آزمایش تصاویر چهره به همراه پایگاه داده مربوطه در بخش سرور قرار دارد.

هدف ارائه سیستم عملیاتی تشخیص چهره تحت وب، ارائه یک راه حل برای مشکل کوچکی از مشکلات کشور در حوزه تخصصی در قالب پایان‌نامه بوده است و بیشتر بحث پیاده سازی در آن مطرح است اما این پایان‌نامه، علاوه بر سیستم مذکور، یک رویکرد نوینی را نیز در باب تشخیص چهره ارائه می‌دهد. ایده این رویکرد از آنجا شکل گرفت که الگوریتم‌های تشخیص چهره موجود، با چهره‌ها همانند اشیاء یا بافت‌ها رفتار کرده اند اما سیستم تشخیص چهره انسانی با چهره با رویکردهای متفاوت دیگری نیز برخورد می‌کند. رویکرد نوین تشخیص چهره، از یک رویکرد آگاهانه سیستم تشخیص چهره انسانی استفاده می‌کند و آن را فرموله می‌کند. در این رویکرد، از توصیفات انسانی برای چهره‌ها استفاده می‌کند. این توصیفات بر اساس تمییزدهنده‌ترین ویژگی‌هاییست که نسبت به چهره میانگین شکل گرفته در ذهن فرد می‌باشد. به طور خاص، مردم در منطقه جغرافیایی خاص در ذهن خود یک چهره میانگینی از

[1] Webcam
[2] Applet

چهره‌های افراد آن منطقه ساخته اند و هرگاه می‌خواهند چهره‌ای را توصیف کنند، به تمییزدهنده‌ترین ویژگی‌های آن چهره نسبت به چهره میانگین اشاره می‌کنند. این رویکرد به صورت ریاضی‌وار فرموله شده است.

برای محک سیستم عملیاتی تشخیص چهره تحت وب از دو پایگاه داده بهره می‌جوییم. پایگاه داده اول، توسط خود سیستم عملیاتی تشخیص چهره تحت وب تهیه شده است و برای محک خود سیستم استفاده می‌شود و پایگاه داده دوم، پایگاه داده Extended Yale Face Database B می‌باشد. پایگاه داده تصاویر IMM Frontal Face را نیز برای آزمایش روش پیشنهادی استفاده می‌کنیم. نتایج آزمایشات به همراه بحث پیرامون آن در فصل پنجم ارائه شده است.

واژه‌های کلیدی: تشخیص چهره، وب، احراز هویت

مقالات مستخرج از پایان‌نامه

- Morteza Zahedi and Iman Firouzian. (2013 Sep.) “Pain Facial Expression Recognition among 8 Facial Expressions using Similarity Measures”, MVIP, Zanjan.
- Morteza Zahedi, Iman Firouzian, “Face Recognition by the Most Discriminant Features”, submitted to Turkish Journal of Electrical Engineering & Computer Sciences.
- مرتضی زاهدی، ایمان فیروزیان، (۱۳۹۲)، " سامانه هدف گیری خودکار سلاح‌های انفرادی با استفاده از تکنیک‌های چهره یابی"، کنفرانس نبردهای هواپایه ،پذیرفته شده برای ارائه ، تهران

فهرست مطالب

فصل اول: مقدمه ۱

۱-۱. مقدمه ۲

۱-۲. احراز هویت ۲

۱-۳. تعریف مسئله ۵

۱-۴. حوزه ۷

۱-۵. هدف ۸

۱-۶. سازمان دهی پایان‌نامه ۸

فصل دوم: مروری بر کارهای پیشین ۹

۲-۱. سیستم‌های بیومتریکی احراز هویت ۱۰

۲-۲. سیستم احراز هویت تحت وب از طریق تشخیص چهره ۱۲

۲-۳. پروژه‌های مشابه ۱۳

۲-۴. الگوریتم‌های موجود در سیستم تشخیص چهره ۱۵

۲-۴-۱. پیش پردازش تصویر ۱۵

۲-۴-۲. چهره یابی ۱۶

۲-۴-۳. همترازی چهره ۱۹

۲-۴-۴. تشخیص چهره ۲۲

۲-۵. تکنیک‌های تشخیص چهره ۲۳

۲-۶. تکنیک‌های استخراج ویژگی چهره ۲۷

۲-۷. نتیجه گیری ۲۸

فصل سوم: سیستم تشخیص چهره عملیاتی ۳۰

۳-۱. سیستم عملیاتی تشخیص چهره ۳۱

۳-۱-۱. مرحله آزمایش ۳۲

۳-۱-۲. مرحله آموزش ۳۴

۳-۲. چهره یابی با الگوریتم Viola-Jones ۳۸
۳-۳. تشخیص چهره با الگوهای دودویی محلی ٤۲
۳-۴. نتیجه گیری ٤۸
فصل چهارم: روش پیشنهادی ٤۹
۴-۱. مقدمه ۵۰
۴-۲. روش پیشنهادی ۵۳
۲-۳. فرایند آزمایش ۵۸
۲-۴. فرایند آموزش ٦٤
۴-۵. ویژگی‌های هندسی کمکی ٦۵
۴-۶. ویژگی‌های کمکی دیگر ٦۸
۴-۷. نتیجه گیری ٦۸
فصل پنجم: نتایج و بحث ۷۰
۵-۱. پایگاه داده‌های مورد استفاده ۷۱
۵-۱-۱. پایگاه داده شماره اول: پایگاه داده تصاویر سیستم تشخیص چهره عملیاتی ۷۱
۵-۱-۲. پایگاه داده شماره دو: پایگاه داده Extended Yale Face Database B ۷۲
۵-۱-۳. پایگاه داده شماره سه: پایگاه داده تصاویر چهره IMM Frontal Face ۷۳
۵-۲. تعاریف مورد استفاده در نتایج ۷٤
۵-۲-۱. اعتبارسنجی ضربدری n-تایی ۷۵
۵-۲-۲. معیارهای ارزیابی ۷٦
۵-۳. نتایج آزمایشات پایگاه داده اول بر روی سیستم عملیاتی ۷۸
۵-۳-۱. بحث پیرامون نتایج آزمایشات پایگاه داده اول ۸۰
۵-۴. نتایج آزمایشات پایگاه داده شماره دوم ۸۱
۵-۴-۱. اعتبار سنجی ضربدری ۶-تایی ۸۱
۵-۴-۲. تأثیر تعداد اشخاص بر عملکرد سیستم ۸۲
۵-۴-۳. تأثیر تصادفی کردن داده‌های مجموعه آموزش و آزمایش ۸٤
۵-۴-۴. انتخاب مقدار مناسب برای فاصله آستانه ۸۵
۵-۴-۵. تأثیر تعداد تصاویر آموزشی در عملکرد سیستم ۸۷

۵-۴-۶. بررسی عملکرد سیستم در یک حالت خاص ۸۸
۵-۵. نتایج آزمایشات پایگاه داده شماره سوم ۹۰
فصل ششم: نتیجه گیری و کارهای آینده ۹۴
پیوست ۹۹
فهرست مراجع و منابع ۱۰۵

فهرست اشکال

شکل ۱-۱: جنبه‌های نفوذپذیری ۳
شکل ۲-۱: سیستم احراز هویت تحت وب از طریق تشخیص چهره ۱۲
شکل ۲-۲: چهره یابی در تصاویر وب کم ۱۷
شکل ۲-۳: مرحله همترازی پس از چهره یابی و پیش از تشخیص چهره ۲۰
شکل ۲-۴: نتیجه متد Face Alignment بر روی مجموعه‌ای از تصاویر. ۲۱
شکل ۳-۱: معماری سیستم تشخیص چهره عملیاتی ۳۱
شکل ۳-۲: نمونه‌ای از اپلت تشخیص کاربر در مرورگر وب برای دریافت تصویر چهره کاربر و ارسال آن به سرور جاوا برای مرحله تشخیص چهره. ۳۳
شکل ۳-۳: نمونه‌ای از اپلت تشخیص کاربر در مرورگر وب در زمانی که چهره یابی انجام شده است ولی چهره کاربر تشخیص داده نشده است. ۳۴
شکل ۳-۴: نمونه‌ای از اپلت افزودن کاربر. در این اپلت نام شخص به همراه ۱۵ تصویر چهره از او برای پایگاه داده ارسال می‌شود. ۳۵
شکل ۳-۵: نمونه‌ای از اپلت افزودن کاربر. این تصویر، اپلت را در زمانی نشان می‌دهد که ۷ تصویر چهره تا به حال دریافت شده است و ۸ تصویر چهره دیگر باید دریافت شود. ۳۶
شکل ۳-۶: نمونه‌ای از اپلت افزودن کاربر. این تصویر، اپلت را در زمانی نشان می‌دهد که ۷ تصویر چهره تا به حال دریافت شده است و ۸ تصویر چهره دیگر باید دریافت شود. ۳۷
شکل ۳-۷: چند نمونه از موجک‌های Haar ۳۹
شکل ۳-۸: حقه تصویر انتگرالی. الف) پس از انتگرال گیری، پیکسل (x,y) حاوی مجموع تمامی پیکسل‌های خاکستری رنگ می‌باشد. ب) مجموع مقادیر پیکسل‌های ناحیه D از فرمول (x1,y1) + (x3,y3) – (x2, y2) – (x4, y4) بدست می‌آید.... ۴۰
شکل ۳-۹: آبشار دسته بندها به عنوان زنجیره فیلترها می‌باشد. زیرناحیه‌هایی از تصویر که از تمام فیلترها عبور کنند به "چهره" دسته بندی می‌شوند. ۴۱
شکل ۳-۱۰: دو ویژگی Haar از زنجیره فیلترها ۴۲
شکل ۳-۱۱: عملگر LBP اصلی ۴۳

شکل ۳-۱۲: نمونه‌هایی از طرح‌های اولیه بافت که با الگوهای دودویی محلی تشخیص داده می‌شوند. پیکسل‌های سفید و سیاه به ترتیب پیکسل‌های سفید و سیاه را مشخص می‌کند. ٤٣
شکل ۳-۱۳: همسایگی دایره‌ای (8, 2). هرگاه نقطه بر روی مرکز یک پیکسل قرار نگیرد، مقادیر پیکسلها به صورت دوخطی درون یابی می‌شوند. ٤٤
شکل ۳-۱۴: الف) نمونه‌ای از تصویر چهره که به پنجره ۷×۷ تقسیم بندی شده است. ب) وزن‌هایی که برای معیار عدم شباهت X2 تعیین شده اند. مربعات مشکی وزن ۰، خاکستری تیره ۱، خاکستری روشن ۲ و سفید ۴ می‌باشد. ٤٦
شکل ۳-۱۵: تصویر چهره به مناطق کوچکی تقسیم می‌شود. هیستوگرام‌های LBP هر منطقه استخراج می‌شود و نهایتاً در یک بردار ویژگی به یکدیگر الحاق می‌شوند. ٤٧
شکل ۴-۱: تصاویر چهره میانگین آقایان و خانم‌های سه ملیت ٥٢
شکل ۴-۲: یک دهان حاشیه نویسی شده از پایگاه دادهIMM Frontal Face ٥٤
شکل ۴-۳: دهان با خطوط مستقیم، دهان شخصی (شخص شکل ۴-۲) از پایگاه داده IMM Frontal Face را مشخص می‌کند. دهان با خطوط شکسته دهان میانگین اشخاص پایگاه داده را مشخص می‌کند. دهان میانگین از میانگین گیری مؤلفه‌های نقاط متناظر تمام اشخاص پایگاه داده بدست می‌آید. ٥٤
شکل ۴-۴: هر شخص از پایگاه داده توسط انحراف ویژگی‌های آن از ویژگی چهره میانگین مربوطه توصیف می‌شود. در این شکل، شما می‌توانید نقاط مشخص کننده دهان شخص ۱ (شکل ۴-۲) را با دایره‌های توخالی و همچنین دهان میانگین پایگاه داده را با دایره‌های توپر مشاهده نمایید. ٥٥
شکل ۴-۵: انحرافات دهان تمام اشخاص از دهان میانگین پایگاه داده. ٥٥
شکل ۴-۶: چگالی تخمین زده شده برای مقادیرانحرافات دهان تمام اشخاص از دهان میانگین پایگاه داده. ٥٧
شکل ۴-۷: تابع چگالی معکوس شده نماینده سطح اهمیت مقادیر مختلف انحرافات خواهد بود. ٥٧
شکل ۴-۸: انحرافات ویژگی چشم راست با تابع تخمین چگالی. ٥٩
شکل ۴-۹: انحرافات ویژگی بینی با تابع تخمین چگالی. ٦٠
شکل ۴-۱۰: انحرافات ویژگی دهان با تابع تخمین چگالی ٦٠
شکل ۴-۱۱: مرز دهان نمونه شکلی را بوجود آورده است؛ مجموعه‌ای از ویژگی‌های هندسی را می‌توان به این شکل منتسب نمود. ٦٥
شکل ۴-۱۲: تابع تخمین چگالی برای ویژگی گریز از مرکز ابروهای راست تمام اشخاص. ٦٧
شکل ۵-۱: نمونه تصاویر یک شخص از اشخاص پایگاه داده اول برای محک سیستم تشخیص چهره عملیاتی ٧٢
شکل ۵-۲: یک نمونه تصویر از پایگاه داده IMM Frontal Face با ۷۳ نقطه کمکی حاشیه گذاری شده که مکان ابروها، چشم‌ها، بینی، دهان و فک را مشخص می‌کنند. ٧٣
شکل ۵-۳: نمودار روند تأثیر افزایش تعداد افراد درون پایگاه داده بر عملکرد سیستم..... ٨٣
شکل ۵-۴: تأثیر حد فاصله آستانه بر روی معیارهای True Positive و True Negative..... ٨٦
شکل ۵-۶: تأثیر افزایش تعداد تصاویر مجموعه آموزش بر روند عملکرد سیستم عملیاتی تشخیص چهره ٨٧
شکل ۶-۱:قطع‌های متفاوت تصویر چهره را مشاهده می‌کنید. تصویر چهره سمت راست خروجی عملیات چهره یابی می‌باشد. تصویر چهره وسط و سمت چپ قطع‌های دیگر تصویر چهره را نشان می‌دهد. ٩٧

شکل ۷-۱: تصاویر برگرفته از مقاله مستخرج از پایان‌نامه، تصویر سمت راست: تصویر چهره به همراه مکان ویژگی‌های چهره. تصویر سمت چپ: تصویر چهره به همراه نقاط ویژه که به کمک ویژگی‌های چهره بدست آمد. ۹۹
شکل ۷-۲: تصویر چپ و وسط: دو فریم متوالی ویدئو با نقاط ویژه مشخص شده. راست: بردارهای جابجایی نقاط ویژه. ۱۰۰
شکل ۷-۳: جابجایی یک نقطه ویژه را می‌توان با جهت جابجایی (Θ) و مقدار مطلق (r) جابجایی محاسبه نمود. ۱۰۰
شکل ۷-۴: بردار مقادیر تتا مربوط به یکی از نقاط ویژه در فریم‌های ویدئو. یک شخص به صورت خط چین و شخص دیگر با خط راست مشخص شده است. ۱۰۱

فهرست جداول

جدول ۲-۱: مقایسه روش‌های بیومتریکی از نظر چهار عامل ۱۱
جدول ۴-۱: امتیازات تمام اشخاص با صفر مقداردهی می‌شود. ۵۸
جدول ۴-۲: دسته بند 1-NN برای ویژگی‌های مستخرج از چهره ۶۱
جدول ۴-۳: دسته بند 1-NN بهبودیافته برای ویژگی‌های مستخرج از چهره ۶۲
جدول ۴-۴: دسته بند پیشنهادی برای ویژگی‌های مستخرج از چهره ۶۳
جدول ۵-۱: نمونه‌ای از یک جدول درهم ریختگی تشخیص تصاویر چهره ۷۶
جدول ۵-۲:معیارهای ارزیابی ۷۷
جدول ۵-۳: جدول درهم ریختگی نتایج اعمال پایگاه داده شماره یک بر روی سیستم تشخیص چهره عملیاتی ۷۹
جدول ۵-۴: نتایج اعمال پایگاه داده دوم بر روی سیستم عملیاتی با اعتبارسنجی ضربدری ۶-تایی ۸۲
جدول ۵-۵: تأثیر تصادفی کردن داده‌های پایگاه داده دوم در نرخ تشخیص سیستم عملیاتی تشخیص چهره ۸۴
جدول ۵-۶: جدول درهم ریختگی نتایج یک حالت خاص برای سیستم عملیاتی تشخیص چهره ۸۹
جدول ۵-۷: نتیجه معیارهای ارزیابی اولیه برای سیستم تشخیص چهره عملیاتی ۹۰
جدول ۵-۸: نتیجه معیارهای ارزیابی نهایی برای سیستم تشخیص چهره عملیاتی ۹۰
جدول ۵-۹: نتایج روش پیشنهادی با دسته بند پیشنهادی ۹۲

فصل اول

مقدمه

۱-۱. مقدمه

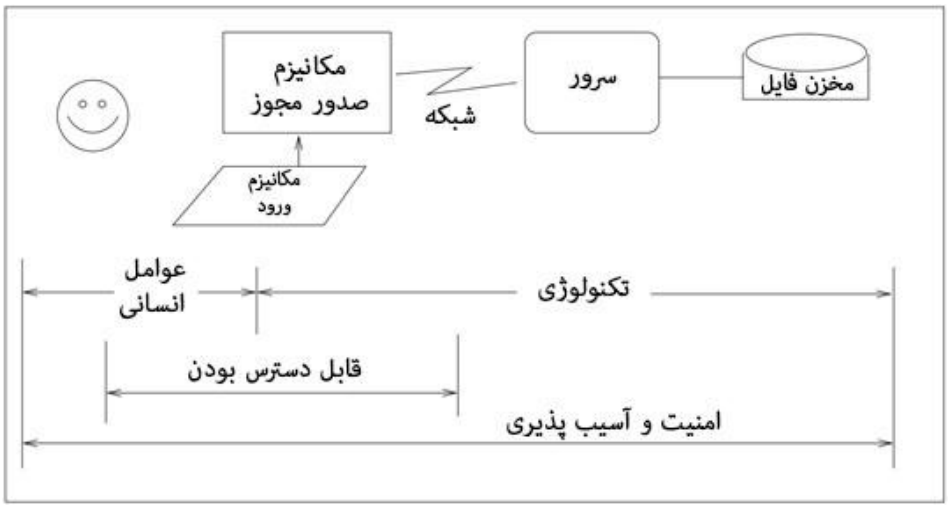
امنیت دیجیتال از نظر تئوری قابل دستیابی است اما در عمل، عامل انسانی نقش بزرگی در تخریب برنامه ریزی‌های کارشناسان امنیت دارد. در امنیت سیستم‌های کامپیوتری، معمولاً از کاربر به عنوان "لینک ضعیف" یاد می‌شود. همانطور که Bruce Schneier [1] اشاره کرده است، مردم عموماً خطرات موجود را نمی‌دانند و گاهی از عملکرد کامپیوترها بی اطلاع هستند. اینترنت میلیون‌ها شخص ناآشنا به کامپیوتر را در معرض خطرات بزرگی قرار داده است که آنها کاملاً در مورد آن بی اطلاع هستند. ویروس‌های اینترنتی از طرق مختلف مانند ایمیل به دلیل عدم آگاهی در مورد امنیت گسترش می‌یابند. ناآگاهی کاربران نسبت به مسائل امنیتی با عدم تطابق اهداف، بیشتر می‌شود. مکانیزم احراز هویت تضمین می‌کند که هر کاربر نتواند از هویت کاربر دیگری استفاده کند، در حالی که هدف کاربر مجاز دسترسی به سیستم نرم افزاری تحت وب برای انجام عمل خاصی است. متأسفانه، هرچقدر مکانیزم احراز هویت قدرتمندتر باشد، سیستم زمان بر خواهد بود و ورود کاربر مجاز به سیستم دشوارتر خواهد شد. بررسی Friedman و همکاران [2] نشان می‌دهد که کاربران بسیار کمی نگران هویت خود در تعاملات دنیای مجازی هستند. نکته قابل توجه اینست که در میان کاربران آگاه به تکنولوژی نیز این نگرانی کم است و در بررسی دیگر نشان داده شده است که تنها ٪۱۴ از این دسته کاربران، نگران هویت اینترنتی خود هستند. بنابراین گستردگی استفاده از دنیای آنلاین، احتمالاً طبیعت اعتماد در وب را تغییر نمی‌دهد. بنابراین اهمیت طراحی و ساخت سیستم مکانیزم امنیتی نزدیک به سطح کاربر بسیار زیاد می‌باشد.

۱-۲. احراز هویت

هدف سیستم‌های امنیتی آنست که به کاربران مجاز، اجازه ورود بدهد (مسئله صدور مجوز) و از ورود کاربران غیرمجاز به سیستم جلوگیری کند (مسئله جلوگیری). علاوه بر این، این سیستم‌ها باید تضمین

کنند که کاربران عملیات‌های مجاز خاص خود را انجام می‌دهند. برای ارضاء این هدف، سیستم باید از کاربر درخواست کند تا خود را معرفی کند. هنگامی که معرفی صورت گرفت، شخص باید هویت خود را اثبات کند که این عمل توسط سیستم احراز هویت انجام می‌پذیرد. هنگامی که احرازهویت کاربر توسط سیستم امنیتی صورت گرفت، کاربران تنها بر حسب سطح دسترسی از امکانات وبسایت استفاده می‌کنند.

در سیستم‌های امنیتی دنیای واقعی، شناسایی، احراز هویت، صدور مجوز دسترسی معمولاً ادغام می‌شود. برای نمونه، یک نگهبان اداره که ورود افراد به ساختمان را کنترل می‌کند، کارمندان آن اداره را می‌شناسد (شناسایی و احراز هویت) و در صورت شناسایی چهره افراد وارد شونده، به آنها اجازه ورود می‌دهد (صدور مجوز ورود). در دنیای دیجیتال این سه گام باید متمایز باشند تا امنیت هر گام تضمین شود. این پایان‌نامه، گام احراز هویت را که پرچالش‌ترین گام می‌باشد، بررسی می‌کند. Schneier [1] بیان می‌کند که احراز هویت نیاز به شناسایی یک شخص تصادفی به عنوان شخص X ندارد اما باید ثابت کند که شخص X همان شخصی است که خود ادعا کرده است.



شکل ۱-۱: جنبه‌های نفوذپذیری

شکل ۱-۱ فرایند احراز هویت را نشان می‌دهد و مهم‌ترین زمینه‌های نفوذپذیری به سیستم را مشخص می‌کند. از جنبه امنیت، چندین نقطه ضعف وجود دارد که تنها دو مورد اول آن به استفاده از مکانیزم احراز هویت مربوط می‌شود:

- کاربر آسیب پذیر: حمله تحقیقاتی، حمله مهندسی تحقیقاتی (همانند phishing) یا ذخیره ناامن کلیدهای احراز هویت.
- ورودی ها: آسیب پذیر در مقابل shoulder surfing، keyboard tapping، شنود و اسب‌های تروجان.
- انتقال: حساس به sniffers و cracking.
- سرور: ذخیره کلید احراز هویت در برابر فرهنگ لغات و حملات brute force آسیب پذیر می‌باشد.

کاربران سیستم‌های دیجیتالی توسط توکن یا رشته متنی همانند آدرس ایمیل یا شماره حساب خود را به سیستم معرفی می‌کنند. احراز هویت آنها توسط آنچه می‌دانند، آنچه دارند یا توسط آنچه هستند، صورت می‌پذیرد؛ هرکدام از گزینه‌ها مسائل خاص خود را دارد. احراز هویت توسط آنچه کاربر می‌داند، مخصوصاً در فضای مجازی به صورت بسیار گسترده صورت می‌پذیرد و مسائل مهمی مرتبط با اجازه و جلوگیری مطرح می‌شود. هنگامی که کاربران کلید ورود را فراموش می‌کنند، اجازه دسترسی به آنها داده نمی‌شود و شخصی که دسترسی به کلید عبور پیدا می‌کند، می‌تواند در محیط کنترل نشده به راحتی از هویت کاربر دیگری برای ورود استفاده نماید. برخی اوقات کاربر کلیدهای عبور سیستم‌های مختلف را که هر از چندگاهی تغییر می‌دهد، یادداشت می‌کند. کاربر این کار را بدلیل دفاع از خود و جهت جلوگیری از ازدست دادن کلیدهای عبور انجام می‌دهد. چنین حفره‌هایی که از طرف عامل انسانی در امنیت بوجود می‌آید، توسط هیچ مکانیزم احراز هویتی جبران نمی‌شود. به همین خاطر، از عامل انسانی در امنیت اطلاعات معمولاً به عنوان "لینک ضعیف" یاد می‌شود.

سیستم‌های بیومتریکی (همانند اثر انگشت، شبکیه چشم و ...) مشکل "لینک ضعیف" را در امنیت اطلاعات برطرف می‌کنند اما از جهتی دیگر، این سیستم‌ها نیاز به سخت افزار خاصی دارند که همه جا قابل دسترس نیست. سیستم بیومتریکی تشخیص چهره از نظر سخت افزاری نیاز به یک وب‌کم دارد که این سخت افزار در بازار کاملاً موجود است و همچنین در اکثر لپ تاپ‌ها از پیش تعبیه شده است. بنابراین، سیستم احراز هویت تحت وب از طریق سیستم بیومتریکی تشخیص چهره راه حل پیشنهادی خوبی در این زمینه محسوب می‌شود.

۱-۳. تعریف مسئله

در سال‌های اخیر، برنامه‌های تحت وب و خدمات اینترنتی همانند سرورهای پست الکترونیک، شبکه‌های اجتماعی، سیستم‌های آموزش الکترونیک و غیره به طور گسترده در سراسر جهان استفاده می‌شوند. احراز هویت در این سیستم‌ها معمولاً از طریق نام کاربری و کلمه عبور انجام می‌شود؛ از طرفی دیگر کلمه عبور ملاک خوبی برای احراز هویت کاربر نمی‌باشد و این تهدیدی برای حریم اطلاعاتی محسوب می‌شود. در سیستم آموزش الکترونیک نیز، هنوز کمبود بسیاری در مکانیزم‌های احراز هویت کاربران، اعم از سیستم ورود کاربران و همچنین احراز هویت در طول جلسه وجود دارد. از این رو، در بهبود امنیت و حریم اطلاعاتی کاربران خاص تحت وب (همانند کاربر ارشد سیستم) از تشخیص چهره برای احراز هویت استفاده می‌کنیم. پیشرفت‌های اخیر در تکنولوژی بیومتریک منجر به استفاده از سیستم‌های اعتبارسنجی بر پایه بینایی ماشین شده است که امن تر هستند و استفاده راحت تری را دارد. با توجه به پیشرفت‌های حوزه تشخیص چهره، احراز هویت از طریق تشخیص چهره در حال تبدیل شدن به یک راه حل قابل اعتماد می‌باشد و از آنجا که برنامه‌های کاربردی تحت وب از راه دور مورد توجه بیشتری قرار می‌گیرند، نیاز به چنین سیستمی افزایش می‌یابد.

سیستم تشخیص چهره یکی از پردازش‌های اطلاعاتی بیومتریک می‌باشد؛ مزایای این سیستم در مقایسه با پردازش‌های اطلاعاتی بیومتریکی دیگر (همانند اثر انگشت، اسکن شبکیه چشم، امضا و ...) قابلیت استفاده آسان و گسترده بودن دامنه کاربرد آن می‌باشد. هدف این پایان‌نامه تحلیل، طراحی، پیاده سازی و تست یک سیستم احراز هویت تحت وب از طریق تشخیص چهره می‌باشد. در این سیستم، کاربران تنها با داشتن یک وب‌کم (که امروزه در بسیاری از لپ تاپ‌ها تعبیه شده) توسط ویژگی‌های چهره احراز هویت خواهند شد. این سیستم را می‌توان به بخش‌های زیر تقسیم نمود:

(۱) دریافت فریم از وب‌کم

(۲) پیش پردازش

(۳) چهره یابی

(۴) هم ترازی چهره

(۵) تشخیص چهره

(۶) شناسایی / اعتبارسنجی چهره

در ابتدا وب‌کم فریم‌های متوالی را تهیه می‌کند اما از آنجا که پردازش‌های سیستم مفروض بر سرور انجام می‌شود و وب‌کم فریم‌های متوالی را دریافت می‌کند، برای کاهش حجم انتقال اطلاعات از کلاینت به سرور باید فریم‌های خاصی را برای پردازش از سمت کلاینت به سرور انتخاب و ارسال نماییم.

در بخش پیش پردازش، با بهبود کیفیت تصویر از طریق اعمال فیلترهای مناسب و عملیات‌های حذف نویز مربوطه می‌توان تاثیر بسزایی در نتیجه نهایی سیستم تشخیص چهره گذاشت.

در چهره یابی دو گام اصلی وجود دارد: الف) آیا چهره انسان در تصویر ظاهر شده است یا خیر؟ ب) چهره‌ها در چه موقعیت‌هایی قرار دارند؟

پس از چهره یابی، تصویر چهره جداسازی می‌شود. جهت بهبود عملکرد سیستم تشخیص چهره، هم ترازی چهره[1] انجام می‌شود تا مقیاس و چرخش تصویر جداسازی شده از چهره را تنظیم نماییم.

یک پایگاه داده برای ذخیره سازی ویژگی‌های تصاویر اشخاص فراهم می‌شود. از هر شخص، چندین تصویر گرفته می‌شود و ویژگی‌های آنها استخراج می‌شود و در پایگاه داده ذخیره می‌شود. سپس، هنگامی که تصویر چهره از ورودی دریافت می‌شود، عملیات چهره یابی و استخراج ویژگی انجام می‌شود و نهایتاً ویژگی‌های آن با هر کلاس چهره موجود در پایگاه داده مقایسه می‌شود. در بحث پیاده سازی از الگوریتم هیستوگرام‌های الگوهای دودویی محلی استفاده می‌شود.

سیستم تشخیص چهره دو کاربرد عمده دارد که عبارتند از شناسایی چهره و اعتبارسنجی چهره می‌باشد. در شناسایی چهره، با مفروض بودن تصویر چهره سیستم باید مشخص کند تصویر چهره مربوط به چه کسی است یا به چه کسی نزدیک تر است اما در اعتبارسنجی چهره، با مفروض بودن تصویر چهره و حدسی از شناسایی چهره، سیستم باید درست یا نادرست بودن حدس را مشخص کند.

۱-۴. حوزه

در این پایان‌نامه، سیستم احراز هویت کاربران به کمک تشخیص بیومتریکی چهره انجام می‌شود. این سیستم در قالب سیستم تحت وب ورود کاربران، هویت کاربران مجاز را از طریق ویژگی‌های چهره تشخیص می‌دهد. تشخیص چهره زمان-واقعی در تصاویر گرفته شده از وب‌کم در شرایط دنیای واقعی اعم از شرایط نوری مختلف انجام می‌گیرد، بنابراین پیش پردازش‌های مناسب و الگوریتم‌های هیستوگرام‌های الگوهای دودویی محلی مورد استفاده قرار می‌گیرد، گرچه طراحی سیستم تشخیص چهره با قابلیت اعتماد بسیار بالا در محیط‌های دنیای واقعی بسیار مشکل می‌باشد. بنابراین، در سیستم ارائه شده تعداد اشخاص

[1] Face Alignment

موجود در پایگاه داده سیستم محدود در نظر گرفته می‌شود و از این سیستم می‌توان در موارد خاصی با تعداد کاربر مشخص استفاده نمود؛ برای مثال کاربران ارشد سیستم.

۱-۵. هدف

اهداف این پایان‌نامه به شرح ذیل است:

i. طراحی و پیاده سازی سیستم احراز هویت کاربران ارشد یک وبسایت از طریق تشخیص چهره.

ii. ارائه بهبودی بر عملکرد یا ارائه یک رویکرد نوین در حوزه تشخیص چهره.

۱-۶. سازمان دهی پایان‌نامه

این پایان‌نامه متشکل از شش فصل می‌باشد. فصل اول خواننده را با مفهوم سیستم احراز هویت و لزوم استفاده از تشخیص بیومتریکی چهره آشنا می‌کند. فصل دوم مروری بر کارهای انجام شده در حوزه سیستم احراز هویت، پیش پردازش تصویر، چهره یابی، هم ترازی چهره و تشخیص چهره را ارائه می‌دهد. در فصل سوم، عملکرد سیستم عملیاتی تشخیص چهره تحت وب تشریح می‌شود. روش پیشنهادی و رویکرد نوین در تشخیص چهره در فصل چهارم ارائه شده است. در فصل پنجم، نتایج اعمال دو پایگاه داده بر روی سیستم عملیاتی تشخیص چهره ارائه شد و همچنین از پایگاه داده دیگری برای محک زدن روش پیشنهادی ارائه شده است. در فصل ششم ، نتیجه گیری و کارهای آینده پایان‌نامه آمده است؛ در این فصل به مشکلات و بهبودهای بیشتر سیستم نیز اشاره شده است.

فصل دوم

مروری بر کارهای پیشین

۲-۱. سیستم‌های بیومتریکی احراز هویت

در سیستم‌های بیومتریکی خود شخص کلید محسوب می‌شود، بنابراین نیاز به بخاطرسپاری کلید یا کارت نمی‌باشد. جسم فیزیکی هر شخص منحصر به فرد است و سیستم بیومتریکی از هویت منحصر به فرد شخص جهت فعالسازی خرید یا ورود به سیستم استفاده می‌کند. سیستم بیومتریکی شامل صدا، رگ، چشم، اثرانگشت، تشخیص چهره و غیره می‌باشد.

- مزایای سیستم بیومتریکی
  - افزایش امنیت؛ با افزودن یک لایه امنیتی راحت و کم هزینه.
  - کاهش کلاهبرداری دیجیتال؛ با بکارگیری تکنولوژی‌های قدرتمند در مقابل جعل هویت.
  - حذف مسائل مرتبط با شناسه‌های گمشده یا کلمه عبورهای فراموش شده با استفاده از خواص فیزیولوژیکی (برای نمونه، جلوگیری از استفاده غیرمجاز از شناسه‌های دزیده شده).
  - کاهش هزینه‌های مدیریت کلمه عبور.
  - جایگزینی برای کلمه عبورهای پیچیده می‌باشد که ممکن است به اشتراک گذاشته شود یا مشاهده شود.
  - ادغام گستره بزرگی از راه حل‌های بیومتریکی و تکنولوژیکی، برنامه‌های کاربردهای و پایگاه داده‌ها به یک راه حل کنترلی قدرتمند برای دسترسی به شبکه و امکانات آن.
  - فراهم نمودن اطلاعاتی همانند اینکه چه کسی، چه کاری را در چه زمانی انجام داده است.
  - صرفه جویی بسیار در هزینه یا افزایش ROI در زمینه‌هایی همانند سیستم حضور.
  - به طور صریح پیوندی را میان یک فرد یا اتفاق ایجاد می‌کند.

- معایب سیستم بیومتریکی

  - تاثیر بیماری‌ها بر خواص بیومتریکی فردی و در نتیجه کاهش دقت عملکرد سیستم.
  - تاثیر سن بر صدا و چهره و دیگر خواص بیومتریکی اشخاص و در نتیجه کاهش دقت عملکرد سیستم.
  - تاثیر نویز محیطی در ارسال سیگنال خواص بیومتریکی.
  - زمان اعتبارسنجی نسبتاً زیاد در برخی سیستم‌های بیومتریکی.
  - تاثیر تمیزی و کثیفی شخص و تاثیر آن به صورت نویز بر روی سیستم.

در ذیل برخی از روش‌های بیومتریکی احراز هویت را در قالب جدولی از نظر چهار عامل بررسی می‌کنیم.

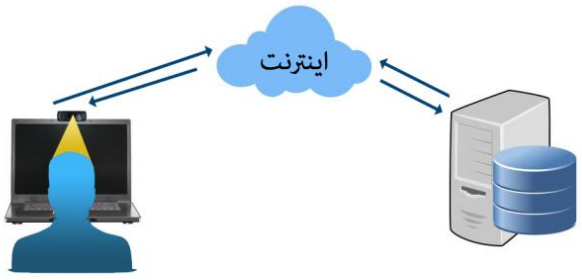
جدول ۲-۱: مقایسه روش‌های بیومتریکی از نظر چهار عامل

تکنولوژی بیومتریکی	دقت	هزینه	دستگاه‌های مورد نیاز	مقبولیت جامعه
DNA	زیاد	زیاد	تجهیزات آزمایشی	کم
تشخیص عنبیه	زیاد	زیاد	دوربین	متوسط به پایین
تشخیص صدا	متوسط	متوسط	میکروفون	زیاد
شکل هندسی دست	متوسط به پایین	متوسط به پایین	اسکنر	کم
اثر انگشت	زیاد	متوسط	اسکنر	متوسط
تشخیص امضا	کم	متوسط	قلم نوری، پانل لمسی	زیاد
تشخیص چهره	متوسط	کم	وب‌کم	زیاد

۲-۲. سیستم احراز هویت تحت وب از طریق تشخیص چهره

در بخش قبل، چند سیستم بیومتریکی از لحاظ دقت، هزینه، دستگاه‌های مورد نیاز و مقبولیت جامعه بررسی گردید. سیستم بیومتریکی مناسب، سیستمی است که تا حد امکان دقت بالایی داشته باشد و هزینه کمی داشته باشد و تجهیزات سخت افزاری مورد نیاز آن قابل دسترس باشد و در جامعه مقبولیت زیادی داشته باشند و از نظر مردم جامعه قابلیت اعتماد داشته باشد. سیستم تشخیص چهره تا حد قابل قبولی این شرایط را ارضاء می‌کند. سیستم احراز هویت از طریق تشخیص چهره یکی از انعطاف پذیرترین ، کاراترین روش‌های تشخیص بیومتریکی است. این سیستم از طریق یک دوربین وب‌کم ویژگی‌های خاصی از چهره را استخراج می‌کند تا عملیات شناسایی را بر اساس آن انجام دهد.

در اینترنت، پروتکل‌های امنیتی در سمت کلاینت هنوز در پی افزودن لایه‌های امنیتی بدون نصب سنسورهای سخت افزاری جدید برای ارتقاء ضریب اطمینان خاطر می‌باشد. برخی مؤسسات وسازمان‌ها همانند بانک‌ها که اطلاعات حساس و خصوصی را از اینترنت قابل دسترس کرده اند، امنیت را افزایش داده و لایه حفاظتی را در مقابل حملات خودکار افزودند. در سیستم‌های آموزش الکترونیکی (مجازی) نیز احراز هویت شخص دانش آموز/ دانشجو اهمیت بسیاری دارد. این نیازها باعث گرایش نهادها و سازمان‌ها به سمت سیستم‌های احراز هویت بیومتریکی تحت وب از طریق تشخیص چهره می‌شود.



شکل ۲-۱: سیستم احراز هویت تحت وب از طریق تشخیص چهره

۲-۳. پروژه‌های مشابه

از آنجا که این پایان‌نامه، یک پروژه کاربردی در بازار هوش مصنوعی می‌باشد، پروژه‌های مشابه نیز توسط شرکت‌های تجاری ارائه شده است. در پروژه‌های تجاری معمولاً الگوریتم‌های مورد استفاده در پروژه نام برده نمی‌شود تا توسط رقیب‌های تجاری مورد سوء استفاده قرار نگیرد. از طرفی دیگر، هیچ کدام از پروژه‌های مشابه، به صورت رایگان در اختیار مردم قرار نگرفته است. در ذیل سه پروژه مشابه را بررسی می‌کنیم.

❖ Luxand

Luxand Inc. یک شرکت خصوصی فناوری است؛ این شرکت در سال ۲۰۰۵ تاسیس شد. فعالیت‌های تحقیقاتی Luxand با استفاده از تکنیک‌های هوش مصنوعی و تکنولوژی‌های شناسایی بیومتریکی آغاز به کار کرد. امروزه، این شرکت گستره بزرگی از مسائل شناسایی ویژگی‌های چهره را در فضای آنلاین پوشش داده است. Luxand FaceSDK توسعه دهندگان برنامه‌های کاربردی تحت وب و برنامه‌های کاربردی تحت ویندوز را قادر می‌سازد تا سیستم خود را مجهز به سیستم تشخیص چهره و یا شناسایی بیومتریکی بر اساس تشخیص چهره نمایند. FaceSDK از طریق webcam عملیات شناسایی، اعتبارسنجی، احراز هویت چهره را انجام می‌دهد.

این پروژه با هزینه $۴۰۰۰ در اختیار برنامه نویسان و توسعه دهندگان وب و نرم افزار قرار می‌گیرد.

وبسایت Luxand ادعا کرده است که این سیستم ویژگی‌های زیر را دارد:

- شناسایی و تشخیص چهره سریع و جامع.
- تشخیص پایدار مستقل از شرایط نوری.

- شناسایی و تشخیص ویژگی‌های چهره.
- پشتیبانی از تمامی webcamهای موجود در بازار.
- پشتیبانی از تصاویر کوچک و بزرگ چند مگاپیکسلی.
- پشتیبانی از جریان‌های زمان-واقعی ویدئویی.

پرواضح است که ویژگی‌های فوق برای جلب مشتری برای خرید می‌باشد و درعمل، تمام این ویژگی‌ها به احتمال زیاد بسیار مورد نقد قرار خواهد گرفت و ضعف‌های بسیاری از آن آشکار خواهد شد.

❖ Face.com

Face.com یک شرکت تکنولوژی است و APIهای تشخیص چهره را برای استفاده عموم فراهم کرده است. این گروه نرم افزاری بزرگترین پلتفرم تشخیص چهره را بر روی cloud پیاده سازی کرده است، اما در سال ۲۰۱۲ شرکت Facebook با خرید امتیاز و مدیریت این گروه نرم افزاری به مبلغ ۵۵ میلیون دلار، تمام فعالیت‌های این گروه نرم افزاری را تحت نظر خود اداره می‌کند و نسخه‌های جدید آن دیگر برای عموم قابل دسترس نیست. الگوریتم‌های تشخیص چهره این شرکت نیز دقیقاً مشخص نیست اما در وبسایت شرکت ادعا شده است که سیستم تشخیص چهره از تکنیک‌های یادگیری ماشین نیز بهره می‌جوید.

❖ Lambda

آزمایشگاه‌های Lambda در بهار سال ۲۰۱۲ تاسیس شد. اولین پروژه این شرکت، API تشخیص چهره تحت وب می‌باشد. این شرکت اخیراً نسخه بتای پروژه تشخیص چهره را با هزینه ۱۲۸$ به فروش می‌رساند.

۲-۴. الگوریتم‌های موجود در سیستم تشخیص چهره

این پایان‌نامه یک سیستم عملیاتی تشخیص چهره تحت وب را طراحی و پیاده سازی می‌کند. این سیستم کامل از چندین بخش تشکیل می‌شود که به صورت گام‌های آبشاری پیاده سازی شده اند. در ادامه، مهم‌ترین الگوریتم‌ها و تکنیک‌های هر بخش را به صورت جداگانه ارائه می‌دهیم. این سیستم را می‌توان به بخش‌های پیش پردازش، چهره یابی، هم‌ترازی چهره، تشخیص چهره تقسیم بندی نمود. در ادامه در مورد هر بخش به صورت مبسوط توضیحاتی ارائه خواهد شد.

۲-۴-۱. پیش پردازش تصویر

تصاویر چهره از طریق وب‌کم دریافت می‌شود. نویز تصاویر وب‌کم اغلب از نوع نویز سفید می‌باشد که به راحتی رفع می‌شود. از طرف دیگر، مشکلی که معمولاً در تصاویر دریافتی از وب‌کم وجود دارد، تغییرات روشنایی تصویر می‌باشد. در ادامه راهکارهای مناسب برای مقابله با این مشکل مورد بحث قرار می‌گیرد.

مسئله تغییرات روشنایی اساساً مسئله تغییرات جزئی ظاهر شیء به علت تغییر شرایط نوری در یک تصویر نسبت به تصویر دیگر است. تغییرات نوری مسئله چالش برانگیزی در زمینه تشخیص چهره می‌باشد و یکی از مهم‌ترین عوامل تشخیص چهره قدرتمند می‌باشد. یک شخص با حالت چهره یکسان در شرایط نوری مختلف بسیار متفاوت به نظر می‌رسد. تغییرات شرایط نوری کاهش قابل توجهی را در عملکرد تشخیص درپی خواهد داشت. یک سیستم تشخیص که براساس فاصله میان تصاویر خاکستریِ پردازش نشده عمل می‌کند، در شناسایی چهره‌های پایگاه داده شکست خواهد خورد و چهره‌ها را اشتباه تشخیص خواهد داد.

رویکردهای این مسئله در سه دسته طبقه‌بندی می‌شوند:

- پیش‌پردازش: در این رویکرد، تصاویر چهره برای نرمال سازی روشنایی پیش پردازش می‌شوند؛ برای نمونه، تصحیح گاما، تبدیل لگاریتم، تعدیل هیستوگرام، تبدیل کسینوسی گسسته و آنالیز مؤلفه اصلی.
- استخراج ویژگی‌های مستقل: این رویکرد ویژگی‌های چهره که مستقل از تغییرات نوری می‌باشد را استخراج می‌کند. برای نمونه، نقشه‌های لبه، مشتقات سطح خاکستری، فیلترهایی همانند گابور و چهره فیشر.
- مدل کردن چهره: تغییرات نوری عمدتاً به دلیل سه بعدی بودن شکل چهره انسان و تحت جهت‌های مختلف نوری می‌باشد. برخی محققان سعی در ساخت مدل چهره سه بعدی و render گرفتن از تصویر چهره با جهت‌ها و شرایط نوری مختلف دارند.

۲-۴-۲. چهره یابی

پس از پیش پردازش، عملیات چهره یابی در سیستم تشخیص چهره انجام می‌شود. الگوریتم‌ها و رویکردهای بسیاری در این زمینه ارائه شده است. الگوریتم Viola-Jones [7] الگوریتم مناسبی برای یافتن چهره در عکس‌ها می‌باشد. چهره انسان در تصاویر ویدئویی می‌تواند جهت‌ها و موقعیت‌های نامحدودی داشته باشد که چالش‌های بسیاری پیش روی محققان بوجود می‌آورد.

شکل ۲-۲: چهره یابی در تصاویر وب‌کم

عملیات چهره یابی در تصاویر ویدئویی عموماً شامل سه رویکرد می‌شود. در رویکرد اول، چهره یابی بر پایه فریم انجام می‌شود. در این رویکرد، متدهای کلاسیک بسیاری را می‌توان معرفی نمود همانند متد مدلینگ آماری [3]، متد بر پایه شبکه عصبی [4]، متد برپایه SVM [5]، متد برپایه HMM [6]، متد BOOST [7] و چهره یابی رنگی [8] و غیره؛ گرچه، عدم استفاده از اطلاعات زمانی مهم‌ترین مشکل این رویکردها می‌باشد. در رویکرد دوم، پس از چهره یابی عملیات ردیابی چهره را نیز می‌توان انجام داد. چهره یابی و ردیابی چهره مستقل از هم می‌باشند، چهره یابی در فریم اول صورت می‌گیرد و سپس در توالی فریم‌های بعدی ردیابی می‌شود. عیب این روش عدم پشتیبانی از چند چهره و همچنین از دست دادن اجتناب ناپذیر اطلاعات می‌باشد. در رویکرد سوم از اطلاعات زمانی و روابط میان فریم‌ها برای تشخیص چند چهره استفاده می‌شود. معمولاً این روش‌ها از دو مرحله تشکیل شده است. مرحله اول با نام "چهره یابی و پیش بینی" و مرحله دوم با نام "بروزرسانی ردیابی" شناخته می‌شود.

۲-۴-۲-۱. روش‌های معمول چهره یابی

در سال ۲۰۰۰، Zhu Lion و Yao Wang [9] یک تطابق الگو با برنامه نویسی تکراری پویا برای چهره یابی از زاویه روبرو و ردیابی چهره‌های در زاویه‌های دیگر با مدل‌های تطبیقی چهره انجام داد. بر

اساس اصلی که بیان می‌کند، تراکم لبه‌ها در مجاورت ویژگی‌های چهره بیشتر می‌باشد و در دیگر مناطق تراکم کمتر است، Li Silva و همکاران [10] متدی با نام شمارش پیکسل‌های لبه ارائه داد که بر طبق آن می‌توان ویژگی‌های چهره را یافت و ردیابی نمود. Han و همکاران [11] تحقیقات و آزمایشاتی برای یافتن و ردیابی چند شیء نامشخص با تعداد مختلف با استفاده از ساختار گراف انجام دادند. در تحقیق [12]، مدل‌های ظاهر بر پایه دسته بندی مناسب در بخش‌های مختلف ویدئویی به صورت اتوماتیک دسته بندی می‌شوند. علاوه براین، برخی روش‌ها ویژگی‌های جهت لبه را برای بهبود کارایی چهره یابی ترکیب می‌کنند[13]. برای استفاده کامل از اطلاعات زمانی توسط ویدئو، [14] یک متد چهره یابی را ارائه دادند که از هیستوگرام‌های محلی ضرایب موجک استفاده می‌کند و آنها را نسبت به یک قاب محوری به نمایش می‌گذارد. Zhenqiu Zhang و همکاران[15] چهره یابی بر پایه Floatboost را برای تصمیم گیری محلی ارائه دادند و از اطلاعات زمانی برای تایید و اعتبارسنجی نتایج استفاده کردند.

۲-۴-۲-۲. روشهای زمان-واقعی و چندگانه چهره یابی

چهره یابی زمان-واقعی و چندگانه در چهره یابی بر پایه ویدئو اهمیت بسیاری دارد. متدهای چهره یابی زمان-واقعی به دو گروه دسته بندی می‌شوند: اولین گروه شامل رویکرد AdaBoost آبشاری است که توسط Viola و Jones [7] پیشنهاد داده شده است و دومین گروه از اطلاعات رنگی برای چهره یابی و اعتبارسنجی چهره انسان استفاده می‌کند[16]. یک تکنیک آماری غیرپارامتریک توسط Bradski و همکاران برای چهره یابی در فضای سه بعدی استفاده می‌شود[17].

Schneiderman و Kanade [18] ادعا کرده اند که برای اولین بار چهره یابی چندگانه را ارائه نموده اند. در طول سال‌های گذشته، پیشرفت بسیاری در این زمینه انجام شده است. دو متد اصلی در

این زمینه وجود دارد: (الف) متد، یک چهره یاب را برای زوایای مختلف چهره طراحی می‌کند؛ (ب) متد، چندین چهره یاب را برای زوایای مختلف صورت طراحی می‌کند.

Feraud و همکاران [19] آرایه‌ای از پنج چهره یاب را برای پنج زاویه دید مختلف طراحی کرده اند. در صورت وجود یک مکانیزم مفید برای تشخیص زاویه چهره، چهره یاب مناسب بر پایه فضای ویژگی PCA خطی انتخاب می‌شود[20]. علاوه براین، معماری چهره یابی هرمی در [21] ارائه شده است که از استراتژی ادغامی از coarse-to-fine و دسته بندی ساده به پیشرفته استفاده می‌کند[22]. برای دستیابی به حداقل خطا، Li and Zhang [23] الگوریتمی را برای ادغام Adaboost آبشاری و آرایه‌ای از چهره یاب‌ها ارائه دادند.

مسئله اصلی این روش‌ها آنست که تغییرات درون کلاسی مجموعه داده چهره از زوایای مختلف، بزرگتر از مجموعه داده چهره از زاویه جلو می‌باشد. گرچه AdaBoost با معماری هرمی این مشکل را حل می‌کند اما پیچیدگی محاسباتی آن بار محاسباتی بسیاری ایجاد می‌کند.

۲-۴-۳. هم‌ترازی چهره

تحقیقات و آزمایشات در زمینه تشخیص چهره بیش از سی سال پیش آغاز گردیده است و متدهای چهره یابی و تشخیص چهره فراوانی ارائه شده است اما در پیاده سازی یک سیستم عملیاتی تشخیص چهره خروجی متد چهره یابی با ورودی متد تشخیص چهره تطابق ندارد و باعث کاهش دقت عملکرد سیستم تشخیص چهره خواهد شد، بنابراین ضرورت وجود گامی برای انجام پردازشهایی در قالب هم‌ترازی چهره مشهود است. هم‌ترازی چهره به دلیل اهمیت آن در دقت عملکرد سیستم اخیراً توجه بسیاری از محققان را به خود جلب نموده است.

بسیاری از الگوریتم‌های تشخیص چهره وابسته به قرار گرفتن دقیق چهره در یک حالت پایه‌ای می‌باشد. بنابراین، موقعیت ویژگیهای چهره نسبت به سیستم مختصات ثابت را باید محاسبه نمود. در حال حاضر، موقعیت‌یابی ویژگی‌های چهره به روش دستی و یا با آموزش یک الگوریتم یادگیری به کمک نمونه‌های کافی که در آنها جهت چهره و و اعضای چهره برچسب گذاری شده اند، صورت می‌گیرد. این نوع ثبت تصویر[1] باعث کاهش واریانس در متد دسته بندی و تشخیص می‌شود. بنابراین متد دسته بندی و تشخیص می‌تواند از موقعیت مکانی حاصل از هم ترازی چهره، برای تحلیل جزئی ساختار چهره مطلوب استفاده نماید. به طور کلی، هرچقدر هم ترازی چهره بهتر انجام شود، متد تشخیص نیز نتایج بهتری به دست خواهد آورد.

لازم به ذکر است که متدهای چهره‌یابی هم‌ترازی چهره را نیز انجام می‌دهند اما هم‌ترازی حاصل از چهره یابی، هم ترازی تقریبی است، که عملکرد سیستم تشخیص چهره را کاهش می‌دهد. برای افزایش دقت عملکرد سیستم نیاز به هم‌ترازی دقیق تری می‌باشد.

شکل ۲-۳: مرحله همترازی پس از چهره یابی و پیش از تشخیص چهره

همانطور که در شکل ۲-۳ مشاهده می‌نمایید، یک مرحله هم ترازی چهره به صورت مستقل برای هم ترازی دقیق چهره باید انجام شود. در اکثر سیستم‌هایی که گام هم ترازی چهره در آن اجرا می‌شود، برخی نقاط ویژگی چهره (همانند میان چشم‌ها یا گوشه‌های دهان)، یا توصیفی از جهت چهره (همانند

[1] Image registration

اطلاعات راستای چهره و موقعیت مکانی) در زمان آموزش سیستم باید به صورت دستی برچسب گذاری شود. برخی متدهای هم ترازی چهره، موقعیت، اندازه و زاویه چرخش را نیز می‌تواند به عنوان خروجی ارائه نماید. با استفاده از زاویه چرخش و اندازه و موقعیت چهره می‌توان ژست پایه‌ای چهره را تشخیص داد. هر چقدر پارامترهای زاویه چرخش و اندازه و موقعیت دقیق تر محاسبه شوند، محاسبه ژست چهره نیز دقیق تر خواهد بود.

متدهای اتوماتیکی نیز برای عملیات هم ترازی چهره نیز ارائه شده است. متد ارائه شده توسط Hung و همکاران [24] تصاویر تقریباً هم تراز شده (همانند خروجی چهره یاب Viola and Jones [7]) را دریافت می‌کند و تصاویر دقیق تر هم ترازشده را به عنوان خروجی برمی‌گرداند. در ذیل نتیجه هم‌ترازی چند تصویر توسط این متد به تصویر کشیده شده است.

شکل ۲-۴: نتیجه متد Face Alignment بر روی مجموعه‌ای از تصاویر.

متد دیگری اتوماتیک دیگری توسط Kim و همکاران [25] برای هم ترازی بکار برده می‌شود اما این متد برای مسئله محاسبه تناظر دو تصویر بسیار مشابه پیشنهاد شده است.

۲-۴-۴. تشخیص چهره

در تشخیص چهره یا شناسایی چهره، تصویر ورودی (Probe) با یک پایگاه داده (گالری) مقایسه می‌شود و میزان تطابق گزارش می‌شود. مفهوم اصلی دسته بندی در تشخیص چهره، شناسایی مجموعه‌ای است که مشاهدات جدید به آن تعلق دارد. در یک دسته بندی کلی، متدهای تشخیص چهره به دو رویکرد تقسیم می‌شود. در رویکرد اول، ویژگی‌های چهره اصلی شناسایی می‌شود و در رویکرد دوم یک دید جامع برای مسئله تشخیص چهره در نظر گرفته می‌شود.

۲-۴-۴-۱. رویکرد جامع

این رویکرد با استفاده از متد آماری، مشخصات آماری را مستقیماً از کل تصاویر نمونه آموزشی استخراج می‌کند و استخراج ویژگی بینی، دهان، یا چشم‌ها به صورت جداگانه انجام نمی‌شود. نمونه‌هایی از متدهای جامع شامل چهره‌های ویژه (پر کاربردترین متد تشخیص چهره)، چهره‌های ویژه احتمالاتی، ماشین بردار پشتیبان، نزدیک‌ترین خطوط ویژگی، آنالیز مؤلفه مستقل می‌باشد.

۲-۴-۴-۲. رویکرد برپایه ویژگی

در رویکردهای برپایه ویژگی، ویژگی‌های محلی چهره همانند بینی، چشم‌ها قطعه بندی می‌شوند و به عنوان داده ورودی به دسته بند ساختاری اعمال می‌شوند. هندسه محض، معماری پیوند پویا، مدل مخفی

مارکوف ازجمله متدهایی هستند که به این دسته تعلق دارند. مطالعات و تحقیق نوروفیزیولوژیکی مشخص کرده است که چشم‌ها، دهان و بینی در میان مهم‌ترین ویژگی‌های تشخیص چهره می‌باشد.[26]

۲-۴-۴-۳. رویکرد ترکیبی

ایده این متد از این حقیقت نشأت می‌گیرد که سیستم بینایی انسان از هر دو ویژگی محلی و کلی استفاده می‌کند. چند نمونه از متدهای این رویکرد شامل چهره‌های ویژه ماژولار، ویژگی محلی ترکیبی، اشکال نرمال شده و متدهای برپایه مؤلفه می‌باشد. ویژگی‌های اصلی چهره انسان نقش بسیار مهمی در درک چهره دارد. ویژگی‌های اصلی چهره همانند چشم‌ها، دهان و بینی خاصیت تفکیک کنندگی را دارد، درحالیکه ویژگی‌های دیگر چهره همانند ویژگی‌های پیشانی، گونه‌ها و چانه این خاصیت تفکیک کنندگی را ندارد.

۲-۵. تکنیک‌های تشخیص چهره

این بخش مروری را بر تکنیک‌های اصلی در تشخیص چهره انسان در تصاویر چهره از جلو ارائه می‌کند و همچنین مزایا و معایب هر متد نیز بیان شده است. این متدها شامل چهره‌های ویژه، شبکه عصبی، معماری پیوند پویا، مدل مخفی مارکوف، تطبیق ویژگی هندسی و تطبیق الگو می‌باشد. این رویکردها در نحوه نمایش و ذخیره سازی چهره متفاوت هستند لذا از این جهت آنالیز می‌شوند.

۲-۵-۱. چهره‌های ویژه

تکنیک چهره‌های ویژه با استفاده از متد آنالیز مؤلفه‌های اصلی (PCA) که با نام Karhunen-Loeve معروف است، برای کاهش ابعاد استفاده می‌شود. به بیان ریاضی، چهره‌های ویژه، مؤلفه‌های اصلی توزیع چهره‌ها یا بردارهای ویژه ماتریس کوواریانس مجموعه تصاویر چهره می‌باشد. Turk و Pentland آنالیز

مؤلفه‌های اصلی را برای چهره یابی و تشخیص چهره اعمال کردند[27]. PCA، چهره‌های ویژه یا بردارهای ویژه مربوط به ماتریس‌های کوواریانس تصاویر آموزشی عمومی را می‌یابد. بردارهای ویژه برای نمایش مقادیر مختلف واریانس میان چهره‌ها مرتب می‌شوند. هر چهره را می‌توان با یک ترکیب خطی چهره‌های ویژه نمایش داد؛ همچنین آن را می‌توان تنها توسط بردارهای ویژه "بهترین" با بالاترین مقادیر ویژه نیز تقریب زد. M تا از بهترین چهره‌های ویژه یک فضای M بعدی با نام "فضای چهره" را می‌سازد. L. Sirovich و M. Kirby [28, 29] از آنالیز مؤلفه اصلی برای نمایش مؤثر تصاویر چهره استفاده نمودند. آنها بیان کردند که هر تصویر چهره را می‌توان با تعداد اندکی وزن برای هر چهره به همراه تصویر چهره استاندارد (تصویر ویژه) تقریباً بازسازی نمود. وزن‌های هر چهره از project کردن هر تصویر چهره بر روی تصویر ویژه بدست می‌آید.

۲-۵-۲. چهره‌های فیشر

R. A. Fisher آنالیز مجزاساز فیشر/خطی (LDA) را در سال ۱۹۳۰ ارائه نمود [30]. متد چهره فیشر، یکی از متدهای FR بر پایه appearance می‌باشد. تکنیک‌هایی که از LDA استفاده می‌کنند نتایج بسیار خوبی را ارائه کردند (Belhumeur et al., 1997; Zhao et al., 1999; Chen et al., 2000; Yu and Yang, 2001; Liu and Wechsler., 2002; Lu et al., 2003a, b; Ye and Li., 2004) . محققان LDA را برای بیشینه سازی نسبت پراکندگی بین کلاسی به پراکندگی درون کلاسی استفاده کردند. یک مشکل LDA در کاربردهای تشخیص چهره آنست که ماتریس پراکندگی درون کلاسی تقریباً همیشه منفرد است زیرا در مواردی که تغییرات زیادی در زاویه چهره یا شرایط روشنایی در تصاویر یک چهره وجود دارد و همچنین تعداد پیکسل‌های تصویر معمولاً بسیار بیشتر از تعداد تصاویر می‌باشد، نرخ خطا در تشخیص افزایش می‌یابد. جهت غلبه بر پیچیدگی ماتریس منفرد، الگوریتم‌هایی پیشنهاد شده است [29-26]. از

آنجا که رویکرد چهره فیشر از اطلاعات درون کلاسی و بین کلاسی بهره می‌جوید، می‌توان بر مشکل تغییرات در تصاویر یک چهره یکسان همانند شرایط مختلف نوری غلبه کرد.

۲-۵-۳. تطابق ویژگی هندسی

تکنیک‌های تطابق ویژگی هندسی بر اساس محاسبه مجموعه‌ای از ویژگی‌های هندسی از تصویر چهره می‌باشد. پیکربندی سراسری را می‌توان توسط یک بردار که شامل موقعیت و اندازه ویژگی‌های اصلی چهره همانند چشم، ابرو، بینی، دهان و فرم کلی چهره می‌باشد، توصیف نمود. هر نوع عنصر جغرافیایی (همانند رود، جاده، خط مرزی و غیره) دارای ویژگی‌های هندسی ذاتی می‌باشد که توسط داده هندسی (مؤلفه) نمایش داده می‌شود. این داده‌های پردازش شده باید ویژگی‌های هندسی را نمایش دهند. با مقایسه ویژگی‌های هندسی داده‌ها با ویژگی‌های ذاتی می‌توان بهترین پیوندها را یافت. بنابراین، ویژگی‌های هندسی عناصر را می‌توان به عنوان داده معتبر در فرایند در نظر گرفت.

۲-۵-۴. شبکه عصبی

شبکه‌های عصبی به طور موفقیت آمیزی به بسیاری از مسائل شناسایی الگو همانند تشخیص کارکتر، تشخیص شیء و ... اعمال شده است. مزیت استفاده از شبکه‌های عصبی در چهره یابی، امکان آموزش سیستم برای یافتن چگالی شرطی پیچیده مربوط به الگوهای چهره می‌باشد. گرچه، یک عیب شبکه‌های عصبی آنست که معماری شبکه (تعداد لایه‌ها، تعداد نودها، نرخ یادگیری و ...) باید به صورت دقیق تنظیم شود تا عملکرد مطلوب حاصل شود [31]. جذابیت استفاده از شبکه‌های عصبی ممکن است مربوط به غیرخطی بودن آن باشد. بنابراین، گام استخراج ویژگی ممکن است از متدهای Karhunen-Loeve خطی کاراتر باشد زیرا کاهش ابعاد با تصویر کردن خطی انجام می‌شود که پراکندگی تمام نمونه‌های تصویر شده را بیشینه می‌سازد [32]. مؤلفان، آزمایشات خود را بر روی پایگاه داده ORL شامل ۴۰۰ تصویر از ۴۰

شخص انجام دادند و درصد تشخیص 96.2% را گزارش کردند. زمان دسته بندی کمتر از ۰.۵ ثانیه می‌باشد اما زمان آموزش ۴ ساعت می‌باشد. این روش مستقل از جابجایی، چرخش، تغییر اندازه و تغییر شکل می‌باشد. گرچه هرگاه تعداد اشخاص افزایش یابد، هزینه محاسبه نیز افزایش خواهد یافت. به طور کلی، رویکردهای شبکه عصبی در مواقعی که تعداد کلاس‌ها (اشخاص) افزایش یابد، با مسائلی مواجه می‌شود. علاوه بر این، شبکه‌های عصبی در آزمایش تشخیص تصویر با یک مدل منفرد مناسب نیستند زیرا آموزش سیستم‌ها و تعیین پارامترهای "بهینه"، نیازمند اعمال تصاویر با چند مدل برای هر شخص می‌باشد.

۲-۵-۵. تطبیق گراف

تطبیق گراف کاربردهایی در زمینه‌های مختلف، از بینایی ماشین تا بیولوژی محاسباتی دارد. در تطبیق گراف، الگوها به عنوان گراف‌ها و مقادیر شناسایی الگو مدل می‌شوند تا ارتباطی میان نودهای گراف‌های مختلف یافت شود [33]. فرمول بندی تطبیق گراف در شناسایی الگو تنها به یافتن بهترین تطبیق میان گراف‌های مدل نمایشی (کلاس) و گراف‌های داده مفروض می‌پردازد. به طور کلی، تطبیق گراف‌ها ممکن است با مقایسه رئوس و لبه‌ها با توجه به سهم هریک در سنجش فاصله نسبی صورت گیرد [34]. M.Lades و همکاران [35] یک ساختار پیوندی پویا را برای تشخیص شیء مستقل از تغییر شکل و تخریب ارائه دادند که از تطبیق گراف کشسانی برای یافتن نزدیک‌ترین گراف ذخیره شده استفاده می‌کند. معماری پیوند پویا یک توسعه‌ای بر شبکه‌های عصبی مصنوعی می‌باشد. فرایند تطبیق، بار محاسباتی بسیاری دارد؛ برای مقایسه ۸۹ شیء ذخیره شده بر روی یک ماشین موازی با ۲۳ انتقال دهنده، حدود ۲۵ ثانیه زمان لازم است. L. Wiskott و همکاران [36] تکنیک را گسترش دادند و چهره‌های ورودی (probe) را با یک گالری از ۱۱۲ چهره خنثی از زاویه جلو مقایسه کردند و تطبیق را انجام دادند. به طور کلی،

معماری پیوند پویا به دلیل ویژگی مستقل از چرخش آن، نسبت به تکنیک‌های دیگر برتری دارد؛ گرچه، فرایند تطبیق آن بار محاسباتی بسیاری دارد.

۲-۶. تکنیک‌های استخراج ویژگی چهره

۲-۶-۱. رویکردهای برپایه ظاهر

مفهوم "ویژگی" در این رویکردها از ویژگی‌های ساده چهره همانند چشم‌ها و دهان متفاوت است. هر مشخصه استخراج شده از تصویر به یک ویژگی رجوع می‌کند. روش‌هایی همانند آنالیز مؤلفه اصلی (PCA)، آنالیز مؤلفه مستقل و موجک‌های گابور [37] برای استخراج بردار ویژگی بکار می‌روند. در مقایسه با تطبیق الگو، یادگیری مدل‌ها (یا قالب ها) ازطریق مجموعه‌ای از تصاویر آموزشی صورت می‌پذیرد که تغییرپذیری نمایشی ظاهر صورت را در نظر می‌گیرد. سپس از این مدل‌های آموزش یافته در یافتن ویژگی‌ها استفاده می‌شود. این روش‌ها اغلب برای چهره یابی طراحی شده اند.

۲-۶-۲. رویکردهای بر پایه قالب

چندین الگوی استاندارد چهره برای توصیف چهره به صورت کلی یا به صورت مجزا برای هر ویژگی چهره ذخیره می‌شود . در چهره یابی، وابستگی‌های میان تصویر ورودی و الگوهای ذخیره شده محاسبه می‌شود. این روش‌ها برای یافتن موقعیت چهره و همچنین چهره یابی استفاده می‌شود. این تکنیک مؤلفه‌های چهره را به قالب‌های از پیش طراحی شده با استفاده از توابع انرژی مناسب تطبیق می‌دهد. در روش پیشنهادی Yuille و همکاران [38]، بهترین تطبیق یک قالب با تصویر چهره، انرژی را کمینه می‌کند؛ این الگوریتم‌ها ابتدا قالب‌ها را مدل می‌کنند و این عمل باعث بوجود آمدن بار محاسباتی می‌شود

که بر عملکرد آنها تأثیر می‌گذارد. الگوریتم‌های ژنتیک برای زمان جستجوی مؤثر در تطبیق الگو نقش دارند.

۲-۶-۳. تکنیک‌های جداسازی رنگ

در تکنیک جداسازی رنگ، از رنگ پوست برای جداسازی چهره استفاده می‌شود. هر منطقه رنگی غیر پوست در چهره به عنوان کاندیدی برای چشم‌ها و/یا دهان در نظر گرفته می‌شود. عملکرد این تکنیک‌ها بر روی پایگاه داده‌های تصویر چهره، با توجه به تنوع پیش زمینه‌های نژادی نسبتاً محدود می‌باشد.[39]

۲-۶-۴. تکنیک‌های برپایه هندسه

ویژگی‌ها توسط اطلاعات هندسه همانند موقعیت‌های نسبی و اندازه‌های مؤلفه‌های چهره استخراج می‌شود. در تکنیک پیشنهادی Kanade [40] چشم‌ها، دهان و بینی توسط نقشه لبه عمودی موقعیت یابی می‌شود. این تکنیک‌ها نیازمند آستانه گیری می‌باشد که در برخی مواقع حساس بر عملکرد حاصل شده تأثیر منفی می‌گذارد.

۲-۷. نتیجه گیری

با رشد روزافزون و گسترده سیستم‌های تحت وب برای کارهای امنیتی، ضرورت وجود یک سیستم احراز هویت تحت وب بیومتریکی کاملاً مشهود است. در این پایان‌نامه، سیستم احراز هویت تحت وب از طریق تشخیص چهره طراحی و پیاده سازی می‌شود. از آنجا که سیستم یک پروژه عملیاتی محسوب می‌شود، دارای بخش‌های مختلفی است، در این فصل، سیستم را به بخش‌های پیش پردازش، چهره یابی، هم‌ترازی چهره، تشخیص چهره تقسیم بندی نمودیم و در هر بخش به توصیف الگوریتم‌ها وتکنیک‌های

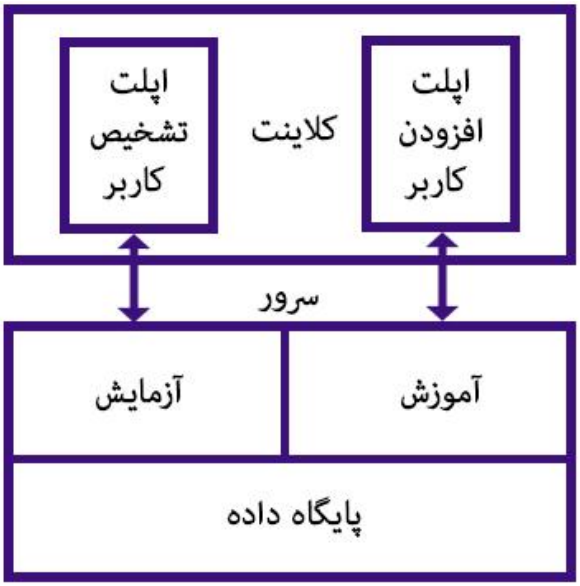
موجود پرداختیم. زمینه تحقیقاتی تشخیص چهره، همیشه مورد توجه محققان بوده است ولی در کمال تعجب تنها سه پروژه مشابه با موضوع پایان‌نامه یافت شد که هرسه به صورت رایگان در اختیار عموم قرار نمی‌گرفت که ارزش انجام و اجرای پایان‌نامه را دوچندان می‌کند.

فصل سوم

سیستم تشخیص چهره عملیاتی

۳-۱. سیستم عملیاتی تشخیص چهره

در این پایان‌نامه، علاوه بر ارائه یک روش نوین در تشخیص چهره (فصل چهارم)، یک سیستم تشخیص چهره تحت وب نیز عملیاتی شده است. این سیستم تشخیص چهره به سه بخش اپلت[1] تشخیص کاربر، اپلت افزودن کاربر و سرور جاوای تشخیص چهره تقسیم بندی می‌شود. از آنجا که این سیستم تشخیص چهره تحت وب می‌باشد، عملیات‌ها باید به صورت بهینه میان کلاینت و سرور توزیع شوند.



شکل ۳-۱: معماری سیستم تشخیص چهره عملیاتی

عملیات اخذ تصویر از وب‌کم و عملیات چهره یابی به همراه پردازش‌های مورد نیاز در اپلت جاوای سمتِ کلاینت صورت می‌گیرد. عملیات‌های اصلی آموزش و آزمایش تصاویر چهره به همراه پایگاه داده مربوطه در بخش سرور قرار دارد. با این تقسیم بندی ارتباطات میان کلاینت و سرور به حداقل می‌رسد. از آنجا که دو اپلت جاوای مجزا برای افزودن کاربر و تشخیص کاربر در نظر گرفته شده است، سرور جاوا ابتدا درخواست‌های دریافتی را بررسی می‌کند و سپس به بخش مرتبط با آن هدایت می‌کند. بنابراین اپلت

[1] Applet

تشخیص کاربر به همراه بخش آزمایش سرور و همچنین پایگاه داده، مرحله آزمایش را تشکیل می‌دهند و از طرفی دیگر، اپلت افزودن کاربر به همراه بخش آموزش سرور و مجدداً پایگاه داده، مرحله آموزش را تشکیل می‌دهند. در شکل ۳-۱ معماری سیستم تشخیص چهره ترسیم شده است.

۳-۱-۱. مرحله آزمایش

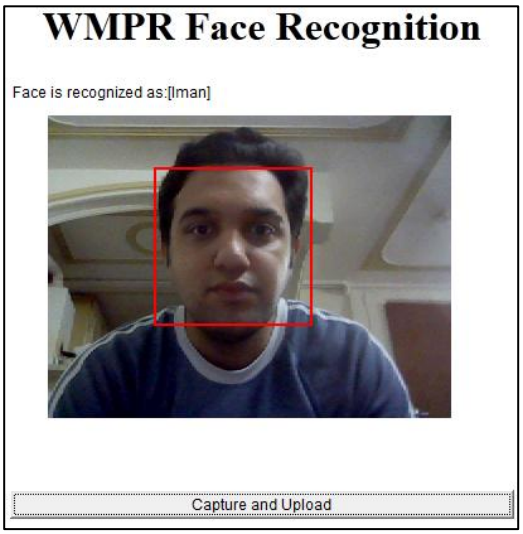
اپلت تشخیص کاربر یک برنامه جاوای سمتِ کلاینتِ قابلِ اجرا در مرورگر می‌باشد که وب‌کم کلاینت را روشن می‌کند و تصویر دریافتی آن را پس از پردازش در اپلت جاوا نمایش می‌دهد. همچنین اپلت تشخیص کاربر، هر فریم را دریافت و سپس پردازش می‌کند. پس از پردازش، تصویر منطقه مطلوب[1] را می‌توان برای سرور جاوا ارسال نمود. در مرحله پردازش، ابتدا تصویر رنگی هر فریم دریافتی به تصویر خاکستری تبدیل می‌شود، سپس معادل سازی هیستوگرام بر روی تصویر خاکستری انجام می‌شود تا نور موجود در تصویر متعادل گردد، سپس عملیات چهره یابی بر روی تصویر خاکستری اعمال می‌شود تا مکان چهره توسط الگوریتم Viola-Jones در تصویر مشخص شود. در صورت یافتن مکان چهره، مکان مورد نظر با یک مستطیل در تصویر نشان داده می‌شود. کاربر می‌تواند پس از تنظیم چهره خود با کلیلک بر روی دکمه "Capture and Upload" تصویر چهره موجود در فریم جاری را با اعمال پیش پردازش‌های فوق الذکر به سرور جاوا ارسال می‌کند. عملیات ارسال شامل دو گام می‌باشد. در گام اول، یک فایل موقتی بر روی سیستم کلاینت با محتوای تصویر چهره کاربر ساخته می‌شود و در گام دوم، این فایل به سرور جاوا با آدرس از پیش تعیین شده ارسال می‌شود.

حال به عملکرد بخش آزمایش سرور جاوا می‌پردازیم. در ابتدای امر، سرور جاوا درخواست را بررسی می‌کند تا مشخص شود که درخواست دهنده اپلت تشخیص کاربر است یا اپلت افزودن کاربر. پس از

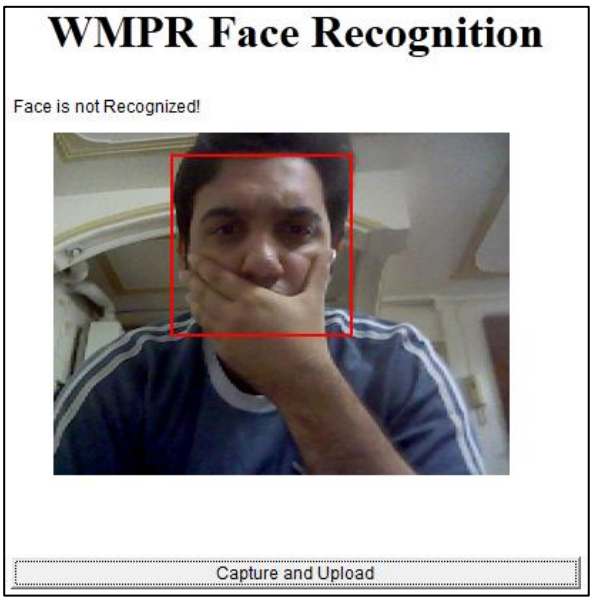
[1] Region Of Interest

تشخیص درخواست دهنده به عنوان اپلت تشخیص کاربر، فایل حاوی تصویر چهره کاربر در سرور ذخیره می‌شود. بخش آزمایش سرور جاوا، الگوریتم تشخیص چهره الگوهای دودویی محلی را بر روی تصویر چهره اعمال می‌کند و با مقایسه هیستوگرام‌های الگوهای دودویی محلی تصویر دریافتی با تصاویر درون پایگاه داده، نتیجه تشخیص چهره را اعلام می‌نماید. نتیجه تشخیص چهره از سمت سرور جاوا به اپلت تشخیص کاربر برگردانده می‌شود و اپلت پاسخ دریافتی را که شامل نام فرد تشخیص داده شده یا عدم تشخیص هیچ فرد باشد را نمایش می‌دهد. همانطور که مشاهده می‌کنید در تصویر شکل ۳-۲، پاسخ بازگشتی از سرور جاوا به اپلت جاوا Iman می‌باشد.

اپلت تشخیص کاربر به صورت مداوم تصویر را از وب‌کم دریافت میکند و پردازش می‌کند. یکی از مزایای این اپلت آنست که بار پردازشی و محاسباتی بر روی سیستم کلاینت قرار می‌گیرد و نیازی به برقراری ارتباط پیوسته با سرور نیست. ارتباط اپلت با سرور فقط برای ارسال تصویر چهره کاربر و دریافت متن کوتاهی به عنوان پاسخ می‌باشد که بهینه‌ترین حالت ممکن برای استفاده از پهنای باند در یک سیستم تشخیص چهره تحت وب می‌باشد.



شکل ۳-۲: نمونه‌ای از اپلت تشخیص کاربر در مرورگر وب برای دریافت تصویر چهره کاربر و ارسال آن به سرور جاوا برای مرحله تشخیص چهره.



شکل ۳-۳: نمونه‌ای از اپلت تشخیص کاربر در مرورگر وب در زمانی که چهره یابی انجام شده است ولی چهره کاربر تشخیص داده نشده است.

اپلت تشخیص کاربر به همراه بخش آزمایش سرور جاوا، در اصطلاح مرحله آزمایش سیستم تشخیص چهره را انجام می‌دهد.

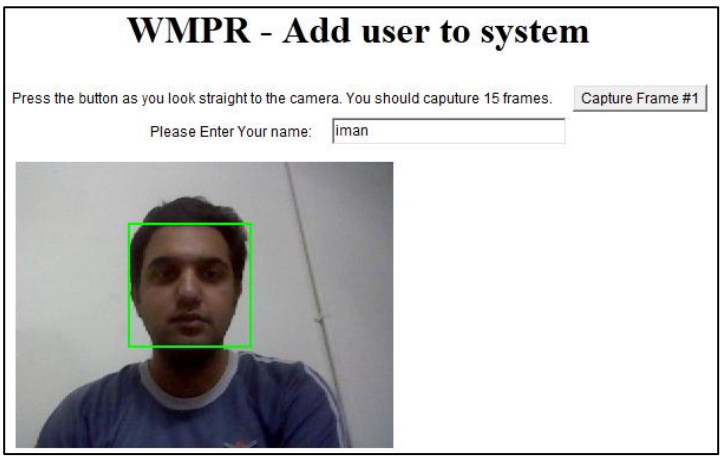
۳-۱-۲. مرحله آموزش

سیستم تشخیص چهره علاوه بر مرحله آزمایش نیاز به مرحله آموزش نیز دارد که چهره‌های جدیدی به پایگاه داده افزوده شود. مرحله آموزش در قالب اپلت افزودن کاربر به همراه بخش آموزش سرور جاوا پیاده سازی شده است.

اپلت افزودن کاربر، برای افزودن چهره شخص جدید به سیستم بکار می‌رود. در سیستم تشخیص چهره از هر شخص ۱۵ تصویر اخذ می‌شود، هر چند که تعداد تصاویر برای هر شخص را می‌توان تغییر داد و مقداری بیشتر و یا کمتر انتخاب نمود. با افزایش تعداد تصاویر به ازای هر شخص، عملکرد سیستم تشخیص چهره بهتر خواهد شد اما از طرفی دیگر، با کاهش تعداد تصاویر به ازای هر شخص، پهنای باند

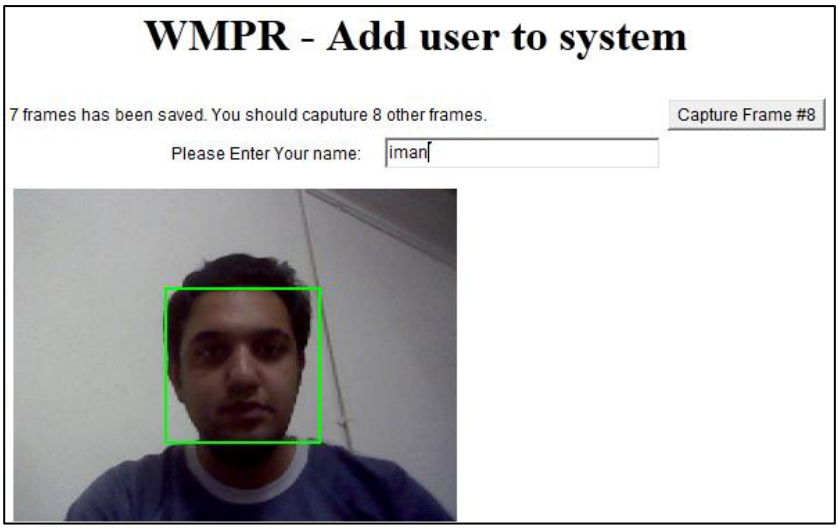
ارتباطی میان اپلت و سرور جاوا به صورت بهینه استفاده خواهد شد. این مقدار ۱۵ به صورت تجربی تعادلی را میان پهنای باند تصاویر ارسالی از اپلت به سرور و صحت عملکرد سیستم تشخیص چهره برقرار می‌کند.

اپلت افزودن کاربر یک برنامه جاوای سمتِ کلاینتِ قابلِ اجرا در مرورگر می‌باشد که وب‌کم کلاینت را روشن می‌کند و تصویر دریافتی آن را پس از پردازش در اپلت جاوا نمایش می‌دهد. همچنین اپلت افزودن کاربر، هر فریم را دریافت و سپس پردازش می‌کند. در مرحله پردازش، ابتدا تصویر رنگی هر فریم دریافتی به تصویر خاکستری تبدیل می‌شود، سپس معادل سازی هیستوگرام بر روی تصویر خاکستری انجام می‌شود تا نور موجود در تصویر متعادل گردد، سپس عملیات چهره یابی بر روی تصویر خاکستری اعمال می‌شود تا مکان چهره توسط الگوریتم Viola-Jones در تصویر مشخص شود. در صورت یافتن مکان چهره، مکان مورد نظر با یک مستطیل در فریم نشان داده می‌شود. کاربر می‌تواند پس از تنظیم چهره خود با کلیلک بر روی دکمه "Capture Frame" تصویر چهره موجود در فریم جاری را با اعمال پیش پردازش‌های فوق الذکر به صورت موقتی در سیستم خود ذخیره کند و در نهایت ۱۵ تصویر چهره را به همراه نام دریافتی به سرور جاوا ارسال کند.



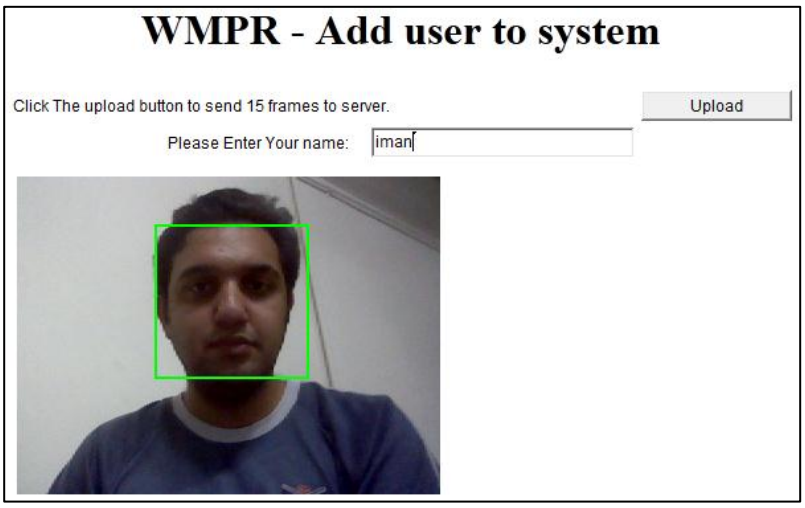
شکل ۳-۴: نمونه‌ای از اپلت افزودن کاربر. در این اپلت نام شخص به همراه ۱۵ تصویر چهره از او برای پایگاه داده ارسال می‌شود.

در اپلت افزودن کاربر، کاربر پس از تنظیم چهره خود در صفحه و رویت مستطیل سبز، می‌تواند دکمه "Capture Frame #1" را بفشارد. در هر مرحله از دریافت ۱۵ تصویر چهره، راهنمایی‌هایی در اپلت نشان داده می‌شود که چه تعداد تصویر چهره دریافت شده است و چه تعداد تصویر چهره دیگر برای دریافت باقی مانده است.



شکل ۳-۵: نمونه‌ای از اپلت افزودن کاربر. این تصویر، اپلت را در زمانی نشان می‌دهد که ۷ تصویر چهره تا به حال دریافت شده است و ۸ تصویر چهره دیگر باید دریافت شود.

کاربر پس از اخذ ۱۵ تصویر چهره در اپلت افزودن کاربر، دکمه‌ای با نام Upload را مشاهده می‌کند. با کلیک بر روی این دکمه، ۱۵ تصویر چهره به همراه نام دریافتی از کاربر برای سرور جاوا ارسال می‌شود.



شکل ۳-۶: نمونه‌ای از اپلت افزودن کاربر. این تصویر، اپلت را در زمانی نشان می‌دهد که ۷ تصویر چهره تا به حال دریافت شده است و ۸ تصویر چهره دیگر باید دریافت شود.

حال به عملکرد بخش آموزش سرور جاوا می‌پردازیم. در ابتدای امر، سرور جاوا درخواست دهنده را بررسی می‌کند تا مشخص شود که درخواست از اپلت تشخیص کاربر ارسال شده است یا اپلت افزودن کاربر. پس از تشخیص درخواست دهنده به عنوان اپلت افزودن کاربر، بخش آموزش سرور ۱۵ تصویر چهره را به همراه نام کاربر را ذخیره می‌کند. پس از آن سیستم تشخیص چهره با آن تصاویر چهره به همراه نام دریافتی به اصطلاح آموزش داده می‌شود. پاسخ سرور در نهایت به اپلت افزودن چهره بازگردانده می‌شود.

همانطور که پیش تر اشاره شد، پایگاه داده این سیستم تشخیص چهره در سرور قرار دارد تا در امنیت و به دور از دسترس کاربران غیرمجاز قرار داشته باشد. نکته‌ای که لازم به ذکر است، آنست که منظور از آموزش سیستم و ذخیره تصاویر چهره جدید در سیستم، محاسبه هیستوگرام‌های الگوهای دودویی محلی تصاویر چهره جدید می‌باشد و سپس ذخیره آن هیستوگرام‌ها می‌باشد. بنابراین مقادیر پیکسل‌های تصاویر چهره جدید در پایگاه داده ذخیره نمی شود بلکه هیستوگرام‌های الگوهای دودویی محلی تصاویر چهره جدید در پایگاه داده ذخیره می‌شود.

رویکردی در سیستم‌های تشخیص چهره مطرح است که سرور در اختیار یک شرکت قرار گیرد و متقاضیان سیستم همگی به آن سرور متصل شوند. از آنجا که در این رویکرد، متقاضیان مختلف به یک سرور متصل می‌شوند، پهنای باند آن سرور اشغال خواهد شد و بر روی عملکرد آن تاثیر خواهد گذاشت. از طرفی دیگر، برخی از متقاضیان سیستم حاضر به استفاده از این سیستم نخواهند بود، زیرا محرمانه بودن اطلاعات و تصاویر چهره برای آنها اهمیت دارد. یکی از مزایای سیستم تشخیص چهره عملیاتی این پایان‌نامه آنست که این مشکل را برطرف نموده است. ساختار و پیاده سازی این سیستم بگونه‌ای طراحی شده است که می‌توان مجموعه کدهای کامپایل شده سرور و کلاینت را می‌توان در اختیار متقاضیان سیستم گذاشت.

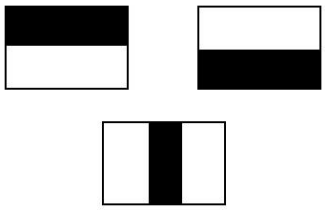
از آنجا که دو عملیات اصلی این سیستم تشخیص چهره، چهره یابی و تشخیص چهره می‌باشد، در ادامه فصل نحوه عملکرد متد چهره یابی Viola-Jones و همچنین متد تشخیص چهره الگوهای دودویی محلی توضیح داده خواهد شد.

۳-۲. چهره یابی با الگوریتم Viola-Jones

رویکرد چهره یابی viola-Jones در یافتن اشیا در تصویر از چهار مفهوم کلیدی بهره می‌برد:

- ویژگی‌های مستطیلی ساده همانند ویژگی‌های Haar
- یک تصویر انتگرالی برای یافتن سریع ویژگی
- متد یادگیری ماشین Adaboost
- دسته بند آبشاری برای ترکیب مؤثر ویژگی ها

ویژگی‌های Viola-Jones بر اساس موجک‌های Haar می‌باشد. موجک‌های Haar موج‌های مربعی تک موجی شامل بازه‌های بالا و پایین می‌باشد؛ در دو بعد، یک موج مربعی متشکل از منطقه‌های روشنایی و تاریکی است.

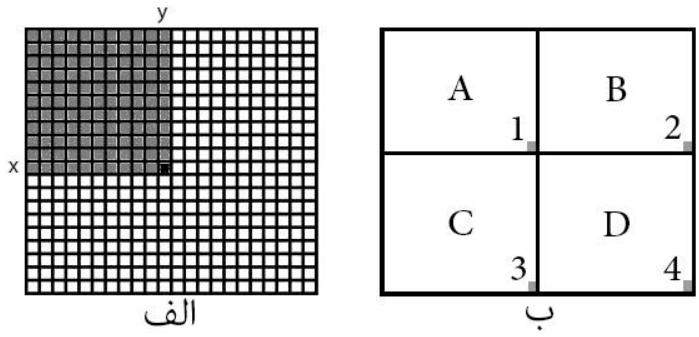


شکل ۳-۷: چند نمونه از موجک‌های Haar

وجود یک ویژگی Haar از تفریق مقدار میانگین پیکسل‌های منطقه تاریکی از مقدار پیکسل‌های منطقه روشنایی تعیین می‌شود. اگر اختلاف بالاتر از یک حد آستانه باشد (مقدار آستانه در حین یادگیری مقداردهی می‌شود)، آن ویژگی Haar موجود می‌باشد.

برای تعیین وجود یا عدم وجود صدها ویژگی Haar در هر مکان تصویر و در چندین مقیاس، در متد Viola-Jones از تصویر انتگرالی استفاده می‌شود. در حالت کلی، "انتگرال گیری" به معنی افزودن واحدهای کوچک به یکدیگر می‌باشد. در این مورد خاص، واحدهای کوچک همان مقادیر پیکسل می‌باشد. مقدار انتگرال برای هر پیکسل مجموع تمام پیکسل‌های بالای آن و چپ آن می‌باشد. پیمایش از بالا و چپ شروع می‌شود و به راست و پایین ختم می‌شود، کل تصویر با چند عملیات صحیح در هر پیکسل انتگرال گرفته می‌شود.

همانطور که در شکل ۳-۸ مشاهده می‌کنید، پس از انتگرال گیری، مقدار پیکسل در مکان (x, y) حاوی مجموع مقادیر تمامی پیکسل‌ها در یک ناحیه مستطیلی است که یک گوشه آن در بالا چپ تصویر و گوشه دیگر آن در مکان (x, y) می‌باشد. برای محاسبه مقدار میانگین پیکسل‌ها در این مستطیل، تنها بایستی مقدار (x, y) را بر مساحت مستطیل تقسیم کرد.



شکل ۳-۸: حقه تصویر انتگرالی. الف) پس از انتگرال گیری، پیکسل (x,y) حاوی مجموع تمامی پیکسل‌های خاکستری رنگ می‌باشد. ب) مجموع مقادیر پیکسل‌های ناحیه D از فرمول (x4, y4) – (x2, y2) – (x3,y3) + (x1,y1) بدست می‌آید.

اگر مجموع مقادیر ناحیه دیگری همانند D مطلوب باشد. شکل ۳-۸-ب راه حل این مسئله را نشان می‌دهد. شما می‌توانید آن را به صورت مجموع مقادیر پیکسل در یک مستطیل مرکب (A+B+C+D) منهای مجموع مستطیل‌های A+B و A+C به علاوه مجموع مقادیر پیکسل ناحیه A محاسبه کنید. به عبارت دیگر:

$$D = A+B+C+D - (A+B) – (A+C) + A. \qquad (۱-۳)$$

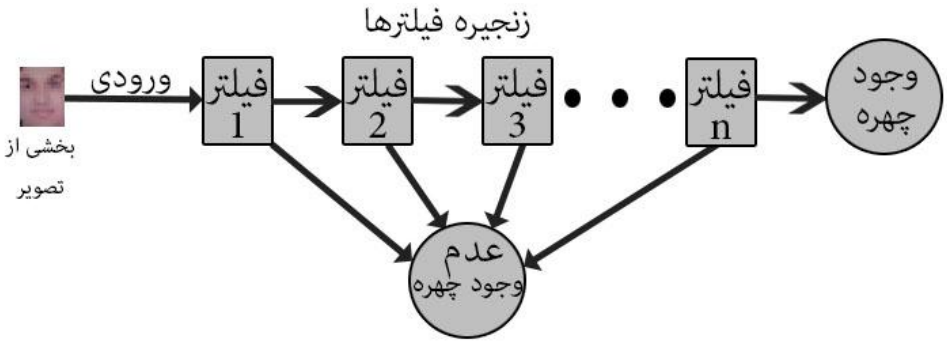
به بیان ساده تر، A+B+C+D مقدار تصویر انتگرال در مکان ۴ می‌باشد، A+B مقدار آن در مکان ۲ می‌باشد، A+C مقدار آن در مکان ۳ می‌باشد، و A مقدار آن در مکان ۱ می‌باشد. بنابراین، با داشتن تصویر انتگرالی، می‌توان مجموع مقادیر پیکسل‌های هر ناحیه تصویر اصلی را تنها با ۳ عمل صحیح محاسبه نمود:

$$(x4, y4) – (x2, y2) – (x3,y3) + (x1,y1) \qquad (۲-۳)$$

Viola-Jones از یک متد یادگیری ماشین همانند Adaboost برای انتخاب ویژگی‌های Haar خاص و همچنین برای تعیین مقدار سطوح آستانه استفاده می‌کند. Adaboost چند کلاسیفایر "ضعیف" را ترکیب می‌کند تا یک کلاسیفایر "قوی" ایجاد کند. "ضعیف" در اینجا به معنی دسته بندی است که تعداد پاسخ‌های صحیح آن بیشتر از دسته بند تصادفی است. در صورتیکه از ترکیب بسیاری از دسته بندهای

ضعیف استفاده شود و هرکدام با احتساب یک وزن مشخص برخی از پاسخ‌های درست نهایی ارائه دهند، مجموعاً دسته بند ترکیبی موزون قوی بدست خواهد آمد.

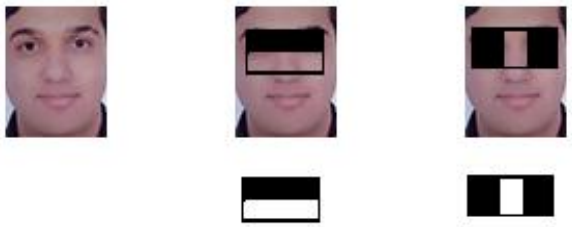
Viola-Jones یک سری از دسته بندهای Adaboost را همانند یک زنجیره فیلتر ترکیب می‌کند (شکل ۳-۹) این عمل برای دسته بندی نواحی تصویر مؤثر است. هر فیلتر یک دسته بند Adaboost جداگانه با تعداد نسبتاً کمی از دسته بندهای ضعیف می‌باشد.



شکل ۳-۹: آبشار دسته بندها به عنوان زنجیره فیلترها می‌باشد. زیرناحیه‌هایی از تصویر که از تمام فیلترها عبور کنند به "چهره" دسته بندی می‌شوند.

آستانه پذیرش در هر سطح آنقدر پایین انتخاب می‌شود تا تقریباً همه نمونه‌های چهره در مجموعه آموزش را عبور دهد. فیلترهای هر سطح آموزش می‌بینند تا تصاویر آموزشی را که تمام مراحل پیشین را گذرانده اند را دسته بندی کنند (مجموعه آموزشی یک پایگاه داده بزرگی از چهره هاست، شاید به هزار یا بیشتر برسد). هنگام استفاده، اگر هر کدام از این فیلترها نتواند یک منطقه تصویر را عبور دهد، آن منطقه فوراً به عنوان "غیر چهره" دسته بندی می‌شود. هرگاه یک فیلتر یک منطقه از تصویر را عبور دهد، آن منطقه تصویر به فیلتر بعدی در زنجیره ارسال می‌شود. مناطقی از تصویر که از تمام فیلترهای زنجیره عبور کنند، به عنوان "چهره" دسته بندی می‌شوند. این زنجیره را می‌توان به نام آبشار نیز شناخت.

ترتیب فیلترهای آبشار بر اساس اولویتی است که Adaboost مشخص کرده است. ابتدا فیلترهای سنگین تر قرار می‌گیرند، تا مناطق "غیر چهره" تصویر به سرعت حذف کنند. شکل ۳-۱۰ نشان می‌دهد که دو ویژگی اول از آبشاری Viola-Jones بر روی چهره من قرار گرفته اند. اولی بر روی گونه‌ها و چشم‌ها قرار گرفته است و چون گونه‌ها روشن تر از چشم‌ها هستند این تطبیق صورت گرفته است. دومی هم از روشنایی پل بینی نسبت به چشم‌ها استفاده می‌کند. در نهایت، مکان چهره با استفاده از تطبیق ویژگی‌های Haar یافت می‌شود.

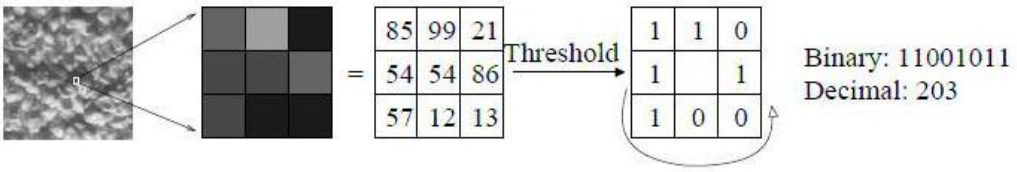


شکل ۳-۱۰: دو ویژگی Haar از زنجیره فیلترها

۳-۳. تشخیص چهره با الگوهای دودویی محلی[1]

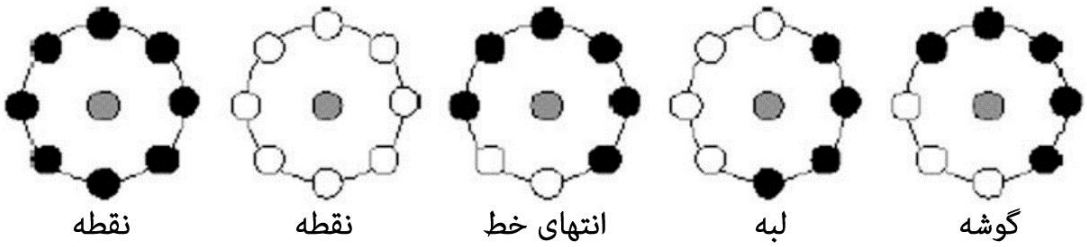
عملگر LBP اصلی، که توسط Ojala و همکاران [۴۱] معرفی گردید، روش بسیار قدرتمندی برای توصیف بافت در تصاویر خاکستری می‌باشد. عملگر در ابتدا پیکسل‌های یک تصویر را با در نظر گرفتن همسایگی ۳×۳ هر پیکسل نسبت به مقدار مرکزی آستانه گیری می‌کند و سپس با نتیجه حاصله که یک عدد دودویی باشد، پیکسل مربوطه را برچسب گذاری می‌کند. سپس هیستوگرام ۲۵۶-ستونی برچسب‌ها را می‌توان به عنوان یک توصیف گر بافت استفاده نمود. شکل ۳-۱۱ نمایشی از عملگر LBP ساده را ارائه می‌دهد.

[1] Local Binary Patterns



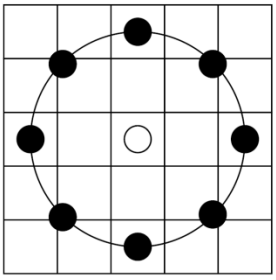
شکل ۳-۱۱: عملگر LBP اصلی

اعداد دودویی مستخرج (کدهای الگوهای دودویی محلی نامیده می‌شوند) طرح‌های ساده محلی شامل انواع لبه‌های منحنی، نقطه‌ها، مناطق تخت و غیره را رمزگذاری می‌کنند. این طرح‌ها در شکل ۳-۱۲ به نمایش گذاشته شده است. بنابراین رمز LBP را می‌توان به عنوان یک ریزبافت در نظر گرفت.



شکل ۳-۱۲: نمونه‌هایی از طرح‌های اولیه بافت که با الگوهای دودویی محلی تشخیص داده می‌شوند. پیکسل‌های سفید و سیاه به ترتیب پیکسل‌های سفید و سیاه را مشخص می‌کند.

کوچک بودن اندازه همسایگی ۳×۳ محدودیتی برای عملگر محسوب می‌شود، زیرا نمی تواند ویژگی‌های غالب با اندازه بزرگ را در نظر بگیرد. پس از آن، از اندازه‌های مختلف همسایگی‌ها برای عملگر استفاده شد. با استفاده از همسایگی‌های دایره‌ای و درون یابی دوخطی مقادیر پیکسل‌ها، هر شعاع و تعداد پیکسلی را در همسایه می‌توان پوشش داد. برای همسایگی‌ها از نماد (P, R) استفاده خواهیم کرد که P به معنی نقاط نمونه بر روی دایره‌ای به شعاع R می‌باشد. شکل ۳-۱۳ نمونه‌ای از همسایگی (8, 2) دایره‌ای می‌باشد.



شکل ۳-۱۳: همسایگی دایره‌ای (8, 2). هرگاه نقطه بر روی مرکز یک پیکسل قرار نگیرد، مقادیر پیکسلها به صورت دوخطی درون یابی می‌شوند.

عملگر $LBP_{P,R}$ تعداد 2^P مقدار مختلف را تولید می‌کند، که مربوط به 2^P الگوی دودویی مختلف می‌باشد و توسط P پیکسل در مجموعه همسایه تشکیل شده است.

توسعه دیگری برای عملگر اصلی وجود دارد که با نام الگوهای یکنواخت [1]شناخته می‌شود [۴۲]. اگر الگوی دودویی محلی با فرض دایره وار بودن، حداکثر دو تغییر بیتی از ۰ به ۱ یا بالعکس داشته باشد، الگوی دودویی محلی را یکنواخت گویند. برای نمونه الگوهای ۰۰۰۰۰۰۰۰، ۰۰۰۱۱۱۱۰ و ۱۰۰۰۰۰۱۱ الگوهای یکنواخت هستند. Ojala و همکاران [۴۱] متوجه شدند که در آزمایشات تصاویر بافت با همسایگی (8, 1)، الگوهای یکنواخت کمی کمتر از ۹۰٪ تمام الگوها را تشکیل می‌دهند و در همسایگی (16, 2) این مقدار به تقریباً ۷۰٪ می‌رسد.

با انباشت الگوهای دارای بیش از دو تغییر بیتی در یک ستون، یک عملگر LBP با تعداد ستون کمتر از 2^P بوجود می‌آید. ما نماد $LBP_{P,R}^{u2}$ را برای عملگر LBP استفاده می‌کنیم. زیرنویس عملگر نشان دهنده همسایگی (P,R) می‌باشد. بالانویس u2 نشان دهنده آن است که الگوهای غیریکنواخت با یک برچسب، برچسب گذاری شود و برای هر الگوی یکنواخت برچسب مجزا قرار داده خواهد شد. برای نمونه، در LBP

[1] Uniform Patterns

استاندارد، تعداد برچسب‌های همسایگی ۸ پیکسلی برابر ۲۵۶ می‌باشد اما این تعداد در LBP^{u2} برابر ۵۹ می‌باشد.

پس از برچسب گذاری تصویر با عملگر LBP، هیستوگرام تصویر برچسب گذاری شده $f_l(x,y)$ را می‌توان به صورت زیر تعریف نمود:

$$H_i = \sum_{x,y} I\{f_l(x,y) = i\},\ i = 0, \dots, n-1 \qquad (۱-۳)$$

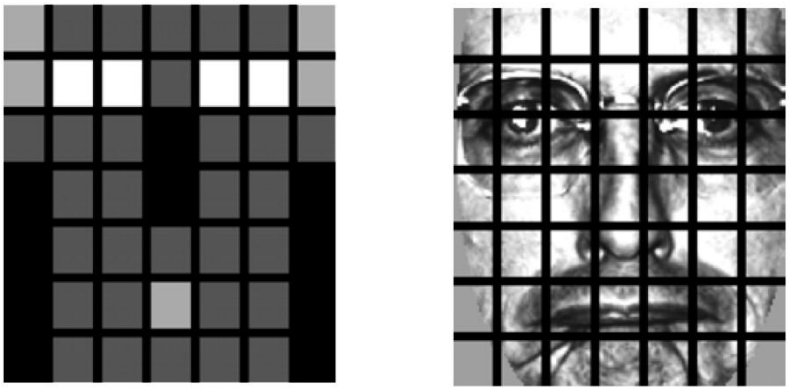
که در آن n تعداد برچسب‌های مختلف می‌باشد که با عملگر LBP تولید شده است و

$$I\{A\} = \begin{cases} 1, A \text{ is true} \\ 0, A \text{ is false} \end{cases}$$

این هیستوگرام حاوی اطلاعاتی در مورد توزیع الگوهای کوچک، همانند لبه‌ها، نقطه‌ها و مناطق تخت در کل تصویر می‌باشد. برای نمایش کارای چهره، باید اطلاعات مکانی را نگهداری شود. برای این هدف، تصویر به مناطق $R_0, R_1, \dots, R_{m-1}$ تقسیم بندی می‌شود (شکل ۳-۱۵) و هیستوگرام بهبودیافته به صورت زیر تعریف می‌شود

$$Hi,j = \sum_{x,y} I\{fl(x,y) = i\}\, I\{(x,y) \epsilon Rj\}, i = 0, \dots, n-1, j = 0, \dots, m-1 \qquad (۲-۳)$$

در این هیستوگرام، به طور مؤثر توصیفی از چهره را در سه سطح مختلف محلی داریم: برچسب‌های هیستوگرام حاوی اطلاعاتی درمورد الگوهای در سطح پیکسلی است، برچسب‌های یک منطقه کوچک با هم جمع می‌شود تا اطلاعاتی را در حد منطقه‌ای تولید کند و هیستوگرام‌های منطقه‌ای به یکدیگر الحاق می‌شود تا یک توصیف سراسری از چهره ارائه دهد.



شکل ۳-۱۴: الف) نمونه‌ای از تصویر چهره که به پنجره ۷×۷ تقسیم بندی شده است. ب) وزن‌هایی که برای معیار عدم شباهت X2 تعیین شده اند. مربعات مشکی وزن ۰، خاکستری تیره ۱، خاکستری روشن ۲ و سفید ۴ می‌باشد.

از دید دسته بندی الگو، یک مسئله معمول در تشخیص چهره زمانی است که کلاس‌ها بسیار زیاد باشد و از طرف دیگر، تنها یک مقدار بسیار کم، احتمالاً تنها یک نمونه آموزشی برای هر کلاس است. به همین دلیل، دسته بندهای پیچیده تر مورد نیاز نیستند و دسته بند نزدیک‌ترین همسایگی استفاده می‌شود. چندین معیار عدم شباهت ممکن برای مقایسه هیستوگرام‌ها پیشنهاد شده است:

اشتراک هیستوگرام:

$$D(S, M) = \sum_i \min(S_i, M_i)$$

آمار لگاریتمیک:

$$L(S, M) = -\sum_i S_i \log M_i$$

آمار Chi-square:

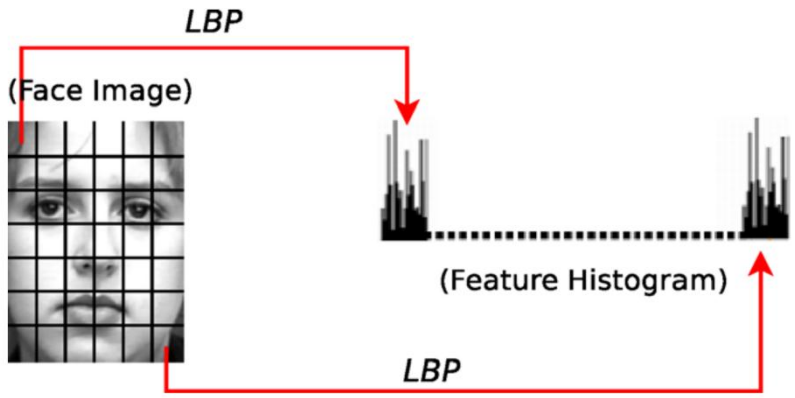
$$X2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i}$$

تمام این معیارها را با انتگرال گیری بر روی i و j می‌توان به هیستوگرام بهبودیافته توسعه داد. هرگاه تصاویر به مناطقی تقسیم بندی شدند، انتظار می‌رود برخی مناطق اطلاعات بیشتری نسبت به بقیه داشته باشد و در تمایز بین افراد تأثیر بیشتری بگذارد. برای نمونه، چشم‌ها راهنمای مهمی در تشخیص چهره انسانی خواهد بود [۴۳، ۴۴]. برای بهره گیری از این موضوع، وزنی برای هر منطقه بر اساس اهمیت اطلاعاتی آن تعیین می‌شود. برای نمونه، آمار X2 موزون به صورت زیر خواهد شد:

$$X_w^2 2(S, M) = \sum_i w_j \frac{(S_{i,j} - M_{i,j})^2}{S_{i,j} + M_{i,j}}$$

که در آن w_j وزنی برای منطقه j می‌باشد.

بنابراین، در الگوریتم هیستوگرام‌های الگوهای دودویی محلی، تصویر به مناطقی همانند شکل ۳-۱۵ تقسیم بندی می‌شود. الگوهای دودویی محلی هر منطقه محاسبه می‌شود و از هر منطقه یک هیستوگرام بدست می‌آید. با الحاق هیستوگرام‌های حاصل از هر منطقه، یک هیستوگرام کلی برای کل تصویر بوجود می‌آید. برای مقایسه و یا دسته بندی تصاویر، هیستوگرام‌های نهایی با استفاده از معیارهای عدم شباهت بررسی می‌شوند.



شکل ۳-۱۵: تصویر چهره به مناطق کوچکی تقسیم می‌شود. هیستوگرام‌های LBP هر منطقه استخراج می‌شود و نهایتاً در یک بردار ویژگی به یکدیگر الحاق می‌شوند.

۳-۴. نتیجه گیری

در این فصل، عملکرد سیستم عملیاتی تشخیص چهره تحت وب تشریح شد. فرایند آموزش چهره‌های جدید در سیستم، از سمتِ کلاینت توسط اپلت افزودن کاربر قابل اجرا است. فرایند آزمایش سیستم نیز از سمتِ کلاینت توسط اپلت تشخیص کاربر قابل اجرا است. دو اپلت مذکور، تصاویر را از دوربین وب‌کم دریافت می‌کنند و فرایند پیش پردازش و چهره یابی را انجام می‌دهند. در نهایت، تصویر چهره را به سرور جاوا ارسال می‌کند. پایگاه داده تصاویر آموزش داده شده، در سرور قرار دارد. سرور جاوا عملیات‌های اصلی تشخیص چهره را بر روی تصویر دریافتی اعمال می‌کند و پاسخ متناسب را به اپلت برمی گرداند. عملیات‌های اصلی سیستم، الگوریتم چهره یابی Viola-Jones و الگوریتم تشخیص چهره هیستوگرام‌های الگوهای دودویی محلی می‌باشد. این دو نیز در بخش‌های ۳-۳ و ۳-۲ تشریح شد.

فصل چهارم

روش پیشنهادی

۴-۱. مقدمه

بیش از سه دهه است که موضوع تشخیص چهره، محققان و علاقه مندان بسیاری را از سراسر کشور جذب نموده است، اما الگوریتم‌های مختلف تشخیص چهره، با چهره‌ها همانند اشیاء یا بافت‌ها رفتار کرده اند. اگر به آنالیز چهره توسط سیستم بینایی انسان توجه نماییم، علاوه بر رویکردهای معمول با رویکرد متفاوتی نیز متوجه خواهیم شد.

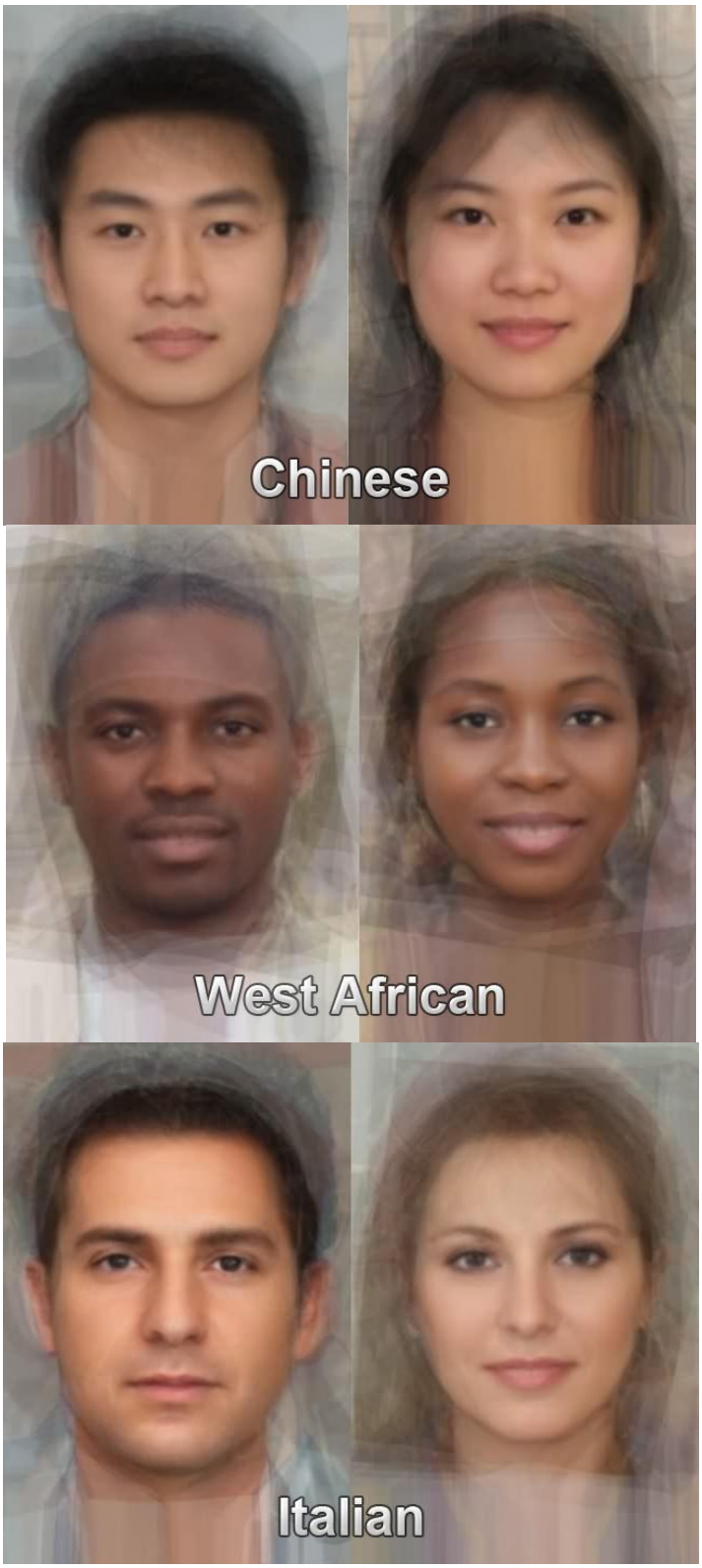
جهت طراحی سیستم تشخیص چهره‌ای که قدرت سیستم تشخیص چهره انسانی را داشته باشد، می‌توان از رویکرد انسانی در طراحی سیستم استفاده کرد. تشریح جزئیات جنبه‌های مختلف مکانیزم تشخیص چهره انسانی مبحث داغی در روانشناسی شناختی است و به طور کامل انجام نشده است. ما متوجه شدیم که مردم به تمییزدهنده‌ترین ویژگی‌های یک چهره در توصیفات خود اشاره می‌کنند. سیستم تشخیص چهره انسانی یک مجموعه از ویژگی‌های اسمی را استخراج می‌کند که برخی از آنها قابل توصیف باشد. برای درک بهتر مطلب، از یک فرد بخواهید تا چهره شخصی را توصیف نماید. مردم معمولاً چهره‌ها را با جملاتی نظیر "دماغ باریکی دارد" یا "چشمان درشتی دارد" یا "دهان کوچکی دارد" و ... توصیف می‌کنند. آنچه ما دریافتیم آنست که این توصیفات بر اساس تمییزدهنده‌ترین ویژگی هاییست که نسبت به چهره میانگین شکل گرفته در ذهن فرد می‌باشد. به طور خاص، مردم در منطقه جغرافیایی خاص در ذهن خود یک چهره میانگینی از چهره‌های افراد آن منطقه ساخته اند و هرگاه می‌خواهند چهره‌ای را توصیف کنند، به تمییزدهنده‌ترین ویژگی‌های آن چهره نسبت به چهره میانگین اشاره می‌کنند. اگر تمام افراد آن منطقه چشم‌های گِرد و دایره‌ای داشته باشند و شخص مورد سؤال نیز چشم‌های گردی داشته باشد، داشتن چشم گِرد برای شخص مورد سؤال، ویژگی تمییزدهنده محسوب نمی شود و نمی توان شخص را با این ویژگی‌ها توصیف نمود اما اگر مردم آن منطقه، بینی باریکی داشته باشند درحالیکه شخص مورد سؤال بینی پهنی داشته باشد، این ویژگی بینی می‌تواند یک ویژگی

تمییزدهنده برای تشخیص آن شخص باشد. بنابراین، تمییزدهنده‌ترین ویژگی‌های چهره نسبت به چهره میانگین، ابزاری برای توصیف چهره یک شخص خواهد بود.

قلب ایده روش پیشنهادی، محاسبه میزان اختلاف و انحراف تصاویر چهره اشخاص مختلف از چهره میانگین می‌باشد. در شکل ۴-۱ تصاویری از چهره میانگین اشخاص چند کشور مختلف آورده شده است. این تصاویر توسط شرکت Face of Tommorrow با میانگین گیری از تصاویر چهره مجموعه‌ای از مردم آن کشور بدست آمده است. هدف از نشان دادن این تصاویر آنست که یک مفهوم عینی از چهره میانگین در ذهن شما خواننده عزیز شکل بگیرد.

در شکل ۴-۱، تصاویر چهره میانگین مجموعه مردها و زن‌ها به صورت جداگانه محاسبه شده است تا ویژگی‌های هر جنسیت حفظ شود. نکته دیگر آنست، که چهره میانگین مد نظر ما، لزوماً به این روش بیان -represent– نمی شود، هر چند از آنجا که روش پیشنهادی انعطاف پذیر طراحی شده است، می‌توان این تصاویر چهره میانگین را نیز بکار برد. تصاویر شکل ۴-۱ تنها ارائه‌ای برای درک عینی چهره میانگین می‌باشد. چهره میانگینی که ما در روش پیشنهادی استفاده نمودیم، قابل به تصویر کشیدن نیست و در قالب فرمول‌های ریاضی بیان می‌شود.

در ادامه فصل، رویکرد را فرموله می‌نماییم و تأکید بسیاری بر روی تمییزدهنده‌ترین ویژگی‌ها قرار می‌دهیم. مراحل استخراج و پردازش ویژگی‌ها و دسته بندی را با جزئیات شرح می‌دهیم. همچنین، مکانیزم آموزش و آزمایش الگوریتم پیشنهادی را نیز شرح داده شده است.



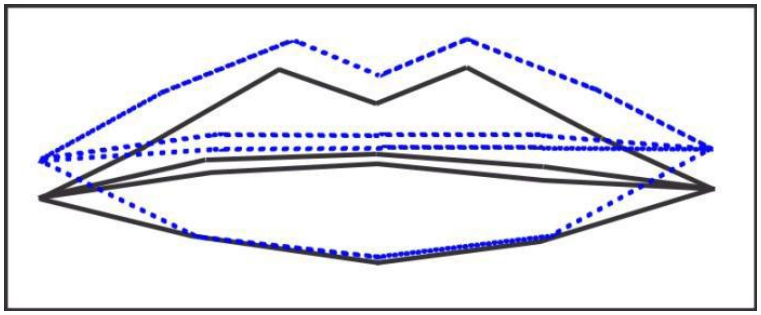
شکل ۴-۱: تصاویر چهره میانگین آقایان و خانم‌های سه ملیت

۴-۲. روش پیشنهادی

از آنجا که هدف آنست که عملکرد سیستم پیشنهادی خود را به سیستم تشخیص چهره انسانی نزدیک نماییم، لذا روش پیشنهادی باید ویژگی‌های اسمی تمییز دهنده را برای رجوع به چهره‌های مختلف فرموله کند. روش پیشنهادی همچنین باید ویژگی‌های چهره (همانند چشم‌ها، بینی، دهان، فک) متناظر از تمام اشخاص پایگاه داده را با یکدیگر مقایسه می‌کنند. بنابراین، چهره‌های پایگاه داده باید آنالیز شود و ویژگی‌های چهره از چهره‌ها باید استخراج شود. با استخراج ویژگی‌های چهره، مرزهای ابروها، چشم‌ها، بینی، دهان و کل چهره تمام اشخاص بدست می‌آید. در گام بعدی، ما از ویژگی‌های چهره متناظر تمام اشخاص میانگین گیری می‌کنیم؛ از نتیجه میانگین گیری، یک چشم راست میانگین پایگاه داده، یک چشم چپ میانگین پایگاه داده، یک بینی میانگین پایگاه داده، یک دهان میانگین پایگاه داده و یک فک میانگین پایگاه داده بدست می‌آید. ویژگی‌های چهره میانگین را به روش‌های مختلفی می‌توان تعریف نمود. در این پایان‌نامه، ما ویژگی‌های چهره را با نقاطی بر روی مرز ویژگی چهره مشخص می‌نماییم. از آنجاکه استخراج ویژگی‌های چهره در روش پیشنهادی بحث اصلی نمی باشد، درگیر جزئیات این کار نمی شویم و به همین خاطر از تصاویر پایگاه داده IMM Frontal Face که حاشیه نویسی شده[1] است، استفاده می‌کنیم (شکل ۴-۲). همانطور که در شکل مشاهده می‌نمایید، نقاط مرز دهان را مشخص می‌کنند. ما میانگین ویژگی‌های چهره را با محاسبه میانگین مؤلفه‌های نقاط متناظر بدست می‌آوریم (شکل ۴-۳). ذکر این نکته ضروری است که روش پیشنهادی مستقل از روش استخراج ویژگی و نوع ویژگی هاست و حتی تعریف میانگین ویژگی‌های چهره نیز قابل انعطاف است و می‌توان تعاریف متعدد دلخواهی را برحسب ویژگی‌های مستخرج ارائه داد. در این پایان‌نامه ما راهی را برای استخراج ویژگی‌ها و محاسبه میانگین ویژگی‌های چهره انتخاب کردیم اما این لزوماً تنها راه موجود و بهترین راه نیست.

[1] Annotated

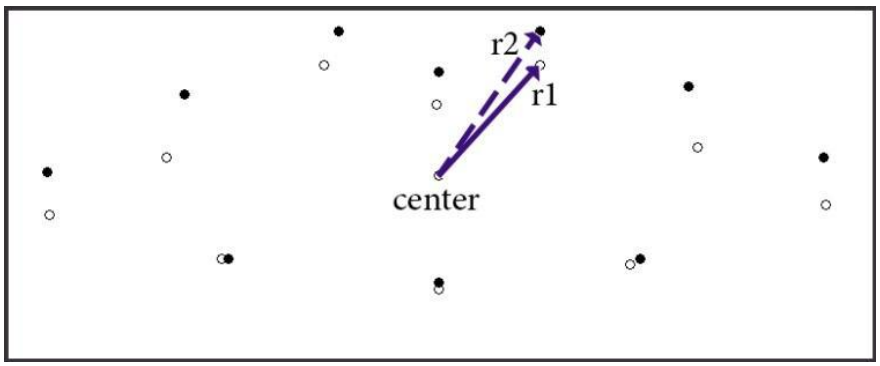
شکل ۴-۲: یک دهان حاشیه نویسی شده از پایگاه داده IMM Frontal Face



شکل ۴-۳: دهان با خطوط مستقیم، دهان شخصی (شخص شکل ۴-۲) از پایگاه داده IMM Frontal Face را مشخص می‌کند. دهان با خطوط شکسته دهان میانگین اشخاص پایگاه داده را مشخص می‌کند. دهان میانگین از میانگین گیری مؤلفه‌های نقاط متناظر تمام اشخاص پایگاه داده بدست می‌آید.

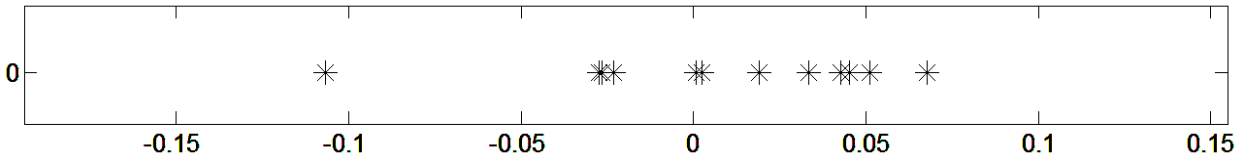
پس از محاسبه ویژگی‌های چهره میانگین، معیار انحراف از ویژگی چهره میانگین را تعریف می‌کنیم؛ با این معیار، انحراف ویژگی چهره (مثلاً دهان) هر شخص از پایگاه داده نسبت به ویژگی چهره میانگین (در این مثال، دهان میانگین) را می‌توان محاسبه نمود و انحرافات را به عنوان ویژگی‌های تمییز دهنده در نظر گرفت. این معیار انحراف را به طرق مختلف می‌توان تعریف نمود. در تعریفی که ما اینجا به کار بردیم، اختلافات میان فواصل اقلیدسی نقاط متناظر از مرکز محاسبه می‌شود؛ منظور از مرکز، نقطه‌ای است که از میانگین گیری تمام نقاط موجود در ویژگی چهره مربوطه بدست می‌آید (شکل ۴-۴). تعریف معیار انحراف بسیار در عملکرد شناسایی و تشخیص چهره نقش دارد. تعریف معیار انحراف باید اطلاعات برجسته از ویژگی چهره همانند کشیدگی، گردی، فرم خاصِ ویژگی‌های چهره را استخراج کند. در اینجا، تعریف

معیار انحراف برای همه ویژگی‌های چهره (چشم‌ها، ابروها، بینی، دهان و فک) یکسان در نظر گرفته شده است اما می‌توان برای هر ویژگی، تعریف منحصر به فردی ارائه نمود.



شکل ۴-۴: هر شخص از پایگاه داده توسط انحراف ویژگی‌های آن از ویژگی چهره میانگین مربوطه توصیف می‌شود. در این شکل، شما می‌توانید نقاط مشخص کننده دهان شخص ۱ (شکل ۴-۲) را با دایره‌های توخالی و همچنین دهان میانگین پایگاه داده را با دایره‌های توپر مشاهده نمایید.

انحراف هر ویژگی چهره (چشم‌ها، ابروها، بینی، دهان و فک) مربوط به هر شخص از ویژگی چهره میانگین مربوطه براساس یک معیار تعریف شده خاص محاسبه می‌شود. معیار تعریف شده یک عدد حقیقی را تولید خواهد کرد. ما انحراف بینی هر شخص پایگاه داده را از بینی میانگین پایگاه داده را محاسبه نموده ایم و نتایج را در شکل ۴-۵ رسم نموده ایم.

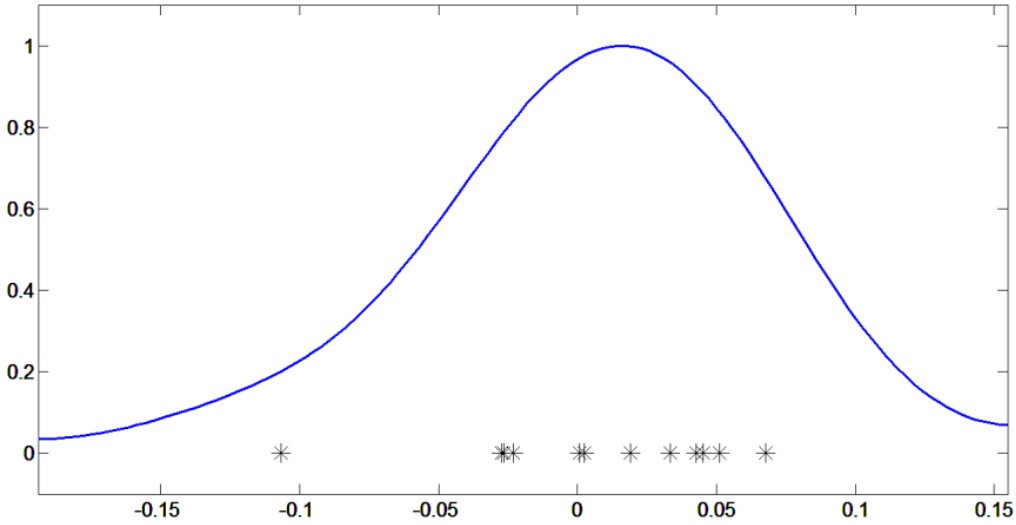


شکل ۴-۵: انحرافات دهان تمام اشخاص از دهان میانگین پایگاه داده.

در شکل فوق، هر ستاره ویژگی یک شخص از پایگاه داده را نشان می‌دهد. توجه داشته باشید مجموع تمام انحرافات لزوماً برابر با صفر نمی شود زیرا در تعریف در نظر گرفته شده برای میانگین، خاصیت تقارنی وجود ندارد. در این پایان‌نامه، ما به مقداری اهمیت می‌دهیم که بیشتر از بقیه از میانگین انحراف

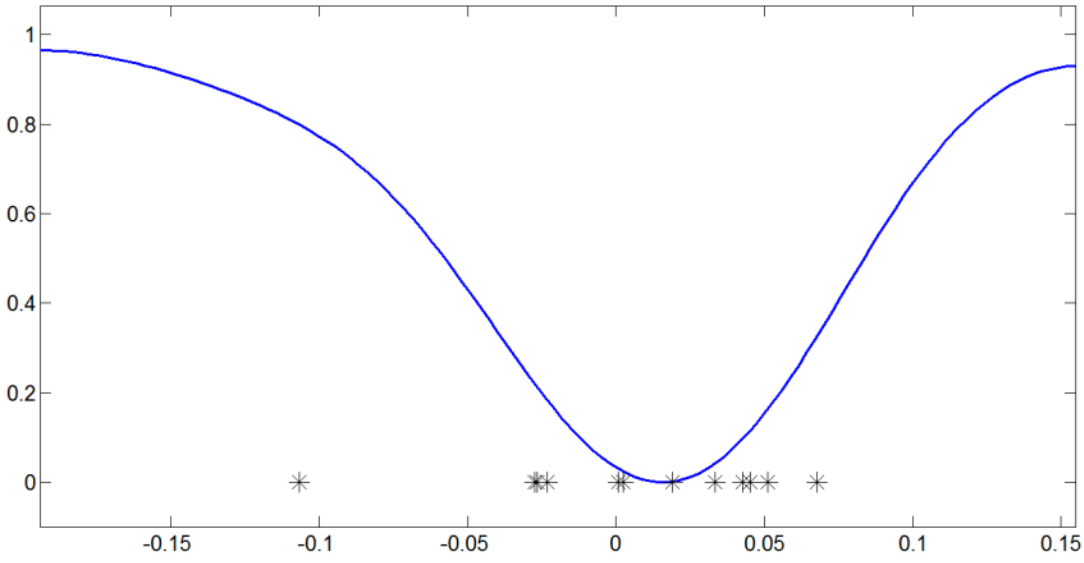
داشته باشد. اگر اختلاف بینی یک شخص از بینی میانگین به طور قابل ملاحظه‌ای زیاد باشد، آنگاه بینی یک ویژگی مناسب برای توصیف آن شخص خواهد بود. همانطور که بخاطر دارید، ما به دنبال شبیه سازی سیستمی هستیم که در آن به دنبال برخی از تمییزدهنده‌ترین ویژگی‌های اشخاص هستیم. شخصی را در نظر بگیرید که چشم‌ها، ابروها، بینی او نزدیک به ویژگی چهره میانگین مربوطه می‌باشد اما ویژگی دهان او به طور قابل توجهی نسبت به دیگر اشخاص از دهان میانگین انحراف دارد. در این حالت، ویژگی دهان یک ویژگی مناسب برای توصیف این شخص خواهد بود.

پس از محاسبه انحرافات از میانگین، ما تابع چگالی مقادیر انحرافات را به فرم غیرپارامتریک تخمین می‌زنیم. تخمین‌های چگالی کرنل، بسیار به هیستوگرام‌ها نزدیک می‌باشد و مشخصاتی همانند نرمی یا پیوستگی را با یک کرنل مناسب می‌توان تنظیم نمود. در اینجا، ما کرنل گاوسی را برای تخمین تابع چگالی انتخاب نمودیم. پارامتر پهنای باند، یکی از پارامترهای نرم کنندگی برای تابع تخمین چگالی است. پارامتر پهنای باند در اینجا به صورت اتوماتیک مقداردهی می‌شود. همانطور که در شکل ۴-۶ مشاهده می‌نمایید، تابع توسط مقادیر انحرافات تمام اشخاص تخمین زده می‌شود. تخمین به ما نشان می‌دهد که تابع در صورت افزایش جمعیت به چه شکل در خواهد آمد. محور عمودی همچنین سطح اهمیت انحرافات از میانگین را نیز نشان می‌دهد. از آنجا که روش پیشنهادی بیان می‌کند که اشخاص با انحراف بیشتر از میانگین، ارزش بیشتری نسبت به اشخاص با انحراف کمتر از میانگین خواهند داشت، نمودار ذیل به نوعی این حقیقت را نمایش خواهد داد.



شکل ۴-۶: چگالی تخمین زده شده برای مقادیرانحرافات دهان تمام اشخاص از دهان میانگین پایگاه داده.

سطح اهمیت را می‌توان با معکوس کردن تابع چگالی تخمین زده شده بدست آورد. همانطور که در شکل ۴-۷ مشاهده می‌کنید، اگر انحراف دهان یک شخص از دهان میانگین افزایش یابد، سطح اهمیت آن متعاقباً افزایش خواهد یافت.



شکل ۴-۷: تابع چگالی معکوس شده نماینده سطح اهمیت مقادیر مختلف انحرافات خواهد بود.

حال به همین روش یک تابع چگالی تخمین زده شده برای هر ویژگی چهره (چشم راست، چشم چپ، ابروی راست، ابروی چپ، دهان و فک) می‌سازیم. توجه داشته باشید که توابع چگالی تخمین زده شده برای ویژگی‌های چهره می‌توانند multimodal باشد و این ویژگی چندمُدی، اختلالی در روند متد پیشنهادی ایجاد نخواهد کرد.

۲-۳. فرایند آزمایش[1]

مکانیزم دسته بند ما، مشارکت[2] اصلی پایان‌نامه می‌باشد و تشریح می‌کند که تصویر چهره آزمایشی چگونه با تصاویر چهره پایگاه داده مقایسه می‌شود. در این بخش، مسئله تشخیص چهره را با دو دسته بند دیگر با نام‌های 1-NN و 1-NN بهبود یافته حل می‌نماییم تا قدرت روش پیشنهادی نشان داده شود. مزایا و معایب سه روش مذکور را بیان می‌کنیم و جزئیات را با یک مثال تشریح می‌کنیم. مقایسه گام به گام عملکرد دسته بند پیشنهادی با 1-NN و نسخه بهبودیافته 1-NN ما را در درک فرایند مکانیزم کمک می‌کند.

در ابتدا، ما جدولی همانند جدول ۴-۱ می‌سازیم که حاوی امتیازات صفر تمام اشخاص می‌باشد. امتیاز هر شخص تنها با توجه به شباهت تصویر آزمایش و شخص مقایسه شونده افزایش/کاهش می‌یابد. در تشریح عملکرد سه متد فوق الذکر، جدول ۴-۱ برای انباشت امتیازات لازم می‌شود.

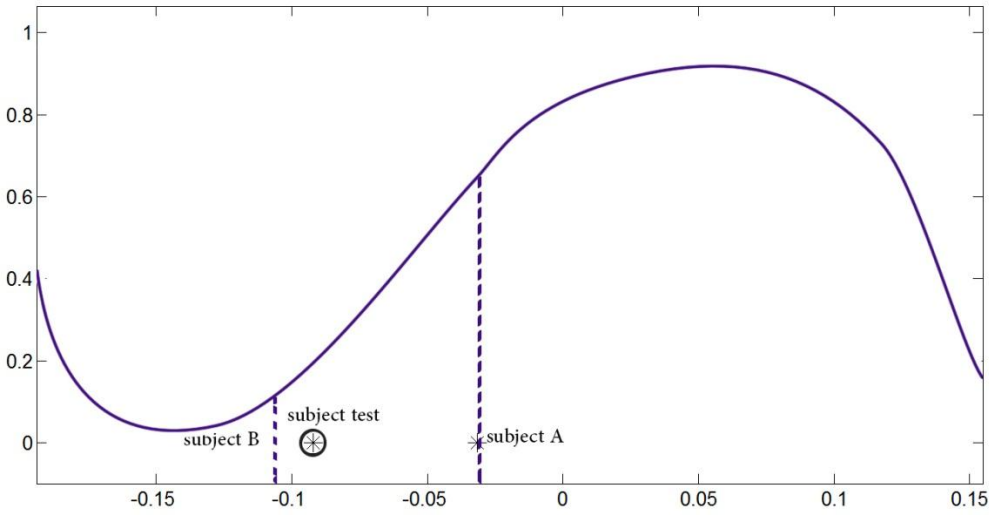
جدول ۴-۱: امتیازات تمام اشخاص با صفر مقداردهی می‌شود.

اشخاص	S_1	S_2	S_3	...	S_{n-1}	S_n
کل امتیازات	0	0	0	...	0	0

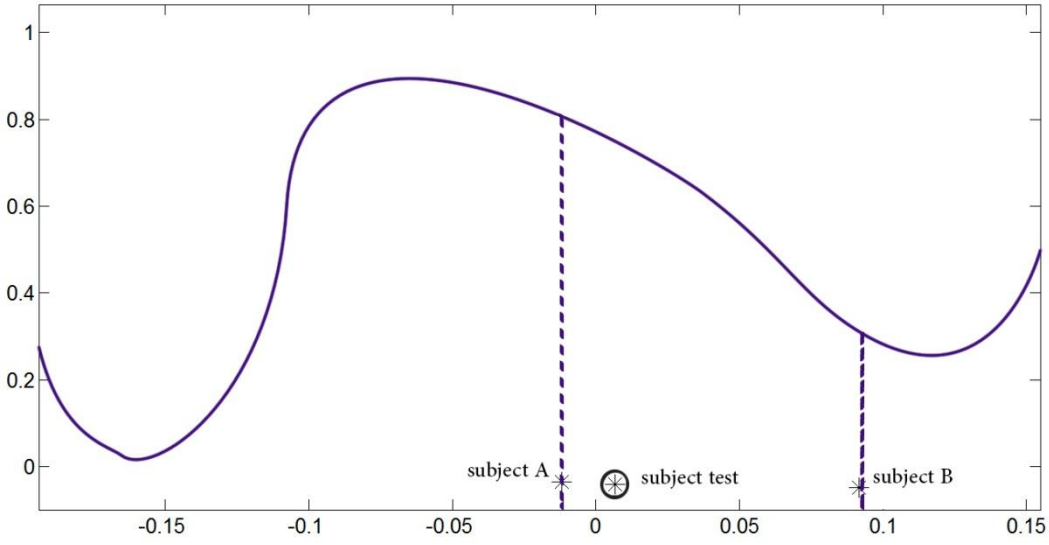
[1] Test
[2] Contribution

پیش از آغاز آنالیز آماری بر روی تصویر چهره آزمایشی، ویژگی‌های چهره تصویر تست باید استخراج گردد. پس از بدست آوردن مرزها یا نقاط ویژگی‌های چهره، انحراف هر ویژگی چهره را از ویژگی چهره میانگین که در مرحله آموزش بدست آمد، محاسبه می‌کنیم.

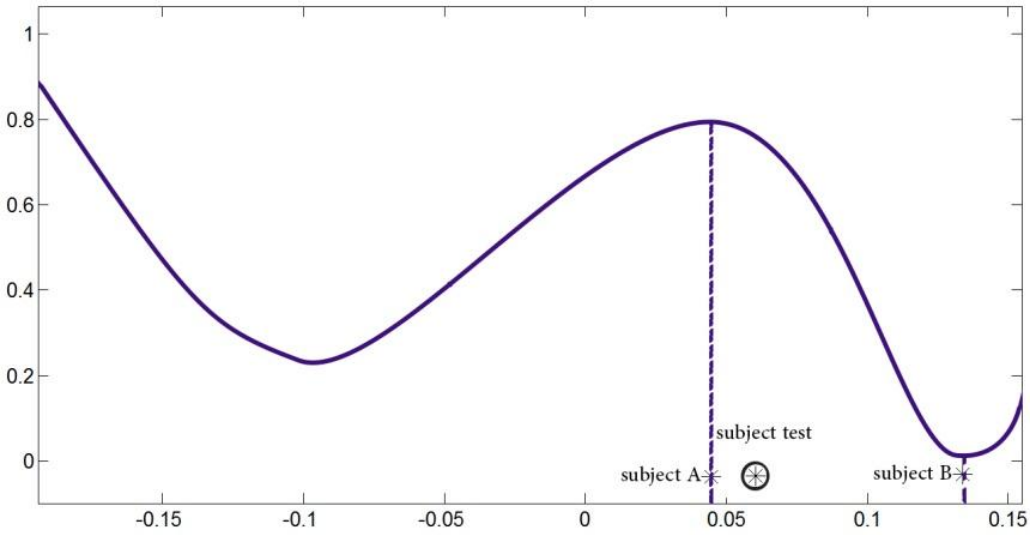
یک سیستم دسته بندی را با ۳ ویژگی (شکل ۴-۸ و شکل ۴-۹ و شکل ۴-۱۰) در نظر بگیرید. برای راحتی در محاسبات، تنها دو شخص در پایگاه داده آموزش درنظرگرفته شده است.



شکل ۴-۸: انحرافات ویژگی چشم راست با تابع تخمین چگالی.



شکل ۴-۹: انحرافات ویژگی بینی با تابع تخمین چگالی.



شکل ۴-۱۰: انحرافات ویژگی دهان با تابع تخمین چگالی

- دسته بند 1-NN ساده

در حالتی که دسته بند اصلی یک دسته بند ساده 1-NN باشد، هرگاه تصویر تست وارد شود، ویژگی‌های چهره شخص تست استخراج می‌شود و با هرکدام از اشخاص پایگاه داده آموزش مقایسه می‌شود. در دسته بند 1-NN، شخص آزمایشی با نزدیکترین فاصله با بیشترین تعداد برچسب گذاری

می‌شود. برای دسته بندی کردن شخص تست، مقایسه شخص آزمایشی و اشخاص درون پایگاه داده باید به صورت ویژگی به ویژگی بررسی شود.

در الگوریتم ساده دسته بندی 1-NN، نزدیکترین شخص آموزشی به شخص تست با بیشترین فراوانی در هر ویژگی مطلوب است. در مثال ارائه شده، نزدیکترین شخص به شخص تست در شکل ۴-۸، شخص B می‌باشد و نزدیکترین شخص به شخص تست در شکل ۴-۹، شخص A می‌باشد و نزدیکترین شخص به شخص تست در شکل ۴-۱۰ شخص A می‌باشد. از آنجاکه شخص تست با نام نزدیکترین شخص آموزشی با بیشترین فراوانی برچسب گذاری می‌شود، نتیجه می‌شود که شخص تست، شخص A شناخته می‌شود.

جدول ۴-۲: دسته بند 1-NN برای ویژگی‌های مستخرج از چهره

شماره ویژگی	نام نزدیکترین شخص آموزشی
ویژگی اول	B
ویژگی دوم	A
ویژگی سوم	A

- دسته بند 1-NN بهبودیافته

از آنجا که در دسته بند 1-NN ساده، تنها تعداد برچسب‌ها مهم است و برخی اطلاعات بدون استفاده باقی می‌ماند، ضرورت استفاده از یک دسته بند دیگر آشکار می‌شود. یک نوع بهبودیافته الگوریتم 1-NN شامل فواصل میان ویژگی‌های شخص تست و تمام اشخاص آموزشی هم می‌شود. در اینجا، برای فواصل

امتیاز منفی بکار می‌بریم؛ فواصل علامت منفی دارند زیرا هرچقدر فاصله‌ها کمتر باشند، مقدار منفی کمتری به آن اختصاص می‌یابد. ما فرم منفی فواصل را به امتیازات کلی اشخاص آموزشی می‌افزاییم.

هرگاه یک تصویر تست وارد می‌شود، فاصله شخص تست با هر شخص آموزشی در هر ویژگی محاسبه می‌شود. شخص تست با نام شخص آموزشی با کمترین مجموع امتیازات منفی برچسب گذاری می‌شود. همانطور که در جدول ذیل مشاهده می‌کنید، در ستون دوم، فرم منفی فواصل میان شخص تست و هر شخص آموزشی (در این مثال، ۲ شخص آموزشی) وجود دارد. در ستون سوم، نتیجه انباشتی از ستون دوم وجود دارد. در ردیف آخر، A مجموع امتیاز منفی کمتری نسبت به B دارد، بنابراین، شخص تست به عنوان شخص A شناخته می‌شود.

جدول ۴-۳: دسته بند 1-NN بهبودیافته برای ویژگی‌های مستخرج از چهره

شماره ویژگی	فرم منفی فواصل میان شخص تست و اشخاص آموزشی	مجموع نتیجه انباشتی
ویژگی ۱	A = -0.05, B = -0.01	A = -0.05, B = -0.01
ویژگی ۲	A= -0.02, B = -0.085	A= -0.07, B = -0.095
ویژگی ۳	A= -0.16, B = -0.76	A= -0.28, B = -0.865

- دسته بند پیشنهادی

1-NN بهبودیافته با دخالت دادن فواصل میان شخص آزمایشی و اشخاص آموزشی مزیتی نسبت به 1-NN ساده دارد، اما هنوز با مسئله عدم استفاده از برخی داده‌های موجود مواجه است. روش پیشنهادی نه تنها شامل فواصل میان شخص تست و هر شخص آموزشی می‌شود، بلکه شامل تابع تخمین چگالی نیز می‌شود. روش پیشنهادی از این ایده حمایت می‌کند که اگر انحراف یک ویژگی شخص آموزشی از بقیه

بیشتر باشد، به عنوان ویژگی تمییزدهنده درنظرگرفته می‌شود. این سطح اهمیت با معکوس تابع تخمین چگالی تعریف می‌شود. تابع تخمین چگالی معکوس این حقیقت را بیان می‌کند که ویژگی‌های در محدوده با چگالی کم هستند، دارای اهمیت بیشتری نسبت به ویژگی‌های در محدوده با چگالی بیشتر هستند؛ این حقیقت از عامل تمییزدهی که پیشتر مطرح شد، برمی آید.

ازآنجا که ما می‌خواهیم ویژگی‌های مهم تر، امتیاز بیشتری کسب کنند و یا امتیازات منفی کمتری کسب کنند، تابع تخمین چگالی برای هر ویژگی تمام اشخاص پایگاه داده محاسبه می‌شود و در فرم منفی فواصل ضرب می‌شود. برای نمونه، مقدار ویژگی ۱ (شکل ۴-۸) شخص A در تابع تخمین چگالی 0.23 می‌باشد که درفاصله منفی میان ویژگی اول شخص تست و ویژگی اول هر شخص آموزشی ضرب می‌گردد. این عملیات برای ویژگی‌های دیگر هم صدق می‌کند. در پایان، شخص تست با نام شخص آموزشی با بیشترین مجموع مقدار برچسب گذاری می‌شود. در جدول ذیل می‌توان مشاهده نمود که B مجموع امتیاز منفی کمتری نسبت به A دارد و در نتیجه، شخص تست به عنوان B شناخته می‌شود.

جدول ۴-۴: دسته بند پیشنهادی برای ویژگی‌های مستخرج از چهره

شماره ویژگی	فرم منفی فواصل × مقدار تابع چگالی	مجموع نتیجه انباشتی
ویژگی ۱	A = -0.05×0.77, B = -0.01×0.22	A= -0.038,B = -0.0022
ویژگی ۲	A= -0.02×0.924, B = -0.085×0.94	A= -0.0565,B = -0.0822
ویژگی ۳	A= -0.16×0.894 ,B = -0.76×0.1	A= -0.1995 ,B = -0.1582

هدف از آزمایشات فوق، بیان نحوه عملکرد دسته بند پیشنهادی و آشکار کردن تفاوت‌های دسته بند پیشنهادی با دسته بند 1-NN و دسته بند 1-NN بهبودیافته می‌باشد. نتیجه نهایی دسته بند پیشنهادی متفاوت از دسته بند 1-NN و دسته بند 1-NN بهبودیافته می‌باشد. از آنجا که داده‌های استفاده شده در مثال فوق، داده‌های واقعی نیستند، نمی توان جمع بندی نهایی ارائه نمود ولی این موضوع برداشت می‌شود که دسته بند پیشنهادی با استفاده از دو اصل مهم به نتیجه کاملاً متفاوتی دست یافت:

- اصل اول: استفاده از همه اطلاعات داده ها
- اصل دوم: برجسته کردن اطلاعات مهم

برای محک روش پیشنهادی با داده‌های واقعی در فصل پنجم، پایگاه داده IMM Frontal Face را بر روی روش پیشنهادی اعمال می‌کنیم و نتیجه را اعلام می‌نماییم.

۲-۴. فرایند آموزش[1]

آموزش به دو صورت تعریف می‌شود:

- حالتی که تمام اشخاص پایگاه داده به صورت یکجا و دسته جمعی به سیستم آموزش داده می‌شود. این حالت، همان روندی است که پیشتر توضیح داده شد ولی مراحل توضیح داده شده را در قالب سه گام مطرح می‌کنیم:
  ۱. محاسبه تمام ویژگی‌های چهره میانگین با استفاده از مجموعه آموزشی.
  ۲. محاسبه انحرافات تمام ویژگی‌های چهره‌های تمام افراد آموزشی.
  ۳. محاسبه تابع چگالی مقادیر انحرافات.

[1] Training

- در حالتی دیگر، در سیستم تشخیص چهره، پیش از این چندین نفر آموزش دیده اند و چهره میانگین از پیش ساخته شده است و چهره فرد جدیدی برای آموزش به سیستم وارد شده است. در این حالت، مشخصه‌های اصلی سیستم باید تغییر یابد که در سه گام زیر بیان می‌شود:

۱. تمام ویژگی‌های چهره میانگین با استفاده از مجموعه آموزشی و همچنین ویژگی‌های چهره جدید باید مجدداً محاسبه شود.

۲. انحرافات تمام ویژگی‌های چهره‌های تمام افراد آموزشی به همراه فرد جدید باید مجدداً محاسبه شود.

۳. تابع چگالی مقادیر انحرافات باید مجدداً محاسبه شود.

۴-۵. ویژگی‌های هندسی کمکی

تا به اینجا، ما از نقاط مشخص کننده مرز ویژگی‌های چهره به عنوان ویژگی استفاده می‌کنیم (همانند شکل ۴-۲). ما می‌توانیم سیستم خود را با استفاده از دسته دیگری از ویژگی‌ها بهبود ببخشیم. از آنجا که مرزی برای هر ویژگی چهره (دهان، بینی، چشم‌ها، ابروها و فک) در پایگاه داده مورد استفاده مشخص شده است و این مرز یک شکل خاص را می‌سازد، مجموعه‌ای از ویژگی‌های هندسی را می‌توان از این اشکال استخراج نمود تا عمل دسته بندی و شناسایی بهتر انجام گیرد. شکل دهان یک شخص در ذیل به صورت نمونه آورده شده است.

شکل ۴-۱۱: مرز دهان نمونه شکلی را بوجود آورده است؛ مجموعه‌ای از ویژگی‌های هندسی را می‌توان به این شکل منتسب نمود.

همانطور که در تصویر فوق مشاهده می‌نمایید، شکل دهان دارای مجموعه‌ای از ویژگی‌های هندسی است. در این پایان‌نامه از مجموعه از ویژگی‌های شکلی استفاده شده است، در ذیل هرکدام با شرح مختصری آمده است.

۱. مساحت: تعداد دقیق پیکسل‌های ناحیه محاسبه می‌گردد.

۲و۳. مستطیل احاطه کننده: کوچکترین مستطیلی که ناحیه را احاطه می‌کند. این ویژگی دو عدد را مشخص می‌کند که طول و عرض آن مستطیل می‌باشد.

۴و۵. مرکز ثقل: برداری که مرکز وزنی آن ناحیه را مشخص می‌کند. این ویژگی شامل مؤلفه‌های x و y می‌باشد و دو ویژگی محسوب می‌شود.

۶. گریز از مرکز: گریز از مرکز به نسبت فاصله میان کانون بیضی به طول محور اصلی آن اطلاق می‌شود. مقدار آن بین ۰ و ۱ می‌باشد. یک بیضی که گریز از مرکز آن ۰ باشد، دایره محسوب می‌شود، در حالیکه بیضی با گریز از مرکز ۱ یک پاره خط محسوب می‌شود.

۷. حد و اندازه: نسبت پیکسل‌های ناحیه به پیکسل‌های درون مستطیل احاطه کننده شکل می‌باشد. مساحت شکل به مساحت مستطیل احاطه کننده شکل تقسیم می‌شود.

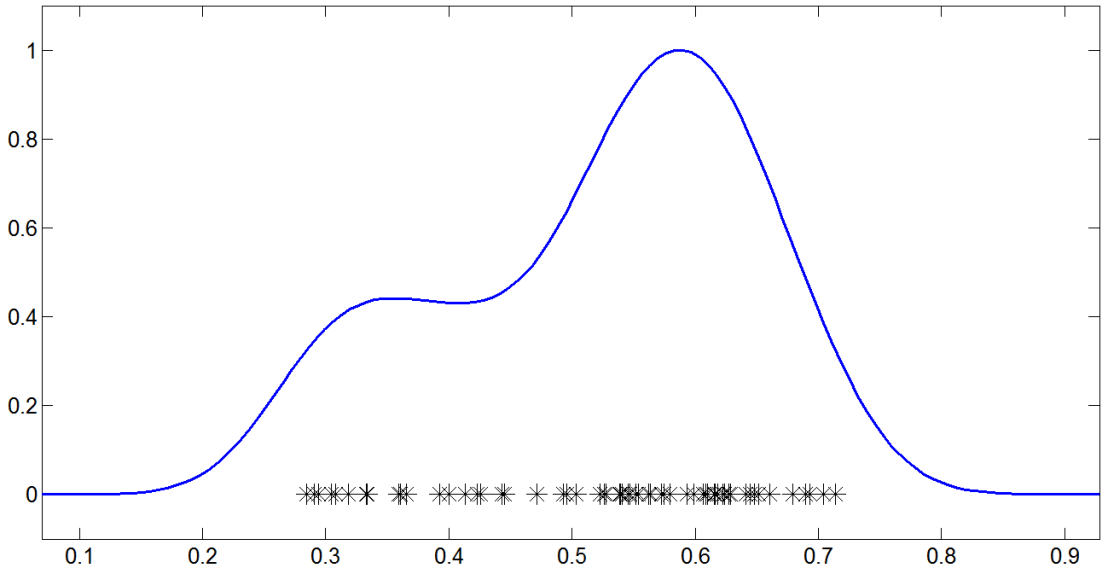
۸. قطر معادل: این پارامتر قطر دایره با اندازه مساوی با ناحیه که با فرمول sqrt(4*area/pi) محاسبه می‌شود، را مشخص می‌کند.

۹. طول محور اصلی: طول محور اصلی بیضی که دارای ممان‌های مرکزی دوم نرمال شده یکسان با ناحیه می‌باشد.

۱۰. جهت: زاویه (به درجه از ۹۰- به ۹۰+) بین محور x و محور اصلی بیضی که ممان‌های دوم برابری با ناحیه دارند.

ما برداری از ۱۰ ویژگی هندسی شکل برای هرکدام از ۷ ویژگی چهره (چشم راست، چشم چپ، ابروی راست، ابروی چپ، بینی، دهان و فک) می‌سازیم؛ بنابراین، هر شخص دارای ۷۰ ویژگی هندسی می‌باشد که به عنوان ویژگی‌های کمکی به دسته بندی و تشخیص بهتر کمک می‌کند. این ویژگی‌ها به همراه ویژگی‌های دیگر که پیشتر ارائه شد، ورودی دسته بند خواهد شد.

برای درک کاربرد ویژگی‌های هندسی، تابع تخمین چگالی ویژگی گریز از مرکز برای ابروهای راست تمام اشخاص پایگاه داده را ترسیم می‌کنیم، ویژگی‌های گریز از مرکز ابروهای راست تنها یک نمونه می‌باشد و برای ۶۹ ویژگی هندسی دیگر نیز این روند برقرار است. ۷۰ تابع تخمین چگالی برای تمام ۷۰ ویژگی هندسی وجود دارد. توجه داشته باشید که برای هرکدام از دوازده شخص مختلف، پنج نمونه وجود دارد.



شکل ۴-۱۲: تابع تخمین چگالی برای ویژگی گریز از مرکز ابروهای راست تمام اشخاص.

هرگاه تصویر شخص تست وارد شود، ویژگی‌های چهره استخراج می‌گردد. سپس، ویژگی‌های هندسی ویژگی‌های چهره باید محاسبه شود. هر ویژگی هندسی شخص تست با ویژگی‌های اشخاص درون پایگاه داده مقایسه می‌شود. فواصل میان شخص تست و اشخاص آموزشی محاسبه می‌شود. فواصل علامت منفی

می‌گیرند و سپس در مقادیر چگالی تخمین زده شده برای داده‌های آموزشی ضرب می‌شوند. شخص تست به عنوان شخص آموزش با کمترین مقادیر منفی (بیشترین امتیاز) شناخته می‌شود.

۴-۶. ویژگی‌های کمکی دیگر

ما متد پیشنهادی را آنقدر انعطاف پذیر طراحی نمودیم که هر نوع ویژگی را می‌پذیرد. به عنوان نمونه، خال‌های روی صورت را می‌توان به عنوان یک ویژگی استفاده نمود؛ سپس یک تعریف برای میانگین خال‌های صورت مطرح می‌کنیم و طبق آن از تمام اشخاص آموزشی میانگین می‌گیریم. این عملیات میانگین گیری را به انواع مختلف می‌توان تعریف نمود. میانگین گیری را می‌توان با محاسبه میانگین مختصات محل و اندازه خال‌ها محاسبه نمود. فرض کنید تنها چهار شخص از سی شخص درون پایگاه داده دارای خال بر روی صورت هستند و باقی آنها خال ندارند؛ همانطور که می‌دانید، متد پیشنهادی بسیار بر روی عامل تمییزدهی متمرکز می‌شود و این ویژگی خال می‌تواند ویژگی تمییزدهنده مناسبی برای سه شخصی که دارای خال هستند، باشد.

خال‌ها تنها یک نمونه از ویژگی‌های کمکی می‌باشد و ویژگی‌های بسیار دیگری همانند چروک بالای بینی، چروک پیشانی، کک چهره و ویژگی‌های قابل استفاده دیگر که دارای پتانسیل تمییزدهی می‌باشند نیز بهره برد.

۴-۷. نتیجه گیری

روش پیشنهادی از حقیقتی استفاده کرد که در آن سیستم تشخیص چهره انسانی بر روی تمییزدهنده‌ترین ویژگی‌ها نسبت به ویژگی‌های چهره میانگین که در ذهن فرد شکل گرفته، تمرکز می‌کند. چهره میانگین از میانگین گیری تمام چهره‌های موجود در یک منطقه جغرافیایی خاص بدست

می‌آید. این حقیقت به صورت ریاضی وار فرموله شده است. ابتدا انواع مختلف ویژگی‌ها از چهره استخراج می‌شود (در این پایان‌نامه، از ویژگی‌های حاشیه نویسی شده و ویژگی‌های هندسی استفاده شد)، سپس میانگین هر ویژگی چهره محاسبه می‌شود، انحرافات ویژگی‌های چهره مربوط به افراد مختلف از میانگین محاسبه می‌شود، سپس تابع تخمین چگالی برای مقادیر انحرافات محاسبه می‌شود. هرگاه یک شخص آزمایشی به سیستم وارد شود، انحرافات ویژگی‌های شخص آزمایشی محاسبه می‌شود و با ویژگی‌های هرکدام از اشخاص آموزشی مقایسه می‌شود. در نهایت برطبق الگویی که پیشتر ارائه شده، اعلام می‌شود که شخص آزمایش به کدام یک از شخص‌های آموزشی نزدیکتر بوده است و عملیات تشخیص انجام می‌شود.

مشارکت اصلی روش پیشنهادی شامل دو بخش می‌شود: پردازش ویژگی‌ها و دسته بندی. برای درک عملکرد دسته بند پیشنهادی، ویژگی‌های مستخرج را با دسته بند 1-NN و دسته بند 1-NN بهبودیافته دسته بندی نمودیم. در فصل پنجم، روش پیشنهادی را با یکی از مشهورترین متدهای تشخیص چهره، Eigenfaces، مقایسه می‌کنیم و نتایج را تشریح خواهیم نمود.

فصل پنجم

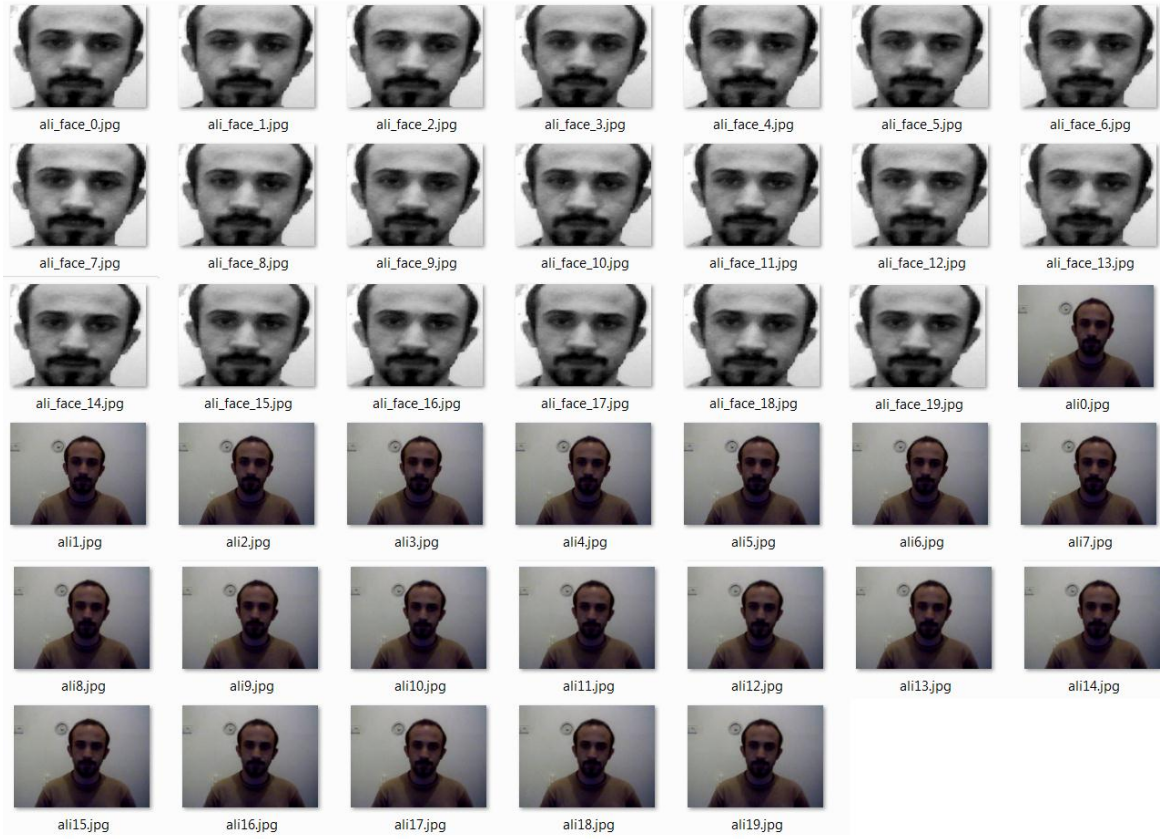
نتایج و بحث

در این فصل به ارائه نتایج آزمایشات و بحث پیرامون آن می‌پردازیم. در این پایان‌نامه، از سه پایگاه داده مختلف برای آزمایش سیستم عملیاتی تشخیص چهره تحت وب (فصل سوم) و همچنین روش پیشنهادی در تشخیص چهره (فصل چهارم) استفاده می‌نماییم. در ابتدا، سه پایگاه داده را تشریح خواهیم کرد و سپس نتایج آزمایشات را ارائه خواهیم نمود. برای محک سیستم عملیاتی تشخیص چهره تحت وب از دو پایگاه داده بهره می‌جوییم. پایگاه داده اول، توسط خود سیستم عملیاتی تشخیص چهره تحت وب تهیه شده است و پایگاه داده دوم، پایگاه داده Extended Yale Face Database B می باشد. پس از آن، پایگاه داده تصاویر IMM Frontal Face، که توسط دانشگاه فنی دانمارک تهیه شده است، را برای آزمایش روش پیشنهادی استفاده می‌کنیم.

۵-۱. پایگاه داده‌های مورد استفاده

۵-۱-۱. پایگاه داده شماره اول: پایگاه داده تصاویر سیستم تشخیص چهره عملیاتی

این پایگاه داده با ایجاد تغییراتی در خود سیستم تشخیص چهره عملیاتی که در فصل سوم ارائه شد، ضبط و اخذ شد. از ۲۰ شخص مختلف، ۲۰ تصویر مختلف از طریق وب‌کم اخذ گردید. ۱۵ تصویر از هرکدام از ۱۵ شخص از اشخاص پایگاه داده را در مجموعه آموزشی طبقه بندی نمودیم، ۵ تصویر باقی مانده از هر یک از ۱۵ شخص آموزشی را در مجموعه آزمایشی طبقه بندی نمودیم. ۵ شخص دیگر با ۲۰ تصویر از هرکدام، را برای محک سیستم تشخیص چهره در مقابل افراد غیر مجاز استفاده می‌کنیم. توجه داشته باشید که برای هر تصویر اخذ شده، نتیجه عمل چهره یابی نیز به صورت جداگانه ذخیره شده است. در ذیل می‌توانید تصاویر پایگاه داده برای یک شخص را مشاهده نمایید. پشت زمینه در تصاویر اخذ شده در افراد مختلف برای نشان دادن قدرت الگوریتم چهره یابی تغییر می‌کند.



شکل ۵-۱: نمونه تصاویر یک شخص از اشخاص پایگاه داده اول برای محک سیستم تشخیص چهره عملیاتی

همانطور که مشاهده می‌کنید، نتیجه عمل چهره یابی تصویر فایل ali0.jpg در فایل ali_face_0.jpg ذخیره شده است. بنابراین نه تنها تصاویر چهره نگهداری می‌شود بلکه تصاویر اصلی اخذ شده از وب‌کم نیز ذخیره می‌شود تا در صورت تغییر الگوریتم چهره یابی، پایگاه داده هنوز قابل استفاده باشد.

۵-۱-۲. پایگاه داده شماره دو: پایگاه داده Extended Yale Face Database B

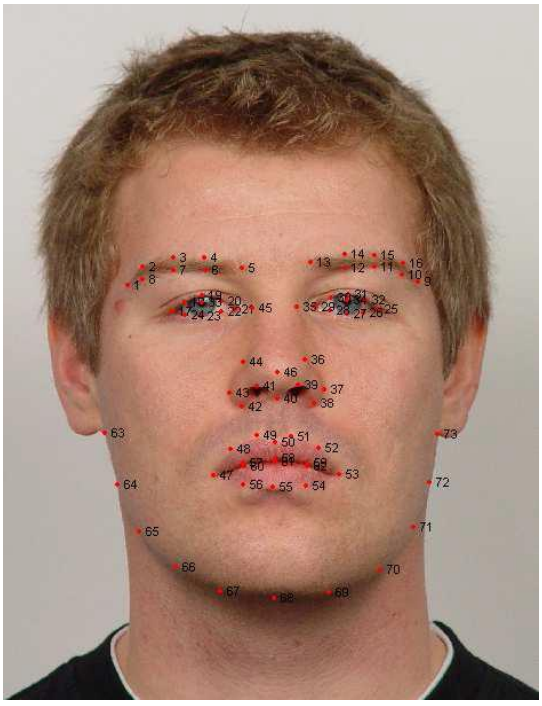
این پایگاه داده اولین بار توسط Kuang-Chih Lee، David Kriegman و Jeffrey Ho در مقاله " Acquiring Linear Subspaces for Face Recognition under Variable Lighting, PAMI, May, 2005" ارائه گردید. این پایگاه داده متشکل از ۳۷ شخص (متشکل از خانم و آقا) می‌باشد و از هر شخص

۳۱ تصویر در شرایط نوری مختلف ضبط شده است. تصاویر در قطع ۱۶۸×۱۹۲ و در مُد خاکستری می‌باشد.

پایگاه داده‌های تشخیص چهره بسیاری تا به حال ارائه شده است اما انتخاب این پایگاه داده از این رو بوده است که شرایط خاص محیط نیمه کنترل شده چهره‌ها در وب‌کم را تا حدی ارضاء می‌کند. شرایط نوری مورد انتظار بر روی چهره‌ها در جلو وب‌کم تا حدی توسط شرایط نوری تصاویر این پایگاه داده شبیه سازی می‌شود.

۵-۱-۳. پایگاه داده شماره سه: پایگاه داده تصاویر چهره IMM Frontal Face

این پایگاه داده توسط دانشگاه فنی دانمارک تهیه شده است. این پایگاه داده شامل ۱۲ شخص (همه مرد) می‌باشد. ۱۰ تصویر چهره از هر شخص دریافت گردیده است.



شکل ۵-۲: یک نمونه تصویر از پایگاه داده IMM Frontal Face با ۷۳ نقطه کمکی حاشیه گذاری شده که مکان ابروها، چشم‌ها، بینی، دهان و فک را مشخص می‌کنند.

پایگاه داده شامل چهره در جهت‌های تقریباً مستقیم در طول یک مدت زمان کوتاه می‌باشد که با تغییرات اندکی در نور و ... همراه است. به همراه تمام عکس‌ها یک دسته اطلاعات جانبی نیز وجود دارد که مکان ابروها، چشم‌ها، بینی، دهان و فک را با مجموعه‌ای از نقاط مشخص می‌کنند. برای نمونه تصویر شکل ۵-۲ این اطلاعات را مشخص می‌کند.

هر شخص ۱۰ تصویر دارد؛ ۶ تا از آنها شخص را در حالت چهره خنثی نشان می‌دهد، ۲ تا از آنها شخص را در حالت چهره شادی ملایم نشان می‌دهد، ۲ تا از آنها شخص را در حالت متفکرانه ملایم نشان می‌دهد. از آنجا که حالت چهره در اینجا مسئله اصلی نمی باشد، ۶ تصویر از هر شخص را استفاده می‌کنیم.

انتخاب این پایگاه داده از این رو بوده است که روش پیشنهادی (فصل چهارم) را بتوان با ویژگی‌های دقیق مورد آزمایش قرار داد. از آنجا که روش پیشنهادی وابسته به خود ویژگی‌ها نیست و اساساً بستری را برای پردازش ویژگی‌ها و دسته بندی ویژگی‌ها فراهم می‌کند و همچنین، دید جدیدی را در تشخیص چهره مطرح می‌کند، عملاً نتایج آزمایشات را درگیر عدم قطعیت برآمده از استخراج ویژگی ننمودیم تا به صورت خاص، روش پیشنهادی مورد آزمایش قرار بگیرد.

در ادامه به شرح معیارهایی می‌پردازیم که معمولاً در بیان نتایج آزمایشات سیستم و روش پیشنهادی بکار می‌آید. این معیارها معمولاً در حوزه شناسایی الگو به صورت استاندارد استفاده می‌شود. بنابراین پیش از ارائه نتایج، این معیارها را همراه با ارائه تعریف تشریح خواهیم کرد.

۵-۲. تعاریف مورد استفاده در نتایج

در این بخش، برخی تعاریف مورد استفاده در تحلیل نتایج را بیان نموده و سپس نتایج را بر این اساس ارائه خواهیم نمود.

۵-۲-۱. اعتبارسنجی ضربدری n-تایی

برای اطمینان از اعتبار نتایج ارائه شده و همچنین اطمینان از اینکه مجموعه داده‌های آموزشی نتایج بخش خاصی از داده‌ها با روابط خاصی نباشد، از اعتبارسنجی ضربدری n-تایی استفاده می‌شود. در این روش، بر اساس مقدار n، مجموعه پایگاه داده را به n بخش تقسیم نموده و n-1 بخش آن را جهت آموزش سیستم و یک بخش آن را جهت آزمایش سیستم استفاده می‌کنیم. جهت اطمینان از نتیجه حاصله در تمامی حالات، این عمل n مرتبه تکرار می‌شود و در هر مرتبه، یکی از n بخش پایگاه داده به عنوان بخش آزمایش سیستم انتخاب می‌شود تا تمامی بخش‌ها در بخش آزمایش سیستم بررسی شوند. نتیجه نهایی این بخش به صورت میانگین نتیجه این n بخش محاسبه و اعلام می‌شود.

برای بررسی سیستم عملیاتی، از اعتبارسنجی ۶-تایی استفاده می‌شود که در هر مرتبه مجموعه آموزش شامل ۲۵ نفر می‌باشد و ۶ نفر باقی مانده در مجموعه آزمایش قرار می‌گیرد. درصد آموزش و آزمایش نیز با توجه به n تعیین می‌شود، به این صورت که درصد داده‌های آموزشی و آزمایشی از فرمول زیر تعیین می‌گردد:

$$\text{Train ration} = \frac{n-1}{n} * 100$$

$$\text{Test ratio} = \frac{1}{n} * 100$$

که Train ratio درصد داده‌های آموزشی نسبت به کل داده‌ها و Test ratio درصد داده‌های آزمایشی نسبت به کل داده‌ها را تعیین می‌نماید که در حالت ۵-تایی نسبت داده‌های آزمایشی به آموزشی ۲۰ به ۸۰ می‌باشد.

۵-۲-۲. معیارهای ارزیابی

برای بیان کمی چگونگی عملکرد سیستم، از چند معیار اندازه گیری برای سنجش درستیِ عمل دسته بندی استفاده می‌شود. برای بیان این معیارها، ابتدا ماتریس درهم ریختگی توضیح داده می‌شود.

در ماتریس در هم ریختگی، که یک ماتریس مربعی با اندازه تعداد کلاس‌های دسته بندی مورد نظر می‌باشد. این ماتریس مشخص می‌کند که چه تعداد از کلاس‌های مورد نظر سیستم، به درستی دسته بندی شده اند. استفاده از این ماتریس همچنین مشخص می‌کند که کدامین کلاس به اشتباه در کدامین کلاس‌های دیگر دسته بندی شده اند. این اطلاعات بسیار در بحث تحلیل خطای سیستم مؤثر می‌باشد. در جدول ۵-۱ نمونه‌ای از یک ماتریس درهم ریختگی ساده شده برای چهره اشخاص جواد، علی و مهدی به ازای هرکدام مشاهده می‌شود.

جدول ۵-۱: نمونه‌ای از یک جدول درهم ریختگی تشخیص تصاویر چهره

		کلاس شناسایی شده		
کلاس اصلی داده ها		مهدی	علی	جواد
	مهدی	۵	۰	۰
	علی	۱	۴	۰
	جواد	۰	۰	۵

برای نمونه، در این جدول مشخص شده است که از ۵ تصویر چهره مربوط به علی که به سیستم برای آزمایش داده شده است، ۴ نمونه به درستی تشخیص داده شده است و یک نمونه به اشتباه با برچسب "مهدی" تشخیص داده شده است. این درحالی است که چهره مهدی و جواد به درستی تشخیص داده شده است.

چهار تعریف مهم بر روی ماتریس همبستگی بوجود می‌آید. این معیارها با توجه به جدول ۵-۲ مشخص می‌شوند. معیار False Negative (FN) تعداد نمونه‌هایی را مشخص می‌کند که بایستی در کلاس مشخص شده قرار می‌گرفتند ولی در کلاس دیگری قرار گرفتند. معیار FP تعداد نمونه‌هایی را مشخص می‌کند که نبایستی در یک کلاس قرار می‌گرفتند ولی به اشتباه در آن کلاس قرار گرفتند [۳۶].

جدول ۵-۲:معیارهای ارزیابی

		خروجی سیستم	
		پاسخ درست	پاسخ غلط
اعتبار داده	داده درست	True Positive (TP)	False Positive (TP)
	داده غلط	True Negative (TN)	False Negative (FN)

معیارهای ارزیابی مختلفی با توجه به جدول فوق و بر اساس تعاریق TP، FN، FP و TN مشخص شده اند که ارزیابی نهایی سیستم بر اساس آنها مشخص می‌شود. مهمترین این معیارها دقت، یادآوری و صحت می‌باشد که از فرمول‌های ذیل بدست می‌آیند [۴۵]:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP+FP})}$$

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

$$\text{Accuracy} = \frac{(\text{TP+TN})}{(\text{TP+TN+FP+FN})}$$

که در آن Precision صحت، Recall یادآوری و Accuracy دقت خروجی‌های دسته بند را مشخص می‌کند. به بیان دیگر، صحت مشخص می‌کند که چند درصد از نمونه‌هایی که در یک کلاس قرار گرفتند، در واقع با آن کلاس تعلق داشتند. همچنین یادآوری مشخص می‌کند که چند درصد از نمونه‌های

آزمایشی یک کلاس به درستی در کلاس خود قرار داده شده اند. معیار دقت نیز به عنوان یک معیار اصلی، نشان می‌دهد که چند درصد از کل نمونه‌های ورودی یک سیستم به درستی در کلاس‌ها قرار داده شده اند یا به عبارتی دیگر میانگین یادآوری کلاس‌ها را مشخص می‌کند.

در ادامه فصل به بررسی نتایج بدست آمده از آزمایش‌های انجام شده بر روی داده‌های پایگاه داده‌های مورد استفاده تحت انتخاب پارامترهای مختلف و مجموعه داده‌های مختلف از این پایگاه داده می‌پردازیم.

۵-۳. نتایج آزمایشات پایگاه داده اول بر روی سیستم عملیاتی

برای آزمایش سیستم تشخیص چهره عملیاتی، پایگاه داده اول را به طور اختصاصی تهیه نمودیم. این پایگاه داده توسط خود سیستم تشخیص چهره عملیاتی البته با تغییراتی در سیستم ضبط و تهیه شده است. در ذیل نتیجه تشخیص چهره تصاویر پایگاه داده شماره دو در قالب جدول درهم ریختگی به نمایش گذاشته شده است. سیستم با موفقیت تمام تصاویر چهره را با برچسب نام درست برگرداند. سیستم با دقت ۱۰۰٪ تمام تصاویر آزمایشی را به نمونه‌های درست مجموعه آموزشی مربوط کرد.

در ادامه آزمایشات همین بخش، پنج شخص غیرمجاز نیز برای سیستم تعریف نمودیم. این پنج شخص جزو مجموعه آموزشی نیستند و سیستم تشخیص چهره نباید هیچ کدام از این پنج اشخاص را به هیچ کدام از پانزده شخص آموزش تطبیق کند. سیستم تشخیص چهره باید پاسخ "Face is not recognized!" را برای تصاویر چهره پنج شخص غیرمجاز برگرداند.

جدول ۵-۳: جدول درهم ریختگی نتایج اعمال پایگاه داده شماره یک بر روی سیستم تشخیص چهره عملیاتی

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
S1	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S2	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S3	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S4	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S5	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S6	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S7	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S8	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S9	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S10	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
S11	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪
S12	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪
S13	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪
S14	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪
S15	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪

حال تصاویر پنج شخص غیرمجاز را به سیستم اعمال کردیم و سیستم برای چهار شخص از پنج شخص، نتیجه "Face is not recognized!" را برگرداند اما چهره شخص غیرمجاز پنجم، به اشتباه با برچسب نام یکی از اشخاص آموزش داده شده به سیستم، تشخیص داده شد. این نتیجه نشان می‌دهد که امنیت سیستم در مقابل اشخاص غیرمجاز آنقدر قابل اعتماد نیست اما نکته قابل توجه آن است که امنیت سیستم قابل ارتقاء است؛ اگر با الگوریتم تشخیص هیستوگرام الگوهای دودویی محلی آشنا باشید، می‌دانید که فاصله هیستوگرام‌ها توسط یک معیار شباهت محاسبه می‌شود و بر حسب نزدیکی و با مقایسه با یک سطح آستانه، چهره شخص آزمایشی برچسب گذاری می‌شود. با کمتر کردن حد آستانه، امنیت

سیستم ارتقاء می‌یابد ولی باید توجه داشت که باعث می‌شود دامنه تشخیص نیز متعاقباً کاهش می‌یابد، بدین معنی که ممکن است تصویر چهره شخصی که در سیستم آموزش داده شده است با پاسخ "Face is not recognized!" مواجه شود. بنابراین پارامتر آستانه برحسب نیاز قابل تغییر است و در نتیجه امنیت سیستم قابل ارتقاء می‌باشد.

۵-۳-۱. بحث پیرامون نتایج آزمایشات پایگاه داده اول

-نکته اول: اکثر تصاویر پایگاه داده در شب تصویر برداری شده اند. هر چند عدم وجود نور کافی (همانند تصویربرداری در روز) ممکن است نرخ تشخیص را کاهش دهد اما این کاهش نرخ مشاهده نشده است و یکی از نقاط قوت سیستم را نشان می‌دهد.

-نکته دوم: در برخی تصاویر، حالت چهره و موقعیت چهره به صورت جزئی در حین تصویر برداری تغییر کرده است، اما این تأثیری در کاهش نرخ تشخیص نداشته است و این نیز یکی از نقاط قوت سیستم می‌باشد.

-نکته سوم: در برخی تصاویر، پشت زمینه پیچیده انتخاب شده است. از سوی دیگر، قطع تصویر چهره توسط الگوریتم چهره یابی مقداری بزرگتر از اندازه دقیق چهره است و بخش کوچکی از پشت زمینه را نیز در برگرفته است. ارسال مستقیم تصویر چهره با همین قطع به الگوریتم تشخیص چهره هیستوگرام‌های الگوهای دودویی محلی موجب وابستگی تشخیص چهره به پشت زمینه هم می‌شود. در پایگاه داده شماره یک، تصاویر هر شخص از مجموعه آموزش و آزمایش دارای پشت زمینه یکسان می‌باشد و موجب افزایش نرخ تشخیص شده است، در حالیکه اگر پشت زمینه تغییر می‌یافت مسلماً در نرخ تشخیص چهره تأثیرگذار خواهد بود.

-نکته چهارم: موفقیت سیستم عملیاتی تشخیص چهره مدیون یکسان بودن شرایط تصاویر مجموعه آموزش و مجموعه آزمایش است. اگر شرایط تصاویر مجموعه آموزش و آزمایش تغییر یابد، نتایج بگونه‌ای دیگر رقم خواهد خورد. بنابراین نیاز به تغییر پایگاه داده و استفاده از پایگاه داده استاندارد دیگری که شرایط خاص را شبیه سازی کند، وجود دارد. پایگاه داده شماره دو در ادامه مجدداً به سیستم عملیاتی تشخیص چهره اعمال خواهد شد.

-نکته پنجم: عدم تشخیص یک نفر از پنج نفر شخص غیر مجاز به عنوان شخص غیر مجاز حاکی از مقدار زیاد حد آستانه برای فاصله تعریف شده در معیار شباهت موجود در سیستم است. این مقدار می‌تواند کاهش یابد.

۵-۴. نتایج آزمایشات پایگاه داده شماره دوم

پایگاه داده شماره دو شامل ۳۱ تصویر از ۳۷ شخص مختلف می‌باشد. در آماده سازی این تصاویر، شرایط نوری مختلف اعمال شده است. حالت چهره و موقعیت چهره به صورت جزئی در تصاویر چهره مختلف متفاوت است. نتایج آزمایشات را با اعتبار سنجی ضربدری ۶-تایی آغاز می‌نماییم.

۵-۴-۱. اعتبار سنجی ضربدری ۶-تایی

مجموعه تصاویر ۱۰ شخص از پایگاه داده آموزشی با اعتبارسنجی ضربدری ۶-تایی به مجموعه‌های آموزش و آزمایش تقسیم بندی نمودیم. نتایج در ذیل ارائه شده است.

جدول ۵-۴: نتایج اعمال پایگاه داده دوم بر روی سیستم عملیاتی با اعتبارسنجی ضربدری ۶-تایی

مرحله اعتبارسنجی ضربدری ۶–تایی	دقت
اعتبارسنجی ضربدری ۶-تایی مرحله اول	۸۷.۲۷٪
اعتبارسنجی ضربدری ۶-تایی مرحله دوم	۹۰.۹۱٪
اعتبارسنجی ضربدری ۶-تایی مرحله سوم	۲۸.۱۸٪
اعتبارسنجی ضربدری ۶-تایی مرحله چهارم	۱۲.۷۳٪
اعتبارسنجی ضربدری ۶-تایی مرحله پنجم	۹۰.۹۱٪
اعتبارسنجی ضربدری ۶-تایی مرحله ششم	۲۴.۴۵٪
اعتبارسنجی ضربدری ۶-تایی (میانگین)	۵۵.۷۴٪

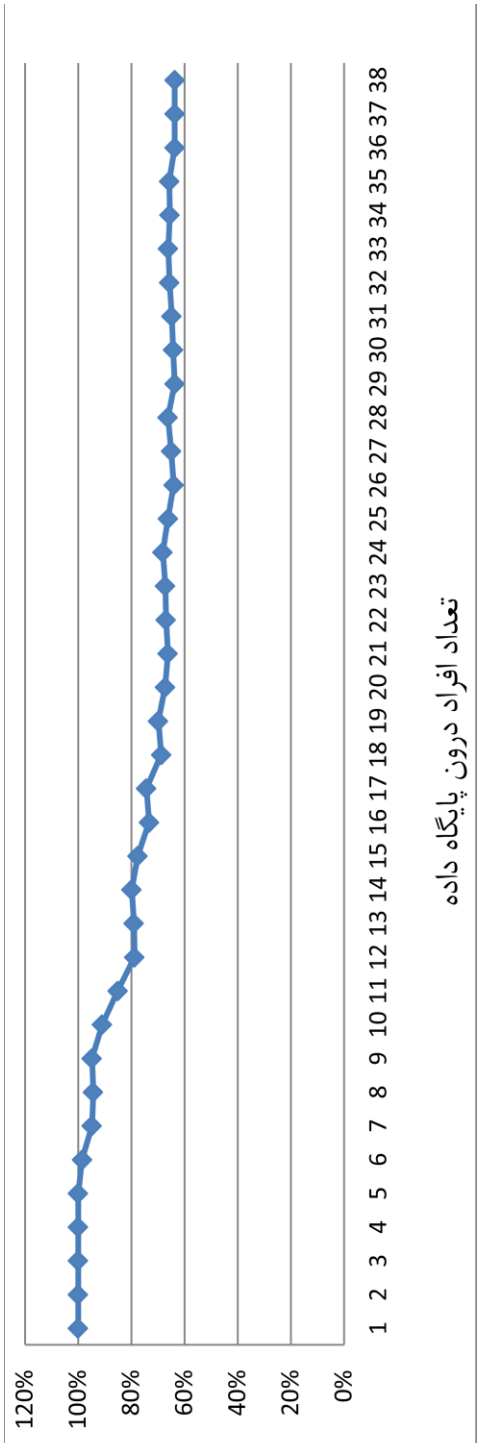
همانطور که پیش تر ذکر شد، اعتبارسنجی ضربدری برای کاهش وابستگی نتایج و دقت تشخیص چهره به مجموعه خاصی از آموزش و آزمایش می‌باشد. میانگین اعتبارسنجی ضربدری ۶-تایی، نرخ تشخیصی نزدیک به نرخ تشخیص واقعی را نشان خواهد داد.

۵-۴-۲. تأثیر تعداد اشخاص بر عملکرد سیستم

یکی از دغدغه‌های طراحان سیستم‌های تشخیص چهره، کاهش عملکرد سیستم تشخیص چهره با افزایش تعداد افراد آموزش دیده در سیستم می‌باشد. در اینجا نیز، برای آزمایش عملکرد سیستم تحت تأثیر تعداد افراد مختلف، شروع به آموزش تک تک اشخاص درون پایگاه داده دوم و برآورد دقت از تشخیص مجموعه داده‌های تست مربوط به افراد آموزش دیده نمودیم.

همانطور که مشاهده می‌شود، عملکرد سیستم تشخیص چهره با افزایش تعداد اشخاص به طور کلی نزولی است، اما به صورت مطلق نزولی نیست. در برخی موارد، با افزایش شخصی دیگر به پایگاه داده، نرخ

تشخیص افزایش خواهد یافت. این نشان می‌دهد که پس از افزودن شخص جدید، نرخ تشخیص برای آن شخص جدید افزوده شده بیشتر از نرخ تشخیص کلی پیش از افزودن شخص جدید می‌باشد و باعث افزایش نرخ تشخیص می‌شود.



شکل ۵-۳: نمودار روند تأثیر افزایش تعداد افراد درون پایگاه داده بر عملکرد سیستم

۵-۴-۳. تأثیر تصادفی کردن داده‌های مجموعه آموزش و آزمایش

در توزیع تصاویر پایگاه داده به مجموعه‌های آموزش و آزمایش، همواره این نگرانی وجود داشته که توزیع داده‌ها در مجموعه آموزش و آزمایش خاص باشد و نرخ تشخیص کاذب بیشتر و یا کمتری نسبت به نرخ تشخیص واقعی تولید نماید. یکی از رویکردهای مواجهه با این مشکل، تصادفی کردن انتخاب برای داده‌های مجموعه آموزش و آزمایش می‌باشد.

در آزمایشی، ۱۰ شخص از اشخاص پایگاه داده را در نظر گرفته و تعداد تصاویر مجموعه آموزش را برابر ۲۰ و تعداد تصاویر مجموعه آزمایش را برابر ۱۱ تعیین می‌نماییم. حال ۳۱ تصویر مربوط به هر شخص را به صورت تصادفی با توزیع‌های مختلف درون مجموعه آموزش و آزمایش قرار می‌دهیم.

جدول ۵-۵: تأثیر تصادفی کردن داده‌های پایگاه داده دوم در نرخ تشخیص سیستم عملیاتی تشخیص چهره

شماره توزیع	نرخ تشخیص (دقت)
۱	۵۶.۳۶٪
۲	۵۰٪
۳	۶۰.۹۱٪
۴	۸۲.۷۲٪
۵	۸۶.۳۶٪
۶	۳۰٪
۷	۷۰٪
۸	۵۰٪
۹	۹۸.۱۸٪
۱۰	۹۳.۶۳٪
میانگین ۱۰ توزیع فوق	۶۷.۸۲٪

۵-۴-۴. انتخاب مقدار مناسب برای فاصله آستانه

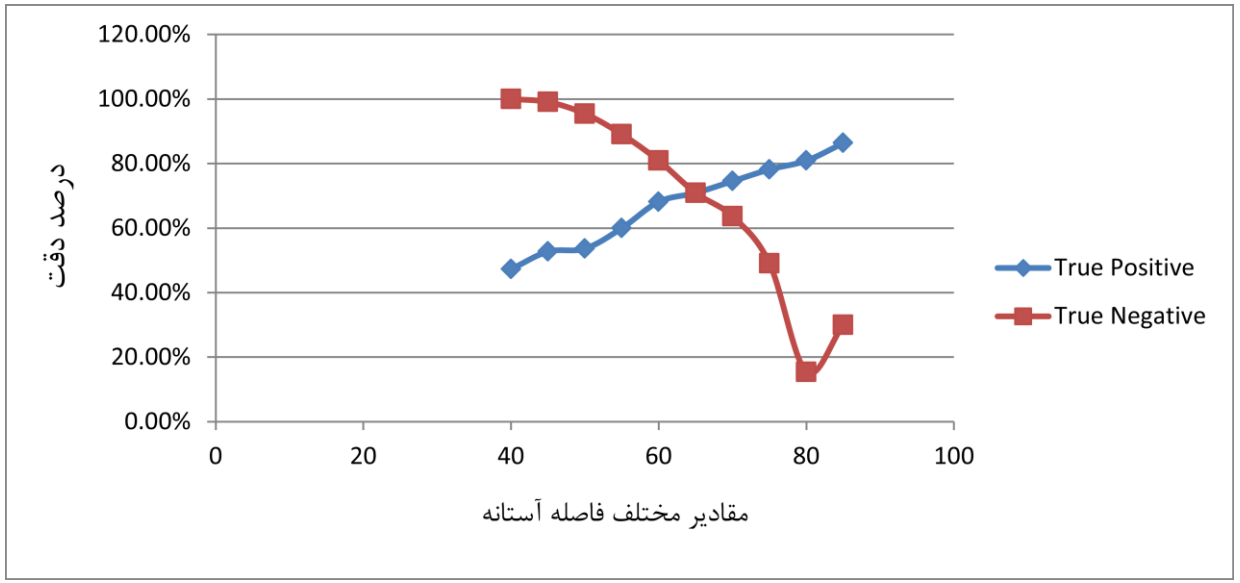
سیستم عملیاتی تشخیص چهره نه تنها باید بتواند تصاویر چهره افراد آموزش دیده در پایگاه داده را به درستی تشخیص بدهد، بلکه تصاویر غیر مجاز و آموزش دیده نشده را نیز باید به درستی به عنوان غیر مجاز تشخیص دهد. بازگرداندن پاسخ درست "Face is not recognized!" برای اشخاص غیر مجاز، نشان از امنیت سیستم می‌باشد و در برخی موارد این امنیت، اهمیت بسیاری دارد.

اگر با الگوریتم تشخیص هیستوگرام الگوهای دودویی محلی آشنا باشید، می‌دانید که فاصله هیستوگرام‌ها توسط یک معیار شباهت محاسبه می‌شود و بر حسب نزدیکی و با مقایسه با یک فاصله آستانه، چهره شخص تست برچسب گذاری می‌شود. با کاهش مقدار فاصله آستانه، امنیت سیستم ارتقاء می‌یابد. از سوی دیگر، کاهش فاصله آستانه، احتمال آنکه برخی تصاویر چهره افراد مجاز و آموزش دیده را نیز غیر مجاز تشخیص دهد، بیشتر می‌دهد.

انتخاب مناسب فاصله آستانه، بر اساس کاربرد و نحوه استفاده از سیستم می‌باشد. اما آزمایشی را تدارک دیدیم که یک حد تعادل برای مقدار فاصله آستانه بیابیم. در این آزمایش، تصاویر ۱۰ شخص از اشخاص پایگاه داده به مجموعه آموزش و آزمایش تقسیم بندی می‌کنیم. مجموعه آموزش شامل ۲۰ تصویر از هر شخص می‌باشد و مجموعه آزمایش شامل ۱۱ تصویر از هر شخص می‌باشد. تصاویر مربوط به ۱۰ شخص دیگر را نیز به عنوان غیر مجاز در نظر گرفتیم. مقادیر TP[1] و TN[2] را برای بازه‌ای از مقادیر مختلف فاصله آستانه محاسبه می‌کنیم.

همانطور که در نمودار مشاهده می‌شود، زمانی که فاصله آستانه برابر با ۶۵ باشد، مقدار TP و TN برابر می‌شود. اگر سیستم تشخیص چهره‌ای با کارایی نرمال مد نظر دارید، این مقدار، مقدار مناسبی است.

[1] True Positive
[2] True Negative

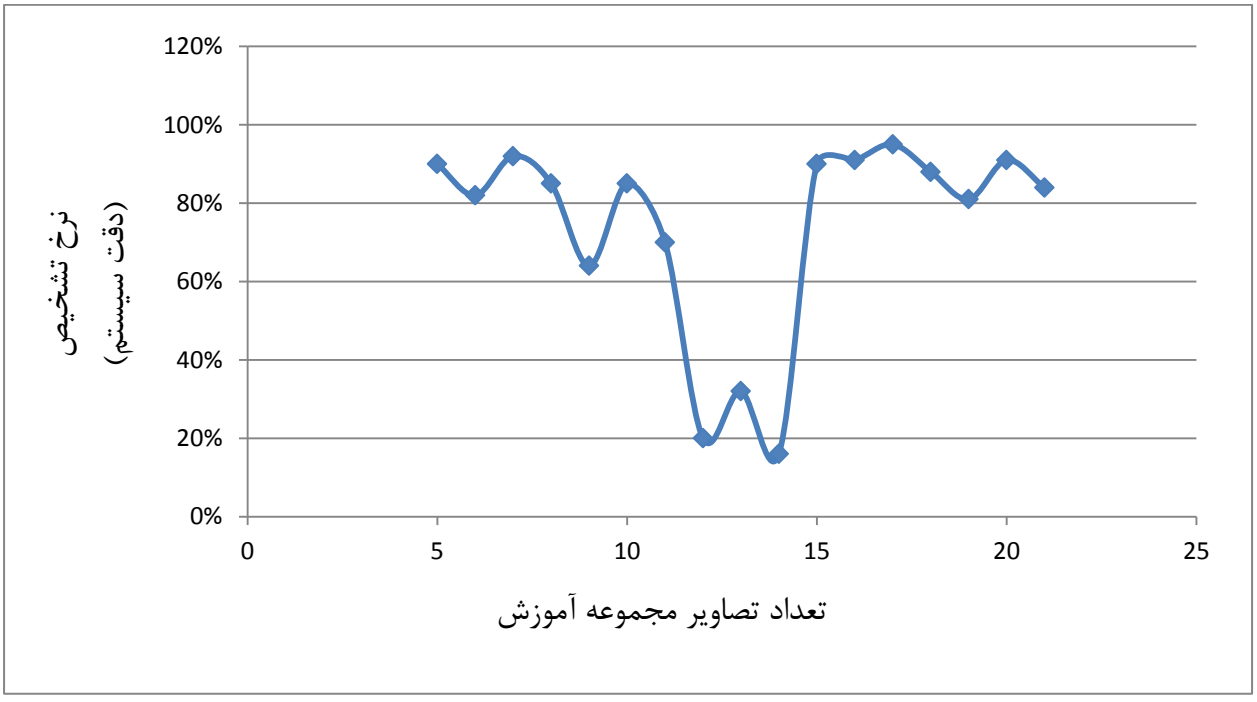


شکل ۵-۴: تأثیر حد فاصله آستانه بر روی معیارهای True Positive و True Negative

۵-۴-۵. تأثیر تعداد تصاویر آموزشی در عملکرد سیستم

در اکثر آزمایشات، تعداد تصاویر مجموعه آموزشی به ازای هر شخص را برابر ۲۰ در نظر گرفتیم و باقی تصاویر را در مجموعه آزمایش قرار دادیم. ممکن است این سؤال در ذهن خطور کند که مجموعه آموزشی با تعداد تصاویر کمتر یا بیشتر، چه تأثیری در نرخ تشخیص خواهد داشت. از دید تئوری، با در نظر گرفتن شرایط ایده آل، قاعدتاً انتظار بهبود نرخ تشخیص با افزایش تعداد تصاویر مجموعه آموزشی را داریم. از سوی دیگر، تعداد تصاویر مجموعه آموزشی نباید به مرز آموزش بیش از حد[1] برسد زیرا پس از آن نرخ تشخیص کاهش خواهد یافت.

در آزمایشی، ۱۰ شخص از اشخاص پایگاه داده را در نظر می‌گیریم. تعداد تصاویر مجموعه آزمایش برابر ۱۰ و تعداد تصاویر مجموعه آموزش از ۵ تا ۲۱ متغیر می‌باشد.



شکل ۵-۶: تأثیر افزایش تعداد تصاویر مجموعه آموزش بر روند عملکرد سیستم عملیاتی تشخیص چهره

[1] Over-train

همانطور که مشاهده می‌کنید، هیچ نظم خاصی در روند افزایش تعداد تصاویر مجموعه آموزش وجود ندارد. این نشان می‌دهد که نرخ تشخیص، بسیار وابسته به تصاویر مجموعه آموزش و آزمایش شخص جدید می‌باشد. هر چقدر سیستم، شخص جدید را بهتر آموزش ببیند و نتیجه بهتری در مجموعه آزمایشی ارائه کند، نرخ تشخیص کلی نیز بهتر خواهد شد. از سوی دیگر، هر چقدر سیستم، آموزش شخص جدید ضعیف باشد و متعاقباً نتیجه ضعیف تری در مجموعه آزمایشی ارائه کند، نرخ تشخیص کلی کاهش خواهد یافت.

۵-۴-۶. بررسی عملکرد سیستم در یک حالت خاص

در آزمایشی، ۱۰ شخص را از اشخاص پایگاه داده در نظر می‌گیریم. مجموعه آموزشی را با ۲۰ تصویر از هر شخص و مجموعه آزمایشی را با ۱۱ تصویر از هر شخص در نظر می‌گیریم. پس از انجام آزمایش، نتایج را در قالب ماتریس درهم ریختگی بر اساس تعداد و توزیع در کلاس‌های مختلف بیان نمودیم که در ادامه آمده است.

در صفحه بعد، جدول درهم ریختگی آورده شده است.

جدول ۵-۶: جدول درهم ریختگی نتایج یک حالت خاص برای سیستم عملیاتی تشخیص چهره

کلاس شناسایی شده		شخص ۱	شخص ۲	شخص ۳	شخص ۴	شخص ۵	شخص ۶	شخص ۷	شخص ۸	شخص ۹	شخص ۱۰	عدم تشخیص
کلاس اصلی داده ها	شخص ۱	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
	شخص ۲	۰٪	۹۰.۹۱٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۹.۰۹٪	۰٪
	شخص ۳	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
	شخص ۴	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
	شخص ۵	۰٪	۰٪	۰٪	۰٪	۹۰.۹۱٪	۰٪	۰٪	۹.۰۹٪	۰٪	۰٪	۰٪
	شخص ۶	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪	۰٪	۰٪
	شخص ۷	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۷۲.۷۳٪	۰٪	۰٪	۹.۰۹٪	۱۸.۱۸٪
	شخص ۸	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪	۰٪	۰٪
	شخص ۹	۰٪	۰٪	۰٪	۰٪	۰٪	۹.۰۹٪	۹.۰۹٪	۵۴.۵۵٪	۰	۱۸.۱۸٪	۹.۰۹٪
	شخص ۱۰	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۰٪	۱۰۰٪	۰٪

جدول ۵-۷: نتیجه معیارهای ارزیابی اولیه برای سیستم تشخیص چهره عملیاتی

نام معیار	تعداد بدست آمده
True Positive (TP)	۹۴
False Positive (FP)	۱۶
True Negative (TN)	۷۸
False Negative (FN)	۳۲

همانطور که در پیش تر ذکر شد، سه معیار مهم دقت[1]، صحت[2] و یادآوری[3] برای میزان قدرت یک روش و سیستم بکار می‌آید. در ذیل نتیجه سه معیار مذکور در قالب جدولی بیان می‌شود.

جدول ۵-۸: نتیجه معیارهای ارزیابی نهایی برای سیستم تشخیص چهره عملیاتی

نام معیار	نتیجه
دقت	۷۸.۲۳٪
صحت	۸۵.۴۵٪
یادآوری	۷۶.۶۶٪

۵-۵. نتایج آزمایشات پایگاه داده شماره سوم

۶ تصویر از هر کدام از ۱۲ شخص داریم که در دو حالت به مجموعه آموزشی و مجموعه تست تقسیم بندی می‌کنیم. در حالت اول، یک تصویر از هر شخص را برای مجموعه تست و ۵ تصویر باقی مانده را برای مجموعه آموزشی در نظر می‌گیریم. در حالت دوم، دو تصویر از هر شخص را برای مجموعه تست و ۴ تصویر باقی مانده را برای مجموعه آموزشی در نظر می‌گیریم.

[1] Accuracy
[2] Precision
[3] Recall

- **حالت اول**

هنگامی که الگوریتم پیشنهادی (ویژگی‌های هندسی مستخرج + دسته بند پیشنهادی) را بر روی پایگاه داده سوم آزمایش نمودیم، دقت آن به ۱۰۰٪ رسید، درحالیکه دقت دسته بند 1-NN با همان ویژگی‌های هندسی به ۷۵٪ رسید. توجه داشته باشید که در هر دو آزمایش، تنها ویژگی‌های هندسی بکار برده شده است. ۲۵٪ اختلاف دقت، اثبات خوبی برای نشان دادن قدرت الگوریتم پیشنهادی در این حالت خاص می‌باشد.

هنگامی که الگوریتم پیشنهادی (ویژگی های حاشیه نویسی شده + دسته بند پیشنهادی) بر روی پایگاه داده سوم آزمایش نمودیم، دقت آن به ۹۱.۶۶٪ رسید و دقت دسته بند 1-NN با همان ویژگی‌ها به ۹۱.۶۶٪ رسید، اما نکته ای که وجود دارد آنستکه چهره‌هایی که روش پیشنهادی اشتباه تشخیص داده بود، از نظر دید انسانی بسیار شبیه به هم بودند، درصورتی که چهره‌هایی که دسته بند 1-NN اشتباه تشخیص داده بود، از نظر دید انسانی مشابه نبودند.

از آنجا که مقالات بسیاری متدهای بسیاری در حوزه تشخیص چهره ارائه داده اند، انتظار می‌رود که نتایج ارائه شده با نتایج یک متد مشهور و مقبول در حوزه تشخیص چهره مقایسه گردد. در اینجا، متد چهره‌های ویژه "Eigenfaces" را بر روی پایگاه داده IMM Frontal Face پیاده سازی نمودیم و نتایج آن را به دست آوردیم. متد "Eigenfaces" به دقت ۶۶.۶۷٪ در rank-1 و به دقت ۸۳.۳۳٪ در rank-2 و به دقت 100% در rank-3 و بالاتر می‌رسد.

- **حالت دوم**

در این حالت روش پیشنهادی برتری قابل توجهی نسبت به روش های موجود ندارد و به نظر می رسد مجموعه داده های آموزشی کمتر از حد مورد نیاز می باشد.

جدول ۵-۹: نتایج روش پیشنهادی با دسته بند پیشنهادی

مجموعه آموزشی و مجموعه آزمایشی	روش	نرخ تشخیص (دقت)
مجموعه آموزشی: ۵ تصویر برای هر شخص مجموعه آزمایشی: ۱ تصویر برای هر شخص	ویژگی‌های هندسی + دسته بند پیشنهادی	100%
	ویژگی‌های هندسی + دسته بند 1-NN	75%
	ویژگی های حاشیه نویسی شده + دسته بند پیشنهادی	91.66%
	ویژگی های حاشیه نویسی شده + دسته بند 1-NN	91.66%
	Eigenfaces	1^{st} rank = 66.67%; 2^{nd} rank = 83.33%; 3^{rd} rank = 100%;
مجموعه آموزشی: ۴ تصویر برای هر شخص مجموعه آزمایشی: ۲ تصویر برای هر شخص	ویژگی‌های هندسی + دسته بند پیشنهادی	91.66%
	ویژگی‌های هندسی + دسته بند 1-NN	62.5%
	ویژگی های حاشیه نویسی شده + دسته بند پیشنهادی	70.83%
	ویژگی های حاشیه نویسی شده + دسته بند 1-NN	91.66%
	Eigenfaces	1^{st} rank = 91.67%; 2^{nd} rank = 91.67%; 3^{rd} rank = 95.83%; 4^{th} rank = 95.83%; 5^{th} rank = 100%;

در آزمایشی دیگر سیستم را مقابل افراد غیر مجاز محک می زنیم؛ در این آزمایش۱۰ شخص از اشخاص پایگاه داده را در سیستم تشخیص چهره آموزش می دهیم و ۲ شخص دیگر را با عنوان اشخاص غیرمجاز، به سیستم اعمال می نماییم. در سیستم، یک حد آستانه ای را تعریف می نماییم تا شخص غیر مجاز از مجاز را تشخیص دهد. این حد آستانه به طور تجربی بگونه ای انتخاب شده است تا به نرخ تشخیص لطمه نزند. پس از انجام آزمایش، نتیجه آن می شود که یکی از افراد غیر مجاز با عنوان غیرمجاز به درستی تشخیص داده می شود ولی فرد دوم متأسفانه با برچسب یکی از اشخاص آموزش دیده در پایگاه داده شناخته می شود که اشتباه می باشد. اگر حد آستانه را از مقدار تنظیم شده کاهش یابد، این اشتباه سیستم برطرف خواهد شد و دو شخص غیرمجاز را به عنوان غیرمجاز خواهد شناخت ولی نرخ تشخیص سیستم کاهش خواهد یافت.

نتایج نشان می‌دهد که روش پیشنهادی، روشی قابل تأمل می‌باشد و توسعه آن ارزشمند است. از آنجا که روش پیشنهادی، در اصل مجموعه‌ای از مفاهیم است که در کنار یکدیگر، عمل تشخیص چهره را انجام می‌دهند، بنابراین توسعه روش پیشنهادی و ارائه این مفاهیم در قالب‌های ریاضی دیگر، ممکن است جنبه‌ها و قابلیت‌های دیگر پنهان این روش را آشکار کند.

فصل ششم

نتیجه گیری و کارهای آینده

قریب به سه دهه است که الگوریتم‌های مختلف تشخیص چهره در حال تولید و تحت محک و امتحان می‌باشد. الگوریتم‌های بسیاری ارائه شده است؛ اما نکته‌ای که وجود دارد آنست که عملیاتی شدن همه این الگوریتم‌ها با چالش‌های بسیاری روبروست، لذا اگر به دنبال انتخاب یک الگوریتم تشخیص چهره مناسب در تصاویر با قابلیت پیاده سازی عملیاتی باشیم که در دهه اخیر هم ارائه شده باشد عموماً به دو الگوریتم می‌رسیم؛ الگوریتم هیستوگرام‌های الگوهای دودویی محلی [1] و الگوریتم بازنمایی‌های پراکنده[2]. توسعه‌های بسیاری نیز برای این دو الگوریتم ارائه شده اند. ما در این پایان‌نامه یک سیستم تشخیص چهره تحت وب را بر پایه هیستوگرام‌های الگوهای دودویی محلی پیاده سازی و عملیاتی نموده ایم. این سیستم تشخیص چهره به سه بخش اپلت[3] تشخیص کاربر، اپلت افزودن کاربر و سرور جاوای تشخیص چهره تقسیم بندی می‌شود. از آنجا که این سیستم تشخیص چهره تحت وب می‌باشد، عملیات‌ها باید به صورت بهینه میان کلاینت و سرور توزیع شوند. عملیات اخذ تصویر از وب‌کم و عملیات چهره یابی به همراه پردازش‌های مورد نیاز در اپلت جاوای سمتِ کلاینت صورت می‌گیرد. عملیات‌های اصلی آموزش و آزمایش تصاویر چهره به همراه پایگاه داده مربوطه در بخش سرور قرار دارد. با این تقسیم بندی، ارتباطات میان کلاینت و سرور به حداقل می‌رسد. از آنجا که دو اپلت جاوای مجزا برای افزودن کاربر و تشخیص کاربر در نظر گرفته شده است، سرور جاوا ابتدا درخواست‌های دریافتی را بررسی می‌کند و سپس به بخش مرتبط با آن هدایت می‌کند. بنابراین اپلت تشخیص کاربر با بخش آزمایش سرور و همچنین پایگاه داده، مرحله آزمایش را تشکیل می‌دهند و از طرفی دیگر، اپلت افزودن کاربر با بخش آموزش سرور و مجدداً پایگاه داده، مرحله آموزش را تشکیل می‌دهند.

[1] Local Binary Pattern Histograms
[2] Sparse Representation
[3] Applet

در بحث بهبود سیستم عملیاتی شده و کارهای آینده، می‌توان به مبحث مهم تغییرات روشنایی اشاره نمود. سیستم مورد نظر از آنجا که تصویر خود را از وب‌کم دریافت می‌کند و وب‌کم محیط کنترل نشده‌ای را تصویربرداری می‌کند، لذا تغییرات روشنایی عامل مزاحمی برای سیستم تشخیص چهره ما خواهد بود. بنابراین، پیشنهاد می‌شود در کارهای آینده تصویر دریافتی از وب‌کم بگونه‌ای پردازش شود که تغییرات نوری تأثیر حداقلی بر سیستم تشخیص چهره داشته باشد.

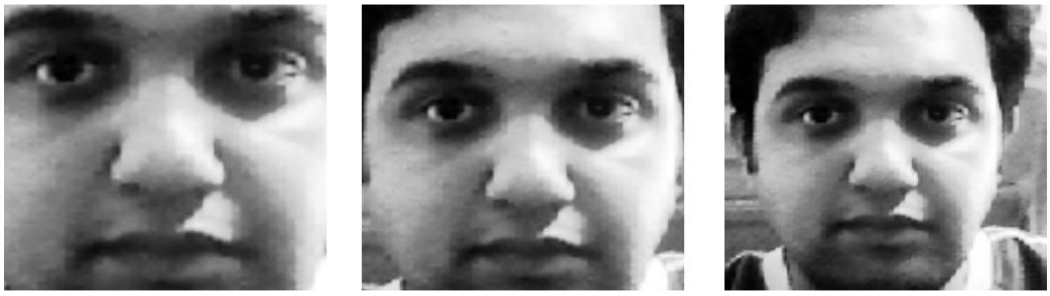
دقت سیستم تشخیص چهره عملیاتی شده جاری، با افزایش تعداد اشخاص موجود در پایگاه داده کاهش می‌یابد. در کارهای آینده، توصیه می‌شود بهبودی صورت گیرد تا بتوان جامعه آماری افراد موجود پایگاه داده افزایش داد و در عین حال دقت تشخیص چهره نیز کاهش نیابد.

الگوریتم چهره یابی استفاده شده در سیستم عملیاتی، الگوریتم Viola-Jones می‌باشد. این الگوریتم یکی از قدرتمندترین و کاراترین الگوریتم‌ها می‌باشد ولی در کارهای آینده می‌توان به جای استفاده از خود آن، از توسعه‌هایی از این الگوریتم استفاده کرد که موجب بهبود بیش از پیش عملکرد چهره یابی شود.

یکی از مواردی که در بهبود عملکرد تشخیص چهره نقش دارد و کمک بسیاری به این سیستم ما می‌نماید، مرحله قطع[1] مناسب تصویر خروجی مرحله چهره یابی است. تصویر خروجی مرحله چهره یابی دارای تصویر چهره است اما اینکه تصویر چهره دقیقاً چه قطعی داشته باشد مسئله بسیار مهمی است، زیرا اکثر الگوریتم‌های تشخیص چهره از تک تک پیکسلهای تصویر خروجی از مرحله چهره یابی استفاده می‌کنند. الگوریتم الگوهای دودویی محلی نیز به همین منوال می‌باشد و قطع مناسب تاثیر بسزایی در نتیجه تشخیص چهره دارد. در کارهای آینده، پیشنهاد می‌شود که قطع تصویر چهره با عملیاتی بهبود یابد تا متعاقباً تاثیر مثبت بسزایی بر روی نتیجه تشخیص چهره بگذارد. در شکل ذیل می‌توانید نمونه‌ای از

[1] Crop

قطع‌های مختلف تصویر را مشاهده کنید. انتخاب قطع مناسب تصویر، نیازمند آزمایشات مختلفی است که کار جداگانه‌ای را می‌طلبد.



شکل ۶-۱:قطع‌های متفاوت تصویر چهره را مشاهده می‌کنید. تصویر چهره سمت راست خروجی عملیات چهره یابی می‌باشد. تصویر چهره وسط و سمت چپ قطع‌های دیگر تصویر چهره را نشان می‌دهد.

ممکن است این ایده به ذهن متبادر شود که سیستم تشخیص چهره لزوماً نیاز به حضور شخص جلو دوربین وب‌کم ندارد و می‌توان با نگه داشتن یک برگه جلو دوربین وب‌کم که چهره فرد بر روی آن چاپ شده است، پاسخ مثبت تشخیص چهره را از سیستم دریافت کرد. راهکاری که برای جلوگیری از این روش نفوذ غیرمجاز پیشنهاد می‌شود آنست که از شخص بخواهیم تا یک حالت چهره خاص از سوی شخص ابراز شود. بنابر این، پس از تشخیص چهره، اعتبارسنجی تصویر چهره شخص کاربر از طریق حالت چهره انجام شود. برای تشخیص حالت چهره، باید چند فریم متوالی دوربین وب‌کم را پردازش نمود. یکی از راه‌های پردازش و تحلیل تصاویر حالت چهره در مقاله‌ای که اینجانب به همراه جناب آقای دکتر مرتضی زاهدی در کنفرانس MVIP 2013 ارائه نمودیم، آورده شده است.

در روش پیشنهادی پایان‌نامه از این حقیقت استفاده شده است که سیستم تشخیص چهره در انسان ویژگی‌هایی را از چهره استخراج می‌کند که این ویژگی‌ها نسبت به چهره میانگینی که در ذهن فرد شکل گرفته است، دارای بیشترین تمایز می‌باشد. چهره میانگین نیز از میانگین گیری چهره‌های دیده شده توسط فرد در یک منطقه جغرافیایی خاص در ذهن فرد تشکیل می‌شود. این حقیقت به صورت ریاضی‌وار

فرموله شده است و انواع مختلف ویژگی‌های چهره استخراج می‌گردد. در این پایان‌نامه ویژگی‌های حاشیه نویسی شده به همراه ویژگی‌های هندسی استخراج گردید؛ برای هر ویژگی چهره (همانند چشم‌ها، ابروها، بینی، دهان و فک) یک ویژگی چهره میانگین محاسبه می‌شود، پس از آن انحرافات ویژگی‌های چهره افراد مختلف از ویژگی‌های چهره میانگین مربوطه محاسبه می‌شود، و در مرحله بعد، از آن تابع تخمین چگالی را برای مقادیر انحرافات بدست می‌آوریم. هنگامی که تصویر چهره شخص تست وارد سیستم می‌شود، انحرافات ویژگی‌های شخص تست از میانگین محاسبه می‌شود و با ویژگی‌های اشخاص آموزش دیده در پایگاه داده مقایسه می‌شود و نهایتاً با راه‌کار تعریف شده سیستم پاسخ نهایی را بازمی‌گرداند.

مشارکت[1] اصلی پایان‌نامه در بخش روش پیشنهادی، شامل پردازش ویژگی و دسته بندی می‌باشد. برای بیان نحوه عملکرد دسته بند پیشنهادی، ویژگی‌های مستخرج را با دسته بند 1-NN نیز دسته بندی نمودیم؛ نتایج گواه بر متفاوت بودن دسته بند پیشنهادی نسبت به دسته بند 1-NN می‌باشد.

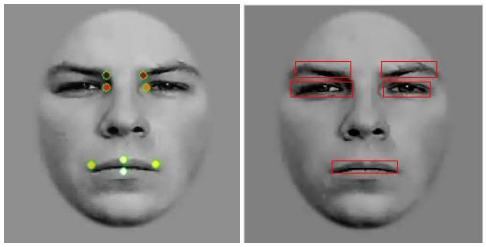
برای نشان دادن قدرت سیستم تشخیص چهره پیشنهادی، نتایج را با نتایج متد مشهور Eigenfaces مقایسه نمودیم. روش پیشنهادی برتری خود را نسبت به روش Eigenfaces در rank-1 و rank-2 اثبات نموده است. در کارهای آینده، فرایند استخراج اتوماتیک ویژگی‌های چهره را دنبال خواهیم کرد تا قابلیت اعمال الگوریتم پیشنهادی بر روی پایگاه داده‌های دیگر فراهم شود. لازم به ذکر است روش پیشنهادی پایان‌نامه کاملاً منعطف بوده، بنابراین در بحث کارهای آینده، پیشنهادهای بسیاری را در بحث انتخاب ویژگی‌های استخراجی، طریقه میانگین گیری، نحوه محاسبه انحراف ویژگی‌ها از میانگین می‌توان ارائه داد.

[1] Contribution

پیوست

تشخیص حالت چهره

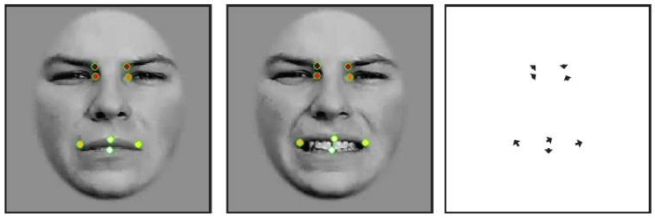
در مقاله " Pain Facial Expression Recognition among 8 Facial Expressions using Similarity Measures" که یکی از مقالات مستخرج از پایان‌نامه می‌باشد، یک روش استخراج ویژگی‌های چهره (اعم از چشم‌ها، ابروها و دهان) ارائه شده است. سپس، بر اساس ویژگی‌های مستخرج، یک مجموعه نقاط ویژه بر حسب مکان ویژگی‌های مستخرج تعیین می‌شود.



شکل ۷-۱: تصاویر برگرفته از مقاله مستخرج از پایان‌نامه، تصویر سمت راست: تصویر چهره به همراه مکان ویژگی‌های چهره. تصویر سمت چپ: تصویر چهره به همراه نقاط ویژه که به کمک ویژگی‌های چهره بدست آمد.

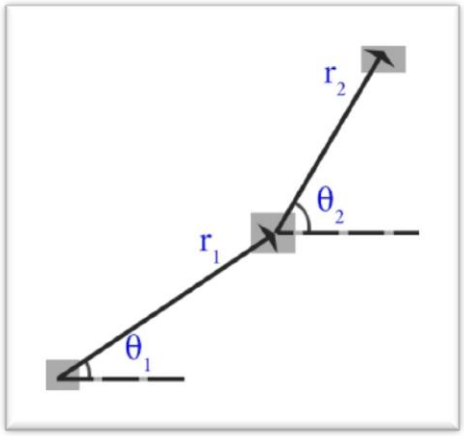
از آنجا که مکان‌های نقاط ویژه باید در هر فریم تعیین شود و از طرفی دیگر، اعمال الگوریتم‌های استخراج ویژگی‌های چهره بر روی هر فریم منطقی نیست، لذا استخراج ویژگی‌های چهره بر روی یک فریم انجام می‌شود و نقاط ویژه بر حسب آن تعیین می‌شود و سپس نقاط ویژه را با الگوریتم Lucas-Kanade [۴۶] ردیابی می‌کنیم.

با ردیابی نقاط ویژه در فریم‌های بعدی، مکان نقاط ویژه در تمام فریم‌ها بدست خواهد آمد. با کمک مکان نقاط ویژه، دو بردار مقادیر مطلق جابجایی‌ها و بردار جهت‌های جابجایی بدست می‌آید.



شکل ۷-۲: تصویر چپ و وسط: دو فریم متوالی ویدئو با نقاط ویژه مشخص شده. راست: بردارهای جابجایی نقاط ویژه.

همانطور که در شکل ۷-۳ مشاهده می‌نمایید، بردارهای جهات جابجایی کلیدی برای تمییز دادن میان انواع مختلف حالات چهره می‌باشد. بردارهای مقادیر مطلق جابجایی‌ها تنها محتوی اطلاعات سرعت هستند و این اطلاعات تنها سرعت شکل گیری حالت چهره را نشان می‌دهد و در دسته بندی حالت چهره تأثیری ندارد.



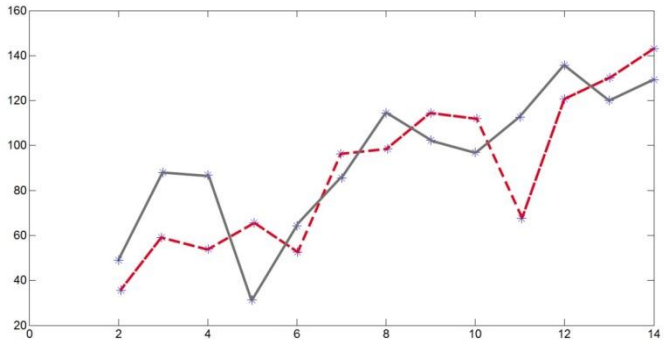
شکل ۷-۳: جایجایی یک نقطه ویژه را می‌توان با جهت جابجایی (Θ) و مقدار مطلق (r) جابجایی محاسبه نمود.

از آنجایی‌که حالت چهره از شخصی به شخص دیگر متفاوت است، جابجایی‌های نقاط ویژه در اشخاص مختلف تقریباً مشابه خواهد بود. بنابراین، برداری از مقادیر تتا ساخته می‌شود که شبیه $(\Theta1, \Theta2, \dots, \Theta n)$ خواهد شد. هر بردار جهات جابجایی (مجموعه تتا) برای هر حالت چهره منحصربه فرد می‌باشد و یک الگوی خاص را پیروی می‌کند. بنابراین، تشخیص حالت چهره را با دسته بندی بردارهای جهات

جابجایی می‌توان انجام داد. برای دسته بندی حالات چهره، نیاز به آماده سازی مجموعه آموزشی و ساخت داده‌های محوری بر اساس مجموعه آموزشی می‌باشد. داده‌های محوری شامل هشت توالی تتا به ازای هر کدام از حالات چهره (درد، خنثی، شادی، تعجب، تنفر، عصبانیت، ترس و غمگینی) می‌باشد؛ هر توالی مربوط به یک نقطه ویژه می‌باشد. هر توالی داده‌های محوری از محاسبه میانگین بردارهای متناظر مجموعه آموزشی برای هر حالت چهره بدست می‌آید.

هرگاه یک ویدئوی حالت چهره برای تست وارد شود، ابتدا باید تصاویر را آنالیز نمود. سپس بردارهای جهات جابجایی مربوط به هر نقطه ویژه محاسبه می‌شود. تعداد بردارها برابر با تعداد نقاط ویژه می‌باشد. برای دسته بندی حالت چهره تست، باید بردارهای تست را با داده‌های محوری مقایسه نمود. مقایسه را می‌توان با یک معیار شباهت انجام داد.

همانطور که در شکل ۷-۴ مشاهده می‌کنید، بردارهای تتای مربوط به یکی از نقاط ویژگی دهان دو شخص مختلف در فریم‌های مختلف رسم شده است. دو توالی مربوط به حالت چهره خوشحالی می‌باشد. همانطور که در شکل ۷-۴ مشاهده می‌نمایید، مقادیر دو توالی، هر دو، صعودی هستند و از یک الگوی خاص پیروی می‌کنند. این مشخصه یک ویژگی تمییزدهنده را برای بردار تتا را به نمایش می‌گذارد.



شکل ۷-۴: بردار مقادیر تتا مربوط به یکی از نقاط ویژه در فریم‌های ویدئو. یک شخص به صورت خط چین و شخص دیگر با خط راست مشخص شده است.

از آنجا که حالات چهره در اشخاص مختلف با سرعت‌های مختلف شکل می‌گیرد، این بردار جهات جابجایی (تتا) با تغییرات مقیاس زمانی. بنابراین، باید معیار شباهتی را انتخاب کنیم که تغییرات شکل در بعد زمان را بپذیرد. در بخش دیگر، معیار شباهت مناسبی را معرفی خواهیم کرد که شرایط را ارضاء نماید.

معیار شباهت مناسب

یک معیار شباهت، نوعی از یک تابع امتیازدهی است که یک امتیاز عددی را به یک جفت از توالی‌ها بر حسب شباهت آنها مقداردهی کند. امتیاز بالاتر دلالت بر شباهت بیشتر دارد. تعریف و تعیین معیارهای شباهت بر اساس نیازهای موجود کار دشواری است. در حال حاضر، معیارهای بسیاری برای تعیین شباهت وجود دارد. این معیارها را می‌توان به سه گروه دسته بندی کرد: معیارهای فاصله، معیارهای شباهت و معیار دودویی.

تغییر شکل‌هایی در سیگنال همانند تغییر مقیاس دامنه، جابجایی دامنه، جابجایی زمانی، تغییر مقیاس زمانی و تأخیر فاز وجود دارند. از آنجا که شباهت دو بردار جهات جابجایی (تتا) را اندازه گیری می‌کنیم، معیار شباهت ما باید تغییر شکل‌هایی همانند جابجایی زمانی و تغییر مقیاس زمانی را باید بپذیرد اما از طرفی دیگر، معیار شباهت انتخابی نباید تغییر شکل‌هایی همانند تغییر مقیاس دامنه و جابجایی دامنه و تأخیر فاز را بپذیرد.

از آنجا که معیار شباهت مطلوب در مطالب موجود[1] یافت نمی‌شود، معیار شباهت اقلیدسی را بگونه‌ای بهبود داده تا تغییرات شکلی همانند شیفت زمانی و تغییر مقیاس زمانی را بپذیرد. نسخه استاندارد فاصله اقلیدسی شیفت زمانی، تغییر مقیاس زمانی، جابجایی دامنه، تغییر مقیاس دامنه و تأخیر فاز را نمی پذیرد. در ذیل فاصله اقلیدسی و سپس نسخه بهبودیافته فاصله اقلیدسی را توصیف می‌کنیم:

[1] Literature

- نسخه استاندارد فاصله اقلیدسی

نسخه استاندارد فاصله اقلیدسی، فاصله میان دو توالی X و Y را با محاسبه ریشه دوم مجموع مجذور اختلافات مقادیر توالی مشخص می‌کند.

$$\text{Euclid}(X, Y) = \sqrt{\sum_i (x_i - y_i)^2}$$

- نسخه بهبودیافته فاصله اقلیدسی

برای آنکه فاصله اقلیدسی، تغییر شکلی جابجایی زمانی را بپذیرد، عملگر جابجایی را به یکی از دو توالی اعمال می‌کنیم. با حرکت دادن ساده یک توالی به تعداد ثابت به راست (یا چپ)، بهترین حالتی را خواهیم یافت که دو توالی بیشترین شباهت را پیدا کنند.

$$X_k = \text{ArithmeticShift}(X, k)$$
$$k = 0,1, \dots n + m$$

برای آنکه فاصله اقلیدسی، تغییر مقیاس زمانی را بپذیرد، عملیاتی را می‌افزاییم که طول توالی را کاهش می‌دهد یا در آن کشیدگی ایجاد می‌کند. فاصله میان توالی جدید تغییر سایز یافته و توالی ثابت با یک فاصله اقلیدسی می‌شود.

$$X_r = \text{Resize}(X, r)$$
$$r = 0.5 : 0.2 : 2$$

نهایتاً، نسخه بهبودیافته فاصله اقلیدسی فاصله میان توالی دستکاری شده X ($X_{k,r}$) و توالی Y دستکاری نشده (Y) را محاسبه می‌کند. $X_{k,r}$ توالی X می‌باشد که توسط شیفت محاسباتی و توابع تغییر سایز دستکاری شده است. فاصله میان دو توالی به عنوان ریشه دوم مجموع مجذور اختلافات مقادیر دو توالی می‌باشد.

$$\text{Improved_Euclid}(X_{k,r}, Y) = \sqrt{\sum_i (X_{r,k\ i} - Y_i)^2}$$

پس از محاسبه خروجی تابع Improved_Euclid(Xk,r, Y) به ازای مقادیر مختلف k و r، مقادیری از k و r که با آنها حداقل مقدار فاصله بدست می‌آید، انتخاب می‌شود.

نتیجه گیری

در این پیوست، روشی با پیچیدگی محاسباتی کم برای آنالیز زمان-واقعی حالت چهره ارائه گردید. از این روش برای تشخیص جعلی بودن تصاویر دریافت شده از دوربین وب‌کم می‌توان استفاده نمود. با پیاده سازی این روش و ترکیب آن با سیستم عملیاتی تشخیص چهره، از کاربر می‌خواهیم تا یک حالت چهره خاص را ابراز کند و سپس آن حالت چهره را توسط سیستم آنالیز می‌نماییم. با اجرای این روش، دیگر با قرار دادن کاغذی که تصویر چهره کاربر بر روی آن چاپ شده باشد، نمی توان سیستم را فریب داد.

فهرست مراجع و منابع

[1] B Schneier. Secrets and Lies. Wiley, 2000.

[2] B Friedman, H Nissenbaum, D Hurley, D C Howe, and E Felten. User's conceptions of risks and harms on the web: A comparative study. In Proceedings of CHI 2002., Minneapolis, Minnesota, April 20-25 2002. ACM.

[3] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997.

[4] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.

[5] B. Heisele, T. Poggio, and M. Pontil, "Face Detection in Still Gray Images," A.I. memo AIM-1687, Artificial Intelligence Laboratory, MIT, 2000.

[6] Ti-QiongXu; Bi-Cheng Li; Bo Wang, Face detection and recognition using neural network and hidden Markov models, Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, pp. 228 - 231,2003.

[7] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proc. Conf. Computer Vision and Pattern

[8] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face Detection in Color Images," Proc. Int'l Conf. Image Processing, pp. 1046-1049, 2001.

[9] Z. Liu and Y. Wang, "Face Detection and Tracking in Video Using Dynamic Programming" Proc. Int'l Conf. Image Processing, 2000.

[10] L. Silva, K. Aizawa and M. Hatori. "Detection and Tracking of Facial Features". Proc. Of SPIE Visual Communications and Image Processing, Taiwan. May, 1995.

[11] M. Han, A. Sethi, and Y. Gong. A detection-based multiple object tracking method. In Proc. Int. Conf. Image Process. (ICIP), pages 3065-3068, Singapore, Oct. 2004.

[12] D. Ramanan and D. Forsyth. Using temporal coherence to build models of animals. In Proc. Int. Conf. Comp. Vision(ICCV), pages 338-346, Nice, France, Oct. 2003.

[13] D. Maio and D. Maltoni. Real-time face location on grayscale static images. Pattern Recognition, 33:1525-1539, September 2000.

[14] Mikolajczyk, K.; Choudhury, R.; Schmid, C. ,Face detection in a video sequence-a temporal approach, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol 2, pp: 96-101, 2001.

[15] Zhenqiu Zhang; Potamianos, G.; Ming Liu; Huang, T., Robust Multi-View Multi-Camera Face Detection inside Smart Rooms Using Spatio-Temporal Dynamic Programming, 7th International Conference on Automatic Face and Gesture Recognition, 2-6 April 2006 Page(s):407 - 412.

[16] J. Yang and A.Waibel. A real-time face tracker. In Proceedings of the Third IEEE Workshop on Applications of Computer Vision, pages 142-147, Sarasota, FL, 1996.

[17] G. Bradski. Computer vision face tracking for use in a perceptual user interface. Technical Report Q2, Intel Corporation, Microcomputer Research Lab, Santa Clara, CA, 1998.

[18] H. Schneiderman and T. Kanade. "A statistical method for 3d object detection applied to faces and cars". In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000.

[19] J. Feraud, O. Bernier, and M. collobert. "A fast and accurate face detector for indexation of face images". In Proc. Fourth IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 52-59, 1998.

[20] S. Gong, S.McKenna, and J.Collins. "An investigation into face pose distribution". In Proc. IEEE International Conference on Face and Gesture Recognition, Vermont,1996.

[21] ZhenQiu Zhang; Long Zhu; Li, S.Z.; HongJiang Zhang,Real-time multi-view face detection, Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp:142-147,2002.

[22] P. Viola and M.J. Jones, "Robust real-time object detection", IEEE ICCV Workshop on Statistical and Computational Theories of Vision. Vancouver, Canada. July 13, 2001.

[23] S. Z. Li, Z. Q. Zhang, "FloatBoost Learning and Statistical Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 26, NO. 9, September, 2004.

[24] Gary B. Huang, Vidit Jain, and Erik Learned-Miller. Unsupervised joint alignment of complex images. International Conference on Computer Vision (ICCV), 2007.

[25] J. Kim, V. Kolmogorov, and R. Zabih. Visual correspondence using energy minimization and mutual information. ICCV, 2003.

[26] H.D. Ellis, "Introduction to aspects of face processing: Ten questions in need of answers", in Aspects of Face Processing, H. D. Ellis, M. Jeeves, F. Newcombe, and A. Young Eds. Dordrecht:Nijhoff, 1986, pp.3-13

[27] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

[28] L. Sirovich and M. Kirby, "Low-Dimensional procedure for the characterisation of human faces," J. Optical Soc. of Am., vol. 4, pp. 519- 524, 1987.

[29] M. Kirby and L. Sirovich, "Application of the Karhunen- Loève procedure for the characterisation of human faces," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, pp. 831-835, Dec.1990

[30] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems", 1936.

[31] Ming-Hsuan Yang, Member, IEEE, David J. Kriegman, Senior Member, IEEE, and Narendra Ahuja, Fellow, IEEE, "Detecting Faces in Images: A Survey," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 1, JANUARY 2002

[32] KIRBY, M. AND SIROVICH, L. 1990. Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Trans. Patt. Anal. Mach. Intell. 12

[33] Tibe´ rio S. Caetano, Julian J. McAuley, Student Member, IEEE, Li Cheng, Member, IEEE, Quoc V. Le, and Alex J. Smola, "Learning Graph Matching," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 31, NO. 6, JUNE 2009.

[34] Linda G. Shapiro and Robert M. Haralick. Structural descriptions and inexact matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 3(5):504 – 519, September 1981.

[35] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. Von Der Malsburg, R.P. Wurtz, and M. Konen, “Distortion Invariant object recognition in the dynamic link architecture,” IEEE Trans. Computers, vol. 42, pp. 300 311, 1993.

[36] Wiskott and C. von der Malsburg, “Recognizing faces by dynamic link matching,”

[37] Y. Tian, T. Kanade, and J.F. Cohn, “Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity”, Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 218 –223, May 2002.

[38] A. Yuille, D. Cohen, and P. Hallinan, “Facial feature extraction from faces using deformable templates”, Proc. IEEE Computer Soc. Conf. On Computer Vision and Pattern Recognition, pp. 104-109, 1989.

[39] T.C. Chang, T.S. Huang, and C. Novak, “Facial feature extraction from colour images”, Proceedings of the 12th IAPR International Conference on Pattern Recognition, vol. 2, pp. 39-43, Oct 1994.

[40] T. Kanade, Computer Recognition of Human faces, Basel and Stuttgart: Birkhauser, 1997.

[41] Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29 (1996) 51–59.

[42] G. Donato, M. Bartlett, J. Hager, P. Ekman, T. Sejnowski, Classifying facial actions, IEEE Transactions on Pattern Analysis and Machine Intelligence 21(10) (1999) 974–989.

[43] B. Fasel, J. Luettin, Automatic facial expression analysis: a survey, Pattern Recognition 36 (2003) 259–275.

[44] Y. Tian, T. Kanade, J. Cohn, Recognizing action units for facial expression analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (2) (2001) 97–115.

[45] Eisner (2012), "Classifier Evaluation (Measuring Prediction Performance)", 2012 [cited August 2012]; Available from: webdocs.cs.ualberta.ca/.

[46] B. D. Lucas and T. Kanade (1981), An iterative image registration technique with an application to stereo vision. Proceedings of Imaging Understanding Workshop, pages 121—130.

Abstract

Recently, some web-based software systems tend to increase security level of user authentication process. Web developers need to authenticate some users in a more secure way; users like web administrators, distance learning users and so on. In this thesis, web-based authentication is done through face recognition using webcam camera. This thesis presents an operational system for web-based face recognition. The system divides into three parts including "Add user applet", "User recognition applet" and "Face recognition Java server". Client-side applets grabs a frame, applies appropriate preprocessing operations, detects the face in the frame and sends the frame to Java server. Train and test operations for face recognition with related database are placed in Java server.

The aim of presenting an operational web-based face recognition was to solve a tiny problem of the country in the field which needed software-oriented implementation. Moreover, we've proposed a novel approach for face recognition. The idea initiated from the fact that human vision system takes some approaches than existing face recognition algorithms. Most of face recognition algorithms have treated faces as an objects or textures, but face is definitely broader than that. Our approach analyzes face descriptions by human beings and concluded that all these descriptions are the most discriminant features with respect to what we call average face. More specifically, people living in a certain geographical region have created an average face in their minds from the people in that area and whenever they want to describe a face, they refer to most discriminant features with respect to the average face. If the people in that region all have long nose while the person in question have broad nose, it can be a discriminant feature to specify the person by. Therefore, the most discriminant face descriptors with respect to regional face average are used to describe and specify a specific person. The approach has been mathematically formulated.

Two databases are used to benchmark the operational web-based face recognition system. One database is prepared and collected by the system itself and the other is Extended Yale Face Database B. IMM Frontal Face database is used to benchmark the proposed method. Results and discussion are fully explained in the fifth chapter.

Keywords: Face recognition, web, authentication.

Shahrood University of Technology

Faculty of Engineering

Computer Engineering and IT Department

Design and Implementation of a Web-based Authentication System through Face Recognition

By:

Iman Firouzian

Under Supervision of:

Dr. Morteza Zahedi

A thesis submitted to the graduate studies office in

partial fulfillment of the requirements for the degree of

Master of Science in **Computer Engineering**

September 2013