بسم الله الرحمن الرحيم

دانشکده مهندسی برق و رباتیک

مهندسی رباتیک گرایش بینایی ربات

پایان‌نامه کارشناسی ارشد

احراز هویت چهره با استفاده از ویژگی‌های مکان و فرکانس تصویر در مجموعه‌های بزرگ

نگارنده: خشایار نوریان

استاد راهنما

دکتر حسین خسروی

شهریور ۱۳۹۶

شماره: ۴۰۵۰۱/۱/۲۰۰۲
تاریخ: ۹۶/۶/۱۴

باسمه‌تعالی

دانشگاه صنعتی شاهرود
مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای **خشایار نوریان چالی** با شماره دانشجویی ۹۳۱۷۸۸۴ رشته **مهندسی رباتیک** گرایش **رباتیک** تحت عنوان: **احراز هویت چهره با استفاده از ویژگی های مکان و فرکانس تصویر در مجموعه های بزرگ** که در تاریخ ۱۳۹۶/۰۶/۱۴ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با امتیاز ۱۸٫۷۵ درجه عالی) ☑ مردود ☐
نوع تحقیق: نظری ☑ عملی ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱- استادراهنمای اول	حسین خسروی	استادیار	
۲- استادراهنمای دوم	—	—	—
۳- استاد مشاور	—	—	—
۴- نماینده تحصیلات تکمیلی	مسعود مهدی زاده	استادیار	
۵- استاد ممتحن اول	علیرضا احمدی فرد	دانشیار	
۶- استاد ممتحن دوم	احمد حسین نایبی آستانه	استادیار	

نام و نام خانوادگی رئیس دانشکده: علیرضا احمدی فرد

تاریخ و امضاء و مهر دانشکده:

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

تقدیم به

پدر و مادر عزیز، دلسوز و مهربانم که برایم آرامش روحی و آسایش فکری فراهم نمودند تا با حمایت های همه جانبه در محیطی مطلوب، مراتب تحصیلی و پایان نامه درسی را به نحو احسن به اتمام برسانم.

تشکر و قدردانی

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمتهای او ندانند و کوشندگان، حق او را گزاردن نتوانند.

بدون شک جایگاه و منزلت معلم، بالاتر از آن است که در مقام قدردانی از زحمات آنها، با زبان قاصر و دست ناتوان چیزی بنگاریم؛ اما بر حسب وظیفه بر خود واجب میدانم از استاد شایسته و فرهیخته جناب آقای دکتر خسروی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این رساله را بر عهده گرفتند کمال تشکر و قدردانی را بجا آورم. همچنین در انتها با تشکر دوباره از پدر و مادر عزیز و مهربانم، از تمامی دوستانم که به نوعی در موفقیت و پیشرفت من نقش داشته اند کمال تشکر را دارم. باشد که این خردترین بخشی از زحمات آنان را سپاس گوید.

تعهد نامه

اینجانب **خشایار نوریان** دانشجوی دوره کارشناسی ارشد رشته مهندسی برق/رباتیک دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود نویسنده پایان نامه **احراز هویت چهره با استفاده از ویژگی‌های مکان و فرکانس تصویر در مجموعه‌های بزرگ** تحت راهنمائی دکتر حسین خسروی متعهد می شوم .

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

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

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

چکیده:

بازشناسی چهره یکی از روش‌های مرسوم در تعیین هویت افراد است که کاربردهای زیادی از جمله ورود به دستگاه‌های شخصی، ردگیری مجرمین، تعامل انسان با کامپیوتر و ... دارد. هدف در بازشناسی چهره تعیین هویت فرد از یک تصویر با استفاده از پایگاه داده است. برای بازشناسی چهره چالش‌های زیادی از جمله تغییرات روشنایی، تغییرات حالات چهره، وجود مانع، تغییرات سن، تعداد کم نمونه‌های آموزش، تعداد افراد زیاد و ... وجود دارد که در این پایان‌نامه به ارائه روشی می‌پردازیم که در برابر تغییرات روشنایی، حالت چهره و تغییر سن مقاوم باشد.

در این پایان‌نامه بعد از آشکارسازی چهره ویژگی‌های محلی HOG، LBP و STFT استخراج می‌شود. بر اساس ویژگی محلی، پیش‌پردازش انجام و پارامترهای آن ویژگی تنظیم می‌شود تا بهترین عملکرد را داشته باشد. سپس با چسباندن بردارهای ویژگی‌های بهینه، یک بردار ویژگی ساخته شده و با استفاده از کلاسه بند نزدیک‌ترین همسایگی و معیار فاصله منهتن به کلاسه‌بندی می‌پردازیم. برای افزایش نرخ بازشناسی نمونه‌های مجازی از تصویر آموزش تولید کردیم. این تصاویر با اعمال تغییرات مات کردن، تیز کردن لبه‌ها و چرخش تصویر می‌باشد سپس تصاویر تولید شده به مجموعه آموزش اضافه می‌شوند. برای ارزیابی روش پیشنهادی، از پایگاه داده FERET استفاده کرده‌ایم که شامل تصاویر حدود ۱۰۰۰ نفر است و به ازای هر فرد بیشتر از یک تصویر چهره در مجموعه آموزش وجود ندارد. لازم به ذکر است که بیشتر روش‌های ارائه‌شده در بازشناسی چهره به حل یکی از چالش‌های موجود پرداخته‌اند و روی پایگاه داده‌های کم‌تر از ۲۰۰ نفر آزمایش‌شده‌اند. نتایج تجربی روی پایگاه داده Feret نرخ بازشناسی ۹۶.۹۶٪ با استفاده از نمونه‌های مجازی و ۹۵.۹۵٪ بدون استفاده از نمونه‌های مجازی، روی مجموعه FB با استفاده از ترکیب دو ویژگی HOG و STFT در زمان ۱۵۸ میلی‌ثانیه برای هر تصویر از مجموعه آزمون بدست آمد.

کلمات کلیدی: بازشناسی چهره، ویژگی‌های محلی، استخراج ویژگی، پایگاه داده حجیم

فهرست مطالب

فهرست شکل‌ها

فهرست جدول‌ها

فصل ۱: مقدمه

۱-۱ تعریف مسئله

بازشناسی چهره یکی از روش‌های زیست‌شناختی[1] برای تعیین هویت افراد است. این موضوع یک مسئله چالش‌برانگیز در زمینه بینایی ماشین، پردازش تصویر و شناسایی الگو می‌باشد. از روش‌های زیست‌شناختی دیگر می‌توان به تشخیص اثرانگشت و شناسایی عنبیه چشم اشاره کرد. در حال حاضر به دلیل چالش‌های زیادی که در بازشناسی چهره وجود دارد کارآمدترین یا قابل‌اعتمادترین نیست اما دو مزیت مهم بازشناسی چهره این است که احتیاج به همکاری فرد مورد آزمایش ندارد و دیگری در مقیاس بزرگ‌تری قابل‌استفاده است.

در مسئله بازشناسی چهره، ورودی سیستم یک تصویر از چهره شناخته نشده است و هدف تعیین هویت این چهره با استفاده از پایگاه داده است.

۱-۲ کاربردهای سیستم‌های بازشناسی چهره

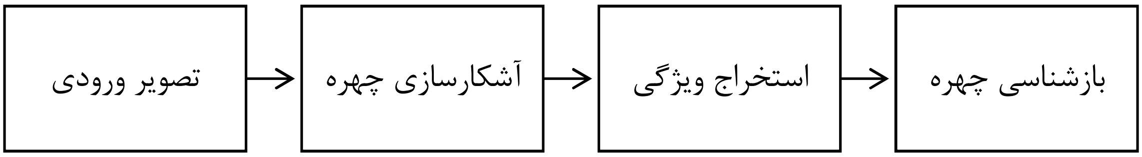
- سرگرمی: بازی‌های کامپیوتری و ارتباط بین انسان و ربات
- کارت‌های هوشمند: تأیید تصاویر گواهینامه و کارت شناسایی
- امنیت اطلاعات: ورود به دستگاه‌های شخصی، امنیت برنامه‌ها و رمزگذاری فایل‌ها
- پلیس و نظارت: نظارت ویدئویی، ردگیری و شناسایی مجرمین

۱-۳ تفاوت بین آشکارسازی و بازشناسی چهره

هدف آشکارسازی چهره[2]، مکان‌یابی و جدا کردن تصویر چهره از تصویر ورودی است اما در بازشناسی چهره[3] هدف تشخیص هویت چهره موجود در تصویر با توجه به پایگاه داده موجود است. به عبارتی آشکارسازی چهره، بخشی از سامانه بازشناسی چهره است.

---
1 Biometric
2 Face detection
3 Face recognition

۱-۴ پیکربندی کلی یک سیستم بازشناسی چهره



شکل ۱-۱ پیکربندی کلی یک سیستم بازشناسی چهره

هر سیستم بازشناسی چهره از ۳ بخش اساسی تشکیل‌شده است[۱]:

۱. قطعه‌بندی چهره‌ها (آشکارسازی چهره):

اولین گام در هر سیستم خودکار بازشناسی چهره، آشکارسازی چهره است یعنی در تصویر ورودی داده‌شده، چهره در کجای تصویر قرارگرفته است.

۲. استخراج ویژگی از نواحی چهره:

بعدازاینکه ناحیه چهره تشخیص داده و جدا شد، پس از بهبود کیفیت به استخراج ویژگی پرداخته می‌شود، ویژگی استخراج شده نرمال‌سازی شده و به‌عنوان ورودی به کلاسه بند داده می‌شود.

۳. شناسایی (احراز هویت[1]) یا تصدیق هویت[2]:

ویژگی‌های استخراج‌شده را به یک کلاسه بند می‌دهیم تا با مقایسه ویژگی استخراج‌شده با پایگاه داده، هویت فرد مورد نظر را شناسایی نماید.

۱-۵ مختصری بر تاریخچه بازشناسی چهره:

شروع کار روی سیستم‌های خودکار بازشناسی چهره از دهه ۱۹۷۰ میلادی است. در ابتدا مسئله بازشناسی چهره به‌عنوان مسئله تشخیص اشیاء سه‌بعدی از روی تصاویر دوبعدی شناخته

1 Authentication
2 Verification

می‌شد و با استفاده از طبقه بندهای کلاسیک و با استفاده از ویژگی‌های فاصله (به‌طور مثال فاصله بین نقاط مهم چهره) به طبقه‌بندی می‌پرداختند[۱].

در حال حاضر، روش‌های مختلف بازشناسی چهره را می‌توان به سه دسته کلی زیر تقسیم‌بندی کرد[۱]:

۱. **روش‌های تطبیق سراسری:** در این روش کل منطقه چهره به‌عنوان ورودی به سیستم داده می‌شود. از معروف‌ترین روش‌ها می‌توان به چهره ویژه[1][۲]، تحلیل جداساز خطی[2][۳]، تحلیل مؤلفه مستقل[۴][3] و غیره اشاره کرد.

۲. **روش‌های مبتنی بر ویژگی محلی:** در این روش از ویژگی‌های محلی استفاده می‌شود. موقعیت و ویژگی‌های آماری چشم‌ها، بینی و دهان در ابتدا استخراج شده، سپس به یک طبقه بند ساختاری داده می‌شود. در این روش چالش اصلی در استخراج ویژگی‌ها است. تفاوت این روش‌ها در نوع ویژگی مورداستفاده است. سه روش تشخیص ویژگی که استفاده می‌شوند عبارت‌اند از:

   ۱. روش‌های مبتنی بر لبه، خطوط و منحنی‌ها.
   ۲. روش‌های مبتنی بر قالب چهره
   ۳. روش‌های تطبیق ساختاری که می‌توان ویژگی‌ها را بر اساس قالب هندسی خاصی استخراج کرد.

۳. **روش‌های ترکیبی:** این روش‌ها ترکیب روش تطبیق سراسری و استخراج ویژگی محلی است.

---

1 Eigenfaces
2 Linear Discriminant Analysis(LDA)
3 Independent Component Analysis(ICA)

۱-۶ چالش‌های سیستم بازشناسی چهره:

- تغییرات روشنایی
- تغییر حالت چهره[1]
- انسداد[2] (پوشیده شدن بخشی از چهره)
- تغییر موقعیت[3]
- چرخش
- تغییرات اندازه
- تغییرات سن
- کم بودن تعداد نمونه‌های آموزش از هر فرد
- تغییرات پس‌زمینه
- زیاد بودن تعداد افراد موجود در پایگاه داده
- شناسایی بلادرنگ

۱-۷ هدف پایان‌نامه

هدف این پایان‌نامه طراحی سیستمی برای بازشناسی چهره با تعداد زیاد افراد پایگاه داده با استفاده از ترکیب ویژگی‌های فرکانسی مثل اطلاعات اندازه و فاز تبدیل فوریه، مثل STFT [4] و ویژگی‌های حوزه مکان، مثل هیستوگرام جهت‌دار گرادیان (HOG [5]) است.

فرض‌های زیر در انجام پایان‌نامه در نظر گرفته‌شده:

- تصاویر چهره از روبرو[6] هستند و تغییر موقعیت نداریم.

1 Facial expression
2 Occlusion
3 Pose variation
4 Short Time Fourier Transform
5 Histogram of Oriented Gradient
6 Frontal Face

- تصاویر در محیط کنترل‌شده گرفته‌شده باشند یعنی تغییرات پس‌زمینه نداریم.

۱-۸ چالش‌های پایان‌نامه

در بین چالش‌های موجود تمرکز بر چالش‌های زیر است:

- تعداد زیاد افراد
- تعداد نمونه‌های آموزش، یک نمونه از هر کلاس است.
- تفاوت سنی بین تصاویر آموزش و آزمون.
- بلادرنگ بودن

۱-۹ ساختار پایان‌نامه

این پایان‌نامه در ۵ فصل تدوین‌شده است. در فصل دوم مروری بر کارهای گذشته روی بازشناسی چهره و معرفی پایگاه داده‌های موجود انجام می‌شود. فصل سوم به طرح مباحث نظری مورداستفاده در این پایان‌نامه می‌پردازد. در فصل چهارم روش پیشنهادی و تأثیر پارامترهای مختلف بر الگوریتم پیشنهادی بررسی می‌شود. فصل پنجم به گزارش نتایج روش پیشنهادی و مقایسه با الگوریتم‌های دیگر و نتیجه‌گیری و پیشنهادهایی برای کارهای آینده می‌پردازد.

فصل ۲: مروری بر کارهای گذشته

۱-۲ مقدمه

در ابتدا از ویژگی‌های هندسی برای بازشناسی چهره استفاده می‌شد، یکی از اولین کارها در بازشناسی چهره در [۵] مطرح شد که از نقاط شاخص (موقعیت چشم‌ها، گوش‌ها، بینی و ...) برای ساخت یک بردار ویژگی (فاصله و زاویه بین این نقاط) استفاده شد سپس با محاسبه فاصله اقلیدسی بین بردارهای ویژگی بین تصاویر آزمون و مرجع، عمل بازشناسی انجام شد. پیدا کردن دقیق نقاط شاخص در این روش به‌صورت دستی انجام شد به دلیل اینکه پیدا کردن خودکار این نقاط کار بسیار مشکل است. یکی از آخرین کارهای انجام‌شده با استفاده از ویژگی‌های هندسی کار انجام شده در [۶] است. در این تحقیق با استفاده از یک بردار ۲۲ بعدی به کلاسه‌بندی پرداخته می‌شود. همچنین این تحقیق نشان داد که روی پایگاه داده‌های بزرگ‌تر از ۴۰ نفر، ویژگی‌های هندسی به‌تنهایی نمی‌توانند عمل بازشناسی را انجام دهند.

روش چهره ویژه که در [۲] مطرح‌شده یکی از روش‌های مبتنی بر تطبیق سراسری است، در این روش هر تصویر چهره به‌عنوان یک نقطه در فضای تصویر با ابعاد بالا(تعداد پیکسل‌های تصویر)، به‌صورت یک نمایش با ابعاد کمتر در فضای ویژگی در می‌آید که کلاسه‌بندی در این فضا آسان‌تر است، زیر فضای کم ویژگی به‌وسیله تحلیل مؤلفه اصلی (PCA) پیدا می‌شود که محورهای این زیر فضا پراکندگی بین داده‌ها را بیشینه می‌کند. هرچند ازنظر بازسازی، این تبدیل، بهینه است اما اطلاعات کلاس‌ها را در این تبدیل لحاظ نمی‌کند. در [۳] یک افکنش[۱] مبنی بر کلاس به همراه تحلیل جداساز خطی (LDA) برای بازشناسی چهره استفاده شد، ایده اصلی این کار کمینه‌سازی واریانس درون کلاسی و بیشینه‌سازی واریانس بین کلاسی به‌صورت هم‌زمان است.

در بیست سال اخیر روش‌های مختلفی مبنی بر استخراج ویژگی‌های محلی برای بازشناسی چهره معرفی شدند، برای اجتناب از ابعاد بالا داده ورودی فقط نواحی محلی تصویر توصیف می‌شوند. به‌طور مثال در[۷] بر اساس ویژگی گابور، در [۸] با استفاده از الگوهای محلی باینری، در [۹]با

1 Projection

استفاده از مدل مخفی مارکوف و تبدیل کسینوسی گسسته، در [۱۰] با استفاده از ویژگی تبدیل هاف، در [۱۱] بر اساس ویژگی SIFT و در [۱۲] بر اساس ویژگی مکان فرکانسی با استفاده از پنجره‌های هم‌مرکز به بازشناسی چهره پرداخته‌اند.

روش‌های ترکیبی به استفاده از ویژگی‌های محلی و سراسری به‌صورت هم‌زمان می‌پردازند، می‌توان برخی از روش‌های ویژگی محلی را نیز در این دسته قرارداد. علاوه بر دسته‌بندی فوق روش‌های مبتنی بر شبکه عصبی[۱۳, ۱۴] و یادگیری عمیق[۱۵] و روش‌های مبتنی بر نمایش تنک [۱۶-۱۸] نیز برای بازشناسی چهره استفاده‌شده است. در ادامه کارهای مهم گذشته در بازشناسی چهره را بر اساس دسته‌بندی مطرح‌شده در فصل اول بررسی می‌کنیم.

۲-۲ روش‌های مبتنی بر تطبیق سراسری

در این روش کل منطقه چهره به‌عنوان ورودی به سیستم داده می‌شود که در زیر روش‌های معروف را بررسی می‌کنیم.

۲-۲-۱ چهره ویژه

یکی از مشکلات تصاویر چهره ابعاد بالا در نمایش تصویر است، یک تصویر دوبعدی با اندازه p*q یک بردار m=pq بعدی را می‌دهد، در این فضای m بعدی، مؤلفه‌هایی مفید هستند که بیشترین واریانس را دارند. در [۲] با معرفی روشی به نام چهره ویژه به حل این مسئله پرداخته و نمایش دیگری را برای تصاویر چهره معرفی کرده است. ایده اصلی این روش یافتن مجموعه کوچکی از مؤلفه‌ها است که بیشترین اطلاعات را داشته باشند، با استفاده از PCA دنبال زیر فضایی هستیم که مؤلفه‌ها در این فضا نسبت به هم، وابستگی نداشته و بیشترین واریانس را داشته باشند، به این مؤلفه‌ها، مؤلفه‌های اصلی گفته می‌شود.

- **جزئیات الگوریتم PCA به شرح زیر است:**

فرض کنیم $X = \{x_1,x_2,\cdots,x_n\}$ برداری تصادفی باشد. هرکدام از x_i ها یک تصویر چهره و n تعداد تصاویر می‌باشد.

۱. با استفاده از رابطه ۱-۲ میانگین μ محاسبه می‌شود.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{۱-۲}$$

۲. با استفاده از رابطه ۲-۲ ماتریس کوواریانس S محاسبه می‌شود.

$$S = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^T \tag{۲-۲}$$

۳. محاسبه مقدار ویژه‌های λ_i و بردار ویژه‌های v_i از ماتریس S با حل رابطه ۳-۲.

$$S\mathrm{v}_i = \lambda_i \mathrm{v}_i \ i = 1,2,\cdots,\mathrm{m} \tag{۳-۲}$$

۴. مرتب کردن بردارهای ویژه به‌صورت نزولی بر اساس مقدار ویژه‌های متناظر و سپس برداشتن K بردار اول $W = \{v_1,v_2,\cdots,v_k\}$.

۵. با استفاده از رابطه ۴-۲ تبدیل به فضای *PCA* و با استفاده از رابطه ۵-۲ بازسازی از فضای *PCA* انجام می‌شود.

$$y = W^T(x - \mu) \tag{۴-۲}$$

$$x = Wy + \mu \tag{۵-۲}$$

روش چهره ویژه برای بازشناسی چهره به شکل زیر عمل می‌کند:

- محاسبه بردار y برای تمامی نمونه‌های آموزش
- محاسبه بردار y برای تصویر آزمون
- پیدا کردن نزدیک‌ترین همسایه بین تصویر آزمون و تصاویر آموزش در فضای *PCA*

این روش روی پایگاه داده‌ای خودساخته متشکل از ۲۵۰۰ تصویر و ۱۶ نفر که از هر شخص ترکیبات ۳ حالت از تغییرات زاویه، اندازه، روشنایی و وضوح[1] که درمجموع ۵۴ تصویر از هر نفر را

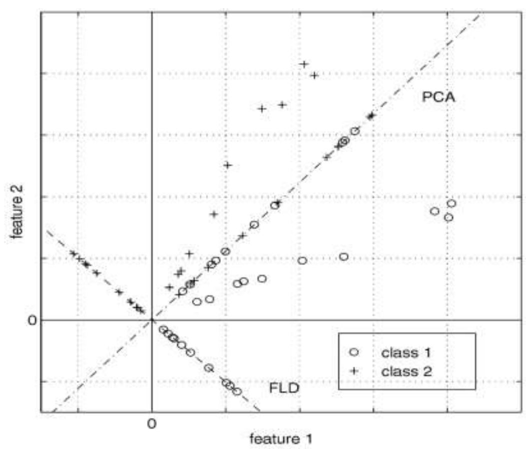
---

1 Resolution

شامل می‌شد آزمایش شد که برای تغییرات روشنایی نرخ ۹۶٪، برای تغییرات زاویه ۸۵٪ و برای تغییرات اندازه به ۶۴٪ نرخ بازشناسی رسید[۲].

۲-۲-۲ چهره فیشر

ایده اصلی روش چهره ویژه، PCA است که ترکیب خطی از مؤلفه‌هایی را پیدا می‌کند که واریانس بین کلاسی را بیشینه می‌کند، اگرچه روشی مؤثری برای نمایش داده‌ها است اما این روش کلاس‌ها را نادیده می‌گیرد درنتیجه ممکن است مؤلفه‌هایی دور ریخته شود که اطلاعات مفیدی برای کلاسه‌بندی در خود دارند. در[۳] پیشنهاد شد که بجای PCA از LDA استفاده شود، در LDA برای پیدا کردن ترکیب مؤلفه‌هایی که بهترین جداسازی را بین کلاس‌ها را انجام می‌دهند بجای بیشینه‌سازی واریانس کلی داده‌ها، نسبت بین واریانس داده‌های درون کلاسی و برون کلاسی بیشینه‌سازی می‌شود. به‌طور مثال درشکل ۲-۱ مقایسه‌ای بین PCA و FLD[1] در جداسازی یک مسئله دو کلاسی با دو بعد را مشاهده می‌کنیم.



شکل ۲-۱ مقایسه PCA و FLD [۳]

- **جزئیات الگوریتم FLD به شرح زیر است:**

1 Fisher Linear Discriminant

فرض کنیم X برداری تصادفی از نمونه‌های استخراج‌شده از C کلاس باشد که X_i برداری از تصاویر چهره کلاس i ام است و x_i تصویر i ام در این مجموعه است.

$$X = \{X_1, X_2, \cdots, X_C\}$$
$$X_i = \{x_1, x_2, \cdots, x_n\}$$

۱. محاسبه ماتریس پراکندگی درون کلاسی S_W و ماتریس برون کلاسی S_B با استفاده از روابط ۶-۲ و ۷-۲.

$$\mathrm{S_B} = \sum_{i=1}^{C} N_i(\mu_i - \mu)(\mu_i - \mu)^T \qquad (۶\text{-}۲)$$

$$\mathrm{S_W} = \sum_{i=1}^{C} \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T \qquad (۷\text{-}۲)$$

که در روابط بالا μ میانگین کلی داده‌ها که با رابطه ۸-۲ و μ_i میانگین کلاس i ام است که با رابطه ۹-۲ محاسبه می‌شود. همچنین N تعداد کل نمونه‌ها و N_i نمونه‌های آموزشی کلاس i ام است.

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad (۸\text{-}۲)$$

$$\mu_i = \frac{1}{|X_i|} \sum_{x_j \in X_i} x_j \qquad (۹\text{-}۲)$$

در الگوریتم کلاسیک *LDA* برای پیدا کردن افکنش *W* که معیار جدا پذیری کلاسی را بیشینه‌سازی می‌کند به‌صورت زیر عمل می‌شد:

$$W_{opt} = \arg\max_{w} \frac{|W^T S_B W|}{|W^T S_W W|} \qquad (۱۰\text{-}۲)$$

در [۳] جوابی برای حل این مسئله بهینه‌سازی با حل معادله مقدار ویژه عمومی به دست می‌آید

$$S_B v_i = \lambda_i S_w v_i \ \Rightarrow S_w^{-1} S_B v_i = \lambda_i v_i \qquad (۱۱\text{-}۲)$$

برای جلوگیری از تکین شدن ماتریس S_W در [۳] ابتدا با PCA داده‌ها را به فضای (N-C) بعدی می‌برند سپس الگوریتم LDA را روی‌داده‌های کاهش داده‌شده اجرا می‌کنند، مسئله بهینه‌سازی را به‌صورت زیر بازنویسی می‌کنیم:

$$W_{pca} = \arg\max_{w} |W^T S_T W| \quad (۱۲-۲)$$

$$W_{fld} = \arg\max_{w} \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|} \quad (۱۳-۲)$$

ماتریس انتقالی که داده‌ها را به فضای (C-1) بعدی می‌برد از معادله زیر به دست می‌آید:

$$W = W_{fld}^T W_{pca}^T \quad (۱۴-۲)$$

روش چهره فیشر سپس به عمل بازشناسی چهره به روش زیر می‌پردازد:

- با استفاده از ماتریس انتقال W داده‌های آموزش را به فضای جدید می‌برد.
- تصویر آزمون موردنظر را با ماتریس انتقال به فضای جدید می‌برد.
- نزدیک‌ترین همسایه را برای تصویر آزمون بین تصاویر آموزش پیدا می‌کند.

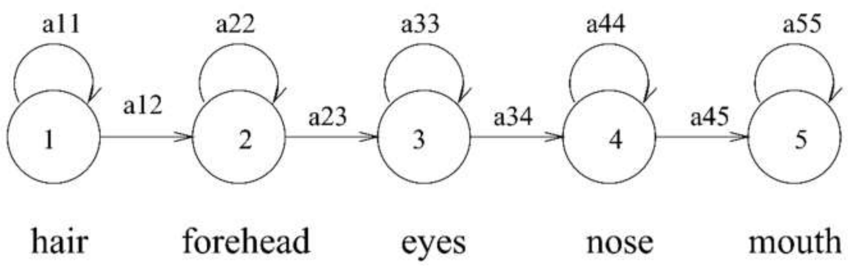
۲-۳ روش‌های مبتنی بر ویژگی‌های محلی

در این روش از ویژگی‌های محلی استفاده می‌شود. ایده اصلی استفاده از ویژگی‌های محلی این است که بجای استفاده از کل تصویر به‌عنوان برداری با ابعاد بالا، فقط ویژگی‌های محلی را مورداستفاده قرار دهیم. ازآنجایی‌که تصاویر ما تغییراتی در روشنایی، انتقال، چرخش و تغییرات اندازه دارند درنتیجه ویژگی‌های محلی نیز باید تا حدودی نسبت به این تغییرات مقاوم باشند. در این روش ویژگی‌های محلی در ابتدا استخراج‌شده سپس به یک طبقه بند داده می‌شود. تفاوت روش‌های مختلف در این دسته در نوع ویژگی مورداستفاده است. در ادامه چند روش معروف را بررسی می‌کنیم.

۲-۳-۱مدل مخفی مارکوف

مدل مخفی مارکوف[1] (HMM) مجموعه‌ای از مدل‌های آماری است که برای توصیف مشخصات سیگنال‌ها مورداستفاده قرار می‌گیرد، در بحث شناسایی گفتار و شناسایی حروف کارایی خیلی خوبی دارد. فرض می‌شود مدل سیستم یک فرآیند مارکوف با پارامترهای ناشناس باشد، هدف پیدا کردن این پارامترها از نمونه‌های مشاهده است. هر حالت فرآیند یک توزیع احتمالاتی روی خروجی احتمالی دارد که هر حالت فرآیند مارکوف قابل‌مشاهده است. در [۹] روشی را برای بازشناسی چهره بر اساس HMM پیشنهاد داده است، این روش بر اساس استخراج ویژگی تبدیل کسینوسی گسسته [2] (DCT) دوبعدی است. تصویر چهره ورودی به بلوک‌هایی تقسیم می‌شود که هر بلوک به یک بردار مشاهده متصل است. HMM از زنجیره‌ای غیرقابل مشاهده ولی متناهی از حالات سیستم، نماد ماتریس احتمالاتی مشاهده B، ماتریس احتمالاتی انتقال A، حالت اولیه π و مجموعه‌ای از توابع احتمالاتی تشکیل می‌شود. یک HMM به شکل یک سه‌تایی $\lambda = (A,\mathrm{B},\pi)$ تعریف می‌شود.

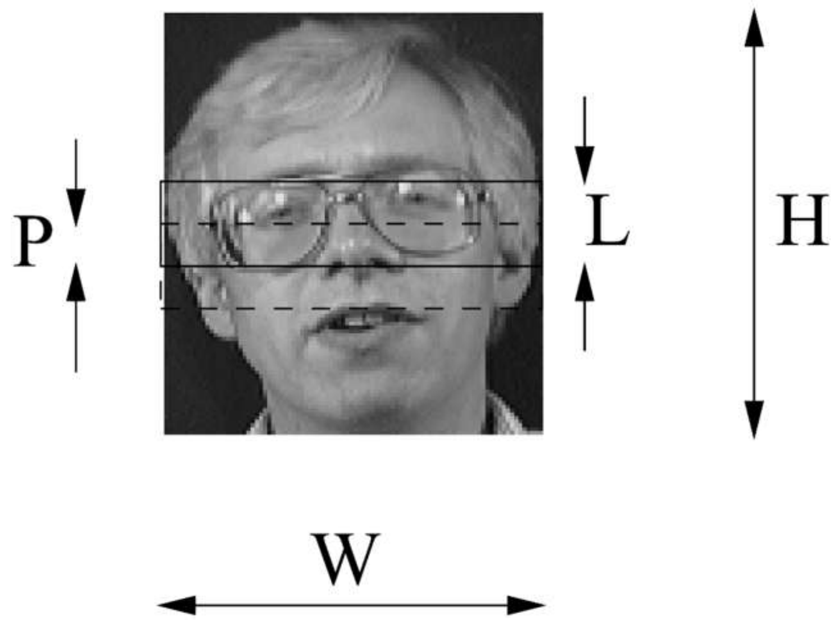
برای تصویر چهره از روبرو، اجزاء مهم چهره به ترتیب بالا به پایین: موها، پیشانی، چشم‌ها، بینی، دهان و چانه است، هرکدام از نواحی چهره به یکی از حالات فرآیند HMM نسبت داده می‌شود، در شکل ۲-۲ احتمال انتقال a_{ij} و ساختار مدل چهره مشخص‌شده است.



شکل ۲-۲ مدل HMM برای بازشناسی چهره[9]

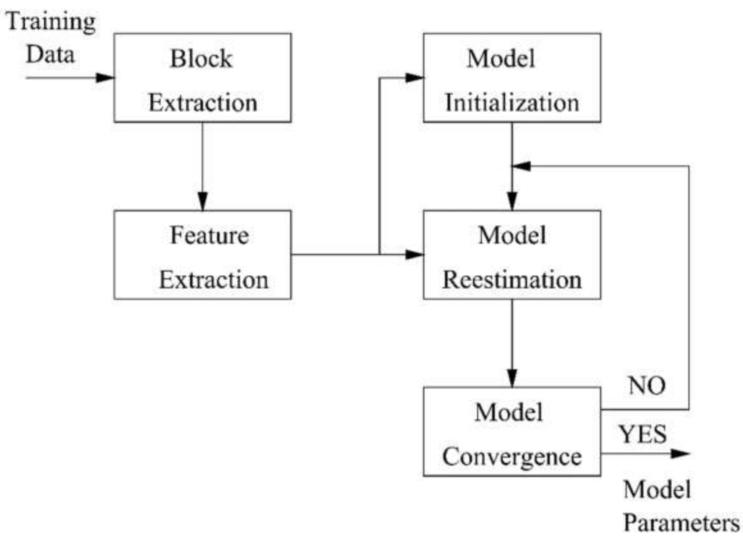
هر تصویر ورودی با عرض W و ارتفاع H به بلوک‌هایی با همپوشانی به ارتفاع L تقسیم می‌شود. در شکل ۲-۳ استخراج بلوک‌ها نشان داده است.

[1] Hidden Markov Model(HMM)
[2] Discrete Cosine Transform



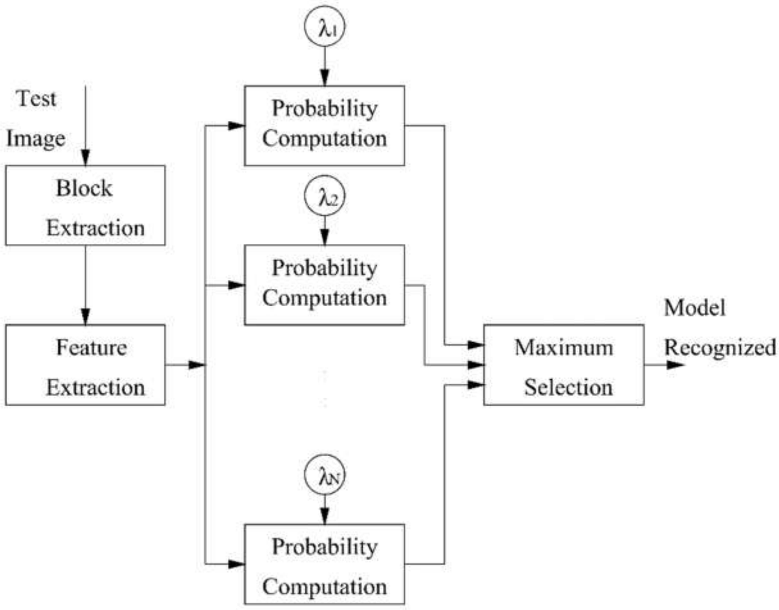
شکل ۲-۳ استخراج بلوک در الگوریتم HMM [۹]

انتخاب مقدار مناسب برای همپوشانی P و طول L تأثیر بسزایی در نرخ بازشناسی دارد چراکه ویژگی‌های استخراج‌شده مستقل از موقعیت عمودی هستند. بعد از استخراج بلوک‌ها، برای هر بلوک تعدادی از ضرایب فرکانس پایین در DCT استخراج می‌شود. در این تحقیق برای کاهش بردار ویژگی از پنجره‌ای به‌اندازه ۱۲*۳ استفاده شد تا اطلاعات مهم را استخراج کند. در مرحله آموزش بعد از جداسازی بخش‌های مختلف چهره که هر بخش متناظر با یک حالت از فرآیند مارکوف است و ساختن A، B و π می‌توان مدل را ساخت، در شکل ۲-۴ فرآیند آموزش را می‌بینیم.



شکل ۲-۴ فرآیند آموزش HMM [۹]

در مرحله آزمون اگر تصویر آزمون توانست با داشتن مدل‌های تصاویر آموزش، احتمال مشاهدات هرکدام از مدل‌ها را بیشینه کند، آن مدل به‌عنوان برنده مشخص می‌گردد، در شکل ۲-۵ مرحله آزمون را می‌بینیم.



شکل ۲-۵ مرحله آزمون HMM [۹]

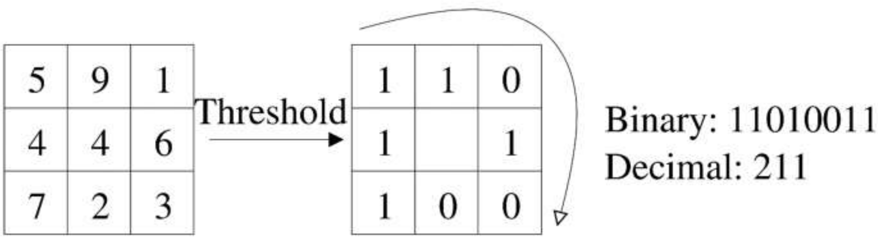
این تحقیق بر روی پایگاه داده ORL این روش و چهره ویژه را پیاده‌سازی کرده است. نتایج آن در جدول ۲-۱ می‌باشند.

جدول ۲-۱ نتایج HMM روی پایگاه داده ORL [۹]

روش	نرخ بازشناسی
HMM	٪۸۴
EigenFace	٪۷۳

۲-۳-۲الگوهای باینری محلی

ریشه استفاده از الگوهای باینری محلی[1] (LBP) در تشخیص و تحلیل بافت است[۱۹]. هدف الگوهای باینری محلی خلاصه کردن ساختارهای محلی در تصویر به‌وسیله مقایسه کردن هر پیکسل با همسایه‌های آن پیکسل است. در [۸] از این ویژگی برای عمل بازشناسی چهره استفاده‌شده است. برای استخراج این ویژگی مانند شکل ۲-۶ عمل می‌شود، برای هر پیکسل یک همسایگی در نظر گرفته می‌شود که پیکسل موردنظر در مرکز این همسایگی قرار می‌گیرد سپس با آستانه گذاری پیکسل‌های همسایه نسبت به این پیکسل اگر پیکسل‌های همسایه مقداری بزرگ‌تر یا مساوی پیکسل مرکزی داشتند مقدار ۱ در غیر این صورت مقدار ۰ را می‌گیرند، با شروع از یک پیکسل قراردادی در همسایگی یک الگوی باینری شکل می‌گیرد. با انجام این عمل برای همه پیکسل‌ها تصویر به‌دست‌آمده را تصویر LBP می‌گوییم.



شکل ۲-۶ اپراتور LBP اولیه روی همسایگی ۳*۳ مربعی

به‌طور مثال شکل ۲-۶، یک همسایگی مربعی ۳*۳ و عملیات انجام‌شده را نشان می‌دهد. تعریف این اپراتور به شکل زیر است.

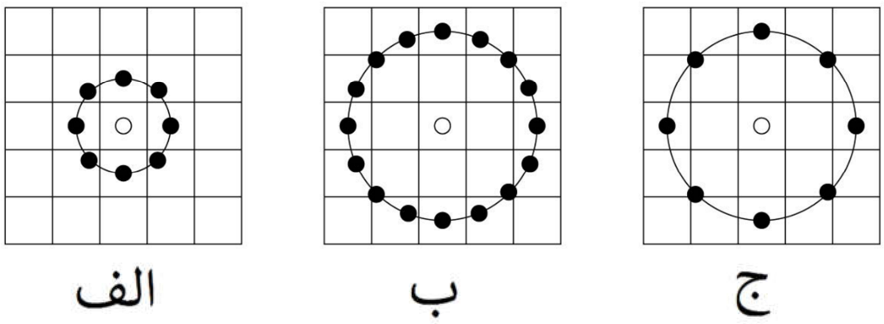
1 Local Binary Pattern(LBP)

$$LBP(x_c,y_c)=\sum_{p=0}^{P-1} 2^p Sgn(i_p - i_c) \qquad (۱۵-۲)$$

که P تعداد همسایه‌ها و i_p مقدار روشنایی همسایه p ام و i_c مقدار روشنایی پیکسل مرکزی است. همچنین:

$$SGN = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \qquad (۱۶-۲)$$

الگوریتم LBP توسعه‌های مختلفی دارد که به نوع همسایگی، نحوه قرارگیری همسایه‌ها، تعداد همسایه‌ها بستگی دارد در [۸] از LBP دایره‌ای استفاده‌شده است. این اپراتور را در شکل ۲-۷ می‌بینید.



شکل ۲-۷ اپراتور LBP دایره‌ای

(الف) شعاع ۱، ۸ همسایه، (ب) شعاع ۲، ۱۶ همسایه و (ج) شعاع ۲، ۸ همسایه

روش استفاده‌شده در این تحقیق برای بازشناسی چهره به شرح زیر است. در ابتدا از تصویر چهره ورودی، تصویر LBP ساخته می‌شود سپس تصویر LBP به m ناحیه محلی تقسیم‌شده، در هر ناحیه هیستوگرام محاسبه و با کنار هم قرار دادن این هیستوگرام‌ها بردار ویژگی ساخته می‌شود. سپس با استفاده از یک کلاسه بند نزدیک‌ترین همسایگی[1] (NN) به کلاسه‌بندی می‌پردازد. برای پیدا کردن نزدیک همسایه بین بردار ویژگی تصویر آزمون و پایگاه داده از معیار فاصله کای دو[2] زیر استفاده‌شده است.

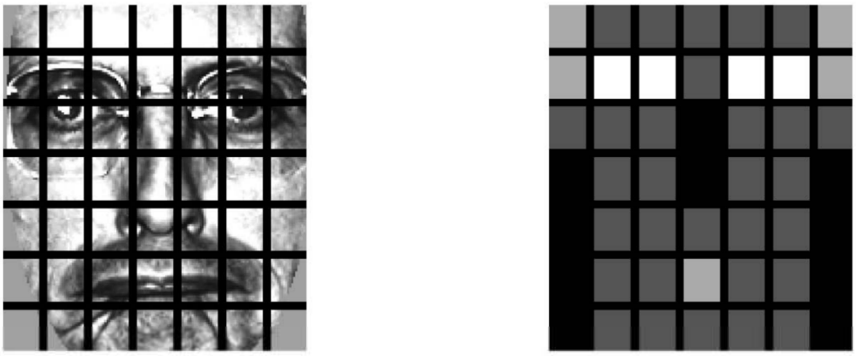
1 Nearest neighbor
2 Chi square

$$\chi^2(S,M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \qquad (۱۷-۲)$$

همچنین در این تحقیق بجای اینکه هر ناحیه تأثیر یکسانی در کلاسه‌بندی داشته باشد قبل از کلاسه‌بندی هر ناحیه در ضریب مخصوصی ضرب می‌شود، ایده این کار به آنجا برمی‌گردد که نواحی خاصی از تصویر اطلاعات مهم‌تری نسبت به بقیه نواحی دارند، به دست آوردن وزن‌ها روی زیرمجموعه‌ای از تصاویر پایگاه داده انجام شد سپس با اعمال این وزن‌ها عمل کلاسه‌بندی انجام شد. در حالت وزن‌دار از معیار فاصله زیر استفاده می‌شود.

$$\chi_w^2(S,M) = \sum_{i,j} W_j \frac{\left(S_{i,j} - M_{i,j}\right)^2}{S_{i,j} + M_{i,j}} \qquad (۱۸-۲)$$

که W_j وزن ناحیه j ام است. نمونه‌ای از این وزن دهی در شکل ۲-۸ قابل‌مشاهده است.



الف ب

شکل ۲-۸ چگونگی وزن دهی به بلوک‌ها

(الف) یک تصویر چهره با نواحی ۷*۷ و (ب) وزن‌های نواحی مشکی ۰، خاکستری روشن ۱، خاکستری تیره ۲، سفید ۴

نتایج اعمال این روش روی پایگاه داده Feret در جدول ۲-۲ قابل‌مشاهده است.

جدول ۲-۲ نتایج [۸] روی پایگاه داده Feret

روش	fb	fc	Dup1	Dup2	میانگین
LBP بی‌وزن	٪۹۳	%۵۱	٪۶۱	٪۵۰	%۷۶
LBP وزن‌دار	٪۹۷	%۷۹	٪۶۶	٪۶۴	%۸۱

۲-۳-۳استفاده از تبدیل هاف

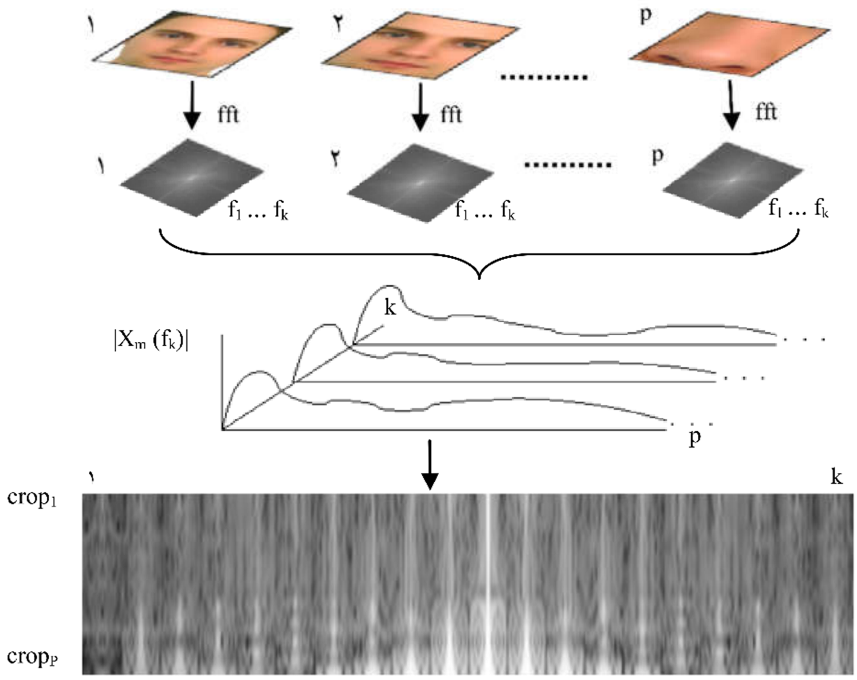
در [۱۰] روشی مبتنی بر استخراج ویژگی هاف برای حل مسئله تغییرات روشنایی پیشنهادشده است. در ابتدا بخش چهره از تصویر ورودی بر اساس روش پیشنهادی در [۲۰] جدا می‌شود و بعد از اعمال پیش‌پردازش (ترکیبی از یک یا چند روش متعادل‌سازی هیستوگرام، تبدیل لگاریتمی، تیز سازی لبه‌ها و تصحیح گاما) با استفاده از الگوریتم تشخیص لبه سوبل لبه‌ها را تشخیص می‌دهد. لبه‌های به‌دست‌آمده توسط اپراتور مورفولوژی گسترش پهن‌تر می‌شوند سپس تصویر بلوک‌بندی و در هر بلوک تبدیل هاف به منظور پیدا کردن خطوط گرفته می‌شود. در فضای هاف نقاطی که تعداد رأی آن‌ها از حد مشخصی بیشتر بود به‌عنوان ویژگی انتخاب می‌شوند. در آخر با الگوریتم بهینه‌سازی انبوه ذرات باینری زیرمجموعه‌ای از ویژگی‌ها برای عمل کلاسه‌بندی انتخاب می‌شوند. این روش روی زیرمجموعه‌ای از هر پایگاه داده آزمایش‌شده که در جدول ۲-۳ به گزارش این نتایج می‌پردازیم.

جدول ۲-۳ نتایج گزارش‌شده در [۹] بر روی پایگاه داده‌های مختلف

نام پایگاه داده	تغییرات	تعداد کلاس‌ها	نمونه‌های هر کلاس	تعداد ویژگی‌ها	زمان آموزش (S)	زمان آزمون (S)	نرخ بازشناسی
Cas-Peal	روشنایی	۴۰	۱۰	۱۷۸۳	۱۹۰	۱۰۱	٪۹۳.۷۵
Extended Yale B	روشنایی	۲۸	۱۹	۱۷۹۳	۲۱۹.۹۴	۹۸.۷۴	٪۹۹.۹۵
Cmu Pie	روشنایی، تغییر حالت چهره	۲۰	۱۰	۱۷۵۸	۱۰۰.۴۲	۱۰۲.۱۵	٪۹۹.۵۰
Feret	روشنایی، تغییر حالت چهره و تغییر اندازه	۳۵	۲۰	۲۰۴۲	۲۴۷.۶۰	۸۵.۰۹	٪۸۰.۸۱

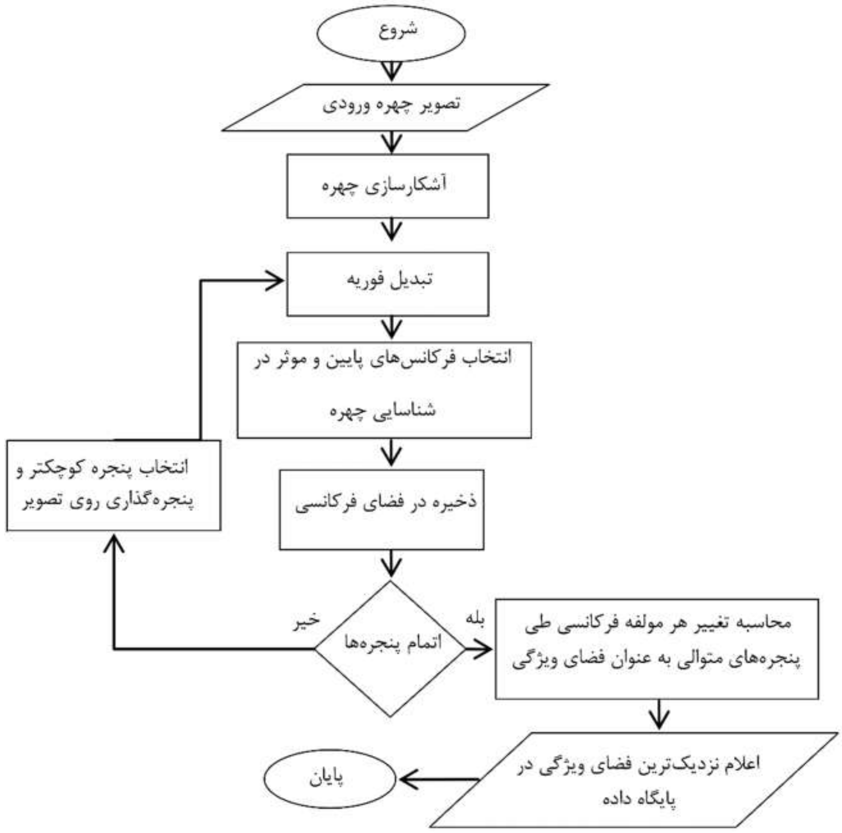
۲-۳-۴استفاده از ویژگی‌های مکان فرکانسی

در [۱۲] با استفاده از ویژگی‌های مکان فرکانسی و با اعمال پنجره‌های هم‌مرکز با ابعاد مختلف روی تصویر چهره و استخراج ویژگی‌های مخصوصی از فضای فرکانسی برای هر پنجره به ساخت بردار ویژگی می‌پردازد، سپس با استفاده از معیار فاصله اقلیدسی کلاسه‌بندی را انجام می‌دهد. در شکل ۲-۹ فرآیند استخراج ویژگی و نمونه‌ای از ویژگی استخراج‌شده از تصویر ورودی را مشاهده می‌شود.



شکل ۲-۹ فرآیند استخراج ویژگی از تصویر چهره ورودی[۱۲]

در این شکل p تعداد پنجره‌ها، k تعداد ویژگی‌های فرکانسی و x_m تبدیل یافته تصویر در حوزه فرکانس است. در شکل ۲-۱۰ فلوچارت این تحقیق را مشاهده می‌کنیم.



شکل ۲-۱۰ فلوچارت روش معرفی‌شده در [۱۲]

نرخ بازشناسی این روش در پایگاه داده Feret روی زیرمجموعه Fb، ۹۰.۴٪ است.

۲-۴ روش‌های ترکیبی

روش‌های ترکیبی به روش‌هایی گفته می‌شود که هم از ویژگی‌های محلی و هم ویژگی‌های سراسری استفاده می‌کنند، اینکه از چه ویژگی‌هایی استفاده شود یا به چه شکلی ویژگی‌های موردنظر باهم ترکیب شوند تا مزیت‌های هرکدام حفظ گردد و معایب آن‌ها خنثی گردد فاکتور مهمی بر کارایی است. در جدول ۲-۴ میزان حساسیت هرکدام از روش‌ها نسبت به تغییرات متغیرهای مختلف بررسی‌شده است.

جدول ۲-۴ میزان حساسیت ویژگی‌های سراسری و محلی به تغییرات[۲۱]

فاکتور متغیر	ویژگی‌های محلی	ویژگی‌های سراسری
تغییرات کم	غیر حساس	حساس
تغییرات زیاد	حساس	خیلی حساس
روشنایی[۲۲]	خیلی حساس	حساس
حالات چهره[۲۳, ۲۴]	غیر حساس	حساس
موقعیت[۲۵]	حساس	خیلی حساس
نویز[۲۶]	خیلی حساس	حساس
انسداد[۲۳, ۲۴]	غیر حساس	خیلی حساس

برای اینکه یک سیستم ترکیبی کارایی خوبی داشته باشد، اجزای این سیستم (ویژگی‌ها و کلاسه بند) باید هرکدام دقیق و ناهمگون باشند. ازآنجایی‌که ویژگی‌های محلی و سراسری خواص متفاوتی دارند می‌توان امیدوار بود با ترکیب این دو بتوان اطلاعات کامل‌تری برای کلاسه‌بندی به دست آورد. با توجه به جدول ۲-۴ می‌توان دید که این ویژگی‌ها هرکدام در بخشی ضعیف‌تر هستند پس با ترکیب این دو ویژگی می‌توان پیچیدگی کلاسه بند نهایی را کاهش داد و عملکرد را بهبود بخشید.

اگرچه این روش مزیت‌های بالقوه‌ای دارد در این دسته کمتر کاری انجام‌شده که به خاطر مشکلات گفته‌شده است و تحقیقات انجام‌شده در این دسته مثل مدل منعطف ظاهر [۲۷] و تحلیل ویژگی‌های محلی [۲۸] برای مسئله یک تصویر برای هر کلاس مناسب نیستند[۲۱].

۲-۵ سایر روش‌ها

۲-۵-۱ روش‌های مبتنی بر شبکه عصبی

این روش‌ها با یادگیری از تصاویر آموزش و تکنیک‌های یادگیری ماشین سعی دریافتن مرتبط‌ترین خصوصیات در تصاویر چهره دارند. این خصوصیات به شکل توابع جداساز (به‌طور مثال سطوح تصمیم غیرخطی) برای بازشناسی چهره استفاده می‌شوند. در ابتدا تصاویر چهره به فضای

ویژگی کم بعدی نگاشت می‌شوند تا سطوح تصمیم غیرخطی با استفاده از یک شبکه عصبی چندلایه یادگیری شوند[۱۳]. مزیت شبکه‌های عصبی این است که با آموزش می‌توان اطلاعات بیشتری را در رابطه با تغییرات مختلف کسب کرد درنتیجه به عمومیت بیشتری دست پیدا کرد. دو عیب اصلی این روش، یکی این‌که نیاز به تعداد نمونه‌های زیادی برای آموزش خوب نیاز دارد و دیگری اینکه نیازمند این است که پارامترهای این روش به‌خوبی تنظیم شوند تا عملکرد خوبی داشته باشد[۱۴]. در بین روش‌های مختلف شبکه‌های عصبی، از MLP[1] با الگوریتم پس انتشار[2] بیشتر استفاده‌شده است اما مشکل این روش همگرایی کند و همچنین اینکه ممکن است کمینه سراسری وجود نداشته باشد از طرفی شبکه‌های مبتنی بر RBF[3] قابلیت یادگیری سریع و خاصیت بهترین تقریب را دارند، درنتیجه در روش‌های مبنی بر شبکه عصبی بیشتر از این روش استفاده‌شده است.

در سال‌های گذشته با استفاده از یادگیری عمیق به نتایج خوبی دست پیداکرده‌اند*[29-32]* .

روش *FaceNet* که در [۱۵] معرفی شد یکی از بهترین روش‌های موجود برای بازشناسی چهره به‌وسیله یادگیری عمیق است. در این روش به‌صورت مستقیم نگاشتی از تصاویر به فضای اقلیدسی فشرده به کمک یادگیری عمیق است. شبکه به‌گونه‌ای آموزش می‌یابد که فاصله نرم *L2* در فضای جدید به‌صورت مستقیم متناظر با مشابهت چهره باشد، یعنی چهره‌های یک فرد فاصله‌ای کم و افراد متفاوت فاصله بیشتری داشته باشند. فرآیند آموزش به‌گونه‌ای انجام می‌شود که خروجی شبکه یک بردار فشرده ۱۲۸ بعدی باشد که با یادگیری به کمک یک تابع خطای ۳ گانه به دست می‌آید. در شکل ۲-۱۱ ساختار مدل را می‌بینیم.

شکل ۲-۱۱ ساختار مدل در [۱۵]

---

[1] Multi-layer Perceptron
[2] Back Propagation
[3] Radial Basis Function

در شکل ۲-۱۱ یک‌لایه ورودی دارد این ورودی‌ها به معماری یادگیری عمیق داده می‌شود، برای معماری یکی از دو معماری معرفی‌شده در [۳۳] یا [۳۴] استفاده می‌شود، خروجی یادگیری عمیق به یک‌لایه نرمال‌سازی L2 داده می‌شود که فضای اقلیدسی موردنظر ماست برای محاسبه خطا از تابع ۳ گانه خطا استفاده می‌شود. بعد از اتمام آموزش و به دست آوردن فضای نهایی با کلاسه بند K نزدیک‌ترین همسایه به عمل بازشناسی می‌پردازد.

این روش روی پایگاه داده LFW به‌دقت ٪۹۹٫۶۳ رسیده است.

روش‌های مبتنی بر یادگیری عمیق دو مشکل عمده همانند روش‌های مبتنی بر شبکه عصبی دارند، نیازمند داده‌های آموزشی زیادی هستند و به دلیل بالا بودن تعداد پارامترها مدت‌زمان زیادی برای همگرایی نیاز دارند.

۲-۵-۲ استفاده از نمایش تنک

هدف استفاده از نمایش تنک این است که با انتخاب زیرمجموعه‌ای از ویژگی‌ها یا مدل از داده‌های آموزش علاوه برافزایش کارایی، کمترین مقدار طول توصیف گر را داشته باشیم. در این روش‌ها نسبت به تغییرات روشنایی و بودن انسداد و مانع کارایی خوبی دارند[۱۷, ۱۸]. در [۱۷] با استفاده از نمایش تنک به حل مسئله بازشناسی چهره پرداخته است، ابتدا n نمونه آموزش کلاس i ام را به شکل یک ستون از ماتریس $A_i = \{v_{i,1}, v_{i,2}, \cdots, v_{i,n}\} \in \mathbb{R}^{m*n}$ درمی‌آوریم. هر تصویر با ابعاد m=w*h به شکل یک بردار m بعدی تبدیل می‌شود، با فرض اینکه داده‌های آموزشی برای هر کلاس در یک زیر فضای خطی قرار دارند می‌خواهیم ترکیب خطی از داده‌های آموزشی پیدا کنیم که تصویر آزمون y را تخمین بزنیم.

$$y = \alpha_{i,1} v_{i,1} + \alpha_{i,2} v_{i,2} + \cdots + \alpha_{i,\mathrm{n}} v_{i,\mathrm{n}} \qquad (۲-۱۹)$$

$$\alpha_{i,\mathrm{n}} \in \mathbb{R}, \mathrm{j}=1,2,\cdots \mathrm{n}$$

شبه کد الگوریتم پیشنهادی به‌صورت زیر است.

Algorithm 1. Sparse Representation-based Classification (SRC)

1: **Input:** a matrix of training samples $A = [A_1, A_2, \ldots, A_k] \in \mathbb{R}^{m \times n}$ for k classes, a test sample $\boldsymbol{y} \in \mathbb{R}^m$, (and an optional error tolerance $\varepsilon > 0$.)

2: Normalize the columns of A to have unit ℓ^2-norm.

3: Solve the ℓ^1-minimization problem:

$$\hat{\boldsymbol{x}}_1 = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad A\boldsymbol{x} = \boldsymbol{y}.$$

(Or alternatively, solve

$$\hat{\boldsymbol{x}}_1 = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|A\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon.)$$

4: Compute the residuals $r_i(\boldsymbol{y}) = \|\boldsymbol{y} - A\,\delta_i(\hat{\boldsymbol{x}}_1)\|_2$ for $i = 1, \ldots, k$.

5: **Output:** identity$(\boldsymbol{y}) = \arg\min_i r_i(\boldsymbol{y})$.

روش پیشنهادی در این تحقیق بر روی پایگاه داده Yale Extended B و زیرمجموعه از پایگاه داده AR آزمایش‌شده، این زیرمجموعه شامل ۱۰۰ نفر و ۱۴ تصویر برای هر فرد است که ۷ تصویر برای مجموعه آموزش و ۷ تصویر برای تصاویر آزمون انتخاب‌شده‌اند. نتایج این روش روی این پایگاه داده AR، ۹۴.۷٪ و روی پایگاه داده Yale Extended B، ۹۸.۱٪ گزارش‌شده است. همچنین این روش نسبت به انسداد و خرابی آزمایش‌شده که در جدول ۲-۵ نتایج را مشاهده می‌کنیم:

جدول ۲-۵ نرخ بازشناسی الگوریتم پیشنهادی [۱۶] در برابر انسداد و خرابی تصویر

درصد انسداد	%۰	%۱۰	%۲۰	%۳۰	%۴۰	%۵۰
نرخ بازیابی	٪۱۰۰	٪۱۰۰	۹۹،۸٪	۹۸،۵٪	۹۰،۳٪	۶۵،۳٪

در [۱۶] بجای ساختن ماتریس A از کل تصاویر آموزش، برای هر کلاس یک ماتریس ساخته می‌شود، سپس نمونه آزمون را به‌صورت ترکیب خطی از هرکدام از این ماتریس‌ها بیان می‌کنیم.

$$Y = DX \qquad (۲-۲۰)$$

در این معادله Y نشان‌دهنده نمونه‌ی آزمون، $X^T = \{x_1, x_2, \cdots, x_n\}$ نشان‌دهنده ضرایب نمونه‌های آموزش و $D = \{d_1, d_2, \cdots, d_n\}$، هرکدام از ستون‌های ماتریس D برابر با یکی از تصاویر نمونه‌های آموزش است. روش پیشنهادی در این تحقیق به‌صورت زیر است.

۱. خوشه‌بندی داده‌های آموزش توسط روش K-means

۲. حل به دست آوردن ضرایب X با حل معادله زیر

$$X = (D^T X + \mu I)^{-1} D^T y \qquad (۲-۲۱)$$

۳. به دست آوردن باقیمانده هر کلاس با حل معادله زیر و تعیین M کلاس برنده (کلاس‌هایی که کمترین مقدار باقیمانده را داشته باشند.)

$$Residual = y - \sum x_i d_i \qquad (۲-۲۲)$$

برای هرکدام از ماتریس‌های برنده یک ماتریس واژگان تشکیل می‌دهد، سپس نمونه آزمون را به‌صورت ترکیب خطی از این ماتریس‌های واژه‌نامه بیان می‌کند.

$$y \approx a_1 t_1 + \cdots + a_n t_n \qquad (۲-۲۳)$$

$$\begin{gathered} y \approx GA \\ A^T = \{a_1. \dots. a_n\} \ . G = \{t_1. t_2. \dots. t_n\} \end{gathered} \qquad (۲-۲۴)$$

۴. تشکیل ترکیب خطی از نمونه‌های آموزش هر کلاس با استفاده از معادله زیر

$$\begin{gathered} y = G_1 A_1 \\ y = G_2 A_2 \\ \dots \\ y = G_M A_M \end{gathered} \qquad (۲-۲۵)$$

۵. به دست آوردن A ها از طریق معادله زیر انجام می‌شود

$$\begin{gathered} A_1 = (G_1^T G_1 + \gamma I)^{-1} + G_1^T y \\ A_2 = (G_2^T G_2 + \gamma I)^{-1} + G_2^T y \\ \dots \\ A_m = (G_m^T G_m + \gamma I)^{-1} + G_m^T y \end{gathered} \qquad (۲-۲۶)$$

۶. به دست آوردن باقیمانده هر کلاس و تعیین کلاس برنده

$$\begin{gathered} Residual_1 = y - \sum t_{1.i} d_{1.i} \\ Residual_2 = y - \sum t_{2.i} d_{2.i} \\ \dots \\ Residual_m = y - \sum t_{m.i} d_{m.i} \end{gathered} \qquad (۲-۲۷)$$

این روش روی پایگاه داده AR موفق به بازشناسی با نرخ ۹۹.۲٪ و روی پایگاه داده ORL به ۹۸.۱٪ رسیده است.

۲-۶ پایگاه داده‌های چهره

بطورخلاصه در جدول ۲-۶ چند پایگاه داده برای چهره معرفی‌شده‌اند.

جدول ۲-۶ معرفی پایگاه داده‌های چهره

نام پایگاه داده	تعداد نفرات	تعداد تصاویر به ازای هر نفر	سایز تصاویر	تغییرات	توضیحات
AT&T(ORL)[35]	۴۰	۱۰	۹۲*۱۱۲	تغییرات روشنایی، حالت چهره، باز و بسته بودن چشم‌ها، داشتن یا نداشتن عینک	پس‌زمینه تاریک و یکنواخت و تصاویر از روبرو
Yale A[3]	۱۵	۱۱	۳۲۰*۲۴۳	تغییرات روشنایی، حالت چهره، داشتن یا نداشتن عینک	
Extended Yale B[36]	۱۰	۵۷۶	۶۴۰*۴۸۰	۶۴ حالت شرایط روشنایی و ۹ موقعیت	
AR [37]	۱۲۶	۲۶	۷۶۸*۵۷۶	حالت چهره، داشتن یا نداشتن عینک و شال	تصویربرداری از روبرو و در دو جلسه
Feret[38]	۹۹۰	متغیر	۷۶۸*۵۱۲	در ۵ موقعیت، تغییرات روشنایی، تغییرات حالت چهره، تغییرات سن، داشتن یا نداشتن عینک و موی صورت	دارای معیار ارزیابی استاندارد
Labeled Faces in the wild[39]	۵۷۴۹	متغیر	۱۵۰*۱۵۰	کنترل نشده	
CAS-PEAL[40]	۱۰۴۰	بین ۹۰ تا ۱۰۰	۴۸۰*۳۶۰	تغییرات روشنایی، حالت چهره، موقعیت و داشتن و نداشتن عینک	

۲-۷ جمع‌بندی

در این فصل به معرفی روش‌های موجود در بازشناسی چهره پرداختیم، اکثر روش‌های موجود بر روی پایگاه داده‌هایی با تعداد افراد کمتر از ۱۵۰ نفر آزمایش‌شده‌اند، همچنین تعداد نمونه‌های

آموزشی اکثر روش‌ها بیشتر از یک نمونه است. با توجه به نتایج گزارش‌شده به نظر می‌رسد روش‌های مبنی بر ویژگی‌های محلی انتخاب بهتری برای بازشناسی چهره باشند. در فصل بعد به معرفی مباحث نظری مورداستفاده ازجمله الگوریتم مورداستفاده برای آشکارسازی چهره، روش‌های استخراج ویژگی محلی، کلاسه بند مورداستفاده می‌پردازیم.

فصل ۳: مباحث نظری

۳-۱ مقدمه

با توجه به پیکربندی معرفی‌شده در فصل اول هر سیستم خودکار بازشناسی چهره از سه بخش اصلی تشکیل می‌شود، در این فصل مباحث نظری مورداستفاده در این پایان‌نامه توضیح داده می‌شود.

۳-۲ آشکارسازی چهره

روش‌های مختلف و متنوعی برای آشکارسازی چهره پیشنهادشده است، [۱۳, ۲۰, ۴۱, ۴۲] از بهترین الگوریتم‌ها در این زمینه هستند. در این پایان‌نامه برای آشکارسازی چهره از روش پیشنهادشده در [۴۱]، معروف به آشکارساز چهره Viola-Jones استفاده کرده‌ایم. در [۴۳] به شکل مفصلی به توضیح این روش پرداخته‌شده که در این بخش به‌صورت خلاصه این روش را مرور می‌کنیم. این روش شامل سه بخش مهم است که عبارت‌اند از:

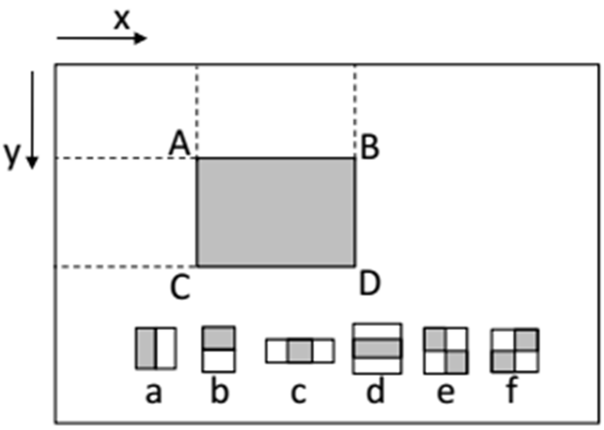
۱. **انتگرال تصویر:**

به‌منظور محاسبه سریع ویژگی‌ها در مقیاس‌های متفاوت، تصویر تجمعی از روی تصویر اصلی تولید می‌شود که تنها یک‌بار نیاز به محاسبه دارد سپس ویژگی‌های مثل هار استخراج می‌شوند، ساخت انتگرال تصویر به شکل زیر است:

$$\mathrm{ii(x,y)} = \sum_{x' \leq x, y' \leq y} i(x', y') \qquad (۳\text{-}۱)$$

با استفاده از انتگرال تصویر محاسبه ویژگی‌ها با جمع چند خانه از تصویر انتگرال قابل‌محاسبه است. در شکل ۳-۱ ویژگی‌های هار[1] و تصویر انتگرال، مشاهده می‌شود.

[1] Haar



شکل ۳-۱ انتگرال تصویر و ویژگی‌های مثل هار
مورداستفاده

محاسبه مجموع مقادیر مستطیل ABCD به کمک انتگرال تصویر به شکل زیر است:

$$\sum_{(x,y)\in ABCD} \mathrm{i}(\mathrm{x},\mathrm{y}) = ii(D) + ii(A) - ii(B) - ii(C) \qquad (۲\text{-}۳)$$

محاسبه ویژگی‌های مثل هار به‌صورت تفاوت روشنایی بین ۲ یا ۴ مستطیل انجام می‌شود که به‌طور مثال برای محاسبه a به ۶ خانه از تصویر انتگرال نیاز دارد.

۲. یادگیری AdaBoost:

در این روش با ترکیب تعداد زیادی کلاسه بند ضعیف به کلاسه‌بندی با دقت بالا می‌رسد. منظور از کلاسه بند ضعیف کلاسه‌بندی است که عملکرد قابل قبولی ندارد. در این تحقیق، هر کلاسه بند ضعیف از یک ویژگی مثل هار در یک مقیاس مشخص استفاده می‌کند. در زیر الگوریتم Adaboost آورده شده است.

Input

- Training example $S = \{(x_i . z_i). i = 1. \dots . N\}$ where $z_i = 0.1$
- T is total number of weak classifiers to be trained

Initialize

- Initialize weights$w_{l.i} = \frac{1}{2\,m} . \frac{1}{2\,l}$for $z_i = 0.1$ where m, l are numbers of positive and negative examples in the training dataset

Adaboost Learning

For t=1,…,T

1. Normalize the weights$w_{l.i} \leftarrow \frac{w_{l.i}}{\sum_{j=1}^{n} w_{l.j}}$,
2. Select the best weak classifier with respect to the weighted error
$$\varepsilon_t = \min_{\mathrm{f.p.\theta}} \sum_i w_i |\, h(x_i . f . p . \theta) - y_i|$$
3. Define$h_t(x) = h(x. f_t . p_t . \theta_t)$ where $f_t . p_t$ *and* θ_t are minimizers of ε_t
4. Update the weights
$$w_{t+1.i} = w_{t.i} \beta_t^{1-e_i}$$

Where$e_i = 0$ if examplex_i is classified correctly and $e_i = 1$ otherwise, and $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$

Output

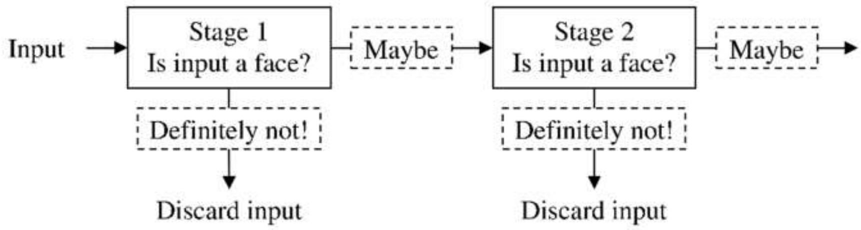
Final Classifier $C(x) = \begin{cases} 1 \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 \quad otherwise \end{cases}$ where $\alpha_t = log \frac{1}{\beta_t}$

۳. کلاسه بند آبشاری:

پایه اصلی آشکارساز چهره Viola-Jones بر اساس پیمایش تصویر به‌دفعات زیاد است، در هر بار پیمایش تصویر با اندازه متفاوتی مورد ارزیابی قرار می‌گیرد. این الگوریتم در اصل به رد کردن زیر پنجره‌هایی که چهره در آن نیست می‌پردازد چراکه رد کردن یک زیر پنجره از یافتن چهره در آن زیر پنجره سریع‌تر است.

کلاسه بند آبشاری ترکیبی از طبقات کلاسه بند است که هر طبقه از یک کلاسه بند قوی تشکیل‌شده است. وظیفه هر طبقه این است که تعیین کند آیا در

زیر پنجره داده‌شده تصویر چهره وجود دارد یا خیر. در شکل ۳-۲ نمایی از کلاسه بند آبشاری نشان داده‌شده است.



شکل ۳-۲ کلاسه بند آبشاری

الگوریتم یادگیری برای ساختن کلاسه بند آبشاری در زیر آمده است:

- User selects values for *f,* the maximum acceptable false positive rate per layer and *d*, the minimum acceptable decision rate per layer
- User selects target overall false positive rate,F_{target}
- *P*=set of positive examples
- *N*=set of negative examples
- $F_0 = 1.0; D_0 = 1.0$
- *i=0*
- while$F_i > F_{target}$
    - $i \leftarrow i + 1$
    - $n_i = 0; F_i = F_{i-1}$
    - While$F_i > f \times F_{i-1}$
        - $n_i \leftarrow n_i + 1$
        - Use *p* and *N* to train a classifier withn_i features using AdaBoost
        - Evaluate current cascaded classifier on validation set to determineF_i andD_i
        - Decrease threshold for the ith classifier until the current cascaded classifier has a detection rate of at least $d \times D_{i-1}$ (this also affects F_i)
    - $N \leftarrow \emptyset$
    - If$F_i > F_{target}$ then evalute the current cascaded detector on the set of non-faces images and put any false detections into the set N

سپس الگوریتم فوق روی تصاویر دارای چهره و بدون چهره آموزش می‌یابد و ساختار آن ذخیره می‌شود. در این پایان‌نامه از ساختار موجود در openCv استفاده‌شده است.

۳-۳ استخراج ویژگی

مرحله دوم در بازشناسی چهره، استخراج ویژگی‌های موردنیاز برای کلاسه‌بندی است. در صورت نیاز به پیش‌پردازش برای بهبود تصویر ورودی برای ویژگی موردنظر عمل پیش‌پردازش نیز صورت می‌گیرد. ورودی این مرحله تصویر برش شده چهره و خروجی این مرحله بردار ویژگی است.

۳-۳-۱ پیش‌پردازش

در صورت نیاز برای بهبود کیفیت تصاویر قبل از استخراج ویژگی از یکی از الگوریتم‌های زیر استفاده می‌شود. برای بهبود روشنایی از یکی از روش‌های یکسان‌سازی هیستوگرام، جبران‌سازی گاما، یکسان‌سازی هیستوگرام وفقی با محدودیت کنتراست (Clahe[1]) استفاده‌شده است. معیار ارزیابی بهبود کیفیت، تأثیر این پیش‌پردازش روی عملکرد کلاسه بند است که در فصل چهار گزارش‌شده‌اند.

- **یکسان‌سازی هیستوگرام**

یکسان‌سازی هیستوگرام روشی برای تنظیم روشنایی تصویر با استفاده از هیستوگرام به‌منظور افزایش تباین[2] تصویر است[۴۴]. اگر تصویر ورودی f دارای r سطر و c ستون باشد و بازه شدت روشنایی آن بین ۰ تا L-1 باشد، p هیستوگرام نرمال شده را به شکل زیر محاسبه می‌کنیم.

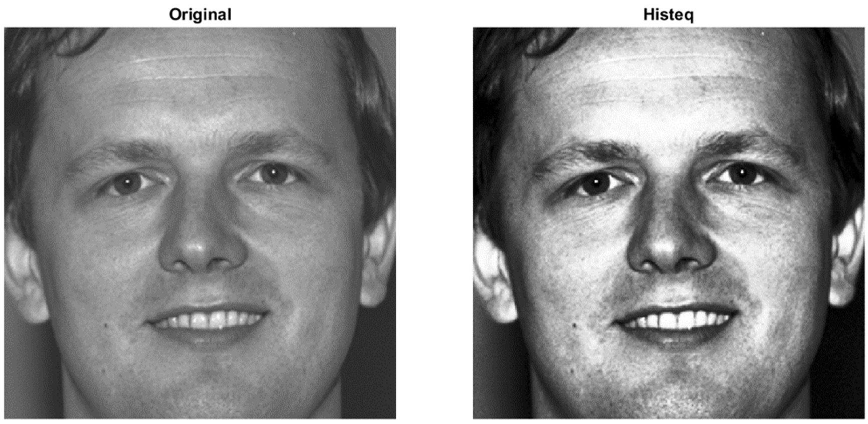
$$P_n = \frac{number\ of\ pixels\ with\ intensity\ n}{total\ number\ of\ pixels} \qquad (۳-۳)$$
$$n = 0,1,\ldots,L-1$$

با استفاده از هیستوگرام نرمال شده، تصویر خروجی G به شکل زیر به دست می‌آید.

1 Contrast Limited Adaptive Histogram Equlazation
2 Contrast

$$G(i,j) = floor((L-1)\sum_{n=0}^{f(i,j)} p_n) \qquad (۴-۳)$$

در شکل ۳-۳ نمونه‌ای از انجام این پیش‌پردازش مشاهده می‌شود.



شکل ۳-۳ تصویر اصلی و تصویر با پیش‌پردازش یکسان‌سازی هیستوگرام

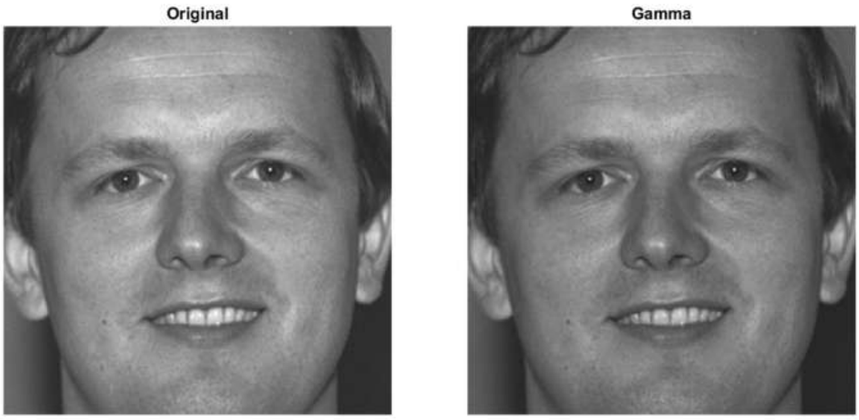
- **جبران‌سازی گاما**

جبران‌سازی گاما یکی دیگر از تبدیل‌هایی است که به‌منظور تنظیم روشنایی و افزایش تباین استفاده می‌شود. می‌توان از یکی از دو تبدیل پیشنهادشده زیر برای این منظور استفاده نمود. که در رابطه ۳-۵ ، max بیشترین مقدار شدت روشنایی و min کمترین مقدار روشنایی و در رابطه ۳-۶ کران بالا و پایین برای تبدیل است.

$$G(i,j) = f(i,j) * ((max - min)/255) \qquad (۵-۳)$$

$$G(i,j) = \begin{cases} 0 \; if \; f(i,j) \leq min \\ 255 \; if \; f(i,j) > max \\ \left[\dfrac{f(i,j) - min}{max - min}\right]^{\gamma} \quad else \end{cases} \qquad (۶-۳)$$

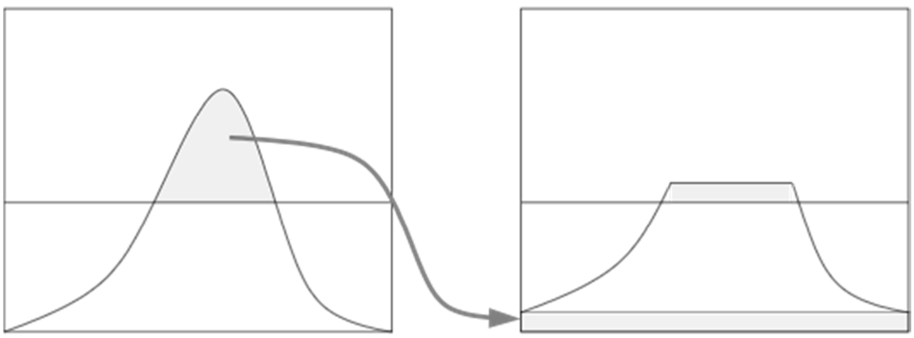
در شکل ۳-۴ نمونه‌ای از انجام این پیش‌پردازش مشاهده می‌شود. اثر این پیش پردازش بیشتر روی مقدار روشنایی پیکسل‌ها می‌باشد و تصویر خروجی تاثیر قابل ملاحظه‌ای از نظر بینایی نمی‌کند.



شکل ۳-۴ تصویر اصلی و تصویر با پیش‌پردازش یکسان‌سازی گاما

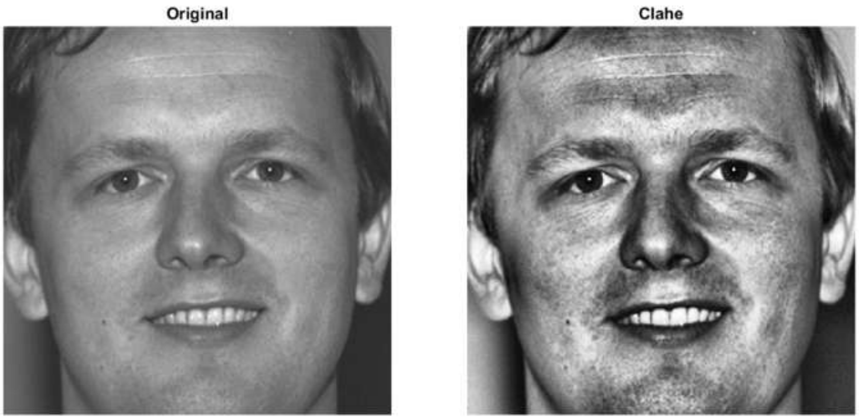
- **یکسان‌سازی هیستوگرام وفقی با محدودیت کنتراست**

تفاوت این روش با یکسان‌سازی هیستوگرام این است که برای بهبود سطوح روشنایی هر پیکسل بجای در نظر گرفتن تمامی پیکسل‌ها، همسایگی از پیکسل را مدنظر قرار می‌دهد، همچنین با گذاشتن محدودیت روی مقدار هر پیکسل و قطع کردن مقدار اضافی و توزیع این مقدار اضافی روی کل تصویر به دست می‌آید[۴۵]. در شکل ۳-۵ مفهوم قطع کردن مقدار اضافی نشان داده‌شده است.



شکل ۳-۵ قطع مقدار هیستوگرام در CLAHE

در شکل ۳-۶ نمونه‌ای از انجام این پیش‌پردازش مشاهده می‌شود.



شکل ۳-۶ تصویر اصلی و تصویر با پیش‌پردازش Clahe

۳-۳-۲ تبدیل فوریه زمان کوتاه

تبدیل فوریه زمان کوتاه (STFT[1]) نسخه اصلاح‌شده تبدیل فوریه است[۴۶]، در تبدیل فوریه سیگنال ورودی به‌صورت مجموعی از سینوس‌ها و کسینوس‌ها با فرکانس‌ها و اندازه‌های متفاوت نشان داده می‌شود درنتیجه تبدیل فوریه نمایش اندازه فرکانس سیگنال ورودی است. تبدیل فوریه ابزار مناسبی برای تحلیل سیگنال‌های غیر ایستا نیست اما دو ابزار موجک و STFT مناسب هستند. به‌طور مثال سیگنال‌های صدا، EEG و تصاویر چهره سیگنال‌های غیر ایستا هستند. در این روش سیگنال ورودی به زیر سیگنال‌هایی که فرض می‌شود که ایستا هستند تقسیم می‌شود. این کار با استفاده از یک پنجره مشخص با عرض دلخواه و ضرب کردن آن با سیگنال اصلی انجام می‌شود. روی این زیر سیگنال‌ها تبدیل فوریه گرفته می‌شود، این تبدیل سیگنال ورودی را به فضای زمان فرکانس نگاشت می‌کند.

برای سیگنال یک‌بعدی $x(t)$ تبدیل STFT به‌صورت زیر تعریف می‌گردد.

$$X(\tau,\omega)=\int_{-\infty}^{+\infty}x(t)*w(t-\tau)e^{-j\omega t}dt \qquad (۷-۳)$$

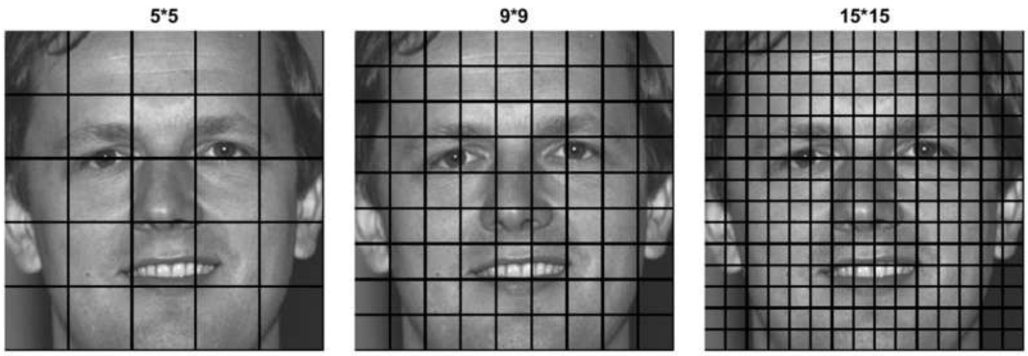
1 Short time fourier transform

در حالتی که سیگنال ورودی دوبعدی باشد، فضای خروجی بجای زمان فرکانس، مکان فرکانس است و به‌صورت زیر تعریف می‌گردد.

$$X(\tau_1,\tau_2,\omega_1,\omega_2)=\iint_{-\infty}^{+\infty} f(x,y)*w(x-\tau_1,y-\tau_2)e^{-j(\omega_1 x+\omega_2 y)}dxdy \quad (۸-۳)$$

روش استخراج ویژگی STFT روی تصویر به‌صورت زیر است.

۱. تصویر ورودی به m*m ناحیه تقسیم می‌شود در شکل ۳-۷ به ازای m=5,9,15 نشان داده‌شده است.



شکل ۳-۷ نحوه بلوک‌بندی برای استخراج ویژگی

۲. برای هر ناحیه تبدیل فوریه گرفته می‌شود به‌گونه‌ای که مؤلفه‌های کم فرکانس‌تر در مرکز تصویر باشند سپس اندازه تبدیل فوریه محاسبه می‌شود.

۳. k المان مرکزی هر ناحیه را به‌عنوان بردار ویژگی آن ناحیه انتخاب می‌کنیم با گذاشتن پنجره‌ای مربعی به مرکز پیکسل مرکزی و اندازه$\sqrt{k}*\sqrt{k}$ این کار انجام می‌شود.

۴. با چسباندن بردارهای ویژگی همه نواحی بردار ویژگی نهایی به دست می‌آید.

در این ویژگی بردار ویژگی به دو پارامتر تعداد المان به ازای هر ناحیه و تعداد نواحی وابسته است.

۳-۳-۳هیستوگرام گرادیان جهت‌دار

هیستوگرام گرادیان جهت‌دار (HOG[1]) در [۴۷] برای شناسایی انسان معرفی شد، این ویژگی مبتنی بر اطلاعات گرادیان است، روش استخراج ویژگی مورداستفاده در این پایان‌نامه کمی با روش استخراج معرفی‌شده در [۴۷] متفاوت است که در زیر آمده است.

۱. از تصویر ورودی گرادیان مرکزی افقی و عمودی گرفته می‌شود.
۲. تصویر ورودی به m*m سلول مطابق شکل ۳-۷ تقسیم‌بندی می‌شود.
۳. برای هر پیکسل در هر سلول مقدار گرادیان و زاویه گرادیان محاسبه می‌شود.
۴. زاویه ۰ تا ۳۶۰ درجه به k دسته تقسیم می‌شود. بر اساس زاویه گرادیان دسته موردنظر انتخاب‌شده و به مقدار آن دسته به‌اندازه مقدار گرادیان افزوده می‌شود.
۵. با انجام عمل ۴ برای همه پیکسل‌های سلول، هیستوگرام جهت‌دار برای سلول ساخته می‌شود. سپس این بردار به‌گونه‌ای نرمال می‌شود که نرم آن یک باشد.
۶. با چسباندن بردارهای سلول‌های مختلف به هم بردار ویژگی ساخته می‌شود.

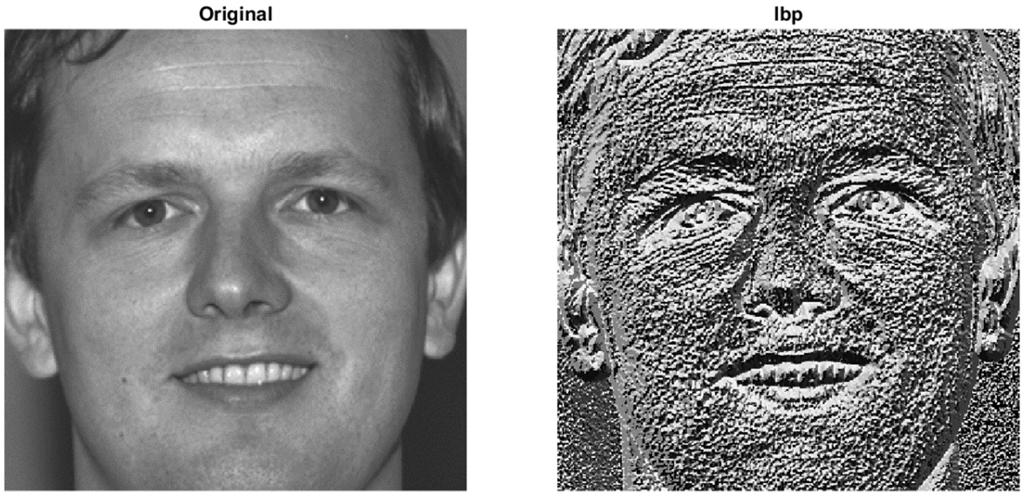
در این ویژگی بردار ویژگی به دو پارامتر تعداد دسته‌ها در هیستوگرام گرادیان جهت‌دار و تعداد سلول‌ها وابسته است.

۳-۳-۴الگوهای باینری محلی

استخراج الگوهای باینری محلی (LBP[2]) مشابه با استخراج انجام‌شده در [۸] است. برای استخراج ویژگی LBP طبق مراحل زیر عمل می‌شود. در [۱۹] نشان داده شد این ویژگی در برابر تغییرات روشنایی و چرخش مقاوم است.

۱. در ابتدا روی تصویر ورودی عملیات LBP انجام می‌شود و خروجی LBP به دست می‌آید در شکل ۳-۸ تصویر چهره‌ای قبل و بعد از اعمال این اپراتور مشاهده می‌شود.

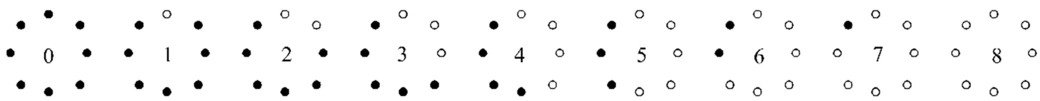
1 Histogram of Oriented Gradiants
2 Local Binary Patterns



شکل ۳-۸ قبل و بعد اعمال اپراتور LBP

۲. تصویر LBP خروجی را به m*m ناحیه مطابق با شکل ۳-۷ تقسیم‌بندی می‌کنیم.

۳. در هر ناحیه هیستوگرام الگوهای یک ریخت را انتخاب می‌کنیم. در [۱۹] نشان داده شد که بین ۷۰ تا ۹۰ درصد از کل تصویر LBP از الگوهای یکنواخت تشکیل‌شده‌اند. در شکل ۳-۹ برای اپراتور LBP با تعداد ۸ خانه در همسایگی با شعاع ۱ این الگوها مشخص‌شده‌اند. در پیاده‌سازی از یک جدول از قبل آماده‌شده با توجه به اپراتور LBP استفاده می‌شود.



شکل ۳-۹ الگوهای یکنواخت در یک اپراتور دایره‌ای با شعاع ۱ و ۸ همسایگی [۱۹]

این هیستوگرام در هر ناحیه ۱۰ خانه دارد که ۹ خانه برای الگوهای یکنواخت و یک خانه برای سایر الگوها است.

۴. با چسباندن هیستوگرام‌های همه نواحی بردار ویژگی نهایی ساخته می‌شود.

۳-۳-۵نرمال‌سازی

در هرکدام از ویژگی‌ها در دو سطح نرمال‌سازی، یکی در سطح ناحیه یکی روی بردار ویژگی کلی انجام می‌شود. روی ناحیه نرمال‌سازی به‌گونه‌ای انجام می‌شود که نرم بردار ناحیه یک باشد و

روی بردار ویژگی کلی به‌گونه‌ای باشد که نرم بردار ۱۰ باشد. این کار به خاطر این است که در مرحله کلاسه‌بندی تأثیر ویژگی‌های استخراج‌شده یکسان باشد.

۳-۴ کلاسه‌بندی

بعد از استخراج ویژگی‌های موردنظر و ترکیب آن‌ها هر نمونه آموزش را با یک بردار ویژگی d بعدی نشان می‌دهیم. با توجه به این‌که کلاسه‌بندی تنها به کمک یک نمونه آموزش استفاده می‌شود، کلاسه بند مورداستفاده نزدیک‌ترین همسایه است. در این روش با داشتن بردارهای ویژگی مجموعه آموزش و نمونه آزمون نزدیک‌ترین داده را در بین مجموعه آموزش پیدا می‌کند[۴۸].

فرض کنیم $D_n = \{x_1, x_2, \dots, x_n\}$، n نمونه در مجموعه آموزش باشند که کلاس آن‌ها مشخص است. هر نمونه با یک بردار ویژگی در فضای d بعدی مشخص می‌شود. اگر نمونه آزمون x باشد که این نمونه هم در فضای d بعدی است برای عمل کلاسه‌بندی کافی است نزدیک‌ترین همسایه را از نمونه‌های آموزش مشخص کنیم و کلاس آن را به نمونه آزمون نسبت دهیم. برای پیدا کردن نزدیک‌ترین همسایه به معیار فاصله مناسبی نیاز داریم که کمترین خطا را داشته باشد.

۳-۴-۱ معیار فاصله

همان‌طور که اشاره شد انتخاب معیار فاصله مناسب در عملکرد کلاسه بند تأثیر زیادی دارد. معیار فاصله خوب باید خاصیت‌های زیر را داشته باشد[۴۸].

- مثبت باشد

$$D(a,b) \geq 0$$

- تنها در صورتی فاصله صفر می‌شود که دو بردار یکسان باشند.

$$D(a,b) = 0 \iff a = b$$

- متقارن باشد.

$$D(a,b) = D(b,a)$$

در این پایان‌نامه از معیارهای فاصله زیر استفاده‌شده است.

۱. فاصله اقلیدسی:

$$D(a,b) = \left(\sum_{i=1}^{d}(a_i - b_i)^2\right)^{\frac{1}{2}} \qquad (۹-۳)$$

۲. آماره کای دو[1]:

$$D(a,b) = \sum_{i=1}^{d}\frac{(a_i - b_i)^2}{a_i + b_i} \qquad (۱۰-۳)$$

۳. فاصله کسینوسی

$$D(a,b) = 1 - \frac{\sum_{i=1}^{d} a_i b_i}{\sum_{i=1}^{d} a_i^2 \sum_{i=1}^{d} b_i^2} \qquad (۱۱-۳)$$

۴. فاصله منهتن:

$$D(a,b) = \sum_{i=1}^{d}|a_i - b_i| \qquad (۱۲-۳)$$

۳-۵ جمع‌بندی

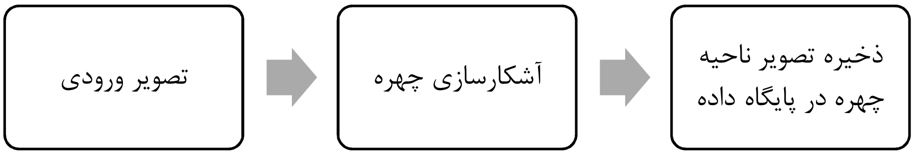
در این فصل مباحث نظری مورداستفاده در این پایان‌نامه، ازجمله آشکارسازی چهره، روش استخراج ویژگی و کلاسه بند مورد استفاده توضیح داده شد. در فصل آینده روش پیشنهادی و نتایج شبیه‌سازی آورده می‌شود.

1 Chi square statistic

فصل ۴: روش پیشنهادی

۴-۱ مقدمه

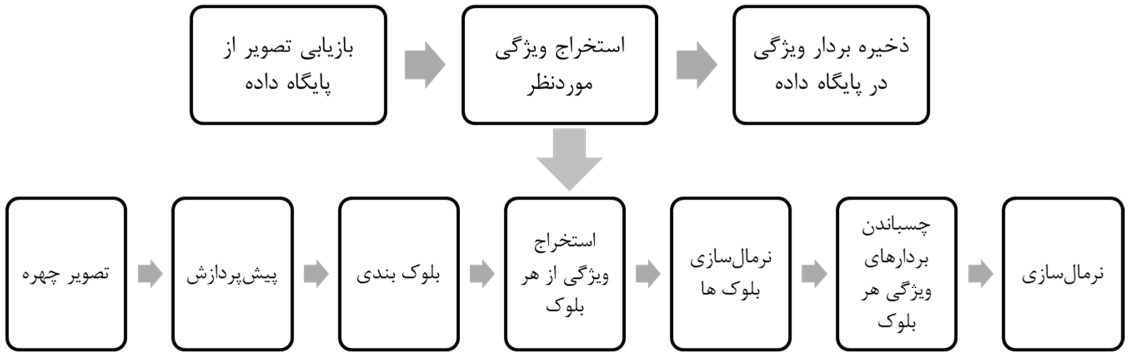
در این فصل روش پیشنهادی معرفی شده و تأثیر پارامترهای مختلف ویژگی‌های استخراج‌شده و عملکرد کلی روش پیشنهادی مورد ارزیابی قرار می‌گیرد. در شکل ۴-۱ فرآیند آشکارسازی چهره در سیستم خودکار بازشناسی چهره مشاهده می‌شود.



شکل ۴-۱ فرآیند آشکارسازی چهره

اولین مرحله در سیستم خودکار بازشناسی چهره این است که چهره از تصویر ورودی جدا شود و اگر چهره‌ای در تصویر نبود مورد ارزیابی بیشتر قرار نگیرد. برای آشکارسازی چهره از آشکارساز چهره Viola-jones [41] که در فصل ۳ موردبررسی قرار گرفت استفاده می‌شود. پس از جداسازی ناحیه چهره از بقیه تصویر، چهره به‌دست‌آمده به ۳۶۰*۳۶۰ تغییر اندازه داده می‌شود. سپس آدرس تصویر در پایگاه داده رابطه‌ای ذخیره می‌شود. علت انتخاب این اندازه تصویر برای امتحان کردن بلوک‌بندی با اندازه‌های متفاوت می‌باشد که می‌توان بلوک‌هایی با اندازه‌های ۲*۲، ۴*۴، ۳*۳، ۵*۵، ۶*۶، ۱۰*۱۰، ۱۲*۱۲، ۲۰*۲۰، ۲۴*۲۴، ۳۶*۳۶، ۶۰*۶۰ را آزمایش کرد.

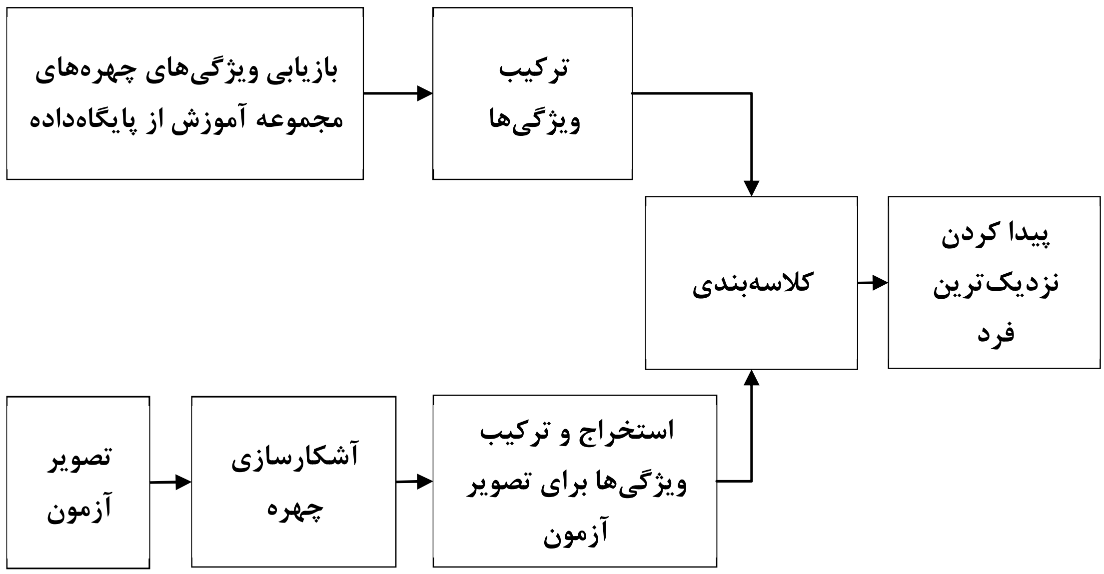
دومین مرحله، استخراج ویژگی‌های موردنظر از تصویر ذخیره‌شده در پایگاه داده است. ۳ ویژگی استخراج‌شده در این روش عبارت‌اند از HOG، LBP و STFT. ویژگی HOG اطلاعات گرادیان را برای هر سلول، ویژگی LBP ساختارهای محلی برای هر ناحیه و ویژگی STFT اطلاعات فرکانسی را در هر ناحیه می‌دهد. روش استخراج ویژگی‌های فوق در فصل ۳ آمده است. در شکل ۴-۲ فرآیند استخراج ویژگی روش پیشنهادی مشاهده می‌شود.



شکل ۴-۲ فرآیند استخراج ویژگی روش پیشنهادی

برای هر ویژگی بالا با تغییر پارامترهای هر ویژگی به دنبال یافتن بهترین عملکرد هستیم. سپس با ترکیب ویژگی‌ها به دنبال یافتن بهترین ترکیب برای کلاسه‌بندی هستیم. ترکیب ویژگی‌های فوق به این شکل است که بردار ویژگی جدید از به هم چسباندن بردارهای ویژگی قبلی به دست می‌آید.

در نهایت با استفاده از یک کلاسه بند نزدیک‌ترین همسایگی، نزدیک‌ترین تصویر را در مجموعه آموزش برای تصویر آزمون پیدا می‌کنیم. در شکل ۴-۳ فرآیند آزمون مشاهده می‌شود.



شکل ۴-۳ فرآیند کلی روش پیشنهادی برای مرحله آزمون

برای بهبود عملکرد الگوریتم پیشنهادی قبل از استخراج ویژگی نمونه‌های مجازی را تولید می‌کنیم. نمونه‌های مجازی تولیدشده شامل تغییرات زیر بر روی تصویر آموزش می‌باشد:

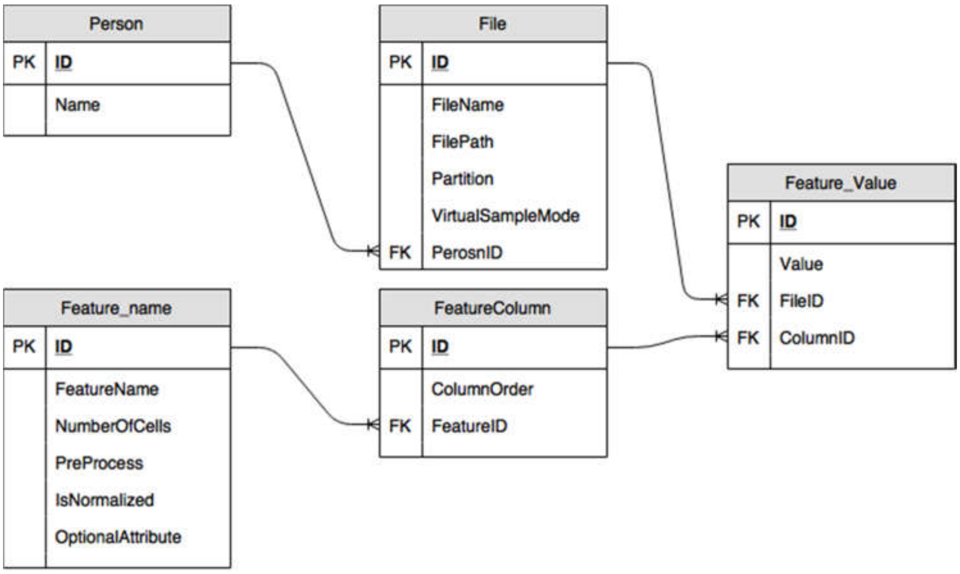
- مات کردن تصویر با کرنل میانگین‌گیر ۳*۳
- تیز کردن لبه‌ها با استفاده از فیلتر لاپلاسین
- چرخش تصویر با زوایای ۳+، ۳-، ۵+ و ۵-

هدف از این تغییرات، نزدیک‌کردن نمونه آموزش به نمونه‌های مورد آزمون می‌باشد. این تغییرات بدون در نظر گرفتن نمونه‌های آزمون انجام شده است تا عمومیت کار حفظ شود.

بعد از تولید این نمونه‌ها مراحل استخراج ویژگی و آزمون انجام می‌شود.

۴-۲ طرح پایگاه داده

برای ذخیره و بازیابی ویژگی‌های استخراج‌شده از پایگاه داده رابطه‌ای استفاده شده است. در شکل ۴-۴ طرح پایگاه داده طراحی‌شده، مشاهده می‌گردد. این پایگاه داده ۵ جدول دارد که با هم مرتبط می‌باشند، جدول File برای ذخیره مشخصات تصویر شامل نام فایل، مسیر فایل ذخیره‌شده، این فایل مرتبط به کدام فرد، کدام حالت از نمونه‌های مجازی و اگر برای پایگاه داده خاصی است جزو کدام زیرگروه از پایگاه داده است، می‌باشد. جدول Person رکوردهای مرتبط با نفرات موجود در پایگاه داده را دارد. جدول Feature_Name مشخصات ویژگی‌های استخراج‌شده شامل نام ویژگی، تعداد سلول‌ها، آیا ویژگی بعد از استخراج نرمال شود، آیا قبل از استخراج ویژگی پیش‌پردازشی انجام شود و خصیصه اختیاری می‌باشد. جدول Feature_Column برای ذخیره ستون‌های ویژگی استخراج‌شده استفاده می‌شود. جدول Feature_Value برای ذخیره مقدار ویژگی استفاده می‌شود که شامل مقدار و چه ستونی از کدام ویژگی و چه فایلی می‌باشد. در همه جداول فوق یک ویژگی شناسه نیز استفاده شده است. طراحی جداول نرمال است. برای اطلاعات بیشتر در رابطه با طراحی پایگاه داده‌های رابطه‌ای به [۴۹] مراجعه شود.



شکل ۴-۴ طرح پایگاه داده استفاده شده برای ذخیره و بازیابی ویژگی‌ها

استفاده از پایگاه داده رابطه‌ای مزیت‌های زیر را دارد:

۱. دسترسی و مدیریت آسان‌تر داده‌ها.
۲. بازیابی و ذخیره سریع رکوردها
۳. کاهش میزان افزونگی[1]
۴. ذخیره دقیق‌تر داده‌ها
۵. عدم نیاز به خواندن تمام فایل برای دستیابی به یک رکورد

۴-۳ پایگاه داده و روش ارزیابی

این پایان‌نامه بر روی پایگاه داده Feret که یکی از پایگاه داده‌های با تعداد افراد بالا است ارزیابی‌شده است، این پایگاه داده از ۹۹۰ نفر تشکیل‌شده است و شامل ۱۳ موقعیت چهره است که در این پایان‌نامه فقط از تصاویر روبرو استفاده‌شده است. تصاویر روبرو شامل تغییرات اندازه، روشنایی، حالت چهره، داشتن یا نداشتن عینک، حالت مو، سن و موی صورت است. این تصاویر در ۴ زیرمجموعه fa,fb,dup1,dup2 توزیع‌شده‌اند.

1 redundancy

- Fa: این تصاویر برای آموزش استفاده‌شده‌اند، به ازای هر نفر یک تصویر وجود دارد.
- Fb: بعد از گرفته شدن تصویر Fa از هر شخص حالت چهره متفاوتی را دارد، تصاویر در این زیرمجموعه فقط تغییرات حالت چهره دارند. به ازای هر نفر یک تصویر در این مجموعه است.
- Dup1: تصاویر در این مجموعه تأثیر افزایش سن، داشتن یا نداشتن عینک، تغییر اندازه، تغییر حالت مو و تغییرات روشنایی را بررسی می‌کند. ۷۳۴ تصویر از ۲۴۹ نفر در این مجموعه است.
- Dup2: این تصاویر زیرمجموعه از Dup1 هستند که تصویر گرفته‌شده بیشتر از ۵۴۰ روز فاصله با تصویر Fa دارند. ۲۲۸ تصویر از ۷۵ نفر در این مجموعه قرار دارد.

در شکل ۴-۵ نمونه‌هایی از این تصاویر مشاهده می‌شوند.

(الف)

(ب)

(ج)

(د)

شکل ۴-۵ نمونه‌هایی از تصاویر موجود در پایگاه داده

(الف) تصاویر Fa، (ب) تصاویر Fb، (ج) تصاویر Dup1، (د) تصاویر Dup2

روش ارزیابی به این شکل است که فرآیند آموزش روی مجموعه Fa انجام می‌شود و روی ۳ مجموعه Fb,Dup1,Dup2 به‌صورت جداگانه آزمایش می‌شود. در گزارش نتایج برای هر مجموعه، بودن تصویر آموزش متناظر تصویر آزمون در بین ۱۵، ۸، ۵، ۳، ۲، ۱ نتیجه با فاصله کمتر گزارش‌شده‌اند.

نرخ بازشناسی p در هر مجموعه به‌صورت زیر محاسبه می‌شود.

$$\mathrm{p}_j = \frac{C_i}{N_j}; i = 1,2,3,5,8,15 \quad j = Fb, Dup1, Dup2 \qquad (۴-۱)$$

که C_i تعداد تصاویر با کلاسه‌بندی صحیح در i نتیجه برتر می‌باشد و N_j تعداد کل تصاویر مجموعه j ام است.

عملکرد میانگین P به‌صورت زیر محاسبه می‌شود.

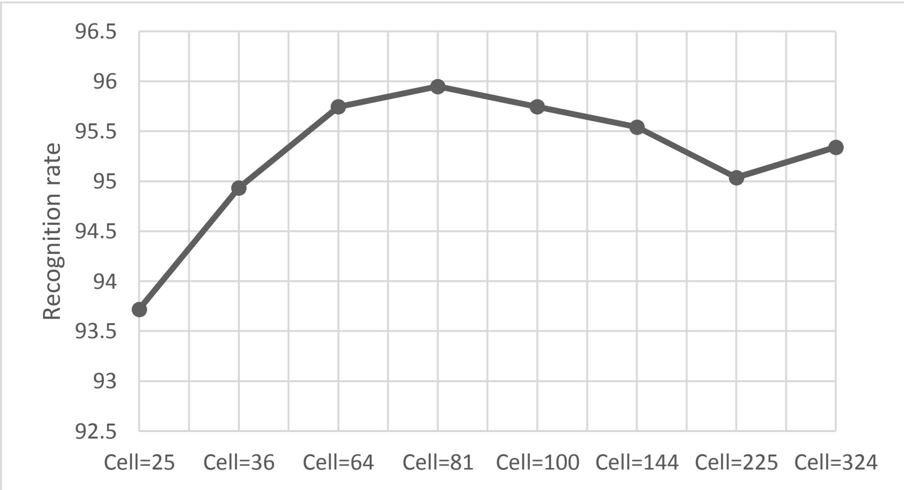
$$P = \frac{p_{Fb} + p_{Dup1} + p_{Dup2}}{3} \qquad (۴-۲)$$

۴-۴ تأثیر تغییر پارامترهای HOG

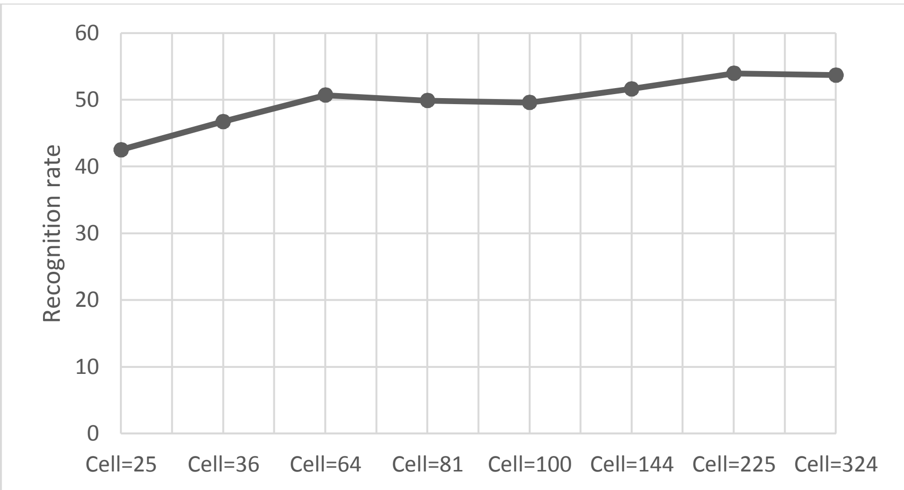
با توجه به الگوریتم استخراج ویژگی HOG که در فصل ۳ معرفی شد، تغییرات دو پارامتر تعداد سلول‌ها و تعداد دسته‌های هیستوگرام را برای یافتن بهترین عملکرد نسبت به این دو پارامتر را موردبررسی قرار می‌دهیم.

۱-۴-۴ تأثیر تعداد سلول‌های HOG

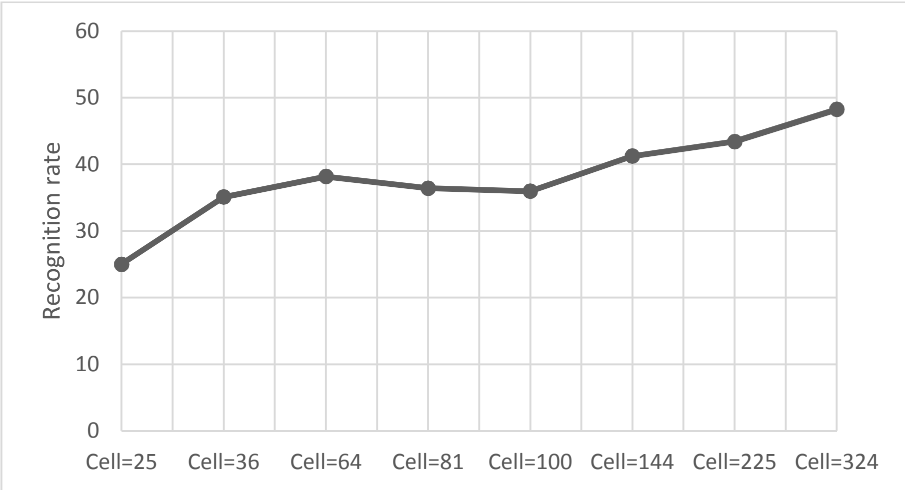
روی هر ۳ مجموعه FB,Dup1,Dup2 تعداد دسته‌ها را ۹ دسته و ثابت فرض می‌کنیم. همچنین قبل از استخراج ویژگی از پیش‌پردازش یکسان‌سازی هیستوگرام و برای کلاسه‌بندی از معیار فاصله منهتن استفاده‌شده است. نتایج گزارش‌شده بر اساس رتبه اول است.



شکل ۴-۶ تأثیر تغییرات تعداد سلول‌های HOG بر مجموعه FB



شکل ۴-۷ تأثیر تغییرات تعداد سلول‌های HOG بر مجموعه Dup1



شکل ۴-۸ تأثیر تغییرات تعداد سلول‌های HOG بر مجموعه Dup2

با توجه به شکل ۴-۶، شکل ۴-۷ و شکل ۴-۸ دو مقدار ۶۴ سلول و ۱۴۴ سلول عملکرد بهتری را دارند. در جدول ۴-۱ نتایج مفصل‌تر این دو مقدار مشاهده می‌شوند. با توجه به شکل‌ها می‌توان دید که با کوچک‌تر شدن سلول‌ها تا مقدار مشخصی سلول اطلاعات جزئی‌تری از گرادیان بدست می‌آید اما بیشتر از این مقدار باعث می‌شود عمومیت اطلاعات بدست آمده کم بشود و ویژگی استخراج‌شده به نمونه آموزشی برازش پیدا کند.

علت اینکه ویژگی HOG با سلول‌های کمتر نتیجه بهتری در مجموعه Dup2 داشته، این است که مجموعه Dup2 شامل تغییرات سن بیشتر از ۵۴۰ روز است و در این مدت چهره تغییرات محلی زیاده داشته در نتیجه در سلول‌های بیشتر با اندازه کوچک‌تر تغییرات بیشتر هستند.

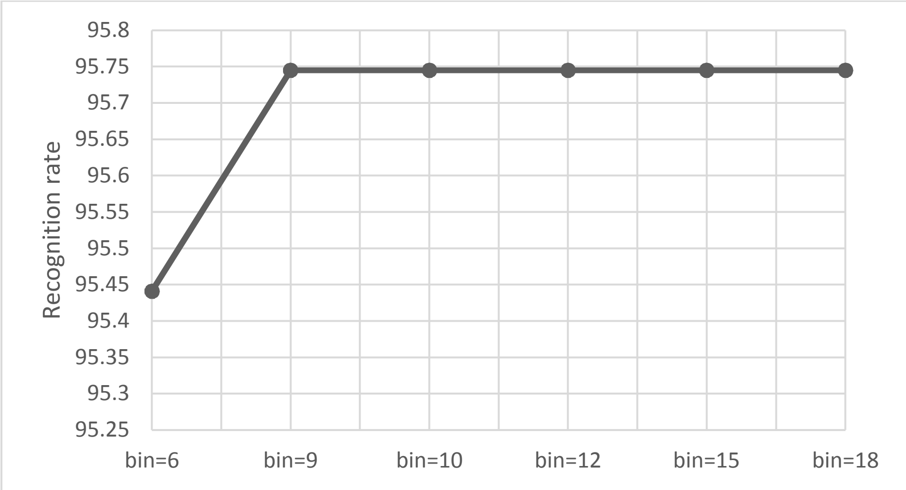
جدول ۴-۱ عملکرد سیستم بازشناسی چهره با استفاده از ویژگی HOG

تعداد سلول		۶۴	۱۴۴
طول بردار ویژگی		۵۷۶	۱۲۹۶
Fb	Rank1	95.74	95.54
	Rank2	96.86	96.86
	Rank3	97.47	97.37
	Rank5	98.07	97.77
	Rank8	98.48	98.38
	Rank15	99.19	98.68
	Mean	97.64	97.43

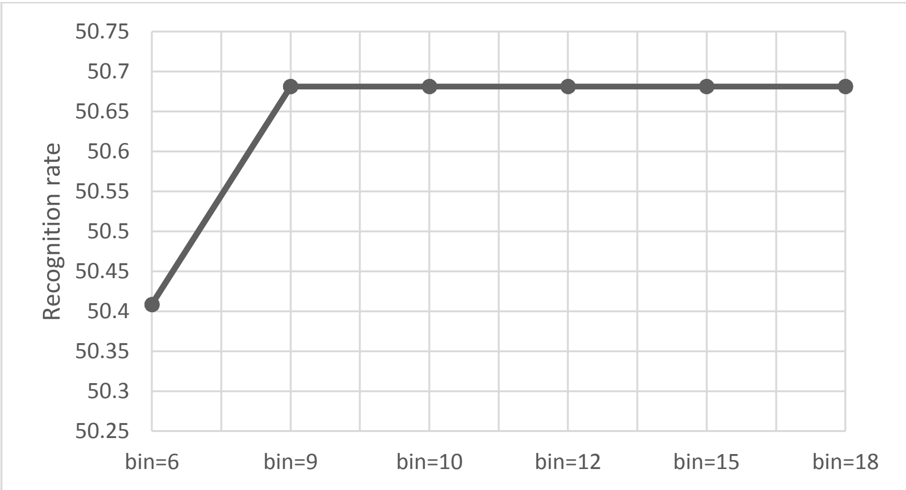
تعداد سلول		۶۴	۱۴۴
Dup1	Rank1	50.68	51.63
	Rank2	56.68	57.63
	Rank3	59.80	61.99
	Rank5	63.76	65.94
	Rank8	67.17	69.75
	Rank15	73.02	74.11
	Mean	61.85	63.51
Dup2	Rank1	41.23	38.16
	Rank2	49.56	46.93
	Rank3	56.58	52.63
	Rank5	63.16	57.46
	Rank8	68.42	60.97
	Rank15	75.44	71.93
	Mean	59.06	54.68

۴-۴-۲ تأثیر تعداد دسته‌های هیستوگرام HOG

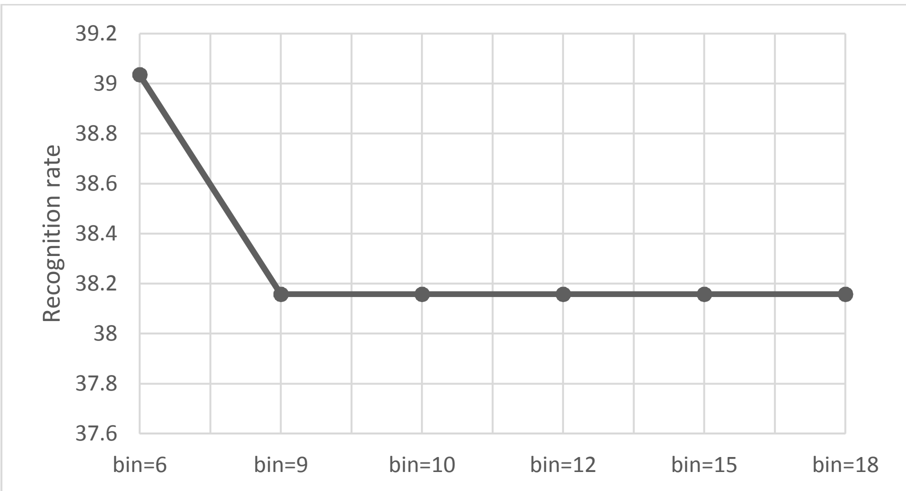
در این آزمایش تعداد سلول‌ها را ثابت و ۶۴ فرض می‌کنیم و مطابق قبل از پیش‌پردازش یکسان‌سازی هیستوگرام و معیار فاصله منهتن استفاده‌شده است.



شکل ۴-۹ تأثیر تغییرات تعداد دسته‌های هیستوگرام HOG بر مجموعه FB



شکل ۴-۱۰ تأثیر تغییرات تعداد دسته‌های هیستوگرامHOG بر مجموعه Dup1

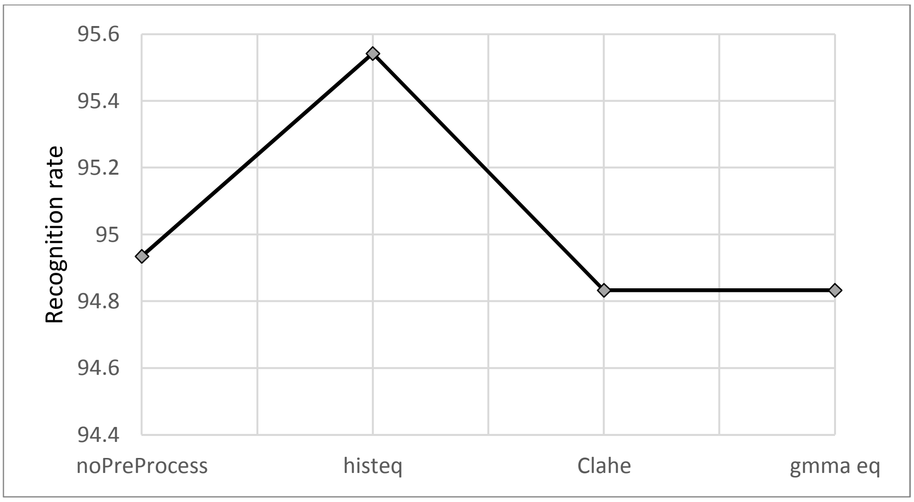


شکل ۴-۱۱ تأثیر تغییرات تعداد دسته‌های هیستوگرام HOG بر مجموعه Dup2

با توجه به شکل ۴-۹، شکل ۴-۱۰ و شکل ۴-۱۱ تعداد دسته‌های هیستوگرام ۹ انتخاب می‌شود. تعداد دسته‌های بالاتر هرچند اطلاعات بیشتری در اختیار ما می‌گذارند اما این اطلاعات بیشتر تأثیری در کلاسه‌بندی ندارند.

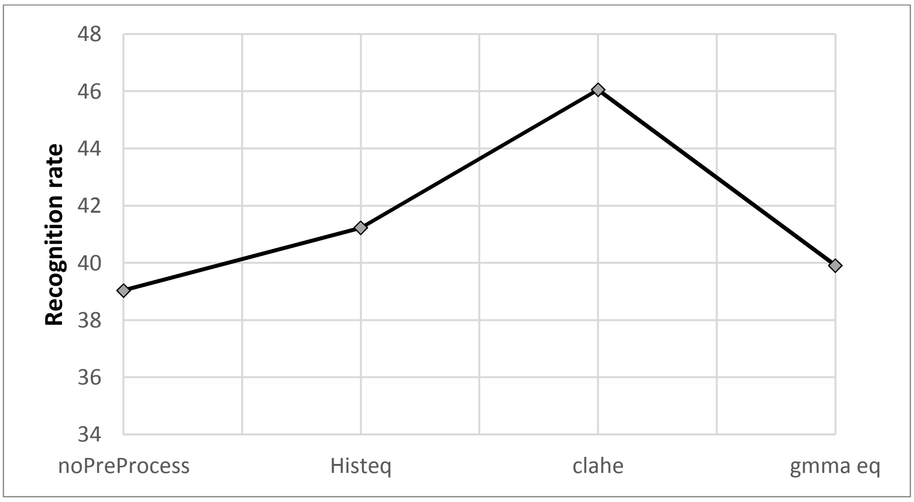
۴-۴-۳ تأثیر پیش‌پردازش بر عملکرد HOG

در این بخش تعداد سلول‌ها را ۱۴۴ و تعداد دسته‌ها را ۹ در نظر می‌گیریم و برنامه را در ۴ حالت بدون پیش‌پردازش، با پیش‌پردازش یکسان‌سازی هیستوگرام، یکسان‌سازی گاما و یکسان‌سازی هیستوگرام وفقی با محدودیت کنتراست اجرا می‌کنیم. شکل‌های زیر نتایج را در ۴ حالت مختلف نشان می‌دهد.



شکل ۴-۱۲ تأثیر پیش‌پردازش‌های مختلف بر عملکرد HOG بر مجموعه Fb

شکل ۴-۱۳ تأثیر پیش‌پردازش‌های مختلف بر عملکرد HOG بر مجموعه Dup1



شکل ۴-۱۴ تأثیر پیش‌پردازش‌های مختلف بر عملکرد HOG بر مجموعه Dup2

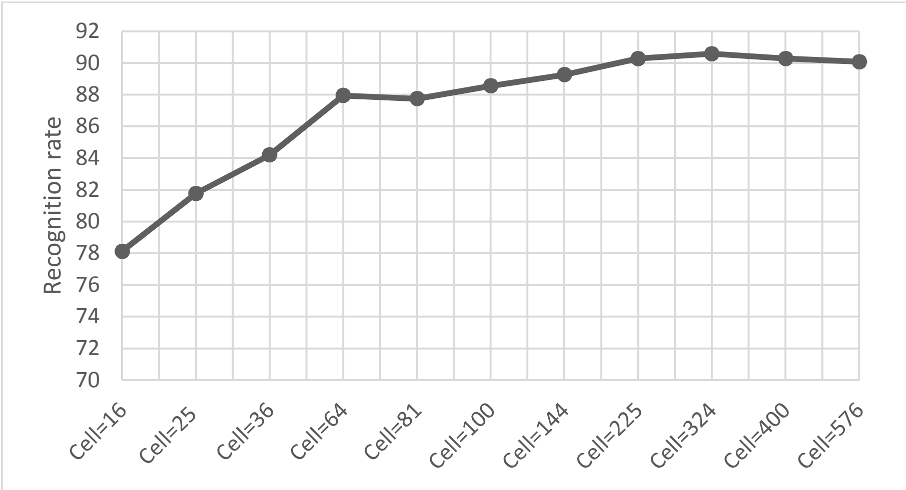
با توجه به شکل ۴-۱۲، شکل ۴-۱۳ و شکل ۴-۱۴ با استفاده از پیش‌پردازش Clahe عملکرد بهتری برای ویژگی HOG به دست می‌آید.

۴-۵ تأثیر تغییر پارامترهای LBP

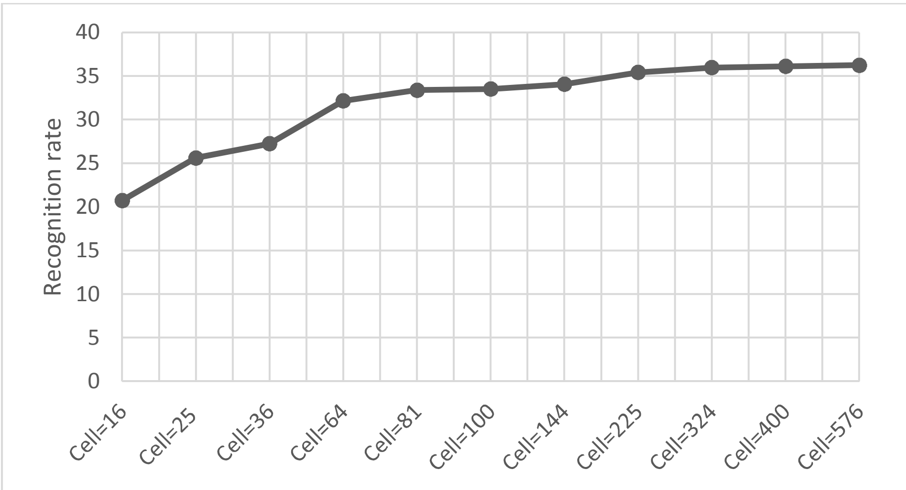
در استخراج ویژگی LBP از اپراتور دایره‌ای با شعاع ۱ و تعداد ۸ همسایه استفاده‌شده است.

۴-۵-۱ تأثیر تعداد سلول‌ها بر عملکرد LBP

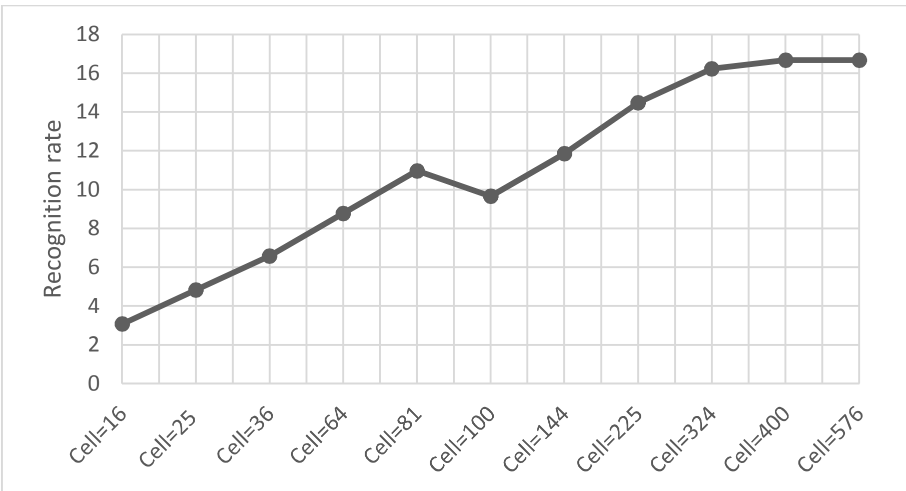
تأثیر تغییر تعداد سلول‌ها را روی عملکرد در شکل‌های زیر نشان داده‌شده‌اند.



شکل ۴-۱۵ تأثیر تغییرات تعداد سلول‌های LBP بر مجموعه Fa



شکل ۴-۱۶ تأثیر تغییرات تعداد سلول‌های LBP بر مجموعه Dup1

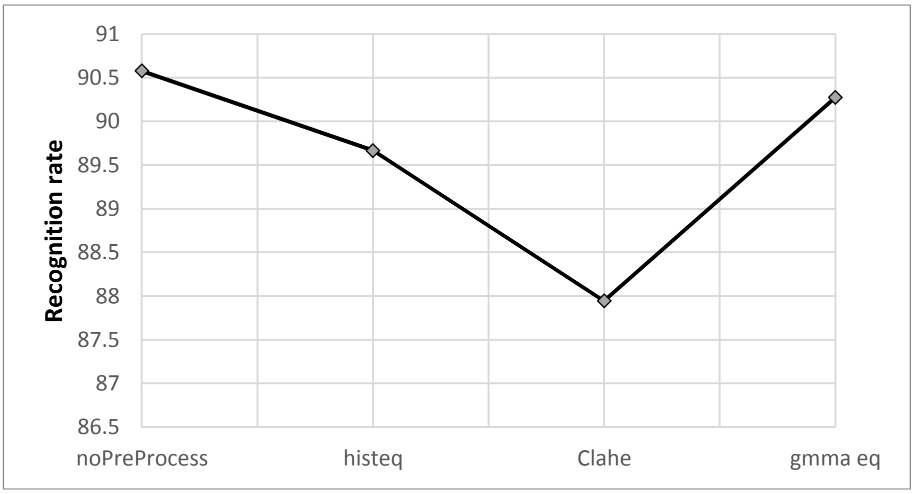


شکل ۴-۱۷ تأثیر تغییرات تعداد سلول‌های LBP بر مجموعه Dup2

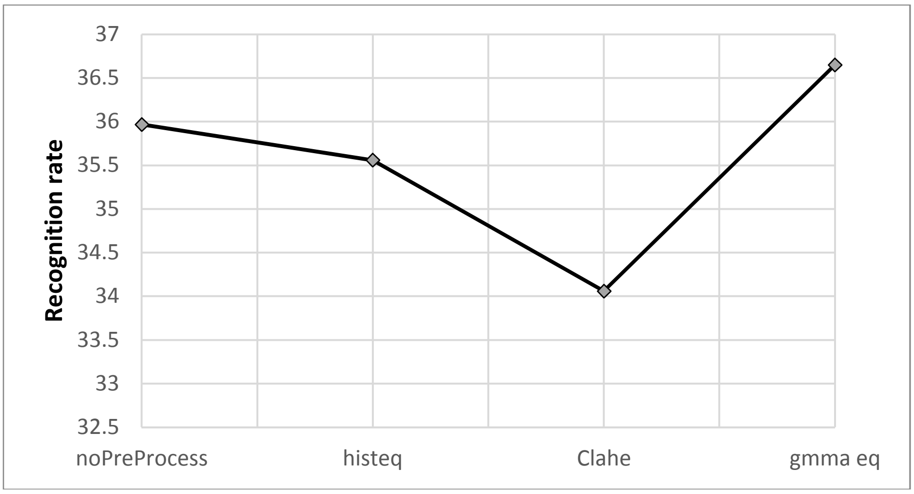
با توجه به شکل ۴-۱۵، شکل ۴-۱۶ و شکل ۴-۱۷ مقدار ۳۲۴ سلول بهترین نتیجه را دارد. علت این اتفاق این است که با افزایش بیشتر تعداد سلول‌ها ساختارهای محلی کوچک‌تری را برای ویژگی در نظر می‌گیرد. اگر مقدار سلول از ۳۲۴ کمتر باشد ساختارهای محلی به‌درستی ذخیره نمی‌شوند و بیشتر از این مقدار طول بردار ویژگی را افزایش می‌دهد که در زمان استخراج ویژگی تأثیرگذار است.

۴-۵-۲ **تأثیر پیش‌پردازش بر عملکرد LBP**

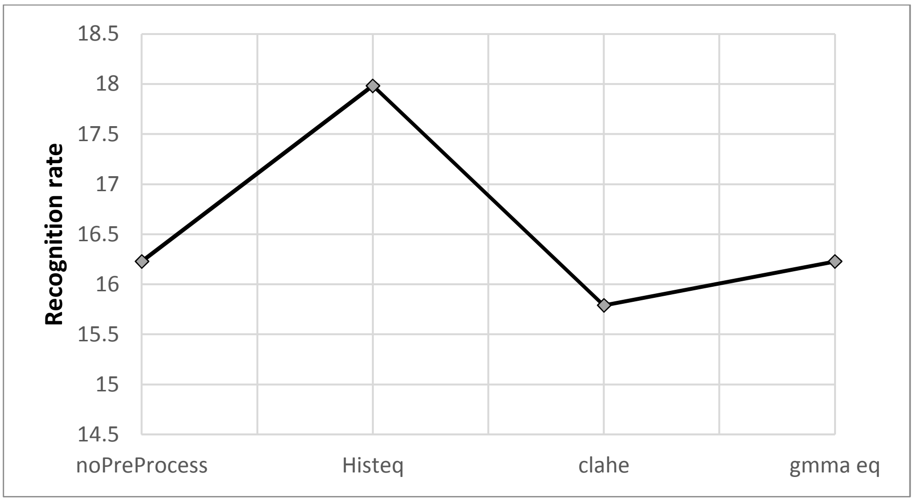
در این بخش تعداد سلول‌ها را ۳۲۴ در نظر می‌گیریم و سپس در ۴ حالت بدون پیش‌پردازش، یکسان‌سازی هیستوگرام، یکسان‌سازی گاما و یکسان‌سازی هیستوگرام وفقی، عملکرد این ویژگی را بررسی می‌کنیم.



شکل ۴-۱۸ تأثیر پیش‌پردازش‌های مختلف بر عملکرد LBP بر مجموعه Fb



شکل ۴-۱۹ تأثیر پیش‌پردازش‌های مختلف بر عملکرد LBP بر مجموعه Dup1



شکل ۴-۲۰ تأثیر پیش‌پردازش‌های مختلف بر عملکرد LBP بر مجموعه Dup2

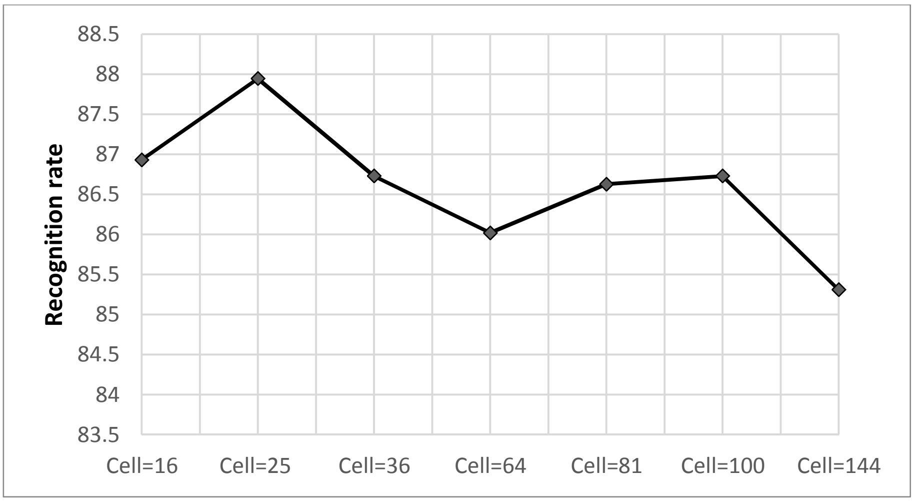
با توجه به شکل ۴-۱۸، شکل ۴-۱۹ و شکل ۴-۲۰ حالت بدون پیش‌پردازش بهترین عملکرد را دارد. دلیل اصلی این اتفاق این است که پیش‌پردازش‌های فوق، درصدد یکسان‌سازی شدت روشنایی می‌باشند اما ویژگی LBP نسبت به تغییرات شدت روشنایی مقاوم است. در پیش پردازش Clahe تغییرات بصورت محلی اعمال می‌شوند که باعث تغییر در ساختارهای محلی می‌شوند. تغییر در ساختارهای محلی باعث کم شدن نرخ بازشناسی می‌شود.

۶-۴ تأثیر تغییر پارامترهای STFT

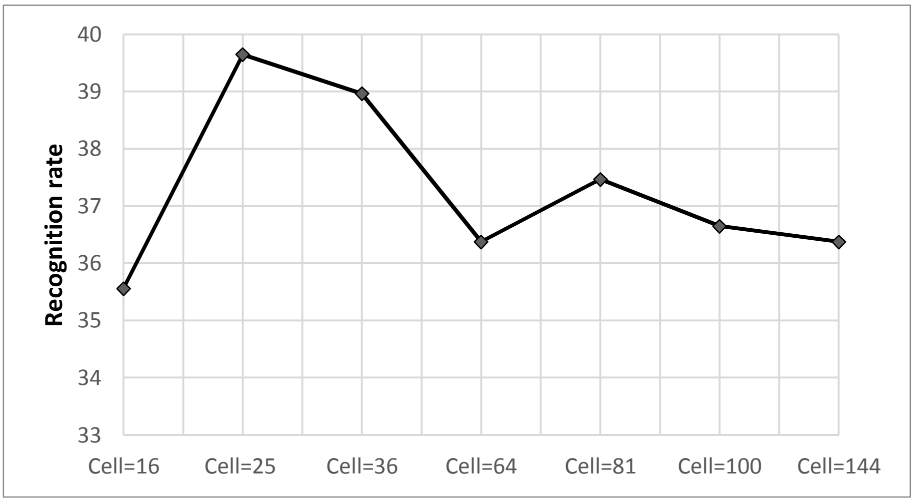
برای یافتن بهترین عملکرد با استفاده از این ویژگی، دو پارامتر تعداد سلول و تعداد ویژگی فرکانسی هر سلول را تغییر می‌دهیم. همچنین تأثیر پیش‌پردازش را نیز بررسی می‌کنیم.

۴-۶-۱ تأثیر تعداد سلول‌های STFT

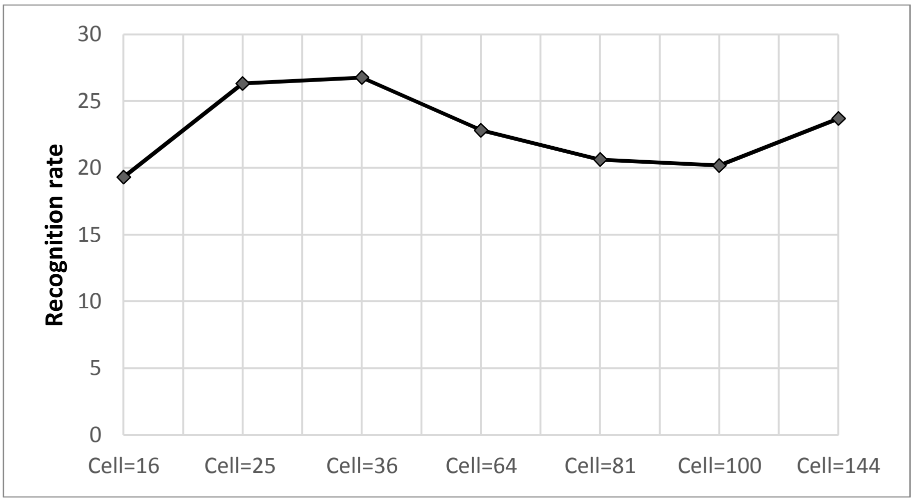
در این بخش تعداد ویژگی‌ها ۲۵ ویژگی فرکانسی به ازای هر سلول در نظر گرفته‌شده، پیش‌پردازشی صورت نگرفته و برای کلاسه‌بندی از معیار فاصله منهتن استفاده‌شده است.



شکل ۴-۲۱ تأثیر تغییرات تعداد سلول‌های STFT بر مجموعه Fa



شکل ۴-۲۲ تأثیر تغییرات تعداد سلول‌های STFT بر مجموعه Dup1



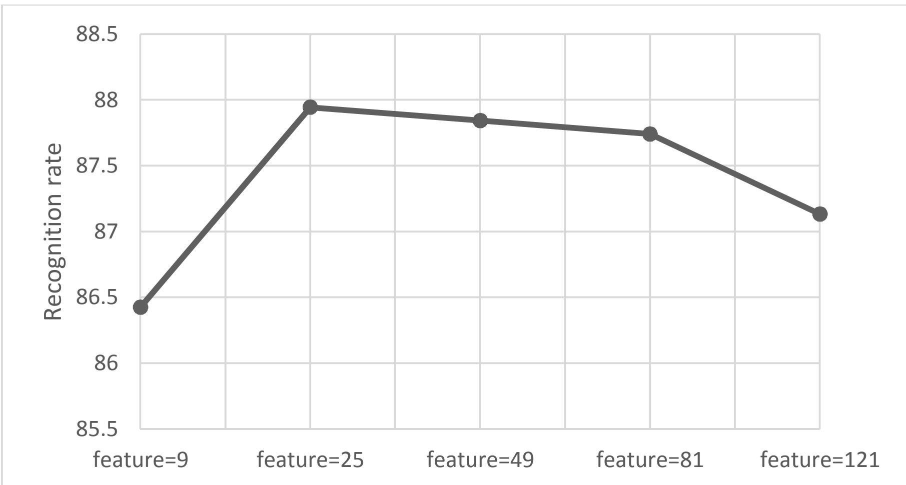
شکل ۴-۲۳ تأثیر تغییرات تعداد سلول‌های STFT بر مجموعه Dup2

با توجه به شکل ۴-۲۱، شکل ۴-۲۲ و شکل ۴-۲۳ تعداد سلول‌ها را ۲۵ سلول انتخاب می‌کنیم.

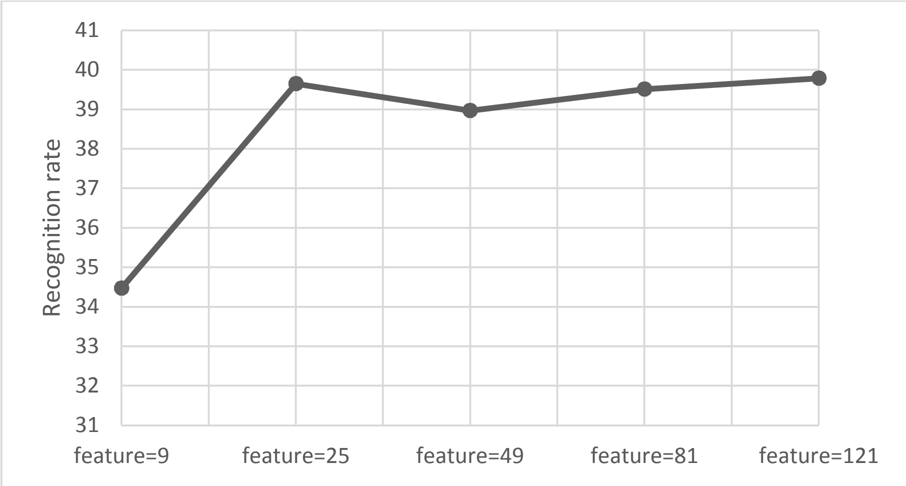
همانند ویژگی HOG تعداد سلول‌های بیشتر از حد خاص باعث برازش به نمونه آموزش شده و نرخ بازشناسی را پایین می‌آورد.

۴-۶-۲ تأثیر تعداد ویژگی فرکانسی هر سلول

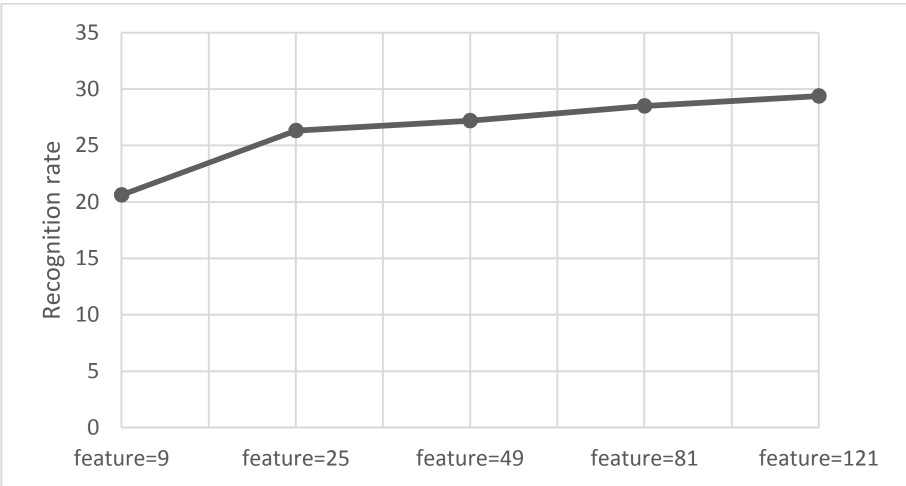
مشابه بخش قبلی این بار تعداد سلول‌ها را ثابت و ۲۵ انتخاب می‌کنیم و تأثیر تعداد ویژگی‌های فرکانسی را بررسی می‌کنیم.



شکل ۴-۲۴ تأثیر تغییرات تعداد ویژگی فرکانسی STFT بر مجموعه Fa



شکل ۴-۲۵ تأثیر تغییرات تعداد ویژگی فرکانسی STFT بر مجموعه Dup1



شکل ۴-۲۶ تأثیر تغییرات تعداد ویژگی فرکانسی STFT بر مجموعه Dup2

با توجه به شکل ۴-۲۴، شکل ۴-۲۵ و شکل ۴-۲۶ تعداد ویژگی‌های ۲۵ و ۸۱ عملکرد بهتری روی‌داده‌ها دارند. درجدول ۴-۲ مقایسه این دو ویژگی آمده است.

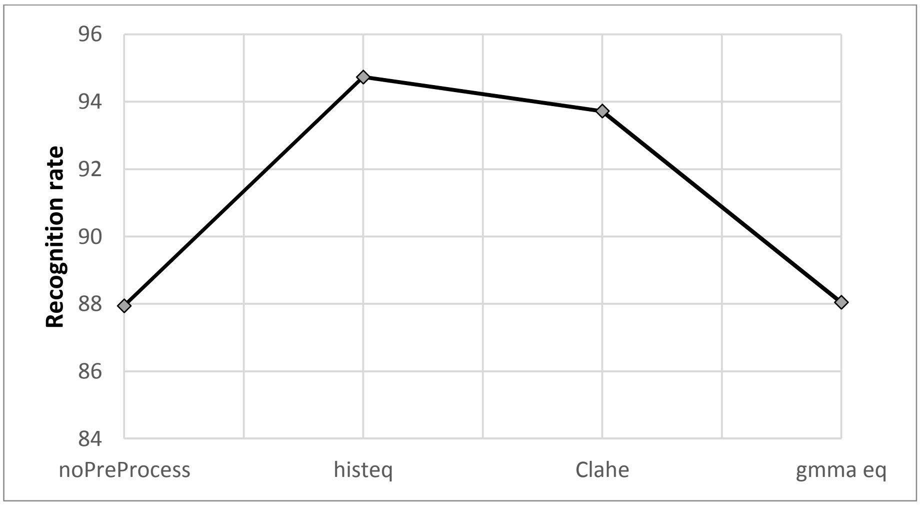
جدول ۴-۲ عملکرد سیستم بازشناسی چهره با استفاده از ویژگی STFT

تعداد ویژگی فرکانسی هر سلول		**۲۵**	**۸۱**
طول بردار ویژگی		۶۲۵	۲۰۲۵
Fb	Rank1	87.94	87.74
	Rank2	90.68	90.07
	Rank3	92.20	91.08
	Rank5	93.21	92.60
	Rank8	94.43	93.72

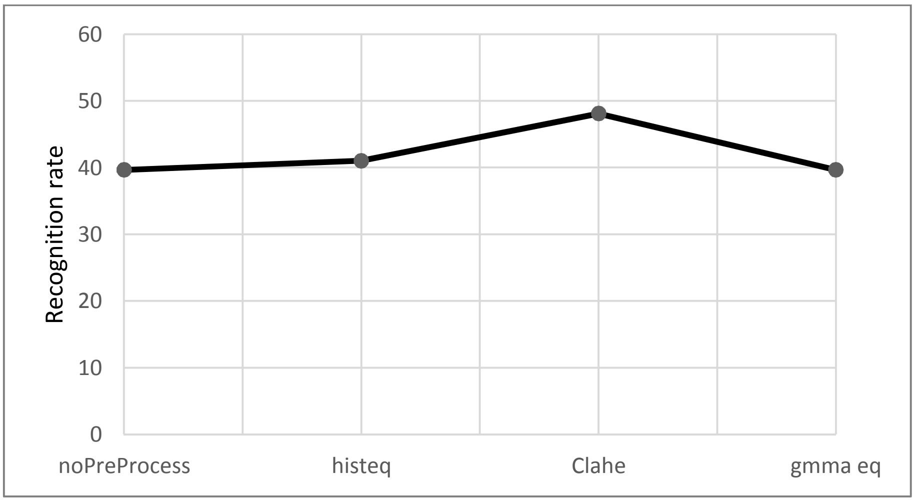
تعداد ویژگی فرکانسی هر سلول		۲۵	۸۱
	Rank15	95.74	95.03
	Mean	92.37	91.71
Dup1	Rank1	39.65	39.51
	Rank2	44.00	45.10
	Rank3	46.59	48.77
	Rank5	50.41	50.41
	Rank8	53.41	53.68
	Rank15	58.45	58.58
	Mean	48.75	49.34
Dup2	Rank1	26.32	28.51
	Rank2	33.33	35.96
	Rank3	37.28	41.23
	Rank5	44.30	44.30
	Rank8	49.12	50.88
	Rank15	59.21	57.46
	Mean	41.59	43.06

۴-۶-۳ تأثیر پیش‌پردازش

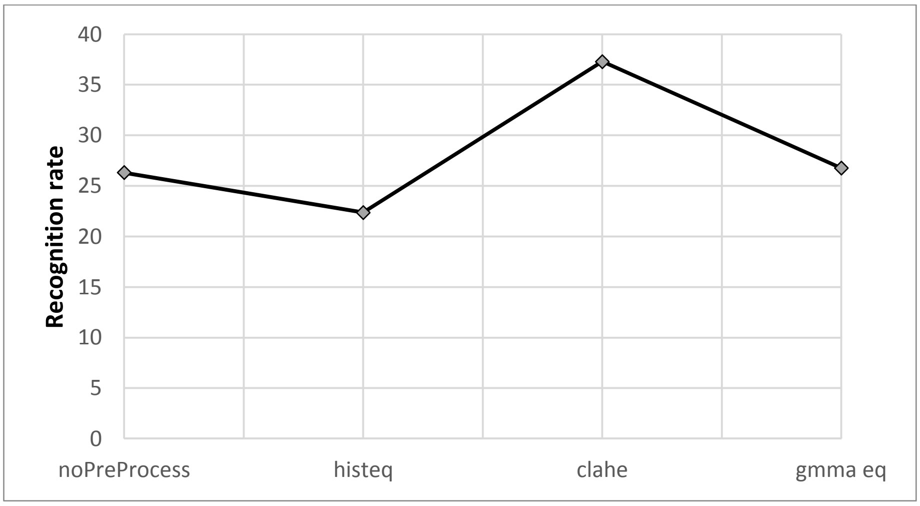
در این بخش تعداد سلول‌ها را ۲۵ و تعداد ویژگی فرکانسی را ۲۵ در نظر می‌گیریم و تأثیر پیش‌پردازش‌های مختلف را روی عملکرد STFT بررسی می‌کنیم.



شکل ۴-۲۷ تأثیر پیش‌پردازش‌های مختلف بر عملکرد STFT بر مجموعه Fb



شکل ۴-۲۸ تأثیر پیش‌پردازش‌های مختلف بر عملکرد STFT بر مجموعه Dup1



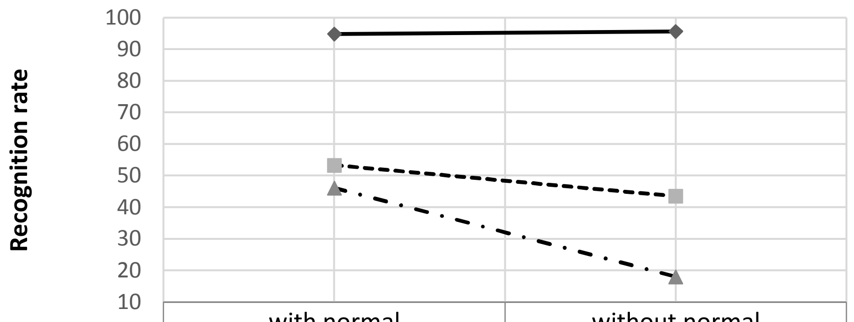
شکل ۴-۲۹ تأثیر پیش‌پردازش‌های مختلف بر عملکرد STFT بر مجموعه Dup2

با توجه به شکل ۴-۲۷، شکل ۴-۲۸ و شکل ۴-۲۹، پیش‌پردازش clahe بهترین عملکرد را داشته است.

۴-۷ نرمال‌سازی ویژگی‌ها

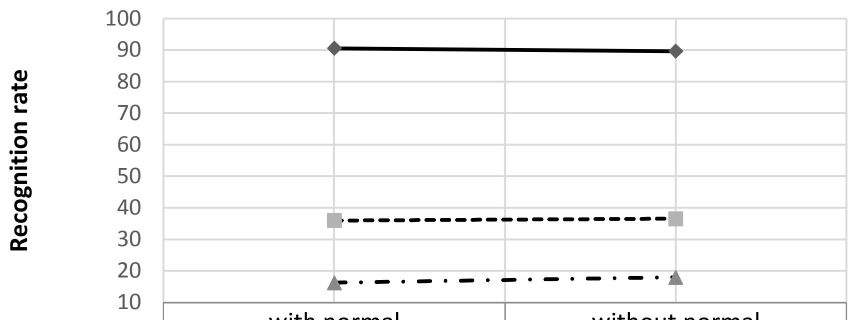
در این بخش تأثیر نرمال‌سازی ویژگی‌ها بر عملکرد را بررسی می‌کنیم، ویژگی HOG با ۱۴۴ سلول و ۹ دسته با پیش‌پردازش Clahe، ویژگی LBP با ۳۲۴ سلول و ۱۰ دسته بدون پیش‌پردازش و

ویژگی STFT با ۲۵ سلول و ۲۵ ویژگی فرکانسی در هر سلول و با پیش‌پردازش Clahe، هرکدام در دو حالت با نرمال‌سازی و بدون نرمال‌سازی آزمایش‌شده‌اند.



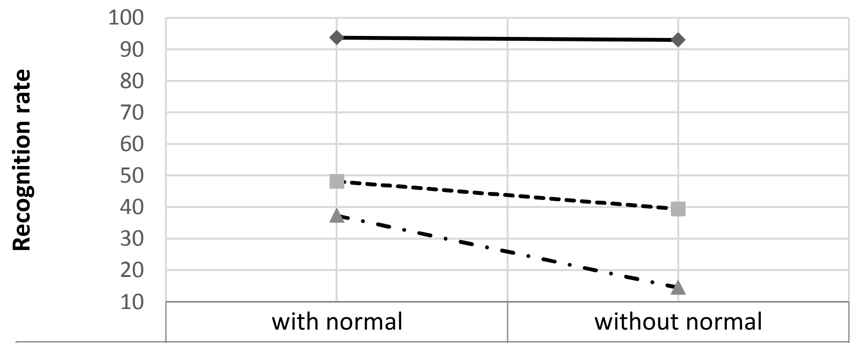
	with normal	without normal
FB	94.83282675	95.64336373
Dup1	53.26975477	43.46049046
Dup2	46.05263158	17.98245614

شکل ۴-۳۰ تأثیر نرمال‌سازی بر Hog



	with normal	without normal
FB	90.5775076	89.6656535
Dup1	35.96730245	36.51226158
Dup2	16.22807018	17.98245614

شکل ۴-۳۱ تأثیر نرمال‌سازی بر LBP



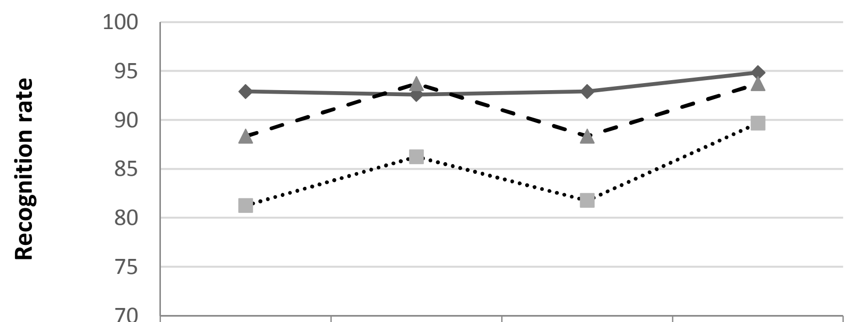
	with normal	without normal
FB	93.7183384	93.00911854
Dup1	48.09264305	39.373297
Dup2	37.28070175	14.47368421

شکل ۴-۳۲ تأثیر نرمال‌سازی بر STFT

با توجه به شکل ۴-۳۰، شکل ۴-۳۱ و شکل ۴-۳۲ مشاهده می‌شود که بدون نرمال‌سازی عملکرد HOG و STFT افت قابل‌ملاحظه‌ای دارد اما در LBP این‌گونه نیست.

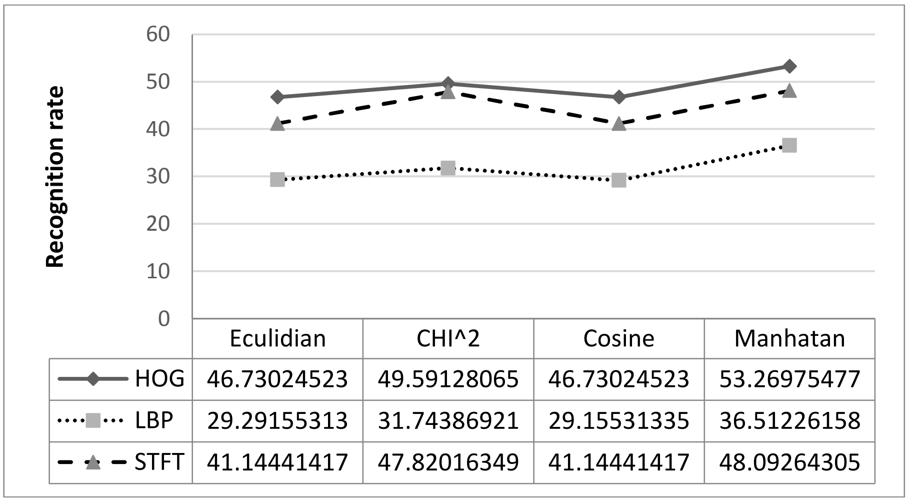
۴-۸ معیار فاصله

در این بخش تأثیر معیار فاصله بر عملکرد کلاسه بند را بررسی می‌کنیم. ویژگی HOG با ۱۴۴ سلول و ۹ دسته با پیش‌پردازش Clahe، ویژگی LBP با ۳۲۴ سلول و ۱۰ دسته بدون پیش‌پردازش و ویژگی STFT با ۲۵ سلول و ۲۵ ویژگی فرکانسی در هر سلول و با پیش‌پردازش Clahe، هرکدام با ۴ معیار ویژگی مختلف آزمایش می‌شوند.



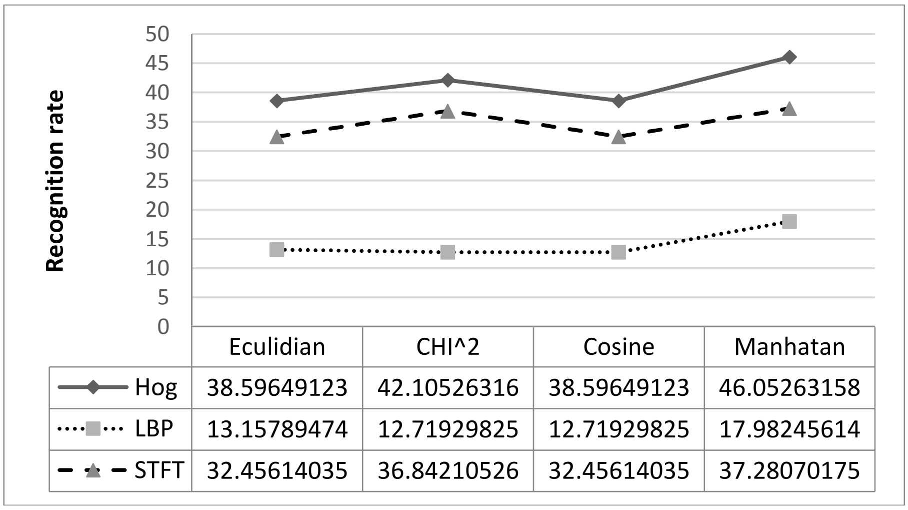
	Eculidian	CHI^2	Cosine	Manhatan
Hog	92.90780142	92.60385005	92.90780142	94.83282675
LBP	81.25633232	86.22087133	81.76291793	89.6656535
STFT	88.3485309	93.7183384	88.3485309	93.7183384

شکل ۴-۳۳ تأثیر معیارهای فاصله متفاوت بر عملکرد کلاسه بند در مجموعه Fb



	Eculidian	CHI^2	Cosine	Manhatan
HOG	46.73024523	49.59128065	46.73024523	53.26975477
LBP	29.29155313	31.74386921	29.15531335	36.51226158
STFT	41.14441417	47.82016349	41.14441417	48.09264305

شکل ۴-۳۴ تأثیر معیارهای فاصله متفاوت بر عملکرد کلاسه بند در مجموعه dup1



	Eculidian	CHI^2	Cosine	Manhatan
Hog	38.59649123	42.10526316	38.59649123	46.05263158
LBP	13.15789474	12.71929825	12.71929825	17.98245614
STFT	32.45614035	36.84210526	32.45614035	37.28070175

شکل ۴-۳۵ تأثیر معیارهای فاصله متفاوت بر عملکرد کلاسه بند در مجموعه dup2

با توجه به نتایج به‌دست‌آمده معیار فاصله منهتن بهتر از بقیه است.

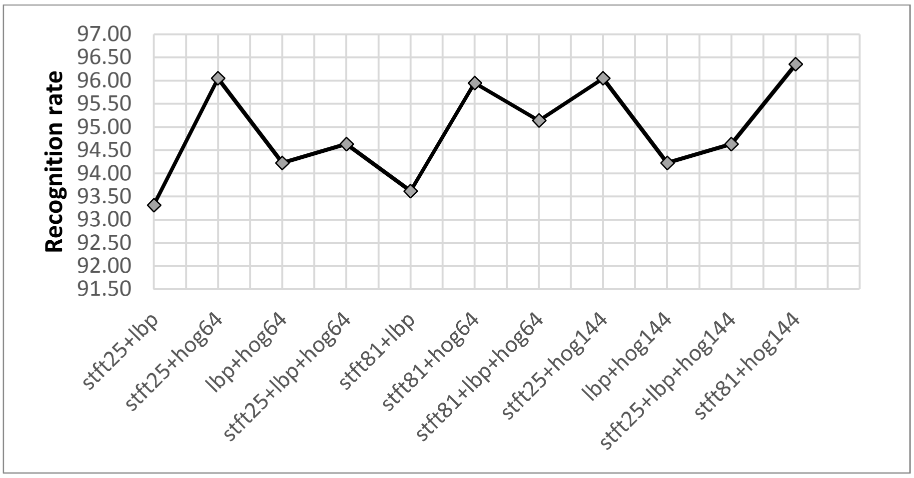
۴-۹ ترکیب ویژگی‌ها

پس از انجام بخش‌های بالا به بهترین کلاسه بند با استفاده از تنها یک ویژگی رسیدیم. در جدول ۴-۳ نتیجه کلاسه‌بندی با استفاده از هر ویژگی آورده شده است. در این بخش با ترکیب ویژگی‌ها به دنبال بهبود عملکرد هستیم.

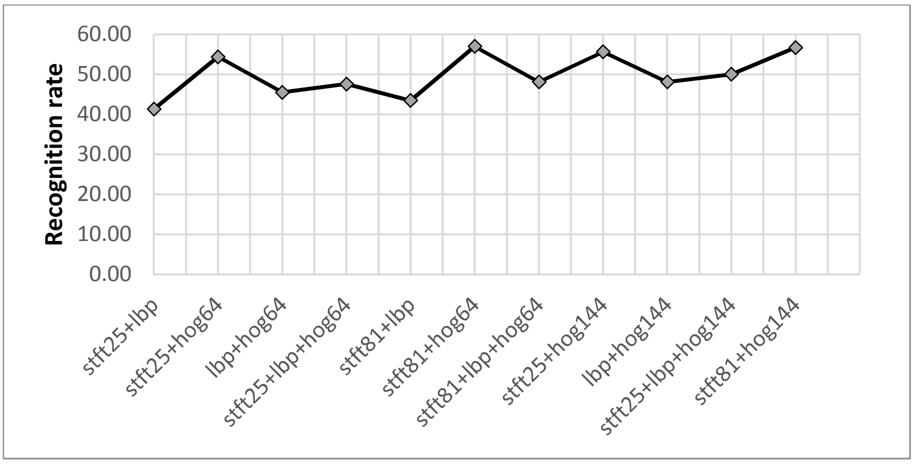
جدول ۴-۳ جزئیات و عملکرد ویژگی‌های نهایی

ویژگی		**HOG**		**LBP**	**STFT**	
تعداد سلول‌ها		64	144	324	25	
تعداد ویژگی در هر سلول		9		10	25	81
طول بردار ویژگی		576	1296	3240	625	2025
نام		HOG64	HOG144	LBP	STFT25	STFT81
Fb	Rank1	**94.33**	**94.83**	**90.58**	**93.72**	**94.43**
	Rank2	96.25	96.45	92.40	95.24	95.95
	Rank3	96.86	96.96	93.41	96.15	96.76
	Rank5	97.26	97.37	94.93	97.06	97.77
	Rank8	97.97	97.97	95.85	97.97	97.87
	Rank15	98.68	98.58	96.56	98.78	98.89
	Mean	96.89	97.03	93.95	96.49	96.94
Dup1	Rank1	**50.68**	**53.27**	**35.97**	**48.09**	**50.41**
	Rank2	57.36	58.86	39.65	52.72	55.86
	Rank3	61.17	62.26	42.64	56.40	58.86
	Rank5	65.40	67.44	45.10	61.04	62.81
	Rank8	69.48	70.84	48.77	64.44	64.85
	Rank15	73.02	74.52	52.86	67.98	69.07
	Mean	62.85	64.53	44.16	58.45	60.31
Dup2	Rank1	**41.67**	**46.05**	**16.23**	**37.28**	**42.54**
	Rank2	52.19	53.95	21.05	44.74	50.44
	Rank3	57.46	58.77	25.88	50.44	53.95
	Rank5	62.28	67.11	30.70	58.77	59.65
	Rank8	66.67	72.37	36.40	64.47	63.16
	Rank15	72.81	78.07	43.86	70.61	70.18
	Mean	58.85	62.72	29.02	54.39	56.65

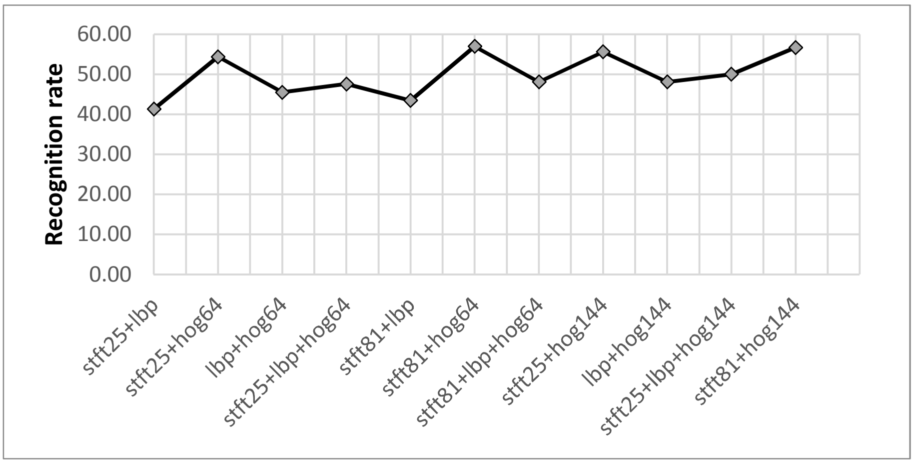
۱۲ حالت ترکیب مختلف از ۵ ویژگی بالا آزمایش‌شده است که نتایج آن در شکل‌های زیر مشاهده می‌شود.



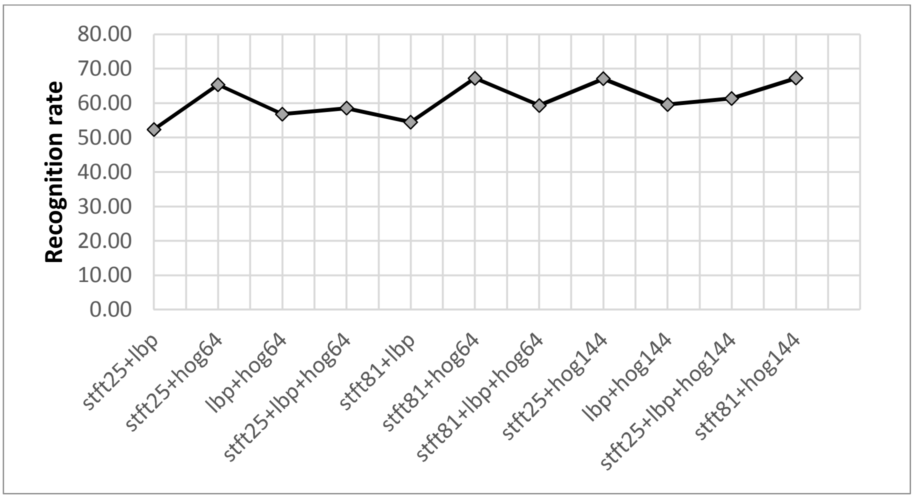
شکل ۴-۳۶ عملکرد ترکیب ویژگی‌های مختلف در مجموعه Fb



شکل ۴-۳۷ عملکرد ترکیب ویژگی‌های مختلف در مجموعه Dup1



شکل ۴-۳۸ عملکرد ترکیب ویژگی‌های مختلف در مجموعه Dup2



شکل ۴-۳۹ میانگین عملکرد ترکیب ویژگی‌ها روی پایگاه داده Feret

با توجه به نتایج به‌دست‌آمده ترکیب دو ویژگی STFT و HOG بهترین عملکرد را دارد.

۴-۱۰ استفاده از نمونه‌های مجازی

برای افزایش نرخ بازشناسی در الگوریتم پیشنهادی یکی از راهکارها تولید نمونه مجازی از نمونه آموزش است. ۶ نمونه مجازی از روی نمونه آموزشی تولید می‌شوند که به‌صورت زیر تولید می‌شوند:

- یک نمونه مجازی: مات کردن تصویر با کرنل میانگین‌گیر ۳*۳
- یک نمونه مجازی: تیز کردن لبه‌ها با استفاده از فیلتر لاپلاسین
- چهار نمونه مجازی: چرخش تصویر با زوایای ۳+، ۳-، ۵+ و ۵-

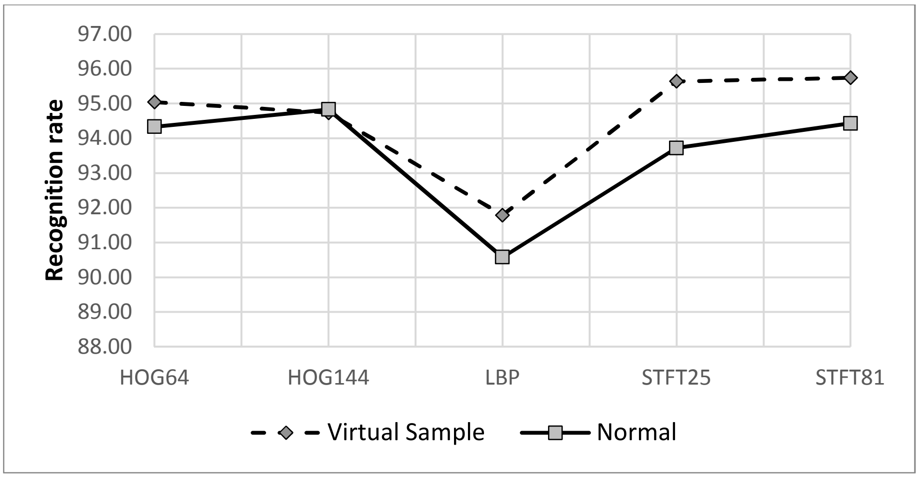
مجموعه جدید آموزش شامل نمونه اصلی و ۶ نمونه مجازی می‌باشد و مجموعه‌های آزمون و روش استخراج ویژگی و روش آزمون مثل بخش قبل است. روی ۵ ویژگی نهایی بخش قبل استفاده از نمونه‌های مجازی را آزمایش کردیم که در جدول ۴-۴ نتایج آن گزارش می‌شود.

جدول ۴-۴ جزئیات عملکرد ویژگی‌های نهایی در حالت نمونه مجازی

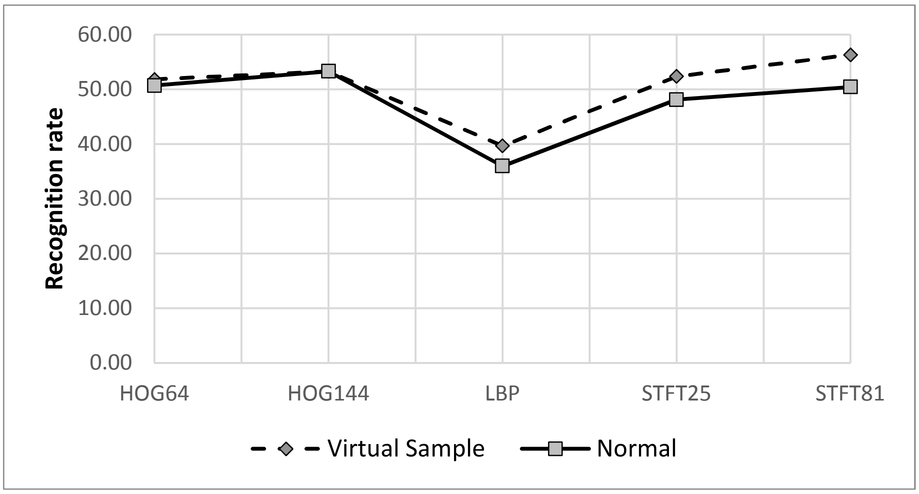
ویژگی		**HOG**		**LBP**	**STFT**	
تعداد سلول‌ها		64	144	324	25	
تعداد ویژگی در هر سلول		9		10	25	81
طول بردار ویژگی		576	1296	3240	625	2025
نام		HOG64VS	HOG144VS	LBPVS	STFT25VS	STFT81VS
Fb	Rank1	**95.04**	**94.73**	**91.79**	**95.64**	**95.74**
	Rank2	95.44	95.74	92.81	96.25	96.35
	Rank3	96.45	96.15	93.31	96.35	96.86
	Rank5	97.16	97.67	93.92	96.76	97.37
	Rank8	97.47	98.07	94.93	97.37	97.67
	Rank15	98.48	98.48	95.74	98.18	98.78
	Mean	96.67	96.81	93.75	96.76	97.13
Dup1	Rank1	**51.77**	**53.27**	**39.65**	**52.32**	**56.27**
	Rank2	54.36	55.31	41.01	55.59	58.72
	Rank3	56.95	57.49	41.83	57.22	60.08
	Rank5	60.76	62.26	45.23	59.81	62.26
	Rank8	64.31	66.35	48.23	61.99	64.99
	Rank15	67.98	70.44	51.23	66.21	67.98
	Mean	59.36	60.85	44.53	58.86	61.72
Dup2	Rank1	**41.67**	**43.86**	**21.49**	**44.30**	**50.88**
	Rank2	43.86	46.93	23.68	51.32	53.51
	Rank3	48.25	51.32	25.88	54.82	54.82
	Rank5	54.39	57.02	32.46	57.89	57.89
	Rank8	60.09	64.47	37.72	59.65	61.84
	Rank15	66.23	69.30	41.67	65.35	66.23
	Mean	52.41	55.48	30.48	55.56	56.65

در شکل ۴-۴۰، شکل ۴-۴۱ و شکل ۴-۴۲ مقایسه بین استفاده از نمونه مجازی و عدم استفاده از نمونه مجازی مشاهده می‌شود. این نمودارها نشان‌دهنده این هستند که استفاده از نمونه مجازی نرخ بازشناسی را در ویژگی‌های فوق افزایش داده‌اند.

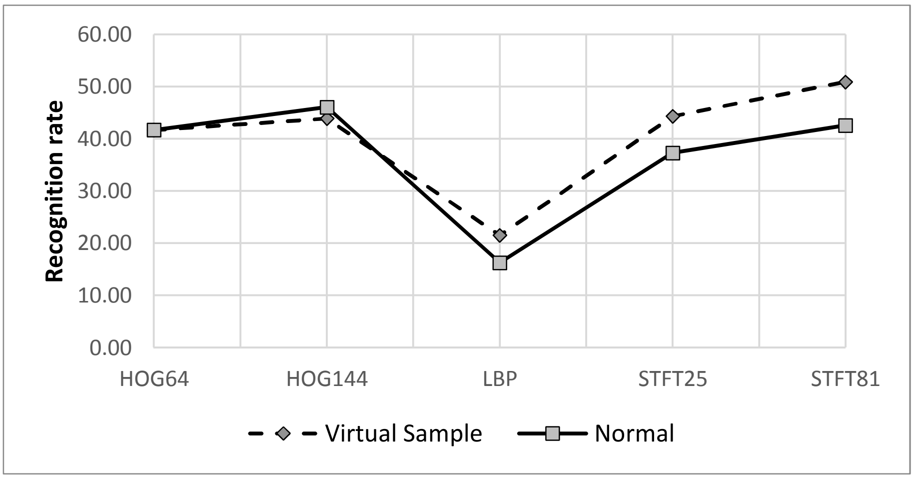
با توجه به شکل ۴-۳۹، ۳ ترکیب ویژگی STFT81+HOG64 و STFT25+HOG144 و STFT81+HOG144 که عملکرد بهتری دارند را نیز با استفاده از نمونه مجازی مورد آزمایش قرار می‌دهیم. در شکل ۴-۴۳ نتایج مقایسه‌ای را مشاهده می‌کنیم. نتایج نشان می‌دهند استفاده از نمونه مجازی باعث بهبود عملکرد ویژگی‌های ترکیبی می‌شود.



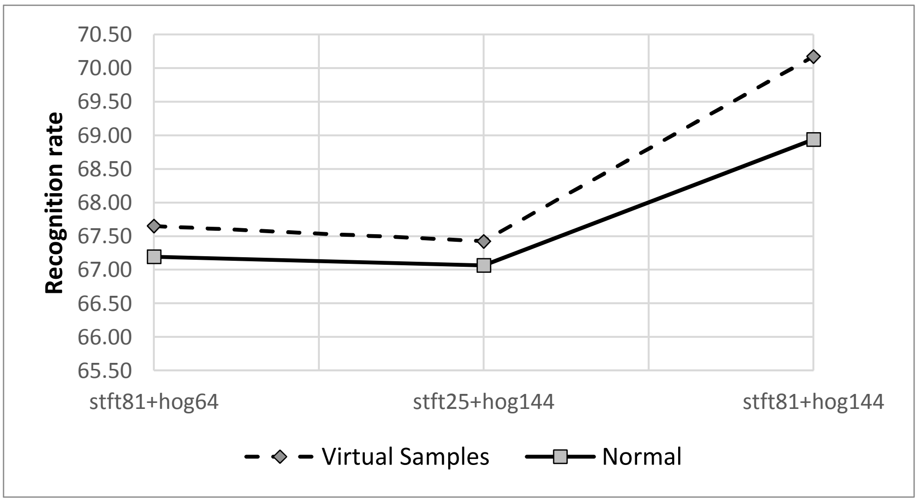
شکل ۴-۴۰ مقایسه بین استفاده از نمونه‌های مجازی با حالت معمولی در مجموعه Fb



شکل ۴-۴۱ مقایسه بین استفاده از نمونه‌های مجازی با حالت معمولی در مجموعه Dup1



شکل ۴-۴۲ مقایسه بین استفاده از نمونه‌های مجازی با حالت معمولی در مجموعه Dup2



شکل ۴-۴۳ مقایسه میانگین عملکرد ویژگی‌های ترکیبی بین استفاده از نمونه‌های مجازی با حالت معمولی

۴-۱۱جمع‌بندی

در این فصل روش پیشنهادی بررسی شد، در ابتدا برای هر ویژگی، پارامترهای آن ویژگی تغییر داده شد به‌گونه‌ای که بهترین عملکرد را بدهد سپس با ترکیب ویژگی‌های بهینه به بهترین کلاسه بند رسیدیم. همچنین با تولید نمونه مجازی و استفاده در مجموعه آموزش دریافتیم که با استفاده از نمونه مجازی عملکرد الگوریتم پیشنهادی بهبود می‌یابد.

فصل ۵: نتیجه گیری

۱-۵ بررسی نتایج

در این پایان‌نامه روشی برای بازشناسی چهره تصاویر از روبرو که به ازای هر فرد یک تصویر در مجموعه آموزش داریم ارائه شد. در این روش ابتدا ۳ ویژگی HOG، LBP و STFT از تصاویر آموزش استخراج شد سپس با تغییر پارامتر هر ویژگی به بهترین عملکرد برای هر ویژگی رسیدیم. از کلاسه بند نزدیک‌ترین همسایگی با معیار فاصله منهتن برای ارزیابی استفاده‌شده است. با امتحان کردن ترکیب‌های مختلف ویژگی‌های بهینه‌شده به بهترین عملکرد با ترکیب این ویژگی‌ها رسیدیم. همچنین با استفاده از نمونه‌های مجازی توانستیم نرخ بازشناسی را بهبود بخشیم. در این پایان‌نامه از پایگاه داده رابطه‌ای برای ذخیره و بازیابی ویژگی‌ها استفاده شد که مدیریت آسان‌تر داده‌ها و سرعت بیشتر در بازشناسی دو مزیت مهم این روش می‌باشند.

در جدول ۵-۱ بهترین عملکرد ارائه‌شده توسط این پایان‌نامه روی پایگاه داده Feret مشاهده می‌شود.

جدول ۵-۱ عملکرد الگوریتم پیشنهادی روی پایگاه داده Feret

ویژگی‌های ترکیبی	طول بردار ویژگی	fb	Dup1	Dup2	میانگین
stft81+hog64	2601	95.95٪	56.95%	48.68%	67.19%
stft81+hog144	3321	95.95٪	57.36%	53.51%	68.94%
stft81VS+hog64VS	2601	**96.66٪**	57.49%	48.81%	67.65%
stft81VS+hog144VS	3321	96.35٪	**58.98%**	**55.18%**	70.17%

۲-۵ بررسی زمانی

الگوریتم پیشنهادی با استفاده از پایگاه داده sqliteو زبان‌های ++C و #C در محیط visual studio 2015 پیاده‌سازی شده است. روی یک لپ‌تاپ cpu: core i7، Ram:8 GB اجراشده که در جدول ۵-۲ زمان مورد نیاز برای هر بخش الگوریتم نشان داده‌شده است. تعداد کل نمونه‌های آموزش در حالت معمولی ۹۹۰ و در حالت استفاده از نمونه مجازی ۶۹۳۰ عدد است.

جدول ۵-۲ عملکرد زمانی الگوریتم پیشنهادی

بخش الگوریتم	مدت‌زمان برای هر تصویر (ms) در حالت معمولی	مدت‌زمان کل در حالت معمول (S)	مدت‌زمان برای هر تصویر (ms) در حالت استفاده از نمونه مجازی	مدت‌زمان کل در حالت استفاده از نمونه مجازی (S)
آشکارسازی و جداسازی چهره	۱۱۷	۱۱۵.۸۳	۱۱۷	۱۱۵.۸۳
استخراج و ذخیره‌سازی ویژگی HOG64	۱۸	۱۷.۸۲	۱۸	۱۲۴.۷۴
استخراج و ذخیره‌سازی ویژگی STFT81	۲۴	۲۳.۷۶	۲۱	۱۴۵.۵۳
استخراج و ذخیره‌سازی ویژگی LBP	۵۳	۵۲.۴۷	۴۲	۲۹۱.۰۶
بازیابی ویژگی از پایگاه داده Hog64+STFT81	۲۱۰۰۰ (باید تمام ویژگی‌های تصاویر مجموعه آموزش فقط یک‌بار بازیابی شوند.)	۲۱ (یک‌بار)	۱۶۲۰۰۰(باید تمام ویژگی‌های تصاویر مجموعه آموزش فقط یک‌بار بازیابی شوند.)	۱۶۲(یک‌بار)
محاسبه نزدیک‌ترین فاصله	۱۰	۹.۹	۱۱	۱۰.۸۹
مدت‌زمان آموزش	۱۵۹	۱۵۷.۴۱	۱۵۶	۱۰۸.۱۱
مدت‌زمان آزمون	۱۵۸+۲۱۰۰۰	۱۵۶.۴۲+۲۱	۱۵۸+۱۶۲۰۰۰	۱۵۷+۱۶۲

۵-۳ مقایسه با سایر کارها

در جدول ۵-۳ مقایسه‌ای بین نتایج الگوریتم پیشنهادی با سایر الگوریتم‌هایی که از پایگاه داده Feret استفاده می‌کنند انجام شده است.

جدول ۵-۳ مقایسه الگوریتم پیشنهادی با سایر الگوریتم‌ها

نام روش	پایگاه داده	زیرمجموعه پایگاه داده	تعداد نفرات مورد آزمایش	تعداد نمونه‌های آموزش	طول بردار ویژگی	نرخ بازشناسی	زمان آزمون برای هر تصویر (ms)
LBP [8]	feret	Fb Dup1 Dup2	1196	1	2301	Fb:97% Dup1:66% Dup2:64%	۸۷۵
[۱۲]	feret	Fb	990	1	گزارش نشده	Fb:90.4%	۸۶۰
هاف [۱۰]	feret	گزارش نشده	990	20	2042	80.81%	۹۰
EBGM [7] گزارش‌شده در [۸]	feret	Fb Dup1 Dup2	1196	1	گزارش نشده	Fb:90% Dup1:46% Dup2:24%	گزارش نشده
روش پیشنهادی	feret	Fb Dup1 Dup2	990	1	مراجعه به جدول ۵-۱		۱۵۸

۵-۴ پیشنهادها و کارهای آینده

- استفاده از الگوریتم بهتری برای آشکارسازی چهره چراکه جداسازی درست چهره از تصویر ورودی تأثیر مستقیمی بر عملکرد نهایی دارد.
- امتحان کردن ویژگی‌های دیگر مثل تبدیل موجک، Dct، سایر اپراتورهای LBP و گابور
- پیش‌پردازش‌های امتحان شده در این پایان‌نامه روی یکسان‌سازی شدت روی روشنایی بین همه نمونه‌ها تمرکز داشتند، احتمال می‌رود پیش‌پردازشی که تصویر چهره را بر اساس موقعیت چشم‌ها و سایر اجزا چهره تراز نماید نیز عملکرد الگوریتم را بهبود ببخشد.

- بجای ترکیب ویژگی‌ها، برای هر ویژگی یک کلاسه بند در نظر گرفته شود و با رأی‌گیری نتایج کلاسه بندها نتیجه نهایی تعیین شود.
- در روش پیشنهادی برای حفظ کردن اطلاعات مکانی از ناحیه بندی استفاده شد که می‌توان بهبودهای زیر را در ناحیه بندی لحاظ کرد.
  - ناحیه‌ها لزوماً هم‌اندازه و هم‌شکل نباشند.
  - می‌توان با وزن دهی به هر ناحیه تأثیر نواحی مهم‌تر را افزایش داد.
  - ناحیه بندی به‌گونه‌ای انجام شود که نواحی مهم چهره مثل چشم‌ها، بینی و ... در نواحی بیشتری باشند یعنی بین نواحی هم‌پوشانی داشته باشیم.
- همان‌طور که مشاهده شد استفاده از نمونه مجازی عملکرد الگوریتم را بهبود می‌بخشد می‌توان با تولید نمونه‌های متنوع، به‌طور مثال تبدیل Affine و یا تعیین اینکه هر نمونه چقدر در دقت اثر دارد به عملکرد بهتری رسید.
- همان‌طور که در فصل ۴ دیده شد، انتخاب معیار فاصله مناسب در عملکرد کلاسه بند تأثیر می‌گذارد، امکان‌پذیر است با تغییر معیار فاصله نتیجه بهتری گرفته شود.
- با استفاده از الگوریتم‌های انتخاب ویژگی، می‌توان طول بردار ویژگی را کاهش داد بطوریکه روی عملکرد تأثیر منفی نداشته باشد.

مراجع

۱ Zhao, W., R. Chellappa, P. J. Phillips and A. Rosenfeld (2003). "Face recognition: A literature survey." ACM computing surveys (CSUR) **35**(4): 399-458.

۲ Turk, M. A. and A. P. Pentland (1991). "Face recognition using eigenfaces". Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on, IEEE.

۳ Belhumeur, P. N., J. P. Hespanha and D. J. Kriegman (1997). "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection." IEEE Transactions on pattern analysis and machine intelligence **19**(7): 711-720.

۴ Yang, M.-H. (2002). "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods". Fgr.

۵ Kanade, T. (1974). "Picture processing system by computer complex and recognition of human faces".

۶ Brunelli, R. and T. Poggio (1992). "Face recognition through geometrical features". Computer Vision—ECCV'92, Springer.

۷ Wiskott, L., N. Krüger, N. Kuiger and C. Von Der Malsburg (1997). "Face recognition by elastic bunch graph matching." IEEE Transactions on pattern analysis and machine intelligence **19**(7): 775-779.

۸ Ahonen, T., A. Hadid and M. Pietikainen (2006). "Face description with local binary patterns: Application to face recognition." IEEE transactions on pattern analysis and machine intelligence **28**(12): 2037-2041.

۹ Nefian, A. V. and M. H. Hayes (1998). "Hidden Markov models for face recognition". Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on, IEEE.

۱۰ Varun, R., Y. V. Kini ,K. Manikantan and S. Ramachandran (2015). "Face Recognition using Hough Transform based Feature Extraction." Procedia Computer Science **46**: 1491-1500.

۱۱ Luo, J., Y. Ma, E. Takikawa, S. Lao, M. Kawade and B.-L. Lu (2007). "Person-specific SIFT features for face recognition". Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, IEEE.

۱۲ بخشی, م. (۱۳۹۵). "استخراج ویژگی از تصاویر چهره جهت بازیابی تصویر از پایگاه داده". کارشناسی ارشد پایان‌نامه, دانشگاه شاهرود.

۱۳ Rowley, H. A., S. Baluja and T. Kanade (1998). "Neural network-based face detection." IEEE Transactions on pattern analysis and machine intelligence **20**(1): 23-38.

۱۴ Yang, F. and M. Paindavoine (2003). "Implementation of an RBF neural network on embedded systems: real-time face tracking and identity verification." IEEE Transactions on Neural Networks **14**(5): 1162-1175.

۱۵ Schroff, F., D. Kalenichenko and J. Philbin (2015). "Facenet: A unified embedding for face recognition and clustering". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

۱۶ سعیدی (۱۳۹۴). "احراز هویت با استفاده از چهره در مجموعه بزرگی از تصاویر افراد" پایان‌نامه, دانشگاه شاهرود.

۱۷ Wright, J., A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma (2009). "Robust face recognition via sparse representation." IEEE transactions on pattern analysis and machine intelligence **31**(2): 210-227.

۱۸ Yang, J., L. Luo, J. Qian, Y. Tai, F. Zhang and Y. Xu (2017). "Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes." IEEE transactions on pattern analysis and machine intelligence **39**(1): 156-171.

۱۹ Ojala, T., M. Pietikainen and T. Maenpaa (2002). "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." IEEE Transactions on pattern analysis and machine intelligence **24**(7): 971-987.

۲۰ Ghazali, K. H. B., J. Ma and R. Xiao (2011). "An innovative face detection based on skin color segmentation." International Journal of Computer Applications **34**(2): 6-10.

۲۱ Tan, X., S. Chen, Z.-H. Zhou and F. Zhang (2006). "Face recognition from a single image per person: A survey." Pattern recognition **39**(9): 1725-1745.

۲۲ Hallinan, P. W., G. G. Gordon, A. L. Yuille, P. Giblin and D. Mumford (1999). "Two–and three-dimensional patterns of the face", AK Peters, Ltd.

۲۳ Martínez, A. M. (2002). "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class." IEEE Transactions on Pattern analysis and machine intelligence **24**(6): 748-763.

۲۴ Tan, X., S. Chen, Z.-H. Zhou and F. Zhang (2005). "Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft k-NN ensemble." IEEE Transactions on Neural Networks **16** :(۴) ۸۷۵-۸۸۶.

۲۵ Heiselet, B., T. Serre, M. Pontil and T. Poggio (2001). "Component-based face detection". Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, IEEE.

۲۶ Costen, N., T. F. Cootes and C. J. Taylor (2002). "Compensating for ensemble-specific effects when building facial models." Image and Vision Computing **20**(9): 673-682.

۲۷ Lanitis, A., C. J. Taylor and T. F. Cootes (1995). "Automatic face identification system using flexible appearance models." Image and vision computing **13**(5): 393-401.

۲۸ Penev, P. S. and J. J. Atick (1996). "Local feature analysis: A general statistical theory for object representation." Network: computation in neural systems **7**(3): 477-500.

۲۹ Taigman, Y., M. Yang, M. A. Ranzato and L. Wolf (2014). "DeepFace: Closing the Gap to Human-Level Performance in Face Verification". Conference on Computer Vision and Pattern Recognition (CVPR), IEEE.

۳۰ Sun, Y., X. Wang and X. Tang (2015). "Deeply learned face representations are sparse, selective, and robust". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

۳۱ Parkhi, O. M., A. Vedaldi and A. Zisserman (2015). "Deep Face Recognition". BMVC.

۳۲ AbdAlmageed, W., Y. Wu, S. Rawls, S. Harel, T .Hassner, I. Masi, J. Choi, J. Lekust, J. Kim and P. Natarajan (2016). "Face recognition using deep multi-pose representations". Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on, IEEE.

۳۳ Zeiler, M. D. and R. Fergus (2013). "Visualizing and understanding convolutional networks." arXiv preprint arXiv:1311.2901.

۳۴ Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich (2015). "Going deeper with convolutions". Proceedings of the IEEE conference on computer vision and pattern recognition.

۳۵ Samaria, F. S. and A. C. Harter (1994). "Parameterisation of a stochastic model for human face identification". Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on, IEEE.

۳۶ Georghiades, A. S., P. N. Belhumeur and D. J. Kriegman (2001). "From few to many: Illumination cone models for face recognition under variable lighting and pose." IEEE transactions on pattern analysis and machine intelligence **23**(6): 643-660.

۳۷ Martinez, A. M. (1998). "The AR face database." CVC technical report.

۳۸ Phillips, P. J., H. Wechsler, J. Huang and P. J. Rauss (1998). "The FERET database and evaluation procedure for face-recognition algorithms." Image and vision computing **16**(5): 295-306.

۳۹ Huang, G. B., M. Ramesh, T. Berg and E. Learned-Miller (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Secondary "Labeled faces in the wild: A database for studying face recognition in unconstrained environments", Technical Report 07-49, University of Massachusetts, Amherst.

۴۰ Gao, W., B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang and D. Zhao (2008). "The CAS-PEAL large-scale Chinese face database and baseline evaluations." IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans **38**(1): 149-161.

۴۱ Viola, P. and M. J. Jones (2004). "Robust real-time face detection." International journal of computer vision **57**(2): 137-154.

۴۲ Zhu, X. and D. Ramanan (2012). "Face detection, pose estimation, and landmark localization in the wild". Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE.

۴۳ Wang, Y.-Q. (2014). "An analysis of the Viola-Jones face detection algorithm." Image Processing On Line **4**: 128-148.

۴۴ Gonzalez, R. C. and R. E. Woods (2006). "Digital Image Processing (3rd Edition)", Prentice-Hall, Inc.

۴۵ Zuiderveld, K. (1994). "Contrast limited adaptive histogram equalization". Graphics gems IV, Academic Press Professional, Inc.

۴۶ Hlawatsch, F. and G .F. Boudreaux-Bartels (1992). "Linear and quadratic time-frequency signal representations." IEEE signal processing magazine **9**(2): 21-67.

۴۷ Dalal, N. and B. Triggs (2005). "Histograms of oriented gradients for human detection". Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, IEEE.

۴۸ Duda, R. O., P. E. Hart and D. G. Stork (2012). "Pattern classification", John Wiley & Sons.

۴۹ Date, C. J. (2006). "An introduction to database systems", Pearson Education India.

Abstract

Face recognition is one of common methods of identifying individuals which has many applications such as personal device login, tracking offenders, human interaction with computers and etc. The objective of face recognition is to identify a person or persons from an image using database. There are many challenges for face recognition such as illumination variation, facial expression variations, occlusion, age variations, low number of training samples, high number of subjects and etc. In this thesis, we propose a robust method against illumination variations, facial expression and age variations.

In this thesis, after face detection local features such as HOG, LBP and STFT are extracted. To achieve best performance, we use preprocessing and tuning feature's parameters for each feature. Then, by concating the vectors of optimal features, a final feature vector is constructed and classifying is done by using the nearest neighborhood classifier and Manhattan's distance criteria. In order to achieve even better performance, we produce virtual samples using image blurring, sharpening and rotation for each train sample and then we add these virtual samples to training set. To evaluate the proposed method, we have used the Feret database, which contains about 1,000 people images, and there is only one face image per person in the training set. It should be noted that most of the methods presented in Face Recognition have addressed one of the existing challenges and have been tested on a database of fewer than 200 people. Experimental results on the Feret database have a recognition rate of 96.96% using virtual samples and 95.95% without using virtual samples with combination of HOG and STFT feature on the fb series with 158 milliseconds per image of the test set.

Keywords: Face Recognition, Local Features, Feature Extraction, Large Database

Shahrood University of Technology
Faculty of Electrical Engineering and Robotic
M.Sc. Thesis in Robotic Engineering

Face authentication using spatial and frequency features of image in large dataset

By: Khashayar Nourian

Supervisor:
Dr Hossein Khosravi

September 2017