

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی ارشد هوش مصنوعی

استفاده از دسته‌بند سلسله‌مراتبی در شناسایی چهره در پایگاه داده حجیم

نگارنده: نوید عباسپور طسمالو

استاد راهنما

دکتر حمید حسن‌پور

استاد مشاور

دکتر محسن بیگری

شهریور ۱۳۹۹

شماره: ۷.۸، رفک
تاریخ: ۲۴، ۱۸، ۹۹

باسمه تعالی



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

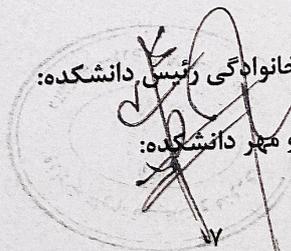
با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد آقای نوید عباسپور طسمالو با شماره دانشجویی ۹۶۱۰۳۶۴ رشته مهندسی کامپیوتر گرایش هوش مصنوعی تحت عنوان استفاده از دسته‌بند سلسله‌مراتبی در شناسایی چهره در پایگاه داده حجیم که در تاریخ ۱۳۹۹/۰۷/۲۲ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می‌گردد:

<input checked="" type="checkbox"/> الف) درجه عالی: نمره ۱۹-۲۰	<input type="checkbox"/> ب) درجه خیلی خوب: نمره ۱۸/۹۹-۱۸
<input type="checkbox"/> ج) درجه خوب: نمره ۱۶-۱۷/۹۹	<input type="checkbox"/> د) درجه متوسط: نمره ۱۴-۱۵/۹۹
<input type="checkbox"/> ه) کمتر از ۱۴ غیر قابل قبول و نیاز به دفاع مجدد دارد	
نوع تحقیق: <input type="checkbox"/> نظری <input type="checkbox"/> عملی	

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	حمید حسن پور		
۲- استاد راهنمای دوم			
۳- استاد مشاور	حسن بهتری		
۴- نماینده تحصیلات تکمیلی	مهرزاد رحادی	مربی	
۵- استاد ممتحن اول	مفتوح نایع	استادیار	
۶- استاد ممتحن دوم	احسان طهرانیان	استادیار	

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:



تقدیم به پدر بزرگوار و مادر مهربانم

آن دو فرشته‌ای که از خواسته‌هایشان گذشتند،
سختی‌ها را به جان خریدند
و خود را سپر بلای مشکلات و ناملایمات کردند
تا من به جایگاهی که اکنون در آن ایستاده‌ام برسم.

سپاس‌گزاری

یگانه آفریدگار من، ترا شاکرم که بر سر این بنده حقیر منت نهادی و توانستم به همت این علم محدود این تحقیق را به سرانجام برسانم.
به مصداق آیه شریفه ”من لم یشکر المخلوق، لم یشکر الخالق“ مراتب قدردانی خود را از محضر استاد راهنمای ارجمند آقای دکتر حمید حسن‌پور اعلام می‌نمایم.
همچنین از زحمات استاد مشاور محترم آقای دکتر محسن بیگلری، نهایت تشکر را دارم.

نوید عباسپور طسمالو

شهریور ۱۳۹۹

تعهد نامه

اینجانب نوید عباسپور طسمالو دانشجوی کارشناسی ارشد رشته هوش مصنوعی مهندسی کامپیوتر دانشگاه شاهرود، نویسنده پایان نامه با عنوان استفاده از دسته‌بند سلسله‌مراتبی در شناسایی چهره در پایگاه داده حجیم، تحت راهنمایی حمید حسن‌پور متعهد می‌شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام “ دانشگاه صنعتی شاهرود “ یا “ Shahrood University of Technology “ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در تمام 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 تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

نوید عباسپور طسمالو

شهریور ۱۳۹۹

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

شناسایی چهره یکی از متداول‌ترین روش‌های بیومتریک احراز هویت است و تحقیقات زیادی در این زمینه برای بهبود عملکرد آن انجام گرفته است. محبوبیت این روش بیومتریک نسبت به روش‌های مبتنی بر عنیبه چشم، اثر انگشت و صدا، به دلیل ساده بودن استفاده از چهره و همچنین دخالت نداشتن انسان در امر احراز هویت است. با این حال چالش‌های زیادی در این موضوع وجود دارند که محققین در حال تحقیق بر روی آن‌ها هستند. یکی از این چالش‌ها حجم بزرگ پایگاه داده است که هر چه تعداد تصاویر بیشتر می‌شود، فاصله بردارهای ویژگی استخراج‌شده به هم نزدیک‌تر و دقت شناسایی شدیداً کاهش می‌یابد. همچنین با افزایش تعداد داده در پایگاه داده سرعت جستجو به شدت کاهش می‌یابد.

هدف این پایان‌نامه ارائه یک روش مقاوم شناسایی چهره است که نسبت به زیاد بودن تصاویر پایگاه داده عملکرد خوبی داشته و بتواند به خوبی عمل شناسایی را انجام دهد. روش ارائه‌شده از یک رویکرد خوشه‌بندی ساده استفاده می‌کند تا به صورت سلسله‌مراتبی تصاویر را به دسته‌های کوچک‌تر تقسیم کند. دقت نهایی در روش ارائه‌شده این پایان‌نامه به دو عامل دقت خوشه‌بندی و کیفیت ویژگی‌های استخراج‌شده در هر مرحله وابسته است.

برای خوشه‌بندی، تصاویر به دو قسمت بالا و پایین تقسیم‌شده و از هر قسمت به‌طور جداگانه با استفاده از روش تجزیه ماتریس نامنفی بردارهای ویژگی استخراج می‌شوند. سپس اختلاف میانگین دو بردار ویژگی محاسبه می‌شود. در نهایت با استفاده از یک حد آستانه تصاویر به دو خوشه تقسیم می‌شوند. داخل هر خوشه از دو روش تجزیه ماتریس نامنفی و توصیفگرهای FREAK برای استخراج ویژگی استفاده می‌شود. روش تجزیه ماتریس نامنفی به صورت سراسری و توصیفگرهای FREAK به صورت محلی استخراج ویژگی می‌کنند. با ترکیب این دو ویژگی می‌توان به‌طور همزمان از خصوصیات محلی و سراسری تصویر بهره برد. برای شناسایی هویت تصویر از روش نزدیک‌ترین همسایه با معیار فاصله بلوک شهری استفاده شده است.

برای ارزیابی روش ارائه‌شده روی پایگاه داده حجیم FERET و همچنین یک مجموعه تصویر جمع‌آوری‌شده که حجم بزرگتری دارد، آزمایش‌هایی انجام گرفت که در نهایت دقت شناسایی $98/36\%$ روی مجموعه تصاویر FERET و $94/54\%$ روی مجموعه تصاویر جمع‌آوری‌شده به دست آمد که حدود ۲ درصد نسبت به بهترین کار قبلی بهبود داشت. نتایج آزمایش‌ها نشان‌دهنده عملکرد مطلوب روش خوشه‌بندی استفاده‌شده و روش شناسایی ارائه‌شده نسبت به روش‌های متداول پیشین است.

کلمات کلیدی: شناسایی چهره، تجزیه ماتریس نامنفی، پایگاه داده حجیم، خوشه‌بندی، توصیفگر FREAK، نزدیک‌ترین همسایه

لیست مقالات مستخرج از پایان نامه

- Abbaspoor. N., Hassanpour. H., "Face Recognition in a large dataset using a hierarchical classifier", *Multimedia Tools and Applications*. (in progress)

فهرست مطالب

ف	فهرست تصاویر
ق	فهرست جداول
۱	۱ سیستم‌های شناسایی چهره
۱	۱.۱ مقدمه
۱	۲.۱ شناسایی چهره و کاربردها
۲	۳.۱ چالش‌ها
۳	۴.۱ روش‌ها
۴	۵.۱ هدف و نوآوری
۶	۶.۱ روش حل
۶	۷.۱ پیش‌فرض‌های مسئله
۷	۸.۱ پایگاه داده
۸	۹.۱ جمع‌بندی و ساختار پایان‌نامه
۹	۲ مروری بر کارهای پیشین
۹	۱.۲ مقدمه
۹	۲.۲ روش‌های محلی
۱۰	۱.۲.۲ روش‌های مبتنی بر ظاهر محلی
۱۳	۲.۲.۲ روش‌های مبتنی بر نقاط کلیدی
۱۴	۳.۲ روش‌های سراسری
۱۴	۱.۳.۲ روش‌های خطی
۲۱	۲.۳.۲ روش‌های غیرخطی
۲۲	۴.۲ روش‌های ترکیبی
۲۵	۵.۲ جمع‌بندی

۲۷	۳	روش پیشنهادی
۲۷	۱.۳	مقدمه
۲۸	۲.۳	پیش پردازش
۲۸	۱.۲.۳	تشخیص چهره
۳۱	۲.۲.۳	بهبود تصاویر
۳۲	۳.۳	خوشه بندی
۳۲	۱.۳.۳	استخراج ویژگی
۳۳	۲.۳.۳	دسته بندی تصاویر
۳۴	۴.۳	شناسایی چهره
۳۵	۱.۴.۳	استخراج ویژگی
۳۹	۲.۴.۳	تطابق تصاویر
۴۰	۵.۳	جمع بندی
۴۱	۴	ارزیابی و بررسی نتایج
۴۱	۱.۴	مقدمه
۴۱	۲.۴	پایگاه داده
۴۲	۱.۲.۴	پایگاه داده FERET
۴۲	۲.۲.۴	مجموعه تصاویر جمع آوری شده
۴۲	۳.۴	تأثیر روش تشخیص چهره
۴۳	۴.۴	تأثیر پیش پردازش
۴۳	۵.۴	بررسی عملکرد روش خوشه بندی ارائه شده
۴۷	۶.۴	اعمال خوشه بندی به صورت سلسله مراتبی
۴۹	۷.۴	تأثیر حجم پایگاه داده
۴۹	۸.۴	پیدا کردن مرتبه در روش NMF
۵۰	۹.۴	تأثیر رزولوشن تصاویر ورودی
۵۱	۱۰.۴	تأثیر معیار فاصله
۵۳	۱۱.۴	تأثیر جداگانه نوآوری ها
۵۵	۱۲.۴	جمع بندی
۵۷	۵	جمع بندی و کارهای آینده
۵۷	۱.۵	خلاصه و نتیجه گیری
۵۸	۲.۵	کارهای آینده
۶۱		مراجع

فهرست تصاویر

۳	مراحل شناسایی چهره	۱.۱
۷	راه حل کلی پیشنهادی	۲.۱
۱۰	روش های شناسایی چهره	۱.۲
۱۱	الگوی دودویی محلی	۲.۲
۱۲	الگوی دودویی سه قلوئی پنج تایی	۳.۲
۱۵	چهره های ویژه استخراج شده	۴.۲
۱۷	مؤلفه های NMF از روی تصاویر چهره	۵.۲
۱۹	فلوچارت روش نیکان [۴۴]	۶.۲
۲۰	فیلتر گابور در ۸ جهت و ۵ اندازه	۷.۲
۲۱	فلوچارت روش بخشی [۴۶]	۸.۲
۲۲	فلوچارت روش بیگری [۱۸]	۹.۲
۲۴	فلوچارت روش آنالاکشمی [۳]	۱۰.۲
۲۹	فلوچارت روش پیشنهادی	۱.۳
۳۰	تصویر انتگرال	۲.۳
۳۳	چهار مرحله اصلی قبل از خوشه بندی	۳.۳
		تشخیص نوع تکه ای از تصویر با استفاده از مقادیر ویژه ماتریس گرادیان	۴.۳
۳۷	کوواریانس	
۳۸	شبکیه چشم	۵.۳
۳۸	مدل نمونه برداری روش FREAK الهام گرفته از توپولوژی شبکیه چشم	۶.۳
		روش پیشنهادی با استفاده از تصاویر پایگاه داده استاندارد FERET و همچنین	۱.۴
۴۳	یک مجموعه تصویر جمع آوری شده از چندین پایگاه داده ارزیابی شده است.	
۴۴	بهسازی تصویر با استفاده از برابری هیستوگرام	۲.۴
۴۴	بهسازی تصویر با استفاده از برابری هیستوگرام وفقی با محدودیت کنتراست	۳.۴
۴۴	بهسازی تصویر با استفاده از تنظیم شدت نور	۴.۴

۵.۴	تأثیر پیش‌پردازش‌ها روی عملکرد روش پیشنهادی روی دو پایگاه داده
۴۵	FERET و مجموعه تصاویر جمع‌آوری‌شده
۶.۴	عملکرد روش خوشه‌بندی پیشنهادی در مقایسه با روش‌های متداول در
۴۶	پایگاه داده FERET
۷.۴	عملکرد روش خوشه‌بندی پیشنهادی در مقایسه با روش‌های متداول در
۴۷	مجموعه تصاویر جمع‌آوری‌شده
۴۸	تأثیر تعداد دفعات خوشه‌بندی در پایگاه داده FERET
۴۸	تأثیر تعداد دفعات خوشه‌بندی در مجموعه تصاویر جمع‌آوری‌شده
۱۰.۴	تأثیر اندازه پایگاه داده روی دقت نهایی روش پیشنهادی در دو مجموعه
۴۹	تصاویر استفاده‌شده
۱۱.۴	تأثیر مرتبه (r) روش NMF و تعداد مؤلفه‌های استخراجی روی دقت نهایی
۵۰	روش پیشنهادی در پایگاه داده FERET
۱۲.۴	تأثیر اندازه تصاویر در استخراج ویژگی NMF و دقت نهایی روی پایگاه داده
۵۲	FERET
۱۳.۴	تأثیر اندازه تصاویر در استخراج توصیفگرهای FREAK و دقت نهایی روی
۵۲	پایگاه داده FERET
۱۴.۴	تأثیر جداگانه بخش‌های مختلف روش ارائه‌شده روی پایگاه داده FERET
	و مجموعه تصاویر جمع‌آوری‌شده. حروف F، N و C به ترتیب نشان‌دهنده
۵۳	ویژگی‌های NMF، توصیفگرهای FREAK و مبحث خوشه‌بندی هستند.

فهرست جداول

۱.۲	نمونه روش‌های شناسایی چهره، تکنیک‌های استفاده‌شده به همراه نقاط قوت و ضعف	۲۵
۱.۴	انواع تصاویر پایگاه داده FERET	۴۲
۲.۴	مقایسه عملکرد دو روش تشخیص چهره	۴۳
۳.۴	تأثیر معیار فاصله در عملکرد روش ارائه‌شده	۵۳
۴.۴	مقایسه عملکرد روش پیشنهادی با کارهای مرتبط	۵۶

فصل ۱

سیستم‌های شناسایی چهره

۱.۱ مقدمه

در دهه‌های اخیر استفاده از روش‌های بیومتریک برای احراز هویت به‌جای استفاده از روش‌های غیر بیومتریک، رو به رشد بوده است. متداول‌ترین روش غیر بیومتریک، استفاده از کلمات عبور^۱ است، که با افزایش استفاده از روش‌های بیومتریک دیگر نیازی به حفظ کردن چندین کلمه عبور متفاوت نیست و برای امر احراز هویت، حضور شخص کافی است. از روش‌های بیومتریک می‌توان به استفاده از عنبیه، اثرانگشت، صوت و چهره اشاره کرد [۱]. در این فصل در رابطه با عامل بیومتریک چهره، سیستم‌های شناسایی چهره، چالش‌های این سیستم‌ها، روش‌های حل و روش پیشنهادی مطالبی را ارائه می‌کنیم.

۲.۱ شناسایی چهره و کاربردها

از میان روش‌های بیومتریک موجود برای احراز هویت، چهره از محبوبیت بیشتری برخوردار است. دلیل این امر سادگی استفاده از عامل چهره و حتی امکان استفاده از فاصله دور و به‌طور پنهانی است [۲]. نظارت ویدیویی، شناسایی مجرم، کنترل دسترسی به ساختمان‌ها، و وسایل نقلیه بدون سرنشین تنها بخشی کوچک از کاربردهای سیستم‌های شناسایی چهره هستند.

^۱passwords

شناسایی چهره به‌طور کلی در دو دسته از وظایف استفاده می‌شود:

۱. تأیید هویت^۱ (تطابق یک به یک): با در دست داشتن تصویر یک چهره، مشخص می‌شود هویتی که شخص ادعا می‌کند مورد تأیید است یا خیر. به عبارتی دیگر، سیستم تأیید هویت باید سؤال "آیا این شخص همان کسی است که ادعا می‌کند؟" را پاسخ دهد.

۲. شناسایی هویت^۲ (تطابق یک به چند): با در دست داشتن یک تصویر چهره مجهول‌الهویه، باید هویت چهره مورد نظر شناسایی و مشخص شود. به عبارتی دیگر، سیستم شناسایی هویت باید سؤال "این تصویر چهره‌ی چه کسی است؟" را پاسخ بدهد.

۳.۱ چالش‌ها

سیستم‌های شناسایی چهره با وجود پیشرفت‌های زیادی که در سالیان اخیر داشتند، همچنان با چالش‌هایی مانند انسداد، تغییر زاویه، تغییرات روشنایی و بقیه عوامل روبرو هستند. به‌عنوان مثال در فضاهایی که از دوربین‌های CCTV برای نظارت استفاده می‌شود، معمولاً تصاویر بریده‌شده اشخاص دارای خرابی‌هایی مانند تاری و نویز می‌باشد [۳]. همچنین به دلیل فاصله شخص با دوربین، ممکن است تصویر چهره استخراج‌شده دارای وضوح مناسب نباشد [۴]. در این مسئله با چالش شناسایی چهره در شرایط کنترل‌نشده و با استفاده از تصاویر وضوح پایین، روبرو هستیم [۵]. به‌طور کلی چالش‌های کلیدی سیستم‌های شناسایی چهره عبارت‌اند از [۶]:

۱. تنوع زیاد در ظاهر صورت: هرچند شکل و ظاهر مشخصه‌های ذاتی چهره هستند، ولی شهود (بافت^۳) یک چهره می‌تواند تحت تأثیر عواملی مانند زاویه صورت^۴، تغییرات روشنایی^۵ و احساسات^۶ قرار بگیرد. همچنین افزایش سن^۷ و عوامل سخت‌افزاری مربوط به دوربین شامل دیافراگم، زمان باز بودن لنز، نوع لنز و نوع حسگر می‌تواند تصویر چهره را تغییر دهد. تغییرات در ظاهر چهره به دلایل تغییر زاویه و روشنایی خیلی متنوع‌تر از تغییرات ظاهر به دلیل تغییر هویت است. [۷]

۲. خمینه‌های^۸ غیرخطی پیچیده: خمینه‌های تصاویر چهره به‌شدت غیرخطی هستند. همین امر باعث عدم کارایی مناسب روش‌های خطی شناسایی چهره می‌شود. روش‌های خطی، فضای با ابعاد بالای تصویر را به‌طور خطی به فضای با ابعاد کم نگاشت می‌کنند.

¹verification

²identification

³texture

⁴facial pose

⁵illumination

⁶expression

⁷aging

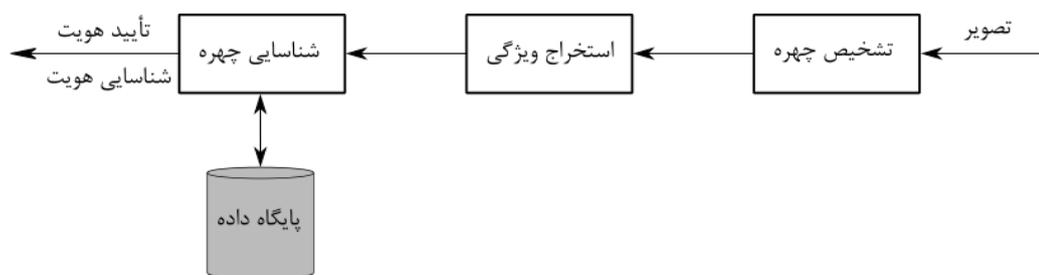
⁸manifolds

در این فضای خطی، برخی از معیارهای فاصله نظیر معیار اقلیدسی و Mahalanobis نمی‌توانند به خوبی تمایز بین داده‌ها (هویت‌های مختلف) را مشخص کنند. این ضعف، عملکرد روش‌های خطی را محدود می‌کند.

۳. ابعاد بالا و تعداد نمونه کم: معمولاً تعداد نمونه‌ها در پایگاه داده خیلی کمتر از ابعاد فضای تصویر است. در این حالت، سیستم آموزش دیده، به خوبی نمی‌تواند نسبت به نمونه‌های مشاهده نشده تعمیم یابد.

۴.۱ روش‌ها

برای پیاده‌سازی یک سیستم شناسایی چهره کارا، نیاز به سه مرحله اساسی است (شکل ۱.۱):
(۱) تشخیص چهره^۱، (۲) استخراج ویژگی، و (۳) شناسایی چهره^۲ [۸]



شکل ۱.۱: مراحل شناسایی چهره

۱. تشخیص چهره: این مرحله به منظور تشخیص و مکان‌یابی چهره اشخاص در تصویر صورت می‌گیرد. تغییرات روشنایی و حالات احساسی صورت ممکن است مانع از تشخیص درست چهره شود [۹]. برای بهبود عملکرد این مرحله از پیش‌پردازش‌هایی استفاده می‌شود.

۲. استخراج ویژگی: پس از استخراج چهره در تصویر و حذف نواحی غیر چهره، بایستی ویژگی‌های معناداری از چهره استخراج شود. می‌توان از این ویژگی‌ها به‌عنوان امضای هر تصویر نام برد. کیفیت ویژگی‌های استخراج‌شده در این مرحله، تأثیر بسزایی در عملکرد مرحله بعدی (شناسایی چهره) خواهد داشت.

۳. شناسایی چهره: با در دست داشتن ویژگی‌های هر چهره، می‌توان با مقایسه تصاویر مختلف، هویت آنها را شناسایی کرد. در این مرحله استفاده از یک معیار فاصله مناسب نقش مهمی در عملکرد سیستم شناسایی دارد.

¹ face detection

² face recognition

با توجه به روش‌های استخراج ویژگی، در حالت کلی سیستم‌های شناسایی چهره به سه دسته اصلی تقسیم می‌شوند [۲]:

۱. روش‌های محلی^۱: این روش‌ها از ویژگی‌های خاص صورت که به صورت محلی استخراج می‌شوند، استفاده می‌کنند. به عبارتی ویژگی‌های کلی تصویر را نادیده گرفته و در مقابل الگوهای موجود در نواحی کوچک‌تر از تصویر را به عنوان ویژگی در نظر می‌گیرند. این روش‌ها به دو دسته مبتنی بر ظاهر^۲ (مانند LBP [۱۰] و HOG [۱۱]) و مبتنی بر نقاط کلیدی^۳ (مانند SIFT [۱۲] و SURF [۱۳]) تقسیم می‌شوند.

۲. روش‌های سراسری (جامع)^۴: این دسته از روش‌ها، کل تصویر را به عنوان ورودی در نظر گرفته، و سپس آن را به زیرفضای کوچک‌تر نگاشت می‌دهند. این روش‌ها به دو دسته خطی^۵ (مانند PCA [۱۴]، LDA [۱۵] و ICA [۱۶]) و غیرخطی^۶ (مانند KPCA [۱۷]) تقسیم می‌شوند.

۳. روش‌های ترکیبی^۷: این روش‌ها با ترکیب روش‌های محلی و سراسری (مانند Gabor-LBP-LDA [۱۸] و DCT-PCA-GA [۱۹])، از مزیت‌های هر دو روش برای بهبود عملکرد سیستم شناسایی استفاده می‌کنند [۲۰].

۵.۱ هدف و نوآوری

برای شناسایی چهره در شرایط کنترل نشده و دنیای واقعی، علاوه بر چالش‌های مطرح شده، یک چالش اساسی دیگر تحت عنوان اندازه پایگاه داده وجود دارد [۲۱]. شناسایی چهره در پایگاه داده‌های حجیم (تعداد تصاویر زیاد) به عنوان یک مسئله بهینه‌سازی چندمتغیره مطرح می‌شود [۶]. یک سیستم بهینه بایستی:

- تعداد مشخصی جستجو
- طی مدت زمانی مشخص
- روی اندازه مشخصی از پایگاه داده
- روی سخت‌افزار مشخص
- و سطح دقت مشخص

¹local approach

²appearance-based

³keypoint-based

⁴holistic approach

⁵linear

⁶non-linear

⁷hybrid approach

عمل شناسایی چهره را انجام دهد. به طور کلی تعریف مسئله در هر کاربردی مشخص می‌کند که کدامیک از شروط ذکرشده متغیر و کدامیک بایستی ثابت در نظر گرفته شوند. هدف ما در اینجا طراحی سیستم شناسایی چهره‌ایست که دقت موجهی داشته و با بالا رفتن اندازه پایگاه داده نیز عملکرد آن افت چشمگیری نداشته باشد.

روش‌هایی که در رابطه با پایگاه داده‌های حجیم ارائه شده‌اند به طور کلی در این سه دسته قرار می‌گیرند:

- روش‌های مستقیم که با ارائه یک روش استخراج ویژگی مناسب این موضوع، تصاویر را به فضایی منتقل می‌کنند که تصاویر با حجم بالا می‌توانند با فاصله‌های مناسب از همدیگر قرار بگیرند.
- روش‌هایی که از یک رویکرد خوشه‌بندی به عنوان پیش‌دسته‌بندی استفاده می‌کنند. این روش‌ها ابتدا تصاویر مشابه را به خوشه‌هایی تقسیم می‌کنند. سپس عمل شناسایی چهره داخل خوشه موردنظر انجام می‌شود.
- روش‌های یادگیری عمیق که اخیراً با پیشرفت‌های زیادی همراه بوده‌اند پتانسیل یادگیری میلیون‌ها تصویر را دارا هستند [۲۲، ۲۳]. البته این روش‌ها برای یادگیری نیاز به تعداد نمونه بالا دارند [۲۴] و از آنجایی که تمرکز این تحقیق بر روی شرایطی است که فقط یک نمونه به ازای هر شخص وجود دارد، نمی‌توان از روش‌های معمول یادگیری عمیق برای این موضوع استفاده کرد. هرچند روش‌های یادگیری عمیقی در این زمینه ارائه شده‌اند که به طور مختصر در فصل‌های بعدی مورد مطالعه و درنهایت با روش پیشنهادی مورد مقایسه قرار خواهند گرفت. از طرفی دیگر آموزش شبکه‌های عمیق نیازمند زمان زیادی بوده و به دلیل نیاز به داده‌های آموزشی زیاد همواره با خطر بیش‌برازش^۱ همراه هستند. در این پایان‌نامه در خصوص چالش‌های مطرح‌شده چندین نوآوری ارائه شده است که عبارت‌اند از:

- **رویکرد شناسایی چهره مقاوم در برابر پایگاه داده‌های حجیم:** روش پیشنهادی، با استفاده از رویکرد خوشه‌بندی و ترکیب دو ویژگی NMF و FREAK طوری عمل می‌کند که قدرت شناسایی بالایی در مواجهه با تعداد تصاویر زیاد دارد که حاکی از ظرفیت بالای این روش در پایگاه داده‌های حجیم است.

- **استفاده از تصاویر با رزولوشن پایین:** روش پیشنهادی برای آموزش نیازی به تصاویر با کیفیت خیلی بالا ندارد. تصاویر با حداقل اندازه 124×124 پیکسل می‌توانند مورد استفاده قرار گیرند. نتایج آزمایش‌های انجام‌شده نشان‌دهنده عملکرد مناسب روی این دسته از تصاویر است.

¹Overfitting

● استفاده از تعداد نمونه‌های کم در پایگاه داده: روش پیشنهادی طوری طراحی شده است که بتواند با تعداد نمونه کم آموزشی هم شناسایی چهره را انجام دهد. حتی با یک تصویر به ازای هر شخص این روش می‌تواند آموزش ببیند. هرچند در روش‌های دیگر ممکن است وجود تعداد کمی نمونه به‌ازای هر شخص باعث افت عملکرد سیستم و یا حتی شکست آن شود، ولی در روش پیشنهادی عدم وجود نمونه‌های بیشتر مشکل‌ساز نیست.

۶.۱ روش حل

نوآوری‌های مطرح‌شده در چارچوب یک سیستم شناسایی چهره در این پایان‌نامه ارائه خواهد شد. چارچوب کلی در شکل ۲.۱ آمده است. سیستم مذکور، ابتدا چهره را در تصاویر استخراج می‌کند. سپس پیش‌پردازش‌هایی به‌منظور بهسازی تصاویر از نظر روشنایی و کنتراست روی تصاویر چهره صورت می‌گیرد. از تصاویر چهره بهسازی شده با استفاده از روش تجزیه ماتریس نامنفی^۱ و یا به‌اختصار NMF [۲۵]، استخراج ویژگی انجام می‌شود. سپس با استفاده از یک رویکرد خوشه‌بندی ساده تصاویر به دو خوشه تقسیم می‌شوند. این خوشه‌بندی می‌تواند به‌صورت سلسله‌مراتبی انجام شده و یک درخت دودویی از خوشه‌ها تشکیل بدهد. درنهایت برای شناسایی چهره داخل هر خوشه، دو دسته ویژگی NMF و توصیفگرهای FREAK [۲۶] استخراج می‌شوند. سپس به‌ازای هر تصویر آزمون، ابتدا به‌صورت سلسله‌مراتبی خوشه‌های آن مشخص گردیده، سپس در همان خوشه فاصله ویژگی‌های تصویر آزمون با تصاویر داخل خوشه محاسبه می‌شود. هویت تصویر آزمون به‌وسیله روش نزدیک‌ترین همسایه و معیار فاصله بلوک شهری مشخص می‌شود.

این روش در مرحله خوشه‌بندی، تصاویر را با دقت بالا به دو خوشه تقسیم می‌کند. سپس به دلیل این‌که در هر مرحله (خوشه) با تعداد کمی از تصاویر روبرو است، تعداد مقایسه‌ها کم شده و می‌تواند با دقت مناسبی عمل شناسایی را انجام دهد. لذا ظرفیت بالایی در مواجهه با پایگاه داده‌های حجیم دارد. از طرفی سیستم ارائه‌شده می‌تواند به‌طور خودکار درخت خوشه‌ها را تشکیل داده و تا حد امکان تعداد تصاویر هر خوشه را به حداقل برساند. درنهایت با اجرای این روش چندین خوشه با اندازه‌های برابر تشکیل خواهد شد.

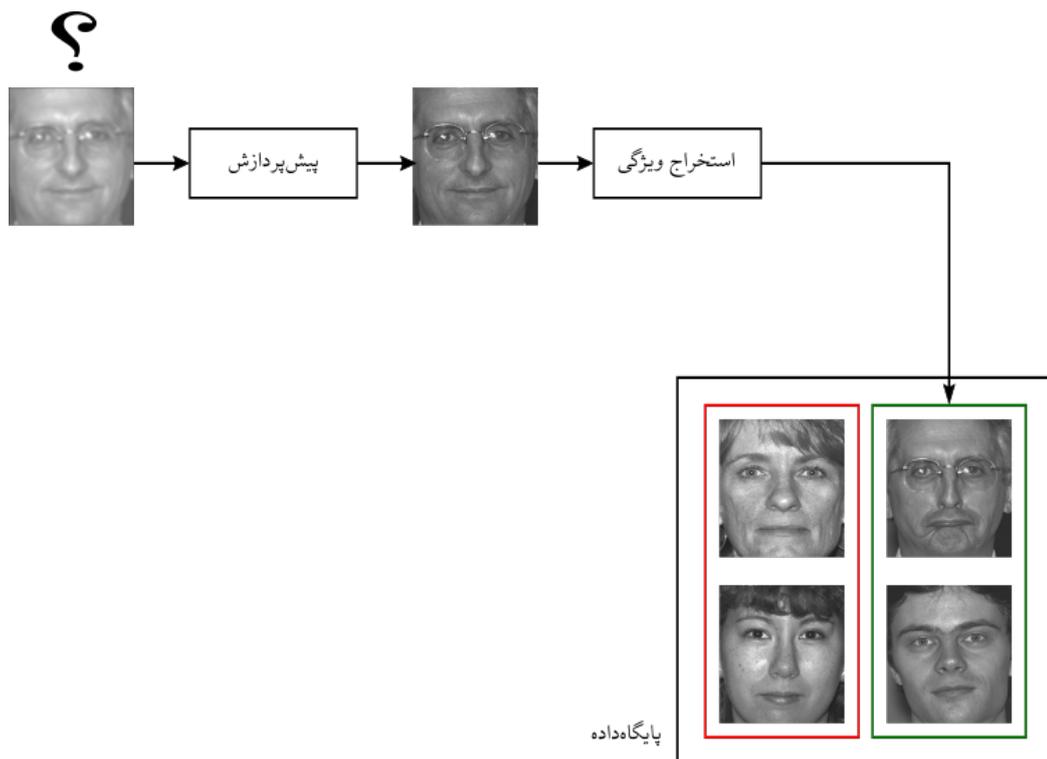
۷.۱ پیش‌فرض‌های مسئله

پیش‌فرض‌هایی که در طراحی این سیستم شناسایی در نظر گرفته شده است، عبارتند از:

- تصاویر چهره استفاده‌شده در دو مرحله آموزش و آزمون، خاکستری^۲ هستند.

^۱Non-negative Matrix Factorization

^۲grayscale



شکل ۲.۱: راه حل کلی پیشنهادی

- زاویه صورت در تصاویر چهره همگی تمام رخ و کامل هستند.
- اشخاص شامل دو جنسیت زن و مرد از ملیت‌های گوناگون بوده و چهره اشخاص ممکن است حاوی عینک، سبیل، ریش، و گوشواره باشد.
- حداقل تعداد نمونه به ازای هر شخص، یک تصویر است.
- اندازه تصاویر ورودی حداقل 124×124 پیکسل است

۸.۱ پایگاه داده

پایگاه داده اصلی استفاده شده در این تحقیق، FERET^۱ [۲۷] نام دارد. این پایگاه داده حاوی تصاویر چهره از ۱۱۹۹ شخص می‌باشد که در زوایای مختلف گرفته شده‌اند. بعضی از تصاویر نیز به فاصله یک سال و یا بیشتر از جلسات قبلی گرفته شده‌اند. برای ارزیابی روش پیشنهادی در این پایان‌نامه از تصاویر ۹۹۰ شخص استفاده شده است. مجموعه تصاویر بزرگ‌تری نیز با استفاده از تصاویر چهره از زاویه روبرو چندین پایگاه داده استاندارد (FERET، MUCT، PICS، FEI، Face94) جمع‌آوری شده که برای ارزیابی روش پیشنهادی مورد استفاده قرار گرفته است.

^۱face recognition technology

در این پایگاه داده تصاویر ۱۶۸۴ شخص موجود است. اطلاعات دقیق‌تر در رابطه با پایگاه داده‌ها به‌طور مفصل در فصل ۴ شرح داده شده است.

۹.۱ جمع‌بندی و ساختار پایان‌نامه

این پایان‌نامه در پنج فصل تدوین شده است. در این فصل با سیستم‌های شناسایی چهره آشنا شدیم. هدف و کاربردهای این سیستم‌ها مطرح شد. چالش‌ها و روش‌های کلی شناسایی چهره و همچنین روش پیشنهادی ما برای حل برخی از این چالش‌ها شرح داده شد. در نهایت ساختار اصلی پایان‌نامه ارائه گردید. در فصل بعدی چندین روش شناسایی چهره مطرح به‌طور مفصل‌تر توضیح داده می‌شود. در فصل سوم روش پیشنهادی این پایان‌نامه، به‌طور کامل ارائه خواهد شد. در فصل چهارم به بررسی آزمایشات انجام‌شده و ارزیابی روش پیشنهادی خواهیم پرداخت. در نهایت در فصل پنجم نتیجه‌گیری نهایی و پیشنهادهایی برای ادامه فعالیت، مطرح خواهد شد.

فصل ۲

مروری بر کارهای پیشین

۱.۲ مقدمه

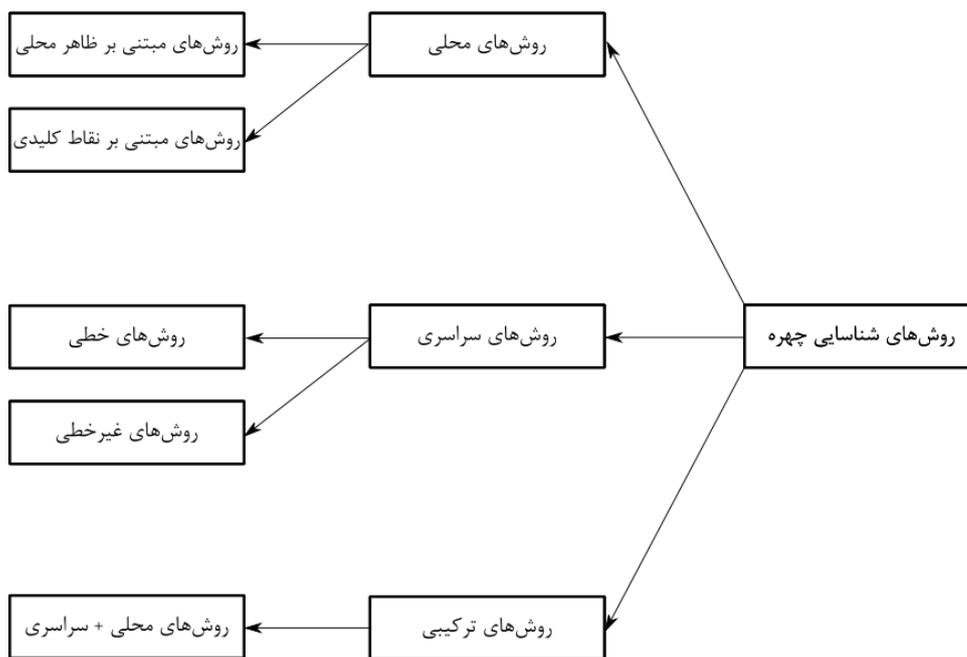
با گسترش استفاده از سیستم‌های شناسایی چهره برای امر احراز هویت، و بروز چالش‌های این مسئله که در بخش ۳.۱ مطرح شد، در طی سالیان متمادی روش‌های متعددی برای حل هرکدام از این چالش‌ها ارائه شده است. همان‌طور که اشاره شد، این روش‌ها به سه دسته کلی روش‌های محلی، روش‌های سراسری (جامع) و روش‌های ترکیبی تقسیم می‌شوند (شکل ۱.۲) [۲]. در این فصل قصد داریم این سه دسته را توضیح داده و همچنین به بررسی چند روش از هر دسته و به‌خصوص مرتبط با هدف این تحقیق بپردازیم.

۲.۲ روش‌های محلی

در سیستم‌های شناسایی چهره که به‌طور محلی عمل می‌کنند، فقط برخی از ویژگی‌های صورت در نظر گرفته می‌شود. این روش‌ها به‌شدت حساس به حالت چهره، انسداد و ژست هستند. به‌طور کلی این روش‌ها به دو دسته تقسیم می‌شوند:

۱. مبتنی بر ظاهر محلی^۱: در این روش‌ها تصویر به نواحی کوچک‌تر تقسیم شده و از هر ناحیه ویژگی‌های محلی استخراج می‌شود.

^۱local-appearance-based



شکل ۱.۲: روش‌های شناسایی چهره

۲. مبتنی بر نقاط کلیدی^۱: در این روش‌ها ابتدا نقاط مورد علاقه کلیدی در تصویر مشخص، سپس ویژگی‌های محلی در این نقاط استخراج می‌شود.

۱.۲.۲ روش‌های مبتنی بر ظاهر محلی

این روش‌ها روی نقاط اساسی صورت همچون بینی، دهان، و چشم‌ها تمرکز دارند تا بتوانند جزئیات بیشتری تولید کنند. به علاوه، این روش‌ها ویژگی‌های محلی را از طریق جهت پیکسل‌ها، هیستوگرام‌ها و مشخصه‌های هندسی توصیف می‌کنند.

- الگوی دودویی محلی (LBP)^۲ [۱۰] و گونه‌های مختلفش: الگوی دودویی محلی یک روش عالی برای استخراج ویژگی بافتی از هر شیئی است [۲۸]. این روش به‌طور گسترده در شناسایی چهره، شناسایی حالت صورت، قطعه‌بندی بافت، و دسته‌بندی بافت استفاده شده و عملکرد خوبی از خود نشان داده است. این روش ابتدا تصاویر را به چند ناحیه تقسیم می‌کند. در هر ناحیه مربع‌های 3×3 تشکیل می‌شود. پیکسل‌های هر مربع به‌صورت ساعت شمار به‌عنوان $p^1 \dots p^8$ نام‌گذاری می‌شوند. پیکسل مرکزی به‌عنوان p^0 و حد آستانه آن مربع در نظر گرفته می‌شود. اگر همسایه‌ای از حد آستانه بزرگ‌تر باشد، مقدار ۱ و اگر کوچک‌تر باشد مقدار ۰ می‌گیرد. بدین‌صورت کد دودویی در هر مربع ایجاد می‌شود. این کد حاوی اطلاعات بافتی هر مربع است. در نهایت در هر ناحیه

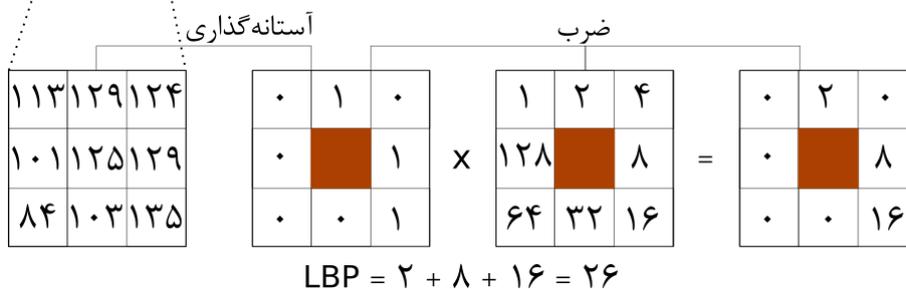
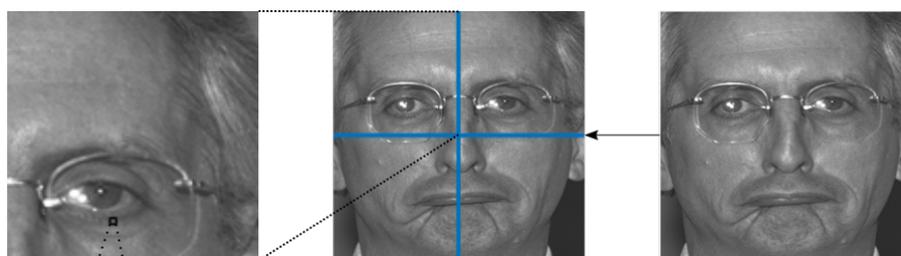
^۱key-points-based

^۲Local Binary Pattern (LBP)

از تصویر هیستوگرامی از این کدها تشکیل می‌شود و همه هیستوگرام‌ها را باهم الحاق کرده، بردار ویژگی نهایی حاصل می‌شود. الگوی دودویی محلی تعریف‌شده در هر یک از مربع‌های 3×3 در رابطه ۱.۲ نشان داده شده است.

$$LBP = \sum_{p=1}^8 2^p s(i_o - i_p), \text{ with } s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (1.2)$$

که i_o و i_p به ترتیب مقدار پیکسل مرکزی و پیکسل همسایه می‌باشد (شکل ۲.۲).



شکل ۲.۲: الگوی دودویی محلی

برایان و همکاران [۲۹] از این روش برای شناسایی چهره بر اساس تصاویر ویدیویی دوربین‌های مداربسته استفاده کرده‌اند.

- هیستوگرام گرادیان‌های جهت‌دار^۱ HOG [۱۱]: یکی از بهترین توصیفگرها برای توصیف شکل و لبه است. روش HOG می‌تواند شکل صورت را با استفاده از توزیع جهات لبه‌ها و یا گرادیان شدت روشنایی توصیف کند. در این روش ابتدا تصویر به مناطق کوچک‌تر تقسیم می‌شود که این مناطق را سلول^۲ می‌نامند. سپس در هر سلول دامنه گرادیان مرتبه اول در دو جهت عمودی و افقی محاسبه می‌شود. معمول‌ترین راه، اعمال یک

^۱Histogram of Oriented Gradients (HOG)

^۲cell

ماسک یک بعدی $[-1, 0, 1]$ است.

$$G_x(x, y) = I(x + 1, y) - I(x - 1, y), \quad (2.2)$$

$$G_y(x, y) = I(x, y + 1) - I(x, y - 1), \quad (3.2)$$

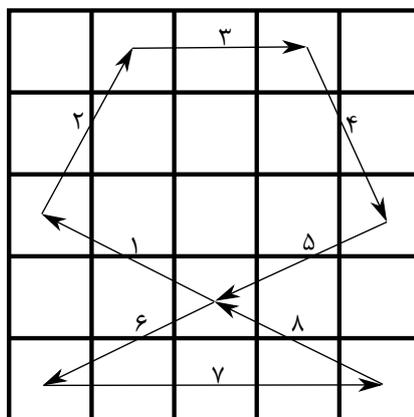
که $I(x, y)$ مقدار پیکسل در نقطه (x, y) و G_x و G_y دامنه گرادیان افقی و عمودی را نشان می‌دهند. اندازه گرادیان و جهت هر پیکسل با استفاده از معادلات ۴.۲ و ۵.۲ محاسبه می‌شوند:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \quad (4.2)$$

$$\theta(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right), \quad (5.2)$$

هیستوگرام اندازه گرادیان‌ها و جهت پیکسل‌های هر سلول تشکیل می‌شود. در نهایت هیستوگرام کل سلول‌ها باهم ترکیب شده تا ویژگی‌های کل تصویر را بسازند [۳۰].

- تانسر و همکاران [۲۰] در یک روش شناسایی چهره جدید از یک الگوی محلی متفاوت نسبت به LBP استفاده کرده‌اند. این روش الگوی دودویی سه‌قلوی پنج‌تایی^۱ نام دارد. این الگو در شکل ۳.۲ نشان داده شده است. برای استخراج این ویژگی از تصویر، ابتدا بلوک‌های 5×5 تشکیل می‌شود، سپس در هر بلوک، این الگو اعمال شده و در نهایت یک هیستوگرام تشکیل می‌شود.



شکل ۳.۲: الگوی دودویی سه‌قلوی پنج‌تایی

¹quintet-triplet binary pattern

۲.۲.۲ روش‌های مبتنی بر نقاط کلیدی

این روش‌ها برای تشخیص ویژگی‌های هندسی خاص بر اساس برخی اطلاعات هندسی سطح صورت، مانند فاصله چشم‌ها، عرض سر، استفاده می‌شوند. این روش‌ها از دو گام اصلی تشکیل می‌شوند: تشخیص نقاط کلیدی، استخراج ویژگی. گام اول روی عملکرد شناساگرهای نقاط کلیدی، و گام دوم روی نمایش اطلاعاتی که این نقاط کلیدی با خود دارند، تمرکز دارد.

- تبدیل ویژگی مستقل از مقیاس (SIFT)^۱ [۱۲]: الگوریتمی است که برای تشخیص و توصیف ویژگی‌های محلی تصویر استفاده می‌شود. ایده اصلی SIFT تبدیل تصویر به نمایشی از ترکیب نقاط موردعلاقه در تصویر است. این نقاط حاوی اطلاعاتی از چهره هستند. SIFT به دلیل سرعت بالا در کاربردهای بلادرنگ به‌وفور استفاده می‌شود، ولی یکی از نقاط ضعف این روش زمان‌بر بودن تطابق نقاط کلیدی در تصاویر است. چهار مرحله اصلی این روش به شرح زیر می‌باشد [۳۱]:

۱. پیدا کردن بیشینه و کمینه نقاط در فضای مقیاس

۲. مکان‌یابی نقاط کلیدی

۳. گرایش‌گماری

۴. تولید توصیفگر از نقاط بحرانی

- ویژگی‌های مقاوم تسریع‌شده (SURF)^۲ [۱۳]: این تکنیک از روش SIFT الهام گرفته است. با این تفاوت که از تبدیل موجک و تقریب دترمینان هسی [۳۲] برای بهبود عملکرد استفاده می‌کند. مزیت اصلی این روش، زمان پردازش آن است که کمتر از زمانی است که SIFT برای استخراج ویژگی استفاده می‌کند [۱۳].

آناند و همکاران در مقاله‌ای برای شناسایی چهره از ویژگی‌های SURF و دسته‌بند SVM استفاده کردند [۳۳]. در این روش پس از تشخیص چهره در تصویر، ویژگی‌های SURF شناسایی و استخراج می‌شوند. سپس به دلیل بالا بودن ابعاد ویژگی‌های استخراج‌شده از دسته‌بند SVM که عملکرد خوبی در داده‌های با ابعاد بالا دارد، برای مرحله تطابق استفاده کرده‌اند. این روش روی تصاویر پایگاه‌های داده UMIST و Yalefaces آزمایش شد که به دقت شناسایی ۹۷/۸۷٪ و ۹۷/۷۸٪ دست‌یافته‌است. این روش برخلاف عملکرد خوب شناسایی در کنار اندازه پایگاه داده کوچک از مشکل زمان اجرای زیاد و ویژگی‌های با ابعاد بالا رنج می‌برد.

^۱Scale Invariant Feature Transform (SIFT)

^۲Speeded-Up Robust Features (SURF)

۳.۲ روش‌های سراسری

روش‌های جامع (سراسری) که بعضاً با اسم زیرفضا^۱ مطرح می‌شوند، کل تصویر را به‌عنوان یک ماتریسی از پیکسل‌ها مورد پردازش قرار می‌دهند، و به نواحی کوچک‌تر در تصویر احتیاجی ندارند. پس از تبدیل ماتریس تصاویر به بردارهای ویژگی، می‌توان آن‌ها را به فضاهای با ابعاد پایین‌تر برد تا مورد بررسی قرار گیرند. این روش‌ها با توجه به طریقه نمایش زیرفضا، به دو دسته کلی خطی و غیرخطی تقسیم می‌شوند.

۱.۳.۲ روش‌های خطی

این روش‌ها به‌صورت خطی فضای تصویر را به فضایی با ابعاد کمتر نگاشت می‌دهند.

- چهره ویژه^۲ [۱۴] و تحلیل مؤلفه‌های اصلی (PCA)^۳: چهره ویژه یکی از محبوب‌ترین روش‌ها برای استخراج ویژگی به شمار می‌رود. این رویکرد که بر اساس روش PCA عمل می‌کند، مؤلفه‌های اصلی به‌دست‌آمده را به‌عنوان چهره ویژه یا قالب چهره در نظر می‌گیرد. PCA چندین متغیر که ممکن است با یکدیگر همبستگی داشته باشند را به تعداد کمتری متغیر نگاشت می‌دهد که مؤلفه‌های اصلی نامیده می‌شوند. هدف اصلی PCA کاهش ابعاد فضای داده است. PCA بردارهای ویژه ماتریس کوواریانس را محاسبه، سپس داده‌ها را به‌وسیله بردارهای ویژه با مقادیر ویژه بزرگ، به فضای با ابعاد کمتر نگاشت می‌دهد. PCA در شناسایی چهره بسیار استفاده شده است، و به آن بردارهای ویژه استخراج‌شده چهره ویژه نیز می‌گویند [۱۴].

تصویر X را با ابعاد $M \times N$ در نظر بگیرید. مجموعه آموزش به‌صورت $\{X_1, X_2, \dots, X_N\}$ نمایش داده می‌شود. چهره میانگین با استفاده از معادله ۶.۲ قابل محاسبه است:

$$X = \frac{1}{N} \sum_{i=1}^N x_i \quad (6.2)$$

سپس ماتریس کوواریانس نسبت به چهره میانگین محاسبه شده، بر اساس معادله ۷.۲ به‌دست می‌آید:

$$Q = \frac{1}{N} \sum_{i=1}^N (X - X_i)(X - X_i)^T \quad (7.2)$$

¹subspace approach

²eigenface

³Principle Component Analysis (PCA)

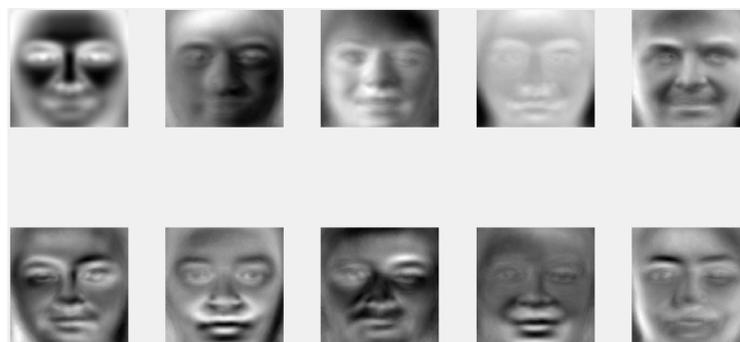
بردارهای ویژه و مقادیر ویژه مربوطه از طریق ۸.۲ محاسبه می‌شوند:

$$CV = \lambda V, (V \in R_n, V \neq 0) \quad (8.2)$$

که V ماتریس بردارهای ویژه با مقادیر ویژه λ می‌باشد. سپس همه تصاویر مربوط به شخص i ام به زیر فضای ویژه نگاشت داده می‌شوند:

$$y_k^i = w^T(x_i), (i = 1, 2, 3 \dots N), \quad (9.2)$$

که y_k^i نشان دهنده نگاشت x است و مؤلفه اصلی و چهره ویژه نیز نامیده می‌شود (شکل ۴.۲).



شکل ۴.۲: چهره‌های ویژه استخراج شده

- نظری و همکاران با استفاده از روش PCA سیستمی برای تشخیص جنسیت از روی تصاویر چهره ارائه کردند [۳۴]. در این روش تصاویر چهره با استفاده از PCA به فضای ویژگی با ابعاد کمتر منتقل شده، سپس با استفاده از خوشه‌بند فازی، به دو خوشه زن و مرد دسته‌بندی می‌شوند. برای این منظور تصاویر گالری پس از نگاشت به فضای ویژگی، میانگین دو دسته زن و مرد محاسبه و مراکز هر دسته مشخص می‌شود. سپس با استفاده از فرمول ۱۰.۲ عضویت تصاویر جدید به هر کدام از دسته‌های زن و مرد مشخص می‌شود.

$$U(i, j) = \frac{(d_{i,j}^2)^{\frac{-1}{m-1}}}{\sum_{k=1}^2 (d_{i,k}^2)^{\frac{-1}{m-1}}} \quad (10.2)$$

در رابطه فوق m پارامتر فازی‌سازی می‌باشد. این روش روی پایگاه داده‌های FG-Net، Stanford و FERET آزمایش شده است و به ترتیب به دقت‌های ۹۱/۸۹٪، ۹۵/۴٪ و ۹۰/۹٪ دست‌یافته است. عملکرد این روش با بالا رفتن تعداد تصاویر پایگاه داده، افت قابل توجهی دارد.

● آنالیز افتراقی خطی (LDA)^۱ [۱۵]: این روش که به نام تشخیص خطی فیشر هم شناخته می‌شود، همانند روش PCA به یافتن یک ترکیب خطی از خصوصیات که به بهترین شکل می‌توانند دو یا چند دسته را از هم جدا کنند، می‌پردازد. این روش نسبت بین واریانس میان دو دسته و واریانس داخل هر دسته را بیشینه می‌کند.

لو و همکاران [۳۵] برای حل مشکل پایگاه داده حجیم از یک روش خوشه‌بندی برای دسته‌بندی تصاویر در k دسته استفاده کرده و در نهایت از روش فیشر برای استخراج ویژگی‌های داخل هر خوشه استفاده می‌کنند.

● تجزیه ماتریس نامنفی (NMF)^۲: به‌طور کلی هدف تمامی روش‌های تجزیه ماتریس، شکستن ماتریس به یک یا چند ماتریس با ابعاد پایین‌تر می‌باشد که در هر روش بنا به فرضیات مسئله، محدودیت‌هایی نیز اعمال می‌شود. در روش تجزیه ماتریس نامنفی، یک ماتریس به دو ماتریس نامنفی که ماتریس‌های عامل نامیده می‌شوند، تجزیه می‌شود [۳۶].

نامنفی بودن ماتریس‌ها خواص اصلی داده در ماتریس اصلی را در ماتریس‌های عامل نیز حفظ می‌کنند که همین امر باعث سهولت در تجسم ستون‌های پایه در ماتریس‌های عامل می‌شود [۳۷]. به عبارتی دیگر مصورسازی داده‌های نامنفی راحت‌تر است. همچنین روش NMF می‌تواند هر شیء را به اجزایش تجزیه کند. به‌عنوان مثال تصویر چهره با استفاده از این روش به اجزای پایه صورت شامل چشم‌ها، بینی، لب‌ها و ... تجزیه می‌شود [۳۸]. همان‌طور که در شکل ۵.۲ مشاهده می‌شود، NMF نسخه‌های مختلفی از اجزای صورت استخراج می‌کند تا بر اساس این مؤلفه‌ها، ویژگی‌های تصاویر جدید محاسبه شود.

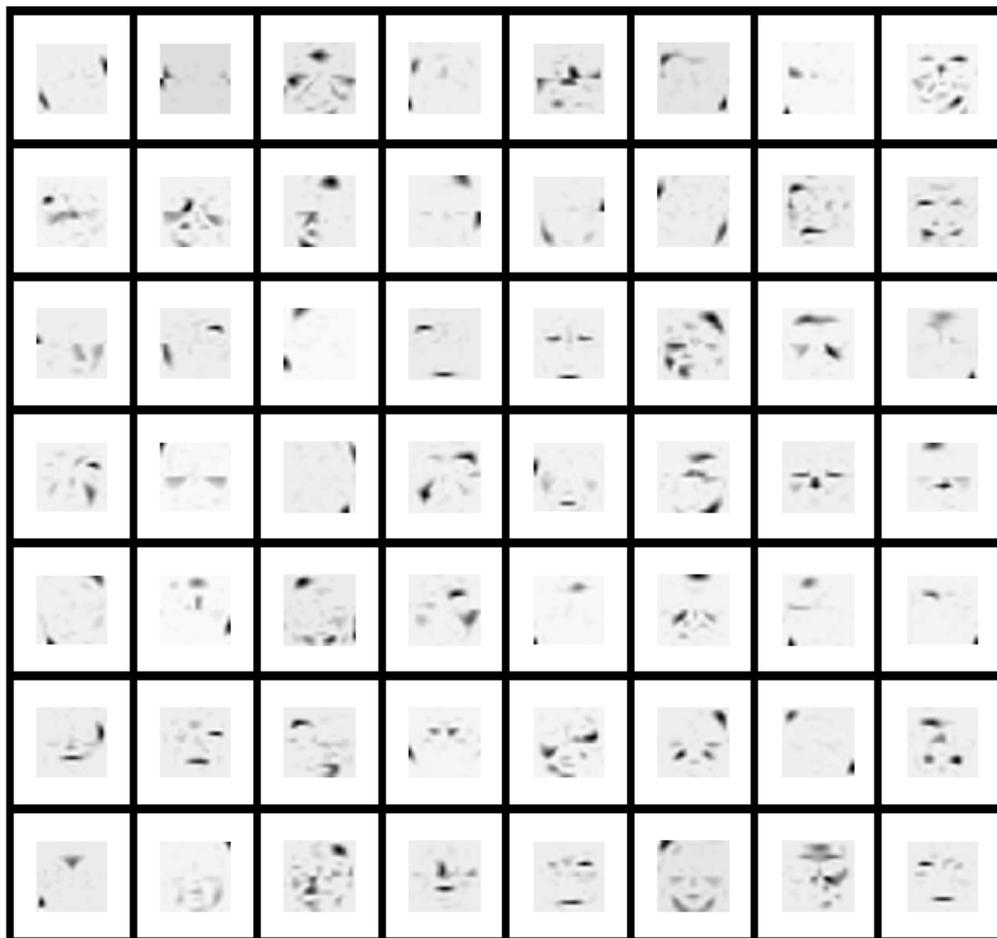
تجزیه ماتریس نامنفی عبارت است از تجزیه ماتریس V با اندازه $m \times n$ به دو ماتریس W و H که ضرب این دو ماتریس تقریبی از ماتریس اصلی V می‌باشد. معادله ۱۱.۲ این تقریب را نشان می‌دهد.

$$V_{m \times n} = W_{m \times k} \times H_{k \times n} \quad (11.2)$$

در این معادله به ماتریس W ماتریس پایه و به ستون‌های آن بردارهای پایه می‌گویند. ماتریس H نیز ماتریس ضرایب نامیده می‌شوند که در مسئله شناسایی چهره پس از تجزیه ماتریس تصاویر چهره این ماتریس رابطه مستقیم و یک به یک با هر یک از تصاویر خواهد داشت [۳۹]. k نیز مرتبه NMF نامیده می‌شود که یک عدد صحیح خیلی کوچک

¹Linear Discriminant Analysis (LDA)

²Non-negative Matrix Factorization



شکل ۵.۲: مؤلفه‌های NMF از روی تصاویر چهره

نسبت به ابعاد ماتریس اولیه است (معادله ۱۲.۲) [۴۰].

$$k \ll \min(m, n) \quad (12.2)$$

ماتریس‌های W و H با استفاده از حل معادله بهینه‌سازی ۱۳.۲ به دست می‌آیند [۲۵]. از آنجایی که این تابع نسبت به هر دو ماتریس W و H نامحدب است، پیدا کردن کمینه سراسری ناممکن است و این یک مسئله NP-Hard به شمار می‌رود. لذا جواب ممکن برای این مسئله، کمینه محلی است [۴۱]. الگوریتم‌های حل مسئله NMF به سه گروه کلی روش‌های مبتنی بر به‌روزرسانی ضربی، روش‌های مبتنی بر کاهش گرادیان و روش‌های مبتنی بر کمترین مربعات متناوب تقسیم می‌شوند [۴۲].

$$\min f(W, H) = \frac{1}{2} \|V - WH\|_F^2 \quad (13.2)$$

الگوریتم ۱ که توسط لی و سانگ [۴۱] معرفی شده است، برای به‌روزرسانی تابع هدف

۱۳.۲ از بهروزرسانی ضربی استفاده می‌کند. این روش برای همگرایی به زمان زیادی نیاز دارد، چراکه تعداد حلقه‌های تکرار زیادی دارد و محاسبات زیادی نیز در هر حلقه انجام می‌شود.

الگوریتم ۱ محاسبه NMF بر اساس بهروزرسانی ضربی

```

W = rand(m, k)
H = rand(k, n)
for i ≤ maxiter do
    H =  $\frac{H(W^T A)}{W^T W H + 10^{-9}}$ 
    W =  $\frac{W(AH^T)}{W H H^T + 10^{-9}}$ 
end for
    
```

در الگوریتم مبتنی بر کاهش گرادیان [۴۳]، نحوه بهروزرسانی ماتریس‌ها بدین صورت است که در هر بار تکرار درصدی از خطای گرادیان هر ماتریس از خود ماتریس کاسته می‌شود تا جایی که تابع همگرا شود. الگوریتم ۲ این روش را نشان می‌دهد. در این الگوریتم ϵ_H و ϵ_W ضریب یادگیری یا همان فاکتور تأثیر میزان گرادیان هر ماتریس در هر بار تکرار است. در روش بهروزرسانی ضربی نیازی به مشخص کردن این پارامترها نبود. این روش نیز مانند روش بهروزرسانی ضربی به دلیل وجود تکرارهای زیاد و محاسبات زیاد کند می‌باشد.

الگوریتم ۲ محاسبه NMF بر اساس کاهش گرادیان

```

W = rand(m, k)
H = rand(k, n)
for i ≤ maxiter do
    H = H -  $\epsilon_H \frac{\partial f}{\partial H}$ 
    W = W -  $\epsilon_W \frac{\partial f}{\partial W}$ 
end for
    
```

در الگوریتم‌های مبتنی بر کمترین مربعات متناوب به‌طور تکراری یک بار ماتریس W را ثابت در نظر گرفته و ماتریس H را محاسبه می‌کنند، و یک بار دیگر ماتریس H را ثابت در نظر گرفته و ماتریس W را محاسبه می‌کنند. در نسخه‌ای از این الگوریتم در هر مرحله مقادیر منفی حاصل از جواب کمترین مربعات با صفر جایگزین می‌شود که روش کاهش مربعات نامنفی متناوب نام دارد. الگوریتم ۳ این الگوریتم را نشان می‌دهد.

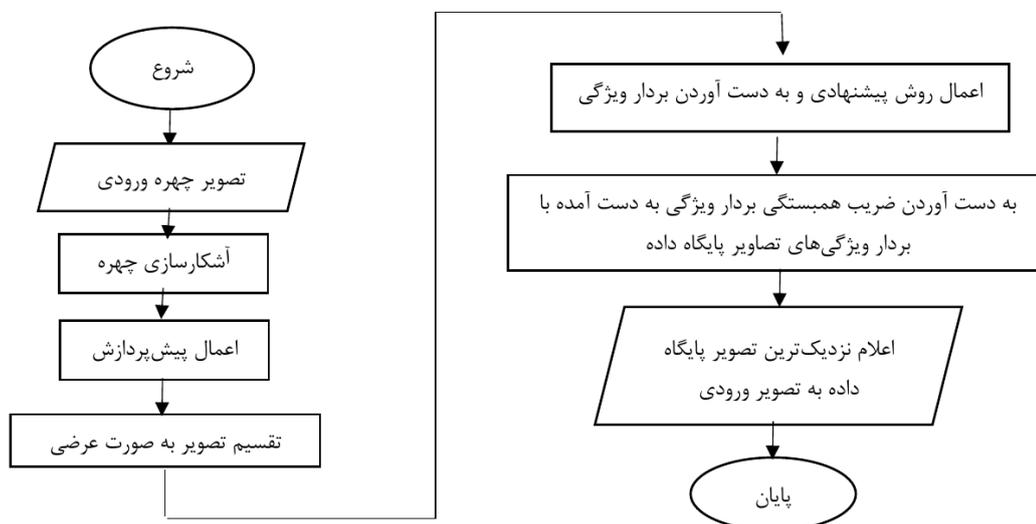
- نیکان و حسن‌پور سیستمی برای شناسایی چهره با استفاده از روش تجزیه ماتریس نامنفی (NMF) ارائه کردند [۴۴]. تمرکز این پایان‌نامه روی چالش تعداد کم نمونه‌های آموزش (یک تصویر به ازای هر شخص)، و سرعت شناسایی چهره می‌باشد. در این روش بعد از تشخیص چهره و انجام پیش‌پردازش‌های برابرسازی هیستوگرام، برابرسازی

الگوریتم ۳ محاسبه NMF بر اساس کمترین مربعات نامنفی متناوب (ANLS)

```

W = rand(m, k)
H = rand(k, n)
for i ≤ maxiter do
    Solve for H in matrix equation WTWH = WTA.
    Set all negative elements in H to 0.
    Solve for W in matrix equation HHTWT = HAT.
    Set all negative elements in W to 0.
end for
    
```

هیستوگرام وفقی با محدودیت کنتراست و تنظیم شدت روشنایی، تصاویر به اندازه ثابت 20×20 تبدیل می‌شوند. تصاویر حاصل به صورت عمودی به دو قسمت تقسیم شده و هر قسمت به صورت یک بردار در می‌آید. تمامی بردارهای مربوط به قسمت بالا و پایین تصاویر تشکیل دو ماتریس می‌دهند. به وسیله روش تجزیه ماتریس نامنفی هرکدام از ماتریس‌های بالا و پایین تصاویر به ماتریس‌های با ابعاد کمتر تجزیه می‌شوند. با اینکار ویژگی‌های مربوط به قسمت بالا و پایین تصاویر جداگانه استخراج می‌شود. در نهایت مقدار همبستگی^۱ قسمت‌های بالا و پایین به طور جداگانه محاسبه شده و با استفاده از ضرایبی باهم ترکیب می‌شوند. دو تصویر با مقدار همبستگی بیشتر، به عنوان تصویر با برجستگی مشترک تشخیص داده می‌شود. این روش روی ۹۹۰ تصویر از پایگاه داده FERET آزمایش شده است و به دقت ۹۲٪ دست‌یافته است (شکل ۶.۲).



شکل ۶.۲: فلوجارت روش نیکان [۴۴]

- فیلترهای گابور^۲: فیلترهای گابور جزو فیلترهای خطی و محلی هستند. هسته کانولوشن فیلتر گابور حاصل ضرب یک تابع نمایی مختلط در یک تابع گوسی است. این فیلترها

^۱correlation

^۲gabor filters

تشابه بسیاری با سیستم بینایی انسان دارند. برای اعمال فیلتر گابور بر روی تصویر، هسته گابور را با اندازه s و جهت o در سطح پیکسل به تصویر ورودی کانوالو می‌گردد.

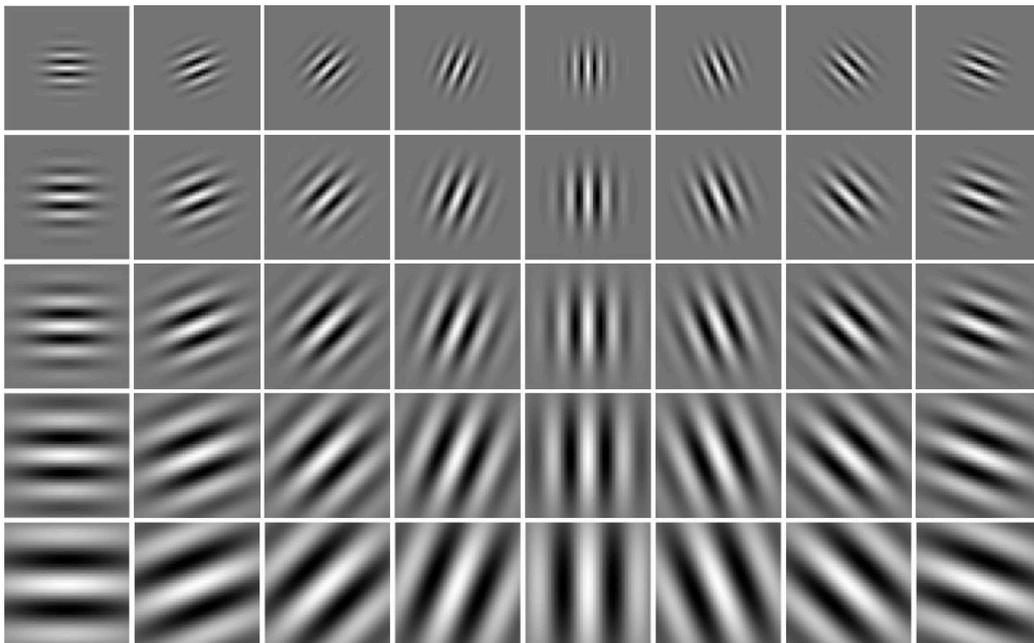
$$G_{s,o}(x, y) = I(x, y) * k_{s,o}(x, y) \quad (14.2)$$

$k_{s,o}$ هسته گابور می‌باشد. با اعمال هسته گابور به هر پیکسل، یک عدد مختلط حاصل می‌شود که بخش حقیقی آن با Re و بخش موهومی آن با Im نشان داده می‌شود. با استفاده از این دو مقدار، بخش‌های اندازه M^1 و فاز P^2 محاسبه می‌شوند:

$$M_{s,o}(x, y) = \sqrt{Im_{s,o}^2(x, y) + Re_{s,o}^2(x, y)} \quad (15.2)$$

$$P_{s,o}(x, y) = \arctan\left(\frac{Im_{s,o}(x, y)}{Re_{s,o}(x, y)}\right) \quad (16.2)$$

در شکل ۷.۲ فیلترهای گابور در ۸ جهت و ۵ اندازه محاسبه و نمایش داده شده است.



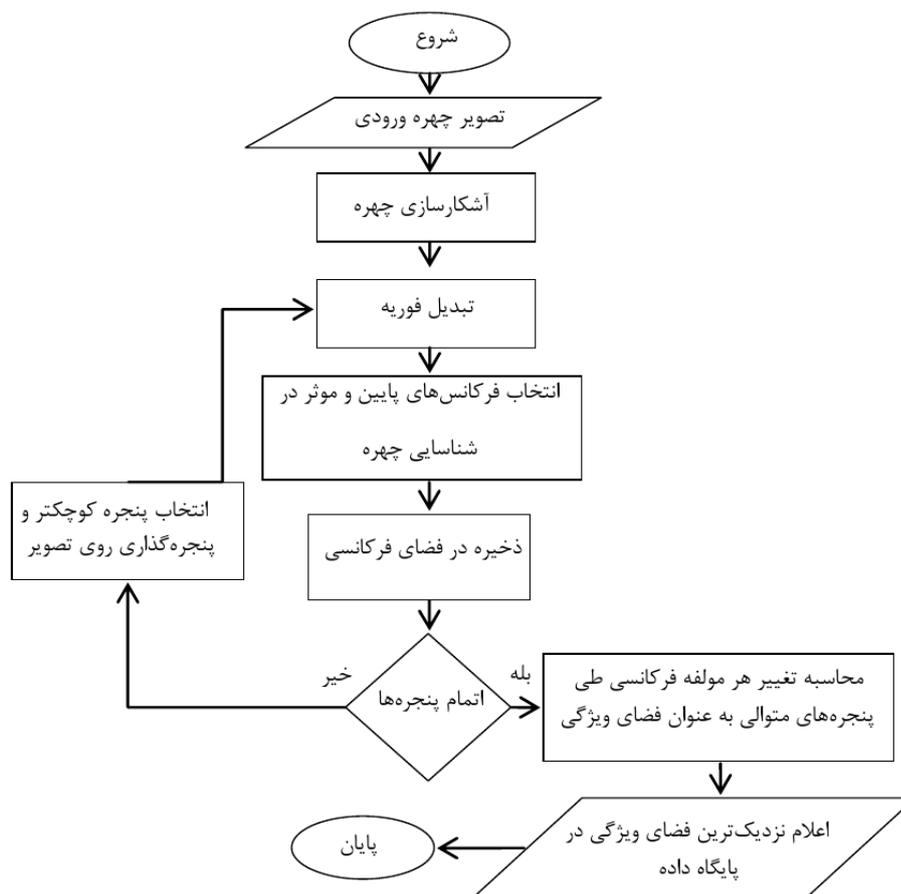
شکل ۷.۲: فیلتر گابور در ۸ جهت و ۵ اندازه

رشید و همکاران [۴۵] روشی برای شناسایی چهره ارائه دادند که پس از پیش پردازش‌های برابرسازی هیستوگرام و فیلتر میانه، با استفاده از فیلتر گابور ویژگی‌هایی استخراج می‌شود. سپس با استفاده از روش PCA ابعاد ویژگی کاهش یافته و در نهایت، یک

¹magnitude
²phase

دسته‌بند ماشین بردار پشتیبان (SVM) ویژگی‌های تصاویر را یاد می‌گیرد.

- بخشی و همکاران در پایان‌نامه خود روشی نوین برای بازیابی تصاویر از پایگاه داده‌های حجیم ارائه کردند [۴۶]. در این پایان‌نامه با استفاده از اعمال پنجره‌های هم‌مرکز با ابعاد مختلف روی تصویر چهره، محتوای هر پنجره به فضای فرکانس منتقل شده و سپس به وسیله فیلتر مناسب، تنها مؤلفه‌های فرکانسی با قابلیت جداسازی بالا بین تصاویر، حفظ می‌شوند. در نهایت با استفاده از ویژگی‌های مکان-فرکانس استخراج شده و معیار فاصله اقلیدسی، تصاویر از پایگاه داده بازیابی می‌شوند. این روش با فرض وجود یک تصویر به ازای هر شخص، روی پایگاه داده FERET با حضور ۹۹۰ شخص آزمایش شد و به دقت ۹۰/۴٪ دست‌یافته است. این روش به دلیل استفاده از ویژگی‌های مکان-فرکانس که نیاز به محاسبات پیچیده دارند، به زمان اجرای بالایی احتیاج دارد (شکل ۸.۲).



شکل ۸.۲: فلوچارت روش بخشی [۴۶]

۲.۳.۲ روش‌های غیرخطی

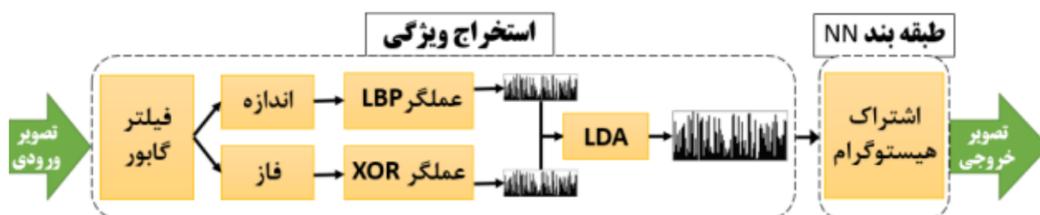
- PCA مبتنی بر کرنل [۱۷]: نسخه‌ای بهبودیافته از PCA است که از روش‌های کرنل

استفاده می‌کند. PCA مبتنی بر کرنل، چهره‌های ویژه را از روی ماتریس کرنل محاسبه می‌کند درحالی‌که روش PCA از ماتریس کوواریانس استفاده می‌کند.

۴.۲ روش‌های ترکیبی

روش‌های ترکیبی از ادغام دو یا چند روش جدا از هم که ممکن است از هر دو روش محلی و سراسری نیز باشند، تشکیل می‌شوند. این روش‌ها از مزیت‌های هر روش در کنار هم استفاده می‌کنند. همین امر پتانسیل این روش‌ها در ساخت سیستم‌های شناسایی کارا تر را بالا می‌برد.

- بیگری و همکاران در مقاله‌ای روشی ترکیبی برای شناسایی چهره با استفاده از ویژگی‌های فیلتر گابور و LBP ارائه کردند [۱۸]. در این روش ابتدا با استفاده از فیلتر گابور دو بخش فاز و اندازه تصاویر استخراج می‌شود. سپس به‌طور جداگانه، عملگر LBP روی بخش اندازه و عملگر LBP+XOR روی بخش فاز به دلیل وجود ناپایداری، اعمال می‌شود. در هر بخش هیستوگرام‌ها باهم ترکیب می‌شوند. در ادامه برای کاهش ابعاد ویژگی از روش LDA استفاده می‌شود. طریقه استفاده از این روش بدین‌صورت است که چون نسبت طول بردار ویژگی به تعداد نمونه‌های هر کلاس و تعداد کلاس‌ها بزرگ است، ابتدا بردارهای ویژگی هر بخش به چند بازه تقسیم می‌شود، سپس در هر بخش به‌طور جداگانه از LDA استفاده می‌شود. در نهایت تمامی ویژگی‌های کاهش‌یافته را باهم ترکیب کرده و بردار ویژگی نهایی حاصل می‌شود. این روش روی تصاویر پایگاه داده YaleB آزمایش شد و به دقت ۹۹/۵۲٪ رسیده است. با توجه به محاسبات پیچیده و استفاده از روش‌های فیلتر Gabor و LDA، این رویکرد از زمان اجرای بالا رنج می‌برد (شکل ۹.۲).



شکل ۹.۲: فلوچارت روش بیگری [۱۸]

- نوریان و همکاران روشی ترکیبی برای استخراج ویژگی از تصاویر در پایگاه داده‌های حجیم با استفاده از روش‌های HOG، LBP و STFT در پایان‌نامه خود ارائه کردند [۴۷]. ابتدا تصاویر چهره با استفاده از روش‌های یکسان‌سازی هیستوگرام، یکسان‌سازی گاما و یکسان‌سازی هیستوگرام افقی پیش‌پردازش می‌شوند. سپس از تصاویر بهبودیافته در سه مرحله جداگانه استخراج ویژگی صورت می‌گیرد. ابتدا برای استخراج ویژگی

فوریه زمان کوتاه (STFT)^۱ تصویر به نواحی جدا تقسیم می‌شود. سپس تبدیل فوریه هر ناحیه محاسبه می‌شود. در تبدیل فوریه، مؤلفه‌های کم‌فرکانس‌تر در مرکز تصویر قرار می‌گیرند. تعداد k مؤلفه مرکزی را به‌عنوان ویژگی هر ناحیه برداشته و ویژگی‌های همه نواحی باهم ترکیب می‌شوند. در مرحله بعدی برای استخراج ویژگی‌های HOG باز هم تصاویر به نواحی جدا تقسیم‌شده، سپس در هر ناحیه سلول‌هایی تشکیل می‌یابند. در هر سلول هیستوگرام اندازه و جهت گرادیان‌ها محاسبه می‌شود. سپس از روی هیستوگرام سلول‌های هر ناحیه، هیستوگرام مربوط به همان ناحیه تشکیل می‌شود. در نهایت هیستوگرام همه نواحی ویژگی نهایی را به وجود می‌آورد. در مرحله آخر برای استخراج ویژگی LBP تصاویر به نواحی جدا تقسیم شده، سپس هر ناحیه به بلوک‌های 3×3 جدا می‌شود و در هر بلوک الگوی دودویی محلی محاسبه می‌شود. سپس هیستوگرام این الگوها در هر ناحیه باهم ترکیب‌شده، ویژگی نهایی را می‌سازد. ویژگی‌های هر سه مرحله در دو گام نرمال‌سازی می‌شوند: یکی در سطح ناحیه، و دیگری در سطح کلی. با ترکیب سه ویژگی STFT، HOG و LBP بردار ویژگی نهایی به‌دست می‌آید. در نهایت با استفاده از دسته‌بند نزدیک‌ترین همسایه، و معیار فاصله منهن، عمل شناسایی صورت می‌گیرد. این روش از روی تصاویر پایگاه داده FERET با وضوح 360×360 پیکسل بردارهای ویژگی به طول ۲۶۰۱ استخراج کرد و به دقت بازیابی ۹۶/۶۶٪ دست‌یافته است.

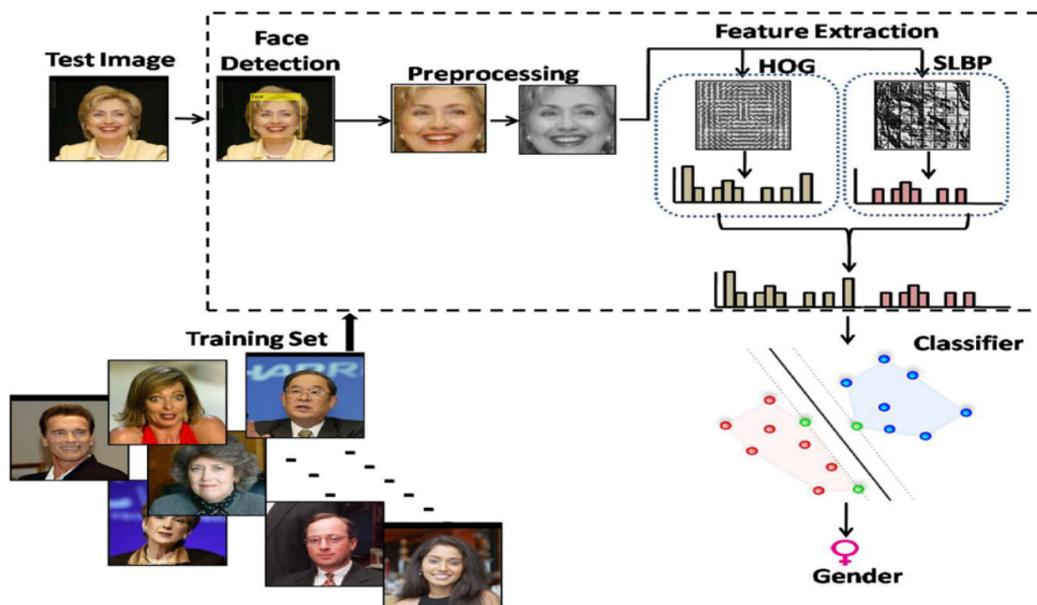
- موسی و همکاران روشی ترکیبی با استفاده از الگوریتم‌های تکاملی برای شناسایی چهره ارائه دادند [۱۹]. در این روش ابتدا با استفاده از تبدیل کسینوسی^۲ تصاویر به فضای فرکانس منتقل می‌شوند. سپس ضرایب DCT استخراج‌شده با استفاده از روش PCA کاهش داده می‌شوند. از میان مؤلفه‌های اصلی به‌دست‌آمده با استفاده از الگوریتم ژنتیک، چند مؤلفه اصلی باارزش انتخاب‌شده، سپس به‌وسیله معیار فاصله اقلیدسی عمل تطابق در پایگاه داده صورت می‌گیرد. این روش روی پایگاه داده‌های ORL دقت ۹۸/۴۵٪، پایگاه داده UMIST دقت ۹۹/۴٪ و پایگاه داده YALE دقت ۹۵/۵٪ را به‌دست آورده است.

- آنالاکشمی و همکاران روشی ترکیبی برای دسته‌بندی جنسیت اشخاص ارائه دادند [۳]. در این روش از ویژگی‌های محلی بهبودیافته الگوی دودویی محلی (SLPB) با ترکیب روش HOG استفاده شده است. ابتدا تصاویر موجود به نواحی کوچک‌تر تقسیم می‌شوند، سپس در هر ناحیه ویژگی‌های بهبودیافته LBP استخراج می‌شوند. برای استخراج ویژگی‌های HOG نیز تصویر به بلوک‌های متصل تقسیم می‌شود، سپس اندازه و جهت گرادیان‌ها در هر سلول از بلوک‌ها محاسبه و هیستوگرام‌ها تشکیل می‌شوند. بعد از استخراج

^۱Short Time Fourier Transform

^۲Discrete Cosine Transform (DCT)

ویژگی‌های بافت از طریق LBP و ویژگی‌های ساختاری از طریق HOG، بردارهای هر دو ویژگی باهم ترکیب و تشکیل یک بردار واحد می‌دهند. برای قسمت تطابق یک دسته‌بند ماشین بردار پشتیبان (SVM)^۱ با هسته مکعبی به‌وسیله بردارهای ویژگی استخراج‌شده آموزش می‌بیند (شکل ۱۰.۲). این روش روی ۴۵۰۰ تصویر از پایگاه داده FERET توانست به دقت ۹۹/۱٪، و با ۱۲۲۳۶ تصویر از پایگاه داده LFW به دقت ۹۵/۷٪ جنسیت تصاویر را از همدیگر تمیز دهد.



شکل ۱۰.۲: فلوجارت روش آنالاکشمی [۳]

- هاو و همکاران [۱۱] یک روش برای بهبود ویژگی‌های HOG ارائه و از ترکیب همین ویژگی‌ها با ویژگی‌های NMF برای شناسایی چهره استفاده کرده‌اند. در این روش ابتدا تصویر به چندین بلوک تقسیم‌شده و آنتروپی اطلاعات^۲ هر بلوک جداگانه محاسبه می‌شود. این آنتروپی اطلاعات به‌عنوان ضریب هر بلوک برای استخراج ویژگی HOG استفاده می‌شود. ابعاد ویژگی‌های HOG با استفاده از روش NMF کاهش یافته، و درنهایت ویژگی‌ها با استفاده از یک شبکه عصبی کانولوشنی دسته‌بندی می‌شوند.

- زین و همکاران [۴۸] یک روش ترکیبی برای شناسایی چهره ارائه کردند. در این روش ویژگی‌های الگوی دودویی گابور محلی (LGBP) از تصاویر استخراج می‌شود. سپس با استفاده از روش فیشر ابعاد ویژگی کاهش می‌یابد. درنهایت از روش نزدیک‌ترین همسایه برای تطبیق تصاویر استفاده می‌شود.

^۱support vector machine

^۲information entropy

۵.۲ جمع‌بندی

در این فصل یک دسته‌بندی کلی برای روش‌هایی که برای شناسایی چهره معرفی شده‌اند، مطرح شد. چند روش از هر دسته به‌طور مفصل توضیح داده شد. بعضی روش‌ها ایرادات و یا حتی نقاط قوت خاصی داشتند که مطرح شد. خلاصه این توضیحات در جدول ۱.۲ گزارش شده است.

جدول ۱.۲: نمونه روش‌های شناسایی چهره، تکنیک‌های استفاده‌شده به همراه نقاط قوت و ضعف

نویسنده (سال)	تکنیک	نقاط قوت	محدودیت‌ها
روش‌های محلی			
برایان و همکاران (۲۰۲۰) [۲۹]	LBP	ویژگی‌های قوی از چهره روبرو	حالات چهره و پوشیده بودن قسمتی از چهره
تانسر و همکاران (۲۰۲۰) [۲۰]	QTBP	سرعت بالا	پارامترهای زیاد مسئله
آناند و همکاران (۲۰۱۶) [۳۳]	SURF	مقاوم در برابر جابجایی، تغییر مقیاس و چرخش	تغییرات روشنایی
روش‌های سراسری			
نظری و همکاران (۲۰۱۶) [۳۴]	PCA-Fuzzy C-Mean	سرعت بالا	مجموعه تصاویر کم
لو و همکاران (۲۰۱۳) [۳۵]	LDA	عدم نیاز به نمونه‌های آموزشی زیاد	مجموعه تصاویر کم
نیکان و حسن‌پور (۲۰۱۹) [۴۴]	NMF	سرعت و دقت بالا	تغییرات زاویه و پوشیده بودن قسمتی از چهره
رشید و همکاران (۲۰۲۰) [۴۵]	Gabor Filters	تغییرات روشنایی و پوشیده بودن	مجموعه تصاویر کم
بخشی و همکاران (۲۰۱۸) [۴۶]	FFT	تمرکز روی پایگاه داده حجیم	محاسبات پیچیده و سرعت کند بازایی
روش‌های ترکیبی و یادگیری عمیق			
بیگلری و همکاران (۲۰۱۴) [۱۸]	Gabor Filters-LBP-LDA	دقت بالا	محاسبات پیچیده و طولانی
نوریان و همکاران (۲۰۱۷) [۴۷]	SFTF-LBP-HOG	دقت بالا	نیاز به تصاویر با رزولوشن بالا و سرعت کند بازایی
موسی و همکاران (۲۰۱۸) [۱۹]	DCT-PCA-GA	دقت بالا	مجموعه تصاویر کم
آنالاکشمی و همکاران (۲۰۱۹) [۳]	SLBP-HOG	دقت بالا و مقاوم در برابر زاویه	محاسبات پیچیده
هاو و همکاران (۲۰۱۹) [۱۱]	HOG-NMF-CNN	مقاوم در برابر روشنایی	عدم توانایی در شرایط یک نمونه به ازای هر شخص
زنگ و همکاران (۲۰۱۸) [۴۹]	DCNN (TDL)	حل مشکل یک نمونه به ازای هر شخص	دقت پایین و مجموعه تصاویر کم
مین و همکاران (۲۰۱۹) [۵۰]	TL-DCNN	حل مشکل یک نمونه به ازای هر شخص	محاسبات پیچیده و طولانی

فصل ۳

روش پیشنهادی

۱.۳ مقدمه

در فصل ۱ به اختصار اشاره شد که هدف از این تحقیق، ساخت و پیاده‌سازی یک سیستم شناسایی چهره است که در حضور تعداد تصاویر زیاد که از هر شخص نمونه تصاویر کمی نیز در اختیار است، بتواند به خوبی تصاویر اشخاص جدید را شناسایی کند. راهکاری که در این تحقیق ارائه می‌شود یک سیستم شناسایی مبتنی بر یک روش خوشه‌بندی ساده است که به صورت سلسله مراتبی تصاویر را به خوشه‌هایی با تعداد تصاویر یکسان تقسیم کرده و در هر خوشه عمل شناسایی را انجام می‌دهد. به دلیل تقسیم تصاویر با تعداد بالا به دسته‌های کوچک‌تر، تعداد مقایسات نهایی کاهش می‌یابد و همچنین سیستم این امکان را پیدا می‌کند که از میان تعداد اشخاص کمتری، هویت چهره خواسته‌شده را پیدا کند. همین امر موجب بهبود عملکرد سیستم می‌شود. در مرحله شناسایی چهره از دو ویژگی NMF و توصیفگرهای FREAK برای بهبود دقت شناسایی درون خوشه‌ها استفاده می‌شود. بدین ترتیب با ترکیب ویژگی‌های سراسری و محلی می‌توان از قدرت متمایزکننده هر دو روش استفاده کرد و به طور همزمان به خصوصیات سراسری و محلی تصاویر اهمیت داد. در نهایت هویت تصاویر با استفاده از روش نزدیک‌ترین همسایه (1NN) [۵۱] و معیار فاصله بلوک شهری مشخص می‌شود. روش پیشنهادی به دلیل ترکیب دو روش متفاوت در دسته روش‌های ترکیبی قرار می‌گیرد.

¹Nearest Neighbor

در این فصل به طول کامل و دقیق تر به توضیح مراحل روش پیشنهادی خواهیم پرداخت. به طور کلی این روش از سه مرحله اساسی تشکیل شده است. هرکدام از این مراحل از گام‌هایی تشکیل شده‌اند که در ادامه به طور مفصل توضیح داده می‌شود.

۱. پیش‌پردازش:

- تشخیص چهره در تصویر

- بهبود تصاویر

۲. خوشه‌بندی تصاویر

- استخراج ویژگی

- دسته‌بندی تصاویر

۳. شناسایی چهره

- استخراج ویژگی

- تطابق تصاویر

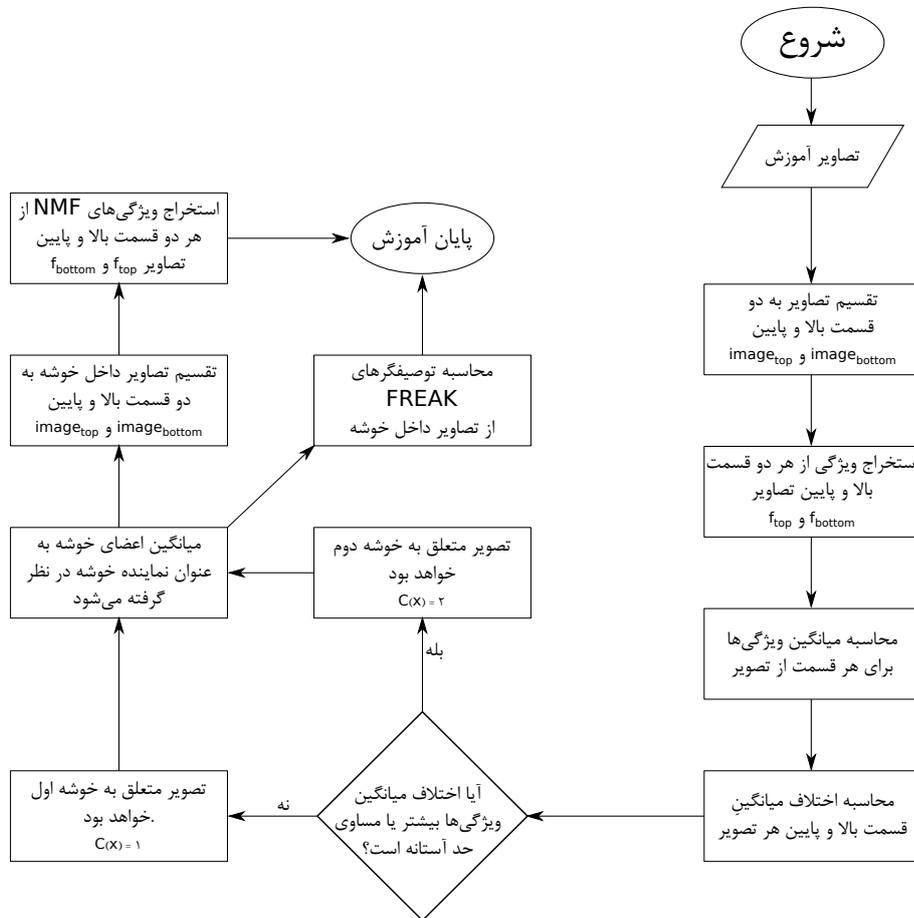
شکل ۱.۳ فلوجارت کلی روش پیشنهادی را نشان می‌دهد.

۲.۳ پیش‌پردازش

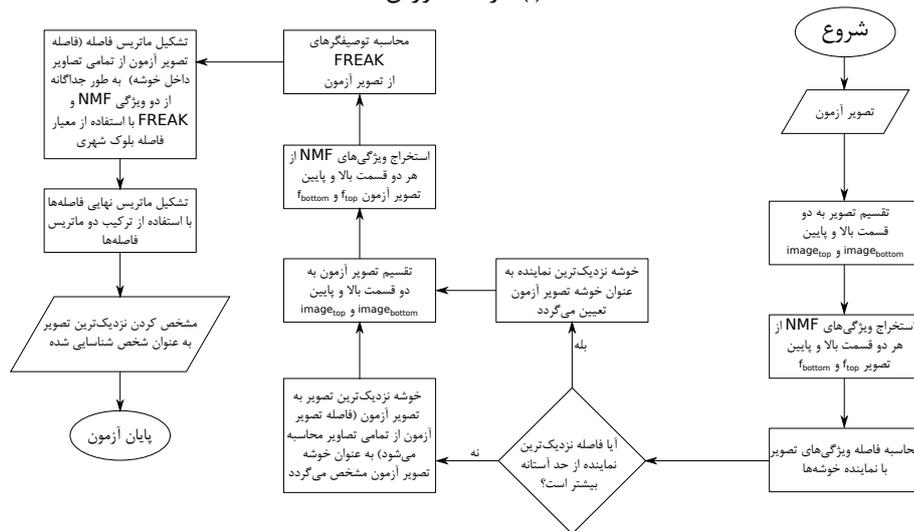
پیش‌پردازش داده‌ها در علم داده‌کاوی از اهمیت بالایی برخوردار است. چراکه وجود داده‌های خراب و یا به اصطلاح کثیف بودن داده‌ها باعث می‌شود تا سیستم دچار اشتباهاتی شود. از این رو قبل از شروع کار بایستی داده‌ها تمیز و آماده ادامه کار شوند. در روش مطرح شده این مرحله از دو بخش جداگانه تشکیل شده است. در بخش اول هدف سیستم جدا کردن قسمت چهره از قسمت‌های غیر چهره از تصویر است. در بخش دوم نیز کیفیت تصاویر چهره استخراج شده با استفاده از چندین روش بهسازی، بهبود می‌یابد.

۱.۲.۳ تشخیص چهره

تصاویر چهره ممکن است حاوی صحنه‌های پس‌زمینه غیر از چهره باشند. این شرایط موجب ایجاد اختلال در روند شناسایی چهره می‌شوند. برای رفع این مشکل بایستی در مرحله اول، محدوده چهره در تصاویر مشخص شده و استخراج شود. برای این منظور از روش متداول Viola-Jones [۵۲] استفاده شده است. این روش در ادامه به طور مفصل توضیح داده شده است.



(آ) مرحله آموزش



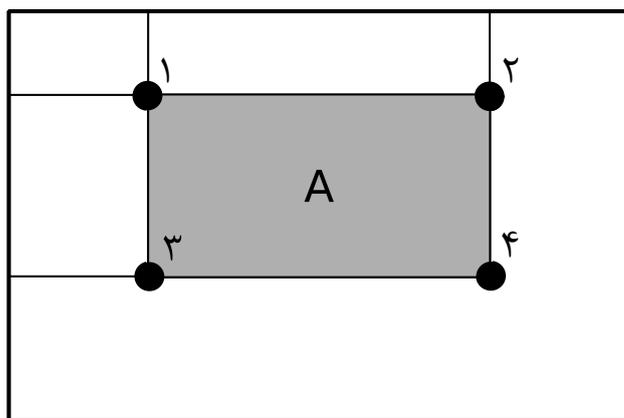
(ب) مرحله آزمون

شکل ۱.۳: فلوجارت روش پیشنهادی

ابتدا تصویر به زیرپنجره‌هایی تقسیم شده و ویژگی‌های هار^۱ از هر زیرپنجره استخراج می‌شود. ویژگی‌های هار مجموعه‌ای از توابعند که به نوعی نشان‌دهنده بخش‌هایی از صورت هستند. تمامی چهره‌های انسانی دارای ویژگی‌های مشترکی هستند که می‌توانند با ویژگی‌های هار تطابق پیدا کنند. به‌عنوان مثال ناحیه چشم‌ها تیره‌تر از ناحیه گونه می‌باشد. با استفاده از همین خصوصیت می‌توان یکی از ویژگی‌های هار را تطبیق داده و این بخش از چهره را مشخص کرد. با توجه به این توضیحات، می‌توان این ویژگی‌ها را به‌صورت مستطیل‌هایی در نظر گرفت که از دو بخش روشن و تیره تشکیل شده‌اند. ارزش هر ویژگی برابر است با اختلاف مجموع پیکسل‌های بخش تیره با مجموع پیکسل‌های بخش روشن. در روش Viola-Jones برای استخراج ارزش ویژگی‌های هار، این مستطیل‌ها روی تصویر قرار می‌گیرند. برای محاسبه ارزش مستطیل‌ها نیاز به انجام تعداد خیلی زیادی عملیات جمع است که باعث بالا رفتن پیچیدگی زمانی می‌شود. برای حل این مشکل مفهومی به نام تصویر انتگرال معرفی شد. در تصویر انتگرال مقدار هر پیکسل برابر است با مجموع تمامی پیکسل‌های سمت چپ و بالای آن پیکسل (معادله ۱.۳).

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1.3)$$

به‌عنوان مثال برای محاسبه مجموع پیکسل‌های ناحیه A در تصویر ۲.۳ به‌جای محاسبات زیاد در تصویر معمولی، می‌توان از تعداد خیلی کمتری جمع و تفریق در تصویر انتگرال استفاده کرد. با استفاده از تصویر انتگرال مجموع پیکسل‌های ناحیه A از معادله ۲.۳ به‌دست می‌آید که به‌مراتب سریع‌تر می‌باشد. بدین ترتیب برای تسریع الگوریتم در استخراج ویژگی‌های هار از تصویر انتگرال استفاده می‌شود.



شکل ۲.۳: تصویر انتگرال

$$\sum = ii(4) + ii(1) - ii(2) - ii(3) \quad (2.3)$$

¹Haar features

به دلیل وجود تعداد بالای ویژگی‌های هار و همچنین انواع مختلفی که از این ویژگی وجود دارد، هنوز پیچیدگی محاسباتی بالاست. در این شرایط برای یادگیری این ویژگی‌ها از چارچوب AdaBoost استفاده می‌شود. این چارچوب برای انتخاب بهترین ویژگی‌های هار و همچنین آموزش یک دسته‌بند برای استفاده از آن ویژگی‌ها به کار می‌رود. در حقیقت این چارچوب با استفاده از ترکیب خطی چند دسته‌بند ضعیف، یک دسته‌بند قوی تشکیل می‌دهد. در ادامه از آنجایی که تقریباً حدود ۱٪/۰ زیرپنجره‌ها مثبت (چهره) هستند، یک معماری آبخاری برای استفاده از الگوریتم AdaBoost پیشنهاد می‌شود. در این معماری، در هر مرحله یک دسته‌بند قوی تشکیل می‌شود که مشخص می‌کند آیا زیرپنجره شامل چهره است یا خیر. همچنین در این معماری مراحل ابتدایی از دسته‌بندهای ساده تشکیل شده‌اند و در مراحل بعدی دسته‌بندها پیچیده‌تر می‌شوند تا دقیق‌تر زیرپنجره را مورد بررسی قرار دهند. در نهایت یک مجموعه‌ای از دسته‌بندهای متوالی تشکیل می‌شود که هر زیرپنجره را فیلتر می‌کنند، و هر زیرپنجره‌ای که به مرحله نهایی در این معماری برسد، به عنوان چهره تشخیص داده می‌شود.

روش دیگری هم برای تشخیص چهره در تصاویر بر اساس شبکه‌های عمیق به نام MTCNN [۵۳] وجود دارد که اخیراً معرفی شده و عملکرد خوبی داشته است. برای پیدا کردن روش بهینه تشخیص چهره در مسئله مورد نظر این تحقیق از بین دو روش Viola-Jones و MTCNN آزمایشی انجام شده است که نتیجه آن نشان‌دهنده عملکرد بهتر روش Viola-Jones نسبت به روش MTCNN است. لذا برای استخراج تصاویر چهره در این تحقیق از روش Viola-Jones استفاده خواهد شد. نتایج این ارزیابی در بخش ۲.۵ شرح داده خواهد شد.

۲.۲.۳ بهبود تصاویر

تصاویر به صورت خام ممکن است دچار خرابی‌هایی ناشی از نویز و تاری و اثر گامای دستگاه‌های عکس‌برداری باشند. این خرابی‌ها اثر مستقیم روی عملیات استخراج ویژگی داشته و در نتیجه عملکرد سیستم را نیز دچار اختلال می‌کنند. در این تحقیق برای رفع این خرابی‌ها از سه روش یکسان‌سازی هیستوگرام [۵۴]، برابرسازی هیستوگرام وفقی با محدودیت کنتراست (CLAHE) [۵۵] و تنظیم شدت نور تصویر استفاده می‌شود.

- یکسان‌سازی هیستوگرام: در این روش بازه پویای تصویر گسترش می‌یابد. منظور از بازه پویای تصویر اختلاف روشن‌ترین نقطه با تاریک‌ترین نقطه در تصویر سطح خاکستری می‌باشد. این روش از یک تابع تبدیل صعودی برای این کار استفاده می‌کند. تابع توزیع تجمعی استفاده‌شده در روش یکسان‌سازی هیستوگرام در معادله ۳.۳ آمده است.

$$p_k = T(q_k) = (L-1) \sum_{i=0}^k d_q(q_i) = \frac{(L-1)}{N} \sum_{i=0}^k n_i \quad k = 0, 1, \dots, L-1 \quad (3.3)$$

در این رابطه q تصویر ورودی، $d_q(q_i)$ مقدار i ام هیستوگرام نرمال‌شده تصویر ورودی،

n_i تعداد پیکسل‌های با مقدار i ، N تعداد کل پیکسل‌ها، L بیشینه مقدار پیکسل‌ها و k مقدار سطح خاکستری در بازی $[0, L - 1]$ است که در انتها p تصویر بهبودیافته حاصل خواهد شد.

- برابری هیستوگرام وقتی با محدودیت کنتراست: برخلاف یکسان‌سازی هیستوگرام که به صورت سراسری توزیع روشنایی تصویر را بهبود می‌دهد، این روش به صورت محلی این کار را انجام می‌دهد. بدین منظور تصویر به بلوک‌هایی تقسیم شده و در هر بلوک هیستوگرام تشکیل می‌شود که با استفاده از تابع تبدیل بهسازی صورت گرفته و کنتراست محلی و کیفیت لبه‌ها بهبود می‌یابد. همچنین در این روش با محدود کردن میزان کنتراست نهایی از تشدید یافتن ناخواسته نویزها در اثر بهسازی جلوگیری می‌شود.
- تنظیم شدت نور تصویر: در این روش به نوعی دامنه تغییرات پیکسل‌های تصویر به محدوده دلخواهی نرمال می‌شود. این روش از معادله ۴.۳ برای این بهبود استفاده می‌کند.

$$A(i, j) = I * \frac{M - m}{255} \quad (4.3)$$

در این معادله I تصویر ورودی، M و m به ترتیب بیشینه و کمینه محدوده مورد نظر است که در این تحقیق به ترتیب از بیشینه و کمینه مقادیر پیکسل‌های تصویر به طور پیش فرض استفاده شده است.

پیش پردازش‌های گفته شده روی تمامی تصاویر ورودی اعمال می‌شود. پس از اعمال این عملیات، تصاویر چهره آماده خوشه‌بندی هستند.

۳.۳ خوشه‌بندی

همان‌طور که قبلاً اشاره شد، در روش ارائه شده برای حل مشکل تعداد زیاد تصاویر پایگاه داده، از خوشه‌بندی برای دسته‌بندی کردن تصاویر در دسته‌هایی با تعداد تصویر کمتر استفاده خواهد شد تا شناسایی چهره در داخل هر دسته و بین تعداد تصاویر کمتر صورت گیرد. این مرحله از دو بخش استخراج ویژگی و سپس خوشه‌بندی تشکیل می‌شود.

۱.۳.۳ استخراج ویژگی

در هر مسئله یادگیری ماشین، استخراج ویژگی و یافتن توصیفگرهای مناسب که با توجه به ماهیت مسئله بتواند خصوصیات داده‌ها را مشخص کند، از اهمیت بالایی برخوردار است. در مسئله خوشه‌بندی تصاویر چهره، استخراج ویژگی نیز بایستی طوری باشد که بتواند خصوصیات

را از چهره‌ها استخراج کند که به‌خوبی نشان‌دهنده تعلق و یا عدم تعلق چهره به یک دسته باشد. به‌عنوان مثال اگر هدف مسئله جداسازی تصاویر چهره بر اساس جنسیت اشخاص باشد، بایستی ویژگی‌هایی استخراج شوند که نشان‌دهنده جنسیت آن چهره‌اند، تا بتوان زن و یا مرد بودن آن‌ها را از هم تمیز داد.

در این تحقیق قصد بر این است که ویژگی‌هایی برای خوشه‌بندی تصاویر استخراج شود که در کنار سادگی، قدرت جداسازی تصاویر را به‌خوبی داشته باشند. از این‌رو برای شروع از تجزیه ماتریس استفاده خواهد شد. به‌طور کلی برای تجزیه ماتریس از روش‌هایی مانند تحلیل مؤلفه اصلی (PCA)، تحلیل مؤلفه مستقل (ICA)، تجزیه مقدرهای منفرد (SVD) و تجزیه ماتریس نامنفی (NMF) استفاده می‌شود.

در این تحقیق از تجزیه ماتریس نامنفی (NMF) و از روشی که بر مبنای کمترین مربعات نامنفی متناوب توسط کیم و همکاران [۲۵] بهبودیافته است، استفاده شده است. عملکرد این روش هم از لحاظ زمان همگرایی و هم از لحاظ دقت نهایی، بهتر از روش‌های دیگر است. برای خوشه‌بندی، ابتدا تصاویر از ارتفاع به دو بلوک غیرهمپوشان مساوی (بالا و پایین) تقسیم می‌شوند. قسمت بالا شامل چشم، ابرو و پیشانی و قسمت پایین شامل بینی، لب‌ها و چانه است. از این دو قسمت به‌طور جداگانه استخراج ویژگی صورت خواهد گرفت. در این صورت هر تصویر دو بردار ویژگی برای قسمت‌های بالا و پایین خود خواهد داشت. نتایج این مراحل در شکل ۳.۳ آمده است



شکل ۳.۳: (آ) تصویر ورودی (ب) تصویر چهره (ج) تصویر بهسازی شده (د) تصویر بریده شده

شکل ۳.۳: چهار مرحله اصلی قبل از خوشه‌بندی

۲.۳.۳ دسته‌بندی تصاویر

پس از بدست آوردن ویژگی‌های هر تصویر، میانگین ویژگی‌هایشان برای قسمت‌های بالا (f_{top}) و پایین (f_{bottom}) محاسبه می‌شود. سپس دو میانگین از هم کم می‌شوند و یک عدد حاصل می‌شود (معادله ۵.۳). با انتخاب یک حد آستانه، می‌توان مشخص کرد که هر تصویر به کدام خوشه تعلق دارد. بدین‌صورت که اگر اختلاف میانگین ویژگی‌های مربوط به دو بخش بالا و پایین تصویری از مقدار حد آستانه بیشتر یا مساوی باشد، متعلق به خوشه اول، و در غیر این صورت (اختلاف میانگین ویژگی‌های دو بخش بالا و پایین تصویر کمتر از حد آستانه باشد)

متعلق به خوشه دوم است (معادله ۶.۳).

$$diff_x = \text{mean}(f_{top}) - \text{mean}(f_{bottom}) \quad (۵.۳)$$

$$C(x) = \begin{cases} 1 & \text{if } diff_x < \tau \\ 2 & \text{if } diff_x \geq \tau \end{cases} \quad (۶.۳)$$

در این تحقیق میانه اختلاف میانگین ویژگی‌های بخش بالا و پایین همه تصاویر به‌عنوان حد آستانه در نظر گرفته شده است. با این کار عملاً تصاویر به دو خوشه با تعداد مساوی تقسیم خواهند شد که همین متعادل بودن تعداد اعضای خوشه‌ها امتیازی برای روش خوشه‌بندی محسوب می‌شود.

اکنون بایستی خوشه تصاویر آزمون مشخص شود. برای این کار ابتدا میانگین اعضای هر خوشه را به‌عنوان نماینده خوشه در نظر می‌گیریم. سپس همانند تصاویر آموزش، تصاویر آزمون نیز به دو قسمت بالا و پایین تقسیم می‌شوند. از طریق NMF از هر دو قسمت استخراج ویژگی صورت می‌گیرد. در این مرحله برای هر تصویر دو بردار ویژگی برای قسمت‌های بالا و پایین وجود دارد.

فاصله تصاویر آزمون با نماینده خوشه‌ها محاسبه می‌شود. برای مشخص کردن خوشه تصویر آزمون از فاصله نزدیک‌ترین خوشه و یک حد آستانه استفاده می‌شود. اگر فاصله نزدیک‌ترین خوشه از حد آستانه تعیین شده بیشتر باشد، به عبارتی دیگر میزان قطعیت بالایی داشته باشد، همان خوشه به‌عنوان خوشه تصویر آزمون تعیین می‌شود. در غیر این صورت از روش نزدیک‌ترین همسایه استفاده خواهد شد. بدین صورت که خوشه نزدیک‌ترین تصویر به‌عنوان خوشه تصویر آزمون تعیین می‌شود (۷.۳).

$$C(y) = \begin{cases} C(c) & \text{if } d_{y,c} \geq \tau' \\ C(\text{argmin}(D)) & \text{if } d_{y,c} < \tau' \end{cases} \quad (۷.۳)$$

در معادله ۷.۳ y تصویر آزمون، c نماینده نزدیک‌ترین خوشه، D ماتریس فاصله تصویر آزمون و تمامی تصاویر، $d_{y,c}$ فاصله تصویر y و نماینده c ، و τ' حد آستانه تعیین شده است.

۴.۳ شناسایی چهره

پس از مشخص شدن خوشه تصاویر، بایستی هویت آن‌ها نیز معین شود. برای این کار برای هر تصویر آزمون، درون خوشه خودش استخراج ویژگی‌های جدیدی صورت می‌گیرد. سپس

با استفاده از معیار فاصله مناسب نزدیک‌ترین تصویر مشخص می‌شود. در مرحله شناسایی چهره از ویژگی‌های NMF و Fast Retina Keypoint (FREAK) استفاده خواهد شد. در نهایت برای تطبیق تصاویر از روش نزدیک‌ترین همسایه استفاده می‌شود.

۱.۴.۳ استخراج ویژگی

برای شناسایی چهره از ترکیب دو ویژگی NMF و FREAK استفاده می‌شود. روش NMF که به‌طور مفصل در بخش ۱.۳.۳ توضیح داده شد برای استخراج ویژگی‌های جدید از تصویر آزمون استفاده می‌شود.

توصیفگر FREAK [۲۶] نیز برای تقویت عملکرد شناسایی‌کننده استفاده خواهد شد. FREAK یک روش مبتنی بر نقاط کلیدی است که به‌صورت محلی ویژگی‌های دودویی را استخراج می‌کند. همان‌طور که قبلاً اشاره شد، این دسته از روش‌ها از دو مرحله شناسایی نقاط کلیدی و استخراج ویژگی حول آن نقطه تشکیل می‌شوند. در این تحقیق از روش تشخیص نقاط گوشه که توسط شی و توماسی [۵۶] ارائه شده است، برای تشخیص نقاط مورد علاقه استفاده شده است. سپس برای استخراج ویژگی حول نقاط گوشه از روش FREAK استفاده می‌شود.

روش شی-توماسی یکی از روش‌های تشخیص نقاط کلیدی است که نقاط گوشه را در تصویر پیدا می‌کند. در حقیقت این روش یک بهبود برای کاری است که هریس [۵۷] برای یافتن نقاط گوشه انجام داده است. به‌طور تئوری برای تشخیص نقاط گوشه هر تکه‌ای از تصویر (پنجره) را کمی در جهت‌های مختلف جابجا می‌کنند. اگر مقدار تغییرات در جهت‌های مختلف ناچیز باشد، آن تکه از تصویر مسطح است. اگر تغییرات در یکی از جهت‌ها خیلی بیشتر از بقیه جهت‌ها باشد، آن تکه از تصویر به‌عنوان لبه شناخته می‌شود. در نهایت اگر تغییرات در تمامی جهت‌ها خیلی زیاد باشد، آن تکه به‌عنوان نقطه گوشه در نظر گرفته می‌شود [۵۸].

برای فرموله کردن این توضیحات بایستی مجموع مربع اختلافات (SSD)^۱ وزنی پنجره انتخاب شده و میزان جابجایی محاسبه شود. بر اساس معادله ۸.۳ مقدار SSD محاسبه می‌شود. در این معادله محدوده (u, v) به‌وسیله بردار (x, y) جابجا می‌شود. وزن این محدوده با $w(u, v)$ نشان داده شده است.

$$S(x, y) = \sum_u \sum_v w(u, v) (I(u+x, v+y) - I(u, v))^2 \quad (۸.۳)$$

تقریبی از مقدار $I(u+x, v+y)$ را میتوان به‌وسیله بسط تیلور محاسبه کرد. اگر I_x و I_y مشتق جزئی تصویر باشند، میتوان به معادله ۹.۳ رسید.

$$I(u+x, v+y) \approx I(x, v) + I_x(u, v)x + I_y(u, v)y \quad (۹.۳)$$

^۱Sum of Squared Differences

با جایگذاری رابطه ۹.۳ در ۸.۳ رابطه ۱۰.۳ حاصل می‌شود.

$$S(x, y) \approx \sum_u \sum_v w(u, v) (I_x(u, v)x + I_y(u, v)y)^2 \quad (10.3)$$

معادله ۱۰.۳ را می‌توان به شکل ماتریسی در معادله ۱۱.۳ نوشت، که در این معادله A ماتریس ممان دوم می‌باشد که از گرادیان ماتریس مشتق می‌شود.

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} \quad (11.3)$$

ماتریس A به صورت معادله ۱۲.۳ نوشته می‌شود.

$$A = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x(u, v)^2 & I_x(u, v)I_y(u, v) \\ I_x(u, v)I_y(u, v) & I_y(u, v)^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix} \quad (12.3)$$

بر اساس حالاتی که برای یک تکه از تصویر با استفاده از میزان تغییرات در جهت‌های مختلف گفته شد، می‌توان تحلیلی روی ماتریس A و مقادیر ویژه آن داشت. متناظراً سه حالت مختلف وجود دارد (شکل ۴.۳):

۱. اگر هر دو مقادیر ویژه ماتریس نزدیک صفر باشند به این معنی است که تغییراتی در این پنجره وجود ندارد، پس این پنجره مسطح است ($\lambda_1 \approx 0$ و $\lambda_2 \approx 0$).

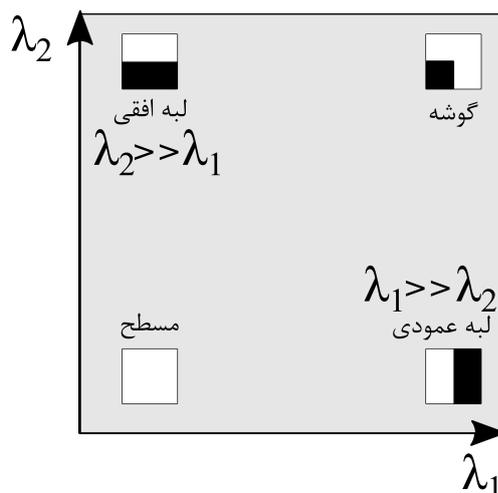
۲. اگر یکی از مقادیر ویژه نزدیک صفر و دیگری مقدار زیادی داشته باشد به معنی وجود تغییرات در یک جهت از پنجره است. پس می‌توان نتیجه گرفت که پنجره لبه می‌باشد ($\lambda_1 \approx 0$ و λ_2 برابر مقدار بزرگ).

۳. اگر دو مقدار ویژه ماتریس هم‌زمان مقدار بزرگی داشته باشند به معنی وجود تغییرات در همه جهات پنجره است، و این دلیلی بر وجود گوشه در این پنجره می‌باشد (λ_1 و λ_2 هر دو مقدار بزرگی دارند).

بر اساس این فرضیات، محققان برای تشخیص گوشه بودن یک پنجره از مقادیر ویژه استفاده می‌کنند. برای اینکار برای هر پنجره یک مقدار پاسخ گوشه^۱ (C) بایستی محاسبه شود. سپس اگر مقدار C از حد آستانه تعیین شده بیشتر باشد، آن پنجره به عنوان نقطه کلیدی (گوشه) انتخاب می‌شود و در غیر این صورت انتخاب نمی‌شود. هریس [۵۷] مقدار پاسخ گوشه را بر اساس معادله ۱۳.۳ محاسبه کرد.

$$C = |A| - k(\text{trace}(A))^2 \quad (13.3)$$

¹corner response



شکل ۴.۳: تشخیص نوع تکه‌ای از تصویر با استفاده از مقادیر ویژه ماتریس گرادیان کوواریانس

شی و توماسی [۵۶] برای ساده‌تر کردن این روابط، کوچک‌ترین مقدار ویژه را به‌عنوان مقدار پاسخ گوشه در نظر گرفتند (معادله ۱۴.۳). این روش در چندین مورد ردیابی اشیاء^۱ [۵۹] و شناسایی فعالیت^۲ [۶۰] کاربرد داشته است.

$$C = \min(\lambda_1, \lambda_2) \quad (14.3)$$

بر اساس توضیحات داده‌شده، نقاط کلیدی تصویر (نقاط گوشه) تشخیص داده می‌شوند. سپس بایستی ویژگی‌های محلی در اطراف این نقاط محاسبه و استخراج شود که در این تحقیق از توصیفگر FREAK استفاده شده است.

روش FREAK [۲۶] از سیستم بینایی انسان و علی‌الخصوص شبکیه چشم^۳ انسان الهام گرفته است. در علوم اعصاب تحقیقات زیادی در مورد چگونگی عملکرد سیستم بینایی انسان و این که تصاویر چگونه به مغز منتقل می‌شوند، انجام گرفته است. آن‌ها بر این باورند که شبکیه چشم انسان جزئیات تصاویر را با استفاده از اختلاف توزیع‌های گوسی (DoG)^۴ با سایزهای مختلف استخراج می‌کند. توپولوژی شبکیه چشم نقش مهمی را در این روند بازی می‌کند. روش FREAK توپولوژی شبکیه چشم انسان را برای استخراج ویژگی تقلید می‌کند.

برای شرح توپولوژی شبکیه چشم بایستی ابتدا ساختمان آن را بدانیم. ابتدا گیرنده‌های نوری^۵ روی یک سلول گانگلیونی^۶ اثر می‌گذارند. ناحیه‌ای که نور بر روی پاسخ سلول‌های گانگلیونی تأثیر می‌گذارد، میدان پذیرایی^۷ نام دارد. سایز این میدان‌ها با فاصله‌اش از قسمت

¹ object tracking

² action recognition

³ retina

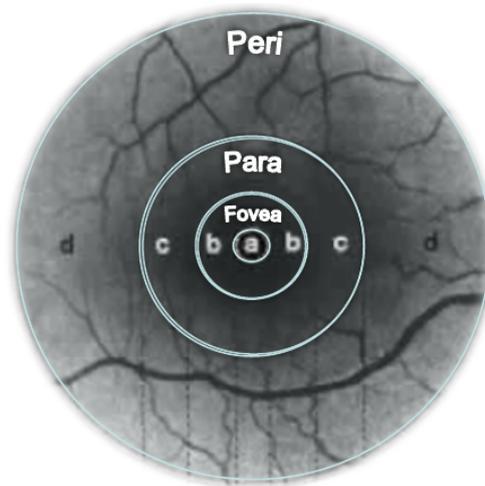
⁴ Difference of Gaussians

⁵ photoreceptor

⁶ ganglion cell

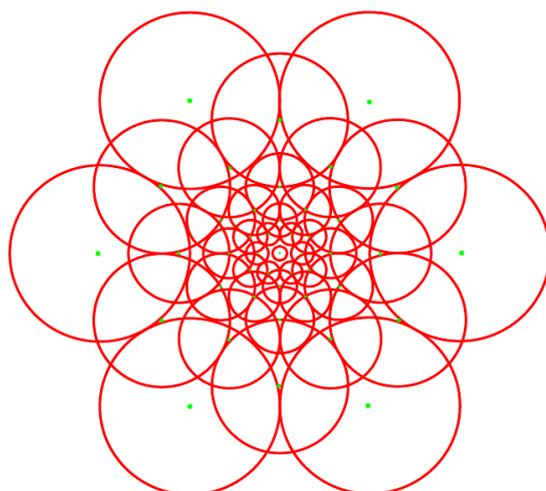
⁷ receptive field

مرکزی (foveal) رابطه مستقیم دارد. همچنین توزیع فضایی سلول‌های گانگلیونی با دور شدن از مرکز به‌طور نمایی کاهش می‌یابد. این توپولوژی در شکل ۵.۳ آمده است.



شکل ۵.۳: شبکیه چشم

برای استخراج ویژگی‌های دودویی از نقاط کلیدی، چندین جفت نقطه به‌طور تصادفی (به‌طور پیش‌فرض) انتخاب‌شده و باهم مقایسه می‌شوند تا یک بردار ویژگی دودویی را بسازند. روش FREAK نیز قصد ساختن ویژگی‌های دودویی دارد، با این تفاوت که برای انتخاب جفت نقاط از شبیه‌سازی سیستم شبکیه استفاده می‌کند. بر این اساس FREAK یک شبکه نمونه‌برداری در شکل ۶.۳ به‌صورت دایره‌ای معرفی می‌کند که همانند سیستم شبکیه چشم چگالی نقاط نزدیک مرکز بیشتر است.



شکل ۶.۳: مدل نمونه‌برداری روش FREAK الهام گرفته از توپولوژی شبکیه چشم

هر نقطه نمونه با استفاده از یک فیلتر گوسی صاف می‌شود تا اثر نویز را کمتر کند. سپس برای تطابق با توپولوژی شبکیه چشم، از هسته‌های با سایز مختلف برای ساخت توابع توزیع

گوسی روی شبکه نمونه برداری استفاده می‌شود. در این شبکه هر دایره نمایانگر یک انحراف معیار از هسته گوسی است که روی نقطه نمونه متناظر اعمال شده است.

پس از طراحی شبکه نمونه برداری، توصیفگر F با استفاده از اعمال یک حد آستانه روی مقدار اختلاف بین دو میدان پذیرایی تعریف می‌شود. این توصیفگر در معادله ۱۵.۳ نشان داده شده است. به عبارتی این توصیفگر یک رشته دودویی از اختلاف توزیع‌های گوسی (Dog) است.

$$F = \sum_{0 \leq a < N} 2^a T(P_a) \quad (15.3)$$

در این معادله P_a جفت میدان پذیرایی و N سایز توصیفگر می‌باشد. $T(P_a)$ نیز پاسخ آزمون جفت میدان پذیرایی است که در معادله ۱۶.۳ آمده است. بنابراین معادله، مقدار صاف شده جفت میدان پذیرایی از هم کم شده و حد آستانه اعمال می‌شود. اگر میزان اختلاف از حد آستانه بیشتر باشد، بیت مربوطه عدد ۱ و در غیر این صورت بیت مربوطه عدد ۰ خواهد بود.

$$T(P_a) = \begin{cases} 1 & \text{if } (I(P_a^{r_1}) - I(P_a^{r_2})) > 0 \\ 0 & \text{if } (I(P_a^{r_1}) - I(P_a^{r_2})) \leq 0 \end{cases} \quad (16.3)$$

۲.۴.۳ تطابق تصاویر

همان‌طور که اشاره شد، برای شناسایی چهره‌ها داخل خوشه‌ها از ترکیب دو ویژگی NMF و FREAK استفاده می‌شود. ویژگی‌های NMF همانند مرحله خوشه‌بندی استخراج شده و توصیفگر FREAK نیز پس از شناسایی نقاط کلیدی تصویر با استفاده از روش شی-توماسی محاسبه می‌شود. در این مرحله بایستی هویت تصویر آزمون با استفاده از ویژگی‌های استخراج شده مشخص شود.

فاصله دو ویژگی NMF و FREAK تصویر آزمون جداگانه با تمامی تصاویر داخل خوشه محاسبه شده و دو ماتریس فاصله برای دو ویژگی به دست می‌آید. این دو ماتریس با استفاده از ضرایبی باهم ترکیب شده و یک ماتریس فاصله نهایی را می‌سازند (معادله ۱۷.۳).

$$D_{total} = \alpha \times D_{NMF} + \beta \times D_{FREAK}, \quad \alpha + \beta = 1 \quad (17.3)$$

در نهایت با استفاده از روش ۱- نزدیک‌ترین همسایه [۵۱] با معیار فاصله بلوک شهری (City-Block) نزدیک‌ترین تصویر شناسایی و هویت آن به عنوان هویت تصویر آزمون تعیین می‌گردد.

۵.۳ جمع‌بندی

در این فصل به‌طور مفصل روش پیشنهادی مورد بررسی قرار گرفت. این روش از یک رویکرد خوشه‌بندی برای تقسیم پایگاه داده به دسته‌های کوچک‌تر استفاده می‌کند تا در نهایت تطبیق تصاویر داخل خوشه‌ها با تعداد تصاویر کمتر صورت بگیرد. روش خوشه‌بندی مطرح‌شده تصاویر را در هر مرحله خوشه‌بندی به دو خوشه با تعداد مساوی تقسیم می‌کند. داخل هر خوشه دو دسته ویژگی NMF و توصیفگرهای FREAK استخراج‌شده و به‌منظور بالا بردن قدرت روش پیشنهادی، باهمدیگر ترکیب می‌شوند. در نهایت هویت تصاویر آزمون با استفاده از روش نزدیک‌ترین همسایه و معیار فاصله بلوک شهری داخل هر خوشه تعیین می‌شود.

فصل ۴

ارزیابی و بررسی نتایج

۱.۴ مقدمه

در فصل قبل به طور مفصل مراحل روش پیشنهادی شرح داده شد. برای بررسی عملکرد هر روش ابداعی و یا بهبودیافته نیاز است که از چندین جنبه مختلف مطرح در مسئله تعریف شده، مورد ارزیابی و آزمایش قرار گیرد. روش ارائه شده در این تحقیق نیز بایستی از جنبه‌هایی مانند عملکرد روش خوشه‌بندی در مقایسه با روش‌های خوشه‌بندی دیگر، دقت نهایی در پایگاه داده‌های حجیم مورد ارزیابی قرار بگیرد. لذا در این فصل آزمایش‌هایی که برای ارزیابی روش ارائه شده انجام شده‌اند، مطرح و بررسی خواهند شد. در این فصل پایگاه داده‌های مورد استفاده و همچنین نکات اضافی پیاده‌سازی روش ارائه شده نیز شرح داده می‌شوند.

۲.۴ پایگاه داده

برای ارزیابی روش ارائه شده و بررسی عملکرد روی تعداد تصاویر زیاد از پایگاه داده حجیم FERET [۲۷] و همچنین یک مجموعه تصویر جمع‌آوری شده از چند پایگاه داده استاندارد کم‌حجم استفاده شده است. در ادامه مختصراً مشخصات و نکات مربوط به این مجموعه تصاویر شرح داده می‌شود.

۱.۲.۴ پایگاه داده FERET

پایگاه داده FERET در بازه زمانی آگوست سال ۱۹۹۳ تا جولای سال ۱۹۹۶ طی ۱۵ جلسه جمع‌آوری شده است. در این پایگاه داده از ۱۱۹۹ شخص عکس برداری شده است که در مجموع ۱۴۱۲۶ تصویر برای اشخاص وجود دارد. اشخاصی که در این پایگاه داده حضور دارند شامل زن و مرد های چندملیتی هستند. هر شخص می‌تواند چندین نوع تصویر داشته باشد. انواع مختلف تصاویر در جدول ۱.۴ آمده است. در ارزیابی‌های انجام شده از ۹۹۰ تصویر fa برای آموزش و تصاویر متناظرشان (fb) برای آزمون استفاده شده است. نمونه‌ای از این تصاویر در شکل ۱.۴ آمده است.

جدول ۱.۴: انواع تصاویر پایگاه داده FERET

نوع	زاویه چهره	توضیحات
fa	روبرو	حالت عادی
fb	روبرو	نسبت به fa از نظر حالت چهره و روشنایی متفاوت است
fc	روبرو	ثبت با دوربین و شرایط روشنایی متفاوت
dup I	روبرو	تصاویر به فاصله یک سال از fa گرفته شده‌اند
dup II	روبرو	تصاویر به فاصله بیشتر یک سال از fa ثبت شده‌اند

۲.۲.۴ مجموعه تصاویر جمع‌آوری شده

برای بررسی بیشتر، از تصاویر چندین پایگاه داده کوچک‌تر برای تشکیل یک پایگاه داده بزرگ‌تر استفاده شده است. بدین منظور تصاویر پایگاه داده‌های FERET، MUCT [۶۱]، FEI [۶۲]، PICS، و Face94 باهم جمع شده و یک پایگاه داده بزرگ‌تر تشکیل می‌شود. تصاویر عمدتاً از زاویه روبرو بوده و در حالت چهره و روشنایی باهم فرق دارند. تعداد کل اشخاص در این مجموعه تصویر به ۱۶۸۴ نفر می‌رسد. نمونه‌ای از تصاویر جمع‌آوری شده در شکل ۱.۴ آمده است.

۳.۴ تأثیر روش تشخیص چهره

همان‌طور که در بخش ۱.۲.۳ گفته شد اولین قدم، تشخیص مکان چهره در تصویر و استخراج تصویر چهره است. برای تشخیص بهتر چهره‌ها در تصاویر، دو روش تشخیص چهره Viola-Jones [۵۲] و MTCNN [۵۳] مورد ارزیابی قرار گرفته است. در شرایط تعریف شده این تحقیق،



(أ) نمونه تصاویر پایگاه داده FERET



(ب) نمونه تصاویر مجموعه جمع‌آوری شده

شکل ۱.۴: روش پیشنهادی با استفاده از تصاویر پایگاه داده استاندارد FERET و همچنین یک مجموعه تصویر جمع‌آوری شده از چندین پایگاه داده ارزیابی شده است.

روش Viola-Jones نتیجه بهتری داشته است. نتایج این آزمایش در جدول ۲.۴ ثبت شده است.

جدول ۲.۴: مقایسه عملکرد دو روش تشخیص چهره

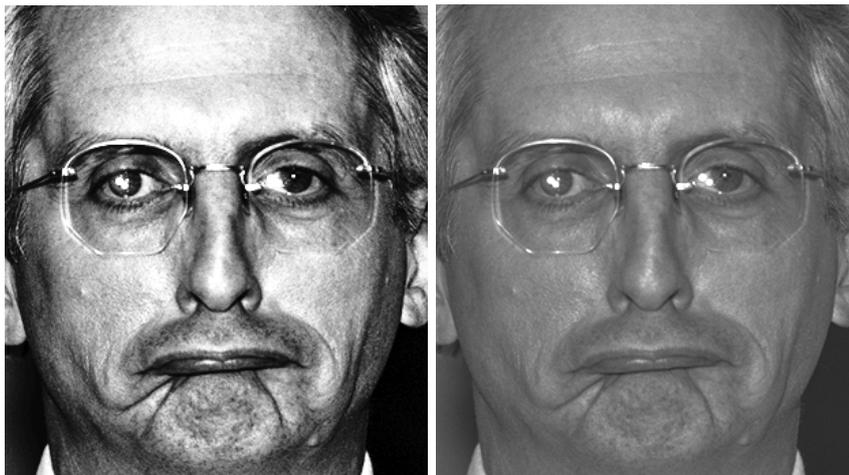
روش	دقت (%)
Viola-Jones	۹۸/۳۶
MTCNN	۹۶/۱۷

۴.۴ تأثیر پیش پردازش

پس از استخراج تصاویر چهره، بایستی از لحاظ روشنایی و کنتراست تصاویر بهبود یابند. برای این کار از روش‌های برابری هیستوگرام [۶۳] (شکل ۲.۴)، برابری هیستوگرام افقی با محدودیت کنتراست [۵۵] (شکل ۳.۴) و همچنین تنظیم شدت نور تصویر [۶۴] (شکل ۴.۴) استفاده می‌شود. برای بررسی تأثیر بهسازی‌ها، روش ارائه‌شده را بدون بهسازی و سپس به همراه بهسازی اجرا می‌کنیم. در حالی که از پیش پردازش استفاده شده است، به ترتیب بهسازی‌های برابری هیستوگرام، برابری هیستوگرام افقی و تنظیم شدت نور روی تصاویر پایگاه داده اعمال شده‌اند. نتایج به دست آمده که در شکل ۵.۴ ثبت شده است، نشان دهنده اثر مثبت اعمال بهسازی‌ها روی تصاویر است.

۵.۴ بررسی عملکرد روش خوشه‌بندی ارائه‌شده

برای بررسی عملکرد روش خوشه‌بندی ارائه‌شده، معیار ارزیابی F1-measure در مرحله خوشه‌بندی، نسبت اندازه خوشه‌ها و دقت نهایی شناسایی چهره روش‌های مختلف خوشه‌بندی مورد مطالعه قرار گرفته است.



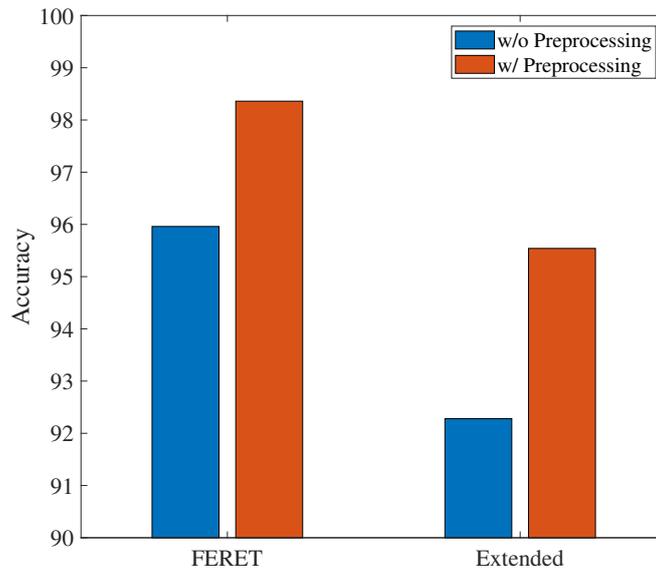
(آ) قبل از پیش‌پردازش
(ب) بعد از پیش‌پردازش
شکل ۲.۴: بهسازی تصویر با استفاده از برابری هیستوگرام



(آ) قبل از پیش‌پردازش
(ب) بعد از پیش‌پردازش
شکل ۳.۴: بهسازی تصویر با استفاده از برابری هیستوگرام وفقی با محدودیت کنتراست



(آ) قبل از پیش‌پردازش
(ب) بعد از پیش‌پردازش
شکل ۴.۴: بهسازی تصویر با استفاده از تنظیم شدت نور



شکل ۵.۴: تأثیر پیش‌پردازش‌ها روی عملکرد روش پیشنهادی روی دو پایگاه داده FERET و مجموعه تصاویر جمع‌آوری‌شده

از معیار F1-measure برای بررسی دو معیار دقت^۱ (p) و بازیابی^۲ (r) توأم استفاده می‌شود. دقت برابر است با تعداد مواردی که به درستی شناسایی شده‌اند نسبت به تعداد کل مواردی که شناسایی شده‌اند (معادله ۱.۴). بازیابی برابر است با تعداد مواردی که به درستی شناسایی شده‌اند نسبت به تعداد مواردی که بایستی شناسایی می‌شدند (معادله ۲.۴). در حقیقت معیار F1-measure میانگین دو پارامتر دقت و بازیابی است که می‌توان به شکل رابطه ۳.۴ نوشت.

$$\text{Precision} = \frac{tp}{tp + fp} \quad (1.4)$$

$$\text{Recall} = \frac{tp}{tp + fn} \quad (2.4)$$

$$\text{F1-measure} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (3.4)$$

در روابط ۱.۴ و ۲.۴ ^۳tp تعداد مواردی که به درستی شناسایی شده‌اند، ^۴fp تعداد مواردی

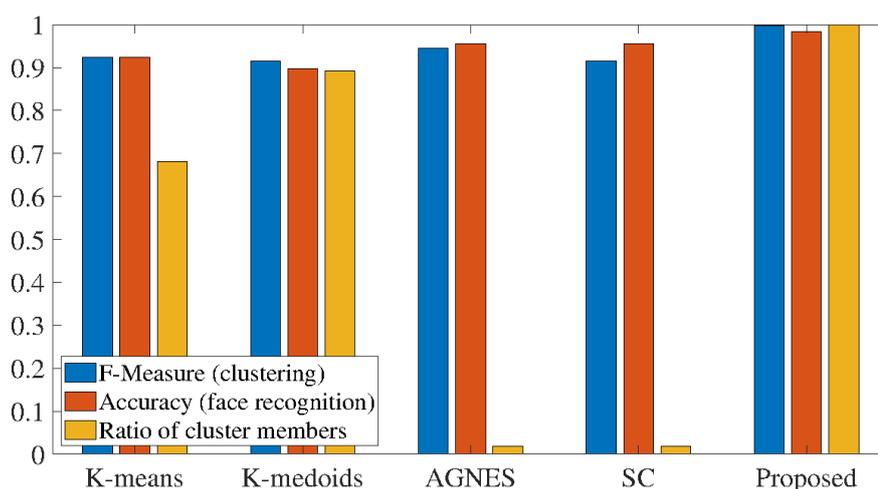
^۱precision
^۲recall
^۳true positive
^۴false positive

که به اشتباه شناسایی شده‌اند و fn ^۱ تعداد مواردی که به درستی شناسایی نشده‌اند را نشان می‌دهند.

روش ارائه شده با روش‌های مرسوم خوشه‌بندی شامل K-means، K-medoids، سلسله مراتبی (AGNES)^۲، DB-Scan و خوشه‌بندی طیفی^۳ مقایسه شده است. البته روش DB-Scan با ویژگی‌های استخراج شده قادر به خوشه‌بندی تصاویر نشد. لذا نتایج این روش مورد بررسی قرار نگرفته است.

بر اساس شکل ۶.۴، روش ارائه شده از نظر معیار F1-measure بهتر عمل کرده است. هرچند مقدار F1-measure و همچنین دقت نهایی دو روش سلسله مراتبی و خوشه‌بندی طیفی نیز بالا می‌باشد؛ ولی خوشه‌هایی که این دو روش می‌سازند، خوشه‌هایی با سایزهای نامتعالی است. به عنوان مثال در یک اجرا به دو خوشه با سایزهای ۲۰ و ۹۷۰ رسیدیم. این درحالی است که روش ارائه شده همیشه تصاویر را به دو خوشه با تعداد مساوی تقسیم خواهد کرد. متعادل بودن هر روش با استفاده از نسبت اندازه خوشه‌ها ارزیابی شده است (معادله ۴.۴). روش خوشه‌بندی پیشنهادی روی مجموعه تصاویر با حجم بالاتر نیز بهتر از بقیه روش‌ها عمل کرده است که نتیجه این ارزیابی در شکل ۷.۴ آمده است.

$$R = \frac{\text{سایز خوشه کوچک‌تر}}{\text{سایز خوشه بزرگ‌تر}} \quad (۴.۴)$$

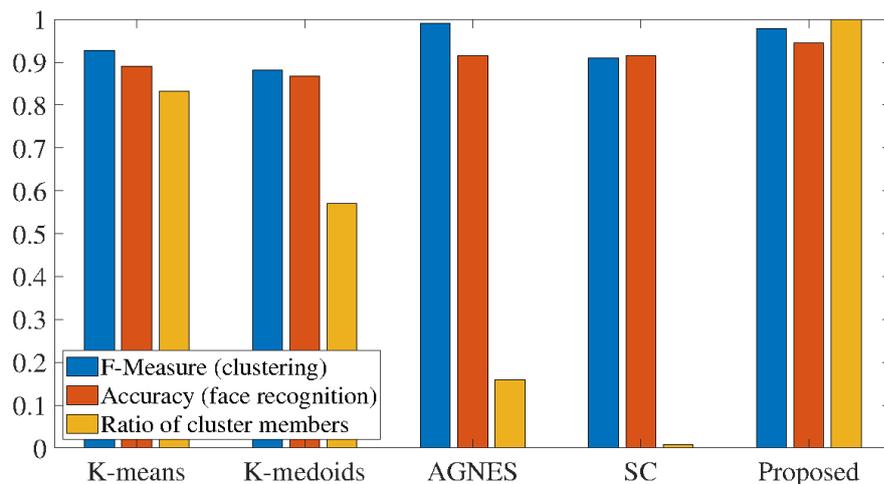


شکل ۶.۴: عملکرد روش خوشه‌بندی پیشنهادی در مقایسه با روش‌های متداول در پایگاه داده FERET

¹false negative

²agglomerative hierarchical clustering

³spectral clustering



شکل ۷.۴: عملکرد روش خوشه‌بندی پیشنهادی در مقایسه با روش‌های متداول در مجموعه تصاویر جمع‌آوری شده

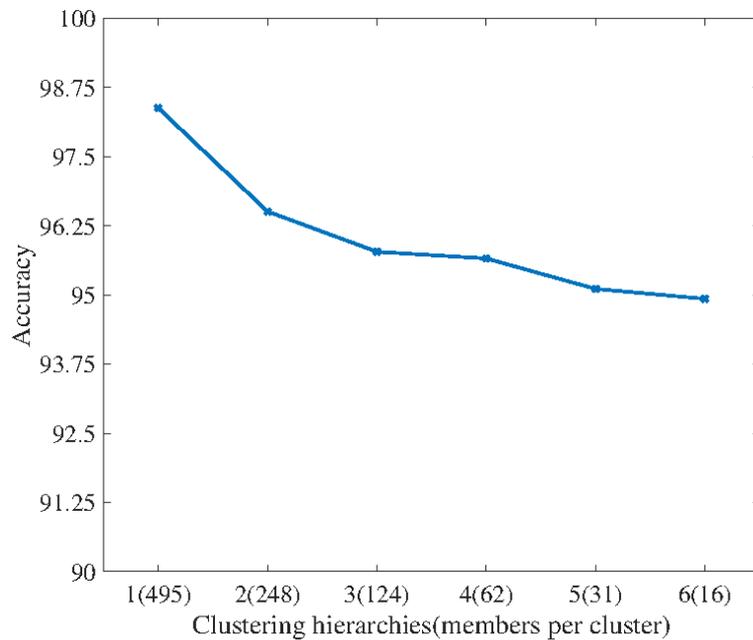
۶.۴ اعمال خوشه‌بندی به صورت سلسله مراتبی

روش خوشه‌بندی ارائه‌شده را می‌توان به صورت سلسله‌مراتبی روی پایگاه داده اعمال کرد. بدین ترتیب می‌توان به تعداد خوشه‌های بیشتری با تعداد اعضای کمتر دست یافت. با اعمال چند مرحله خوشه‌بندی پشت سر هم، یک درخت دودویی از خوشه‌ها تشکیل می‌شود که در هر مرحله هر خوشه به دو خوشه دیگر تقسیم شده است.

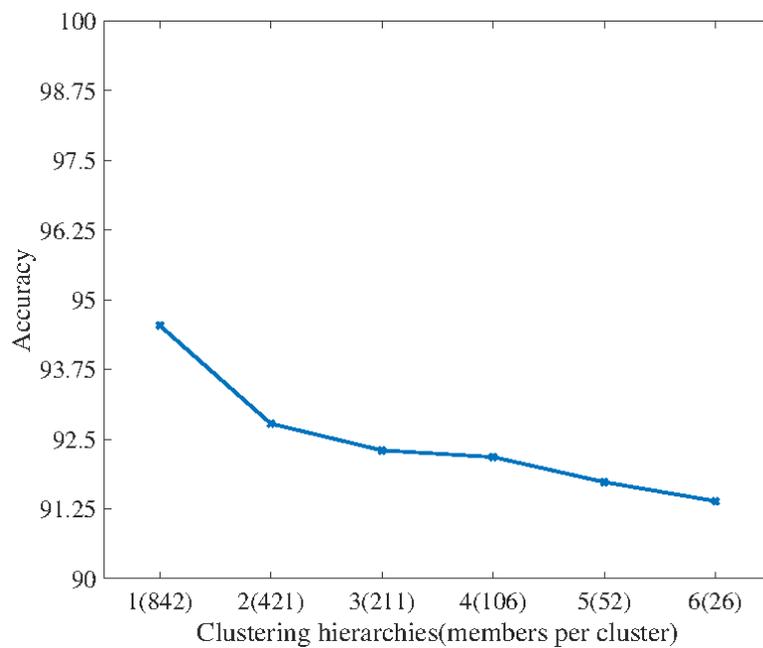
برای بررسی تأثیر استفاده چندمرحله‌ای از خوشه‌بندی، پایگاه داده با تعداد مراحل خوشه‌بندی متفاوت ارزیابی شد که نتایج در شکل‌های ۸.۴ و ۹.۴ آمده است. همان‌طور که مشاهده می‌شود، با افزایش مراحل خوشه‌بندی، دقت نهایی کاهش می‌یابد. بهترین شرایط اجرای روش ارائه‌شده با یک مرحله خوشه‌بندی است.

دقت خوشه‌بندی از اهمیت بالایی برخوردار است، چراکه تأثیر مستقیم روی دقت نهایی شناسایی روش ارائه‌شده دارد. از طرفی بر اساس نتایج به دست آمده هرچه تعداد خوشه‌ها کمتر باشد، دقت ارزیابی افزایش می‌یابد.

با توجه به نتایج به دست آمده در این بخش، و حقیقت این که هرچقدر مراحل خوشه‌بندی را افزایش دهیم، عملکرد نهایی افت می‌کند، سؤالی که مطرح می‌شود این است که آیا اصلاً خوشه‌بندی نیاز است یا خیر. برای پاسخ به این سؤال، ارزیابی دیگری برای بررسی تأثیر جداگانه عوامل اصلی روش پیشنهادی شامل ترکیب ویژگی‌های NMF و FREAK و همچنین خوشه‌بندی انجام شده است که نتایج این ارزیابی در شکل ۱۴.۴ آمده است.



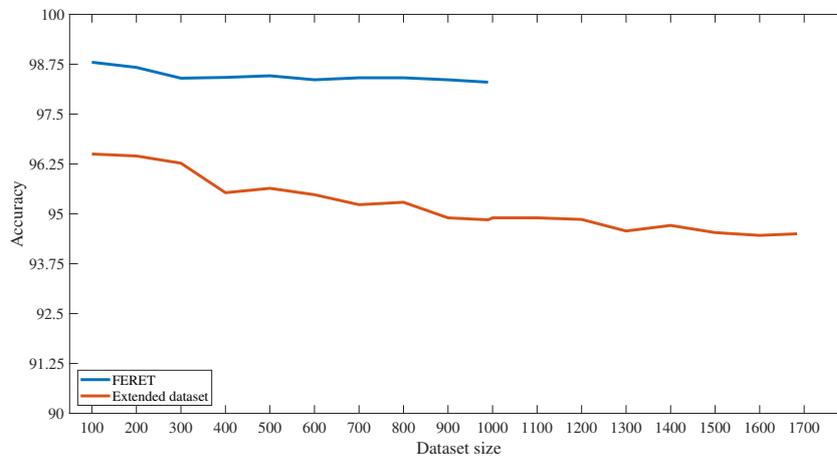
شکل ۸.۴: تأثیر تعداد دفعات خوشه‌بندی در پایگاه داده FERET



شکل ۹.۴: تأثیر تعداد دفعات خوشه‌بندی در مجموعه تصاویر جمع‌آوری شده

۷.۴ تأثیر حجم پایگاه داده

همان‌طور که قبلاً مطرح شد، تمرکز اصلی این تحقیق بر ارائه روشی است که نسبت به حجم پایگاه داده مقاوم بوده و عملکرد خوبی داشته باشد. برای بررسی این موضوع، روش ارائه‌شده با حجم پایگاه داده‌های مختلف ارزیابی می‌شود. ابتدا حجم پایگاه داده ۱۰۰ شخص در نظر گرفته می‌شود. اشخاص به صورت تصادفی از بین تمامی اشخاص پایگاه داده انتخاب می‌شوند. سپس ۲۰ بار این مراحل تکرار می‌شوند. در نهایت میانگین دقت‌های بدست آمده ثبت می‌شود. به ترتیب این عملیات را برای حجم‌های بیشتر نیز اجرا می‌کنیم. در هر مرحله ۱۰۰ نفر به تعداد قبلی اضافه می‌شود؛ در نهایت کل پایگاه داده مورد ارزیابی قرار می‌گیرد.

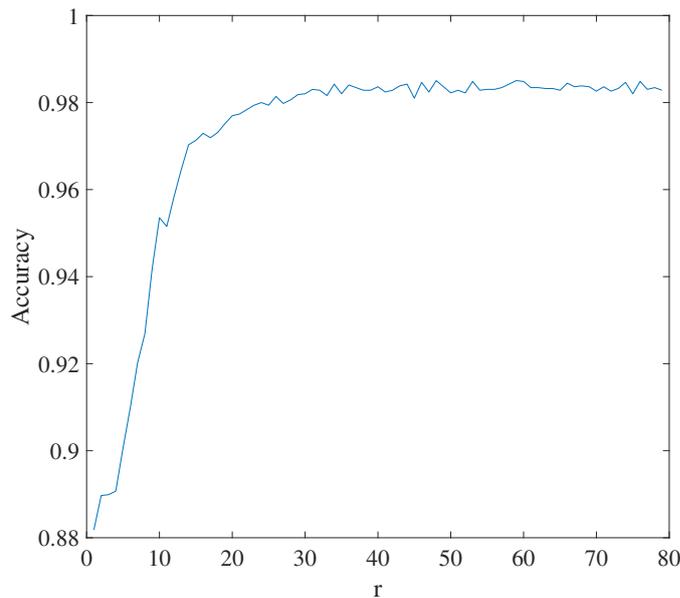


شکل ۱۰.۴: تأثیر اندازه پایگاه داده روی دقت نهایی روش پیشنهادی در دو مجموعه تصاویر استفاده‌شده

با توجه به نتایج این آزمایش که در شکل ۱۰.۴ آمده است، می‌توان قدرت روش پیشنهادی را در برابر حجم زیاد پایگاه داده مشاهده کرد. در هر دو مجموعه تصویر مورد ارزیابی، دقت نهایی روش ارائه‌شده با نرخ خیلی کمی کاهش می‌یابد که همین نکته نشان‌دهنده عملکرد مطلوب این روش در پایگاه داده‌های حجیم است.

۸.۴ پیدا کردن مرتبه در روش NMF

یکی از پارامترهایی که نقش مستقیم در عملکرد روش پیشنهادی دارد، مرتبه روش NMF است. مرتبه همان ابعاد جدید داده است که به مراتب کوچک‌تر از ابعاد داده‌های اصلی می‌باشد. لذا اگر مرتبه استفاده‌شده کمتر و یا بیشتر از مقدار بهینه باشد، روش پیشنهادی عملکرد مناسبی نخواهد داشت. برای پیدا کردن مقدار بهینه این پارامتر، روش پیشنهادی در مقدار مرتبه‌های مختلف چندین بار اجرا شده و در نهایت میانگین دقت به‌دست‌آمده در هر مرتبه ثبت می‌شود. گزارش این آزمایش در شکل ۱۱.۴ رسم شده است.



شکل ۱۱.۴: تأثیر مرتبه (r) روش NMF و تعداد مؤلفه‌های استخراجی روی دقت نهایی روش پیشنهادی در پایگاه داده FERET

با توجه به شکل ۱۱.۴، عملکرد روش پیشنهادی بعد از مقدار مرتبه ۵۵ بهبود چندانی نداشته و می‌توان گفت به مقدار بهینه خود همگرا شده است. نکته دیگری که حائز اهمیت است، مواردی است که در مقدار مرتبه بهینه تأثیر دارند. مواردی همچون تعداد تصاویر و اندازه تصویر ورودی تأثیر مستقیم در مقدار مرتبه بهینه دارند. به همین دلیل در هر مورد، برای پیدا کردن مقدار مرتبه بهینه به صورت سعی و خطا عمل شده است.

۹.۴ تأثیر رزولوشن تصاویر ورودی

رزولوشن تصاویر ورودی به شدت روی عملکرد نهایی تأثیرگذار است. اگر رزولوشن تصویر ورودی خیلی پایین باشد، نمی‌توان اطلاعات مفید و کافی از تصویر استخراج کرد. چراکه با پایین بودن رزولوشن تصویر خیلی از اطلاعات حذف می‌شوند. از طرفی دیگر، اگر رزولوشن تصویر بیش از اندازه بزرگ باشد، میزان اطلاعات نامفید و نویزها در تصویر افزایش می‌یابند. همین امر موجب گمراه شدن دسته‌بند می‌شود. در حالت ایده‌آل رزولوشن تصاویر باید طوری باشد که بتوان به اندازه کافی اطلاعات برای شناسایی هویت از تصویر استخراج کرد.

از آنجایی که رزولوشن تصویر رابطه مستقیم با اندازه آن دارد، برای پیدا کردن رزولوشن بهینه، تصاویر پایگاه داده در اندازه‌های مختلف مورد ارزیابی قرار گرفته‌اند. به عبارتی با تغییر اندازه تصاویر به رزولوشن‌های متفاوتی از آن‌ها می‌رسیم. به دلیل این که در روش پیشنهادی دو نوع ویژگی استخراج می‌شود، به‌طور جداگانه تأثیر اندازه تصاویر برای دو روش NMF و

FREAK مورد ارزیابی قرار گرفته است. در ارزیابی اندازه تصویر بهینه هر ویژگی، اندازه تصاویر برای ویژگی دیگر ثابت فرض شده است. نتایج این ارزیابی برای ویژگی‌های NMF در شکل ۱۲.۴ و برای توصیفگرهای FREAK در شکل ۱۳.۴ آمده است.

همان‌طور که مشاهده می‌شود برای استخراج ویژگی‌های NMF تصاویر در اندازه 20×20 پیکسل اطلاعات کافی را دارا می‌باشند. برای استخراج توصیفگر FREAK نیز تصاویر بایستی در اندازه 124×124 پیکسل باشند. در تصاویر با اندازه کمتر و یا بیشتر، روش استخراج ویژگی قادر به استخراج ویژگی‌های مناسب نخواهد بود.

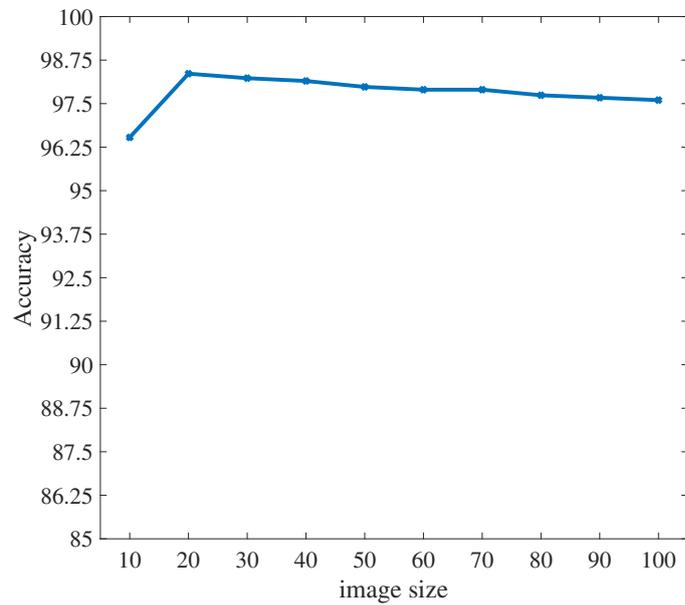
دلیل متفاوت بودن اندازه تصاویر برای دو روش NMF و FREAK ماهیت این دو روش است. روش NMF به دلیل نگرش سراسری که به تصویر دارد، نیازی به تصاویر با رزولوشن بالا نداشته و با رزولوشن پایین می‌تواند به خوبی ویژگی‌های متمایزکننده‌ای استخراج کند. روش FREAK که به صورت محلی استخراج ویژگی انجام می‌دهد، برای محاسبه توصیفگرهای دودویی اطراف هر نقطه کلیدی به مجموعه‌ای از جفت‌نقاط احتیاج دارد. به دلیل این که الگوی نمونه‌برداری این روش یک محدوده 64×64 پیکسلی اطراف هر نقطه کلیدی را شامل می‌شود، برای استخراج توصیفگرهای FREAK به اندازه‌های بیشتر از 64×64 پیکسل نیاز داریم. لذا ماهیت این دو روش باعث می‌شود تا برای استخراج ویژگی از اندازه و رزولوشن‌های متفاوت برای هر کدام استفاده کنیم.

۱۰.۴ تأثیر معیار فاصله

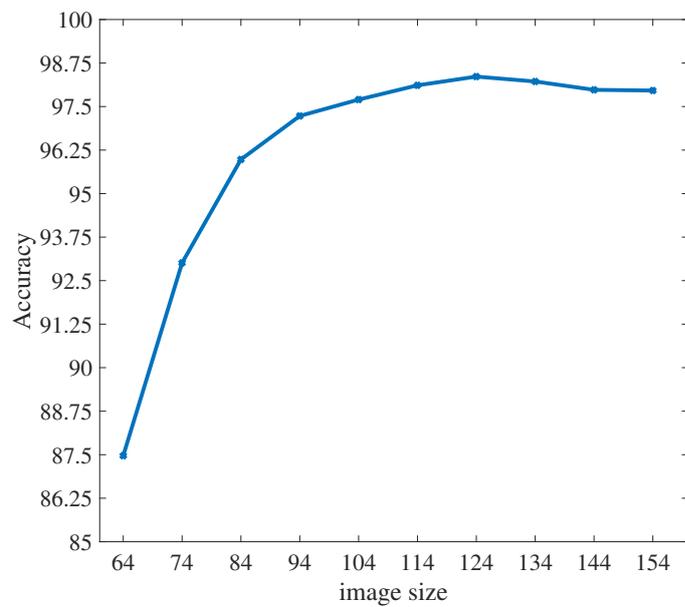
از آنجایی که یکی از مهم‌ترین قدم‌های روش پیشنهادی، انتخاب نزدیک‌ترین همسایه است، بایستی یک معیار فاصله (شباهت) مناسب برای مسئله، مورد استفاده قرار گیرد. معیار فاصله‌ای مناسب‌تر است که بتواند فاصله تصاویر مشابه را کوچک و فاصله تصاویر متفاوت را بزرگ نشان دهد.

به دلیل ماهیت دودویی توصیفگرهای FREAK با استفاده از معیار فاصله همینگ^۱ فاصله تصویر آزمون با تصاویر داخل خوشه محاسبه و ماتریس فاصله برای ویژگی FREAK تشکیل می‌شود. برای محاسبه فاصله ویژگی‌های NMF تصویر آزمون با تصاویر داخل خوشه بایستی یک معیار ویژگی مناسب تعیین شود. در این آزمایش معیارهای فاصله متفاوتی همچون معیار اقلیدسی، همبستگی، کسینوسی، بلوک شهری و مینکوفسکی، مورد ارزیابی قرار گرفتند. بدین منظور در مرحله تطابق تصاویر داخل خوشه‌ها، طی چندین بار اجرا، با استفاده از معیارهای اشاره‌شده، فاصله بین ویژگی‌های NMF تصویر آزمون و تصاویر داخل خوشه محاسبه و ماتریس فاصله‌ها تشکیل می‌شود. نتایج این آزمایش در جدول ۳.۴ آمده است.

^۱Hamming



شکل ۱۲.۴: تأثیر اندازه تصاویر در استخراج ویژگی NMF و دقت نهایی روی پایگاه داده FERET



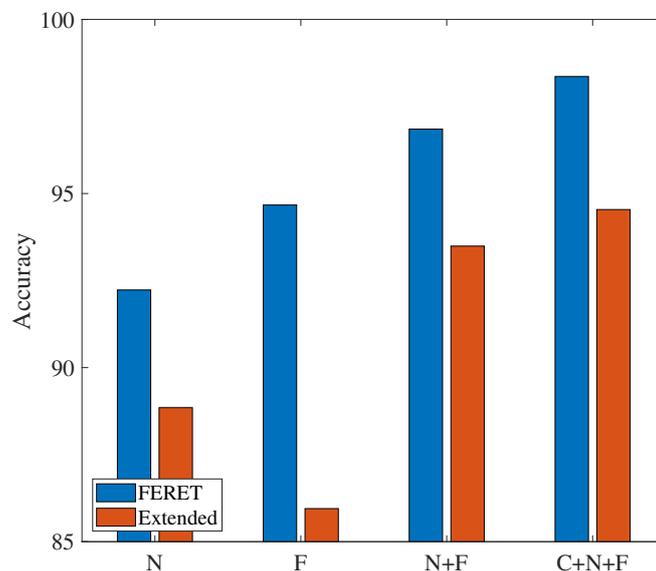
شکل ۱۳.۴: تأثیر اندازه تصاویر در استخراج توصیفگرهای FREAK و دقت نهایی روی پایگاه داده FERET

جدول ۳.۴: تأثیر معیار فاصله در عملکرد روش ارائه‌شده

روش	دقت (%)
اقلیدسی	۹۷/۴۳
همبستگی	۹۷/۵۸
کسینوسی	۹۷/۴۱
بلوک شهری	۹۸/۳۶
مینکوفسکی	۹۷/۴۷

۱۱.۴ تأثیر جداگانه نوآوری‌ها

برای بررسی بیشتر نوآوری‌های روش ارائه‌شده و همچنین پاسخ به سؤال مطرح‌شده در بخش ۶.۴، تأثیر جداگانه هر نوآوری مورد ارزیابی قرار گرفته است. بدین منظور روش ارائه‌شده در شرایط مختلف شامل استفاده جداگانه و یا ترکیبی از ویژگی‌های NMF، FREAK و مبحث خوشه‌بندی ارزیابی شده است. نتایج این آزمایش در شکل ۱۴.۴ آمده است.



شکل ۱۴.۴: تأثیر جداگانه بخش‌های مختلف روش ارائه‌شده روی پایگاه داده FERET و مجموعه تصاویر جمع‌آوری‌شده. حروف N، F و C به ترتیب نشان‌دهنده ویژگی‌های NMF، توصیفگرهای FREAK و مبحث خوشه‌بندی هستند.

با بررسی این نتایج می‌توان به تأثیر جداگانه هر عامل در عملکرد نهایی پی برد. همان‌طور که مشاهده می‌شود، هرکدام از این عوامل به‌طور جداگانه عملکرد ضعیف‌تری داشته‌اند؛ ولی با

ترکیب این عوامل بیشترین بهبود در مسئله حاصل شده و از بیشینه قدرت هر یک از نوآوری‌ها استفاده شده است.

برای توجیه نتایج به دست آمده در ارزیابی بخش ۶.۴ که نشان‌دهنده تأثیر منفی مراحل خوشه‌بندی بیشتر روی عملکرد نهایی بود، و همچنین نتایج به دست آمده در این بخش که نشان‌دهنده تأثیر مثبت خوشه‌بندی روی عملکرد نهایی است، بایستی به دقت خوشه‌بندی و دقت شناسایی داخل هر خوشه توجه کرد. در حالتی که از خوشه‌بندی استفاده نمی‌شود، تعداد مقایسه‌ها بالا و توانایی شناسایی میان تعداد تصاویر زیاد، پایین می‌باشد. زمانی که یک بار خوشه‌بندی انجام می‌شود، می‌توان با دقت نزدیک ۱۰۰٪ تصاویر را به دو خوشه با اندازه‌های مساوی تقسیم کرد. در این حالت دقت شناسایی داخل خوشه نیز بالاتر از زمانی است که شناسایی میان تمامی تصاویر صورت می‌گیرد. ترکیب دقت بالای خوشه‌بندی و دقت بالای شناسایی داخل خوشه‌ها منجر به بهبود در عملکرد نهایی روش پیشنهادی می‌شود. در ادامه با اضافه کردن مراحل خوشه‌بندی به دلیل افت دقت خوشه‌بندی، عملکرد نهایی نیز تحت تأثیر این عامل، افت خواهد کرد.

برای مقایسه نهایی، عملکرد چندین روش دیگر که در موضوع همین تحقیق بودند، جمع‌آوری و گزارش شدند. برخی از این روش‌ها در فصل ۲ مورد مطالعه قرار گرفتند. نتایج این مقایسه در جدول ۴.۴ گزارش شده است. نتایج ثبت شده در این جدول از متن تحقیقات انتشاریافته آورده شده است. نتیجه گزارش شده در روش نیکان و حسن‌پور [۴۴] برای ۹۹۰ تصویر از پایگاه داده FERET بود که ما برای بررسی بیشتر این روش را پیاده‌سازی و روی مجموعه تصاویر جمع‌آوری شده نیز آزمایش کرده و در جدول ثبت کردیم. تعداد تصاویر استفاده شده در عمده این روش‌ها کمتر از تعداد تصاویر استفاده شده توسط روش پیشنهادی است. تصاویر استفاده شده در روش‌هایی که از ۹۹۰ تصویر پایگاه داده FERET استفاده کرده‌اند، یکسان و تصاویر استفاده شده در روش‌هایی که از تعداد تصاویر کمتر از ۹۹۰ تصویر استفاده کرده‌اند، مجموعه انتخابی از تصاویر کل پایگاه داده می‌باشد.

روش‌های SPCA+ [۶۵]، FLDA-SVD [۶۶]، DMMA [۳۵]، Fast-DMMA [۶۷]، LU [۶۸] و LPP [۶۹] رویکردی سنتی^۱ برای حل مشکل یک نمونه به‌ازای هر شخص ارائه کردند که عملکردشان روی تعداد تصاویر کم (۲۰۰ تصویر) قابل قبول بود و با افزایش تعداد تصاویر شکست می‌خوردند. روش TDL [۴۹] که برای حل مشکل یک نمونه به‌ازای هر شخص از یک روش افزایش نمونه^۲ استفاده کرد و سپس با استفاده از نمونه‌های جدید یک شبکه عمیق را fine-tune کردند. روش KCFT [۵۰] نیز ابتدا یک شبکه عمیق را آموزش داده سپس با استفاده از یک روش افزایش نمونه در فضای ویژگی نمونه‌های جدید ایجاد کرده و در نهایت شناسایی را انجام می‌دهد. روش ML-EM [۷۰] برای حل مشکل تعداد زیاد تصاویر در پایگاه داده با استفاده از خوشه‌بندی مبتنی بر روش EM تصاویر را به خوشه‌های کوچک‌تر تقسیم کرد. روش FFT

¹traditional

²sample expansion

[۴۶] با استفاده از فیلتر مناسب در فضای فرکانس ویژگی‌های مفید استخراج کرد. روش STFT-LBP-HOG [۴۷] از بردار ویژگی به‌دست‌آمده از ترکیب ویژگی‌های قوی محلی و سراسری استفاده کردند که بهترین عملکرد را میان روش‌های پیشین داشته است. روش NMF [۴۴] که بر روی سرعت بازیابی نیز تمرکز داشت، روش پایه و اساس روش پیشنهادی این تحقیق است.

روش‌های سنتی به دلیل ضعف در پایگاه داده‌های حجیم، بهترین نتایجشان روی مجموعه تصاویر کوچک‌تری بود. روش‌های یادگیری عمیق نیز به دلیل نیاز به نمونه‌های آموزشی زیاد، قادر به حل مشکل یک نمونه به‌ازای هر شخص نیستند. هرچند نمونه کارهای عمیق ذکرشده در جدول ۴.۴ که به‌منظور حل مشکل یک نمونه به‌ازای هر شخص مطرح‌شده‌اند، با محدودیت‌ها و ضعف‌های زیادی روبرو هستند که یکی از نقاط ضعف، عملکرد پایین روی مجموعه تصاویر بزرگ است. روش پیشنهادی که از ترکیب دو ویژگی NMF و FREAK استفاده کرده‌است، توانسته عملکردی بهتر از ترکیب سه ویژگی قوی STFT، LBP و HOG در روش STFT-LBP-HOG [۴۷] داشته‌باشد. در روش STFT-LBP-HOG تعداد پارامترهای مسئله زیاد بوده و طول بردارهای ویژگی نیز بلند می‌باشد. از طرفی دیگر در روش پیشنهادی با استفاده از ویژگی‌های FREAK به‌ویژگی‌های بهتری نسبت به ویژگی‌های روش STFT-LBP-HOG دست‌یافتیم که همین امر باعث عملکرد بهتر روش پیشنهادی شده‌است. روش پیشنهادی در مقایسه با روش NMF [۴۴] نیز از نوآوری‌هایی مانند ترکیب ویژگی‌های NMF با FREAK و رویکرد خوشه‌بندی استفاده کرده‌است که باعث بهبود عملکرد نهایی شده‌است. در نهایت با مقایسه کلی روش‌های پیشین و روش پیشنهادی، مشاهده می‌شود که در چالش‌های مطرح‌شده روش پیشنهادی توانسته عملکرد بهتری داشته‌باشد. به‌طور همزمان دو چالش پایگاه داده حجیم و یک نمونه به‌ازای هر شخص در روش پیشنهادی حل شده‌است. این نکته نشان‌دهنده قدرت روش ارائه‌شده در شناسایی چهره است.

۱۲.۴ جمع‌بندی

در این فصل، روش پیشنهادی از چندین جنبه مورد ارزیابی قرار گرفت. به لحاظ استفاده از خوشه‌بندی، با استفاده از معیار F1-measure برتری روش خوشه‌بندی پیشنهادی نسبت به روش‌های معمول دیگر نشان داده شد. علاوه بر این، متعادل بودن اعضای خوشه در روش پیشنهادی امتیاز محسوب می‌شود. برای بررسی عملکرد روش ارائه‌شده در پایگاه داده‌های حجیم‌تر یک مجموعه تصویر جمع‌آوری‌شده از چندین پایگاه داده استاندارد کوچک‌تر ارائه شد. سپس بر اساس چندین آزمایش مقاومت روش ارائه‌شده نسبت به پایگاه داده حجیم ارزیابی‌شده و دقت نهایی این روش با روش‌های دیگر مقایسه شد.

جدول ۴.۴: مقایسه عملکرد روش پیشنهادی با کارهای مرتبط

دقت (%)	نمونه به ازای هر شخص	اندازه پایگاه داده	روش
۸۷/۴۹	۱	۲۰۰	[۶۵] SPCA+
۹۵	۱	۶۹۵	[۷۰] ML-EM
۹۰/۵	۱	۲۰۰	[۶۶] FLDA-SVD
۹۳	۱	۲۰۰	[۳۵] DMMA
۹۱	۱	۲۰۰	[۶۷] Fast-DMMA
۹۱/۳۰	۱۰	۴۹۰	[۶۸] LU
۸۹	۱	۲۰۰	[۶۹] LPP
۹۳/۹	۱	۲۰۰	[۴۹] TDL
۹۳/۱۶	۱	۲۰۰	[۵۰] KCFT
۹۰/۴	۱	۹۹۰	[۴۶] FFT
۹۶/۶۶	۱	۹۹۰	[۴۷] STFT-LBP-HOG
۹۲/۷۳	۱	۹۹۰	[۴۴] NMF
۹۸/۳۶	۱	۹۹۰	روش پیشنهادی
۸۹/۸۹	۱	۱۶۸۴	NMF
۹۴/۵۴	۱	۱۶۸۴	روش پیشنهادی

فصل ۵

جمع‌بندی و کارهای آینده

۱.۵ خلاصه و نتیجه‌گیری

در فصل ۱ به‌طور مختصر درباره موضوع شناسایی چهره و چالش‌های موجود در این زمینه صحبت شد. هدف این تحقیق تمرکز روی چالش‌های پایگاه داده حجیم و تعداد نمونه‌های کم برای هر شخص است و ترکیب روش خوشه‌بندی آستانه‌گذاری به همراه دو ویژگی NMF و FREAK به‌عنوان نوآوری‌های روش ارائه‌شده معرفی شدند. به‌طور خلاصه روش‌هایی که برای این مشکلات ارائه‌شده‌اند مطرح شد.

چندین روش مطرح در این زمینه که به حل مشکلات پایگاه داده حجیم و تعداد نمونه کم پرداخته بودند، در فصل ۲ به‌طور مفصل توضیح داده شد. روش پیشنهادی این تحقیق به‌طور مفصل و با جزئیات در فصل ۳ مطرح شد. در روش ارائه‌شده ابتدا تصاویر با استفاده از روش Viola-Jones به دو قسمت چهره و غیر چهره تقسیم می‌شوند. سپس تصاویر چهره استخراج شده از لحاظ روشنایی و کنتراست بهسازی می‌شوند. روش پیشنهادی از یک رویکرد خوشه‌بندی ساده برای تقسیم تصاویر پایگاه داده به چندین دسته استفاده می‌کند.

برای خوشه‌بندی، ابتدا تصاویر آزمون به دو بخش بالا و پایین تقسیم می‌شوند. از هر دو بخش به‌طور جداگانه، ویژگی‌های NMF استخراج می‌شوند. میانگین دو بردار ویژگی بالا و پایین هر تصویر محاسبه‌شده و از هم کم می‌شوند. با استفاده از یک حد آستانه و اختلاف میانگین ویژگی‌های بالا و پایین تصاویر مشخص می‌کنیم که تصویر متعلق به کدام خوشه است.

سپس برای تشخیص خوشه تصویر آزمون نیز، این تصویر به دو بخش بالا و پایین تقسیم‌شده و ویژگی‌های NMF از دو بخش آن استخراج می‌شود. سپس فاصله ویژگی‌های تصویر آزمون با میانگین خوشه‌ها (نماینده) محاسبه و نزدیک‌ترین خوشه مشخص می‌گردد. اگر فاصله نزدیک‌ترین خوشه از حد آستانه کمتر باشد، همان خوشه به‌عنوان خوشه تصویر آزمون تعیین می‌شود. در غیر این صورت از روش نزدیک‌ترین همسایه بر روی ماتریس فاصله تصویر آزمون با تمامی تصاویر استفاده می‌شود. بدین صورت که خوشه نزدیک‌ترین تصویر به‌عنوان خوشه تصویر آزمون تعیین می‌گردد.

پس از مشخص شدن خوشه تصویر آزمون، ویژگی‌های NMF و توصیفگرهای FREAK از تصویر آزمون و تصاویر داخل خوشه استخراج می‌شود. فاصله ویژگی‌های تصویر آزمون با تصاویر داخل خوشه محاسبه و با یکدیگر ترکیب می‌شوند. درنهایت هویت تصویر آزمون با استفاده از روش نزدیک‌ترین همسایه و معیار فاصله بلوک شهری تعیین می‌گردد.

برای ارزیابی روش پیشنهادی دو پایگاه داده حجیم استفاده شدند. پایگاه داده استاندارد FERET که شامل ۹۹۰ شخص است و یک مجموعه تصویر جمع‌آوری شده از پایگاه داده‌های استاندارد دیگر که حجم بیشتری دارد. این مجموعه تصاویر شامل تصاویر از زاویه روبروی ۱۶۸۴ شخص است. عملکرد قسمت خوشه‌بندی روش پیشنهادی و همچنین استخراج ویژگی در شناسایی چهره مورد ارزیابی قرار گرفت. نتایج این آزمایش‌ها در فصل ۴ گزارش شده‌است. بر اساس نتایج به دست آمده، روش خوشه‌بندی ارائه‌شده، بهتر از روش‌های خوشه‌بندی معمول دیگر عمل کرده‌است. از لحاظ شناسایی چهره نیز دقت نهایی روش پیشنهادی، بهتر از روش‌های قبلی است.

۲.۵ کارهای آینده

روش ارائه‌شده با این که در مقابل حجم بالای تصاویر مقاوم است؛ ولی همچنان جای بهبود دارد. این روش روی تصاویری که از روبرو هستند، نتیجه خوبی دارد. درحالی که اگر چهره‌ها دارای زاویه و یا چرخش باشند، روش‌های استخراج ویژگی استفاده‌شده در این روش قادر به استخراج ویژگی مناسب نبوده و درنهایت نمی‌توانند به‌درستی شناسایی انجام دهند. بنابراین تحقیقات برای مقاوم‌سازی روش پیشنهادی نسبت به چرخش صورت و تغییرات زاویه ادامه خواهد داشت.

در آزمایش‌ها دیدیم که اگر چندین مرحله از خوشه‌بندی استفاده کنیم، دقت نهایی کمتر از زمانی است که یک بار خوشه‌بندی استفاده شده‌است. یکی از پیشنهاد‌های آینده این است که روش خوشه‌بندی طوری بهبود یابد که توانایی خوشه‌بندی چندمرحله‌ای به‌صورت سلسله مراتبی با دقت و عملکرد بهتر فراهم شود تا تعداد خوشه‌های نهایی بیشتر و درنتیجه تعداد مقایسه‌های درون خوشه‌ای به‌مراتب خیلی کمتر شود. در این صورت زمان و دقت روش پیشنهادی بهبود می‌یابد. به عبارتی دیگر دقت خوشه‌بندی باید در حالت ایده‌آل نزدیک ۱۰۰٪

باشد. در این صورت با اضافه کردن تعداد خوشه‌ها می‌توان به راحتی دقت نهایی را بهبود داد. در نهایت می‌توان از پتانسیل روش‌های استخراج دیگر نیز در این کار بهره برد. مطالعه روش‌های استخراج ویژگی و روش‌هایی که نتایج مطلوبی دارند ادامه خواهد داشت، تا بتوان از این روش‌ها به صورت ترکیبی با روش پیشنهادی استفاده کرد. در این صورت می‌توان روش بهبودیافته را روی پایگاه داده‌های حجیم‌تر و حتی تصاویری که در شرایط کنترل نشده ثبت شده‌اند، اعمال کرد و انتظار نتیجه خوبی را داشت.

مراجع

- [١] Taskiran, M., Kahraman, N., & Erdem, C. E. (2020). "Face recognition: Past, present and future (a review)". *Digital Signal Processing*, 106.
- [٢] Kortli, Y., Jridi, M., Al Falou, A., and Atri, M. (2020), "Face Recognition Systems: A Survey". *Sensors*, 20, 2, p. 342.
- [٣] Annalakshmi, M., Roomi, S. M. M., and Naveedh, A. S. (2019), "A hybrid technique for gender classification with SLBP and HOG features". *Cluster Computing*, 22, S1, pp. 11–20.
- [٤] Zou, W. and Yuen, P. C. (2011), "Very Low Resolution Face Recognition Problem". *IEEE Transactions on Image Processing*, 21, pp. 327-340.
- [٥] Wang, Z., Wang, Z., Jonathan Wu, Q. M., Wan, Y., and Tang, Z. (2014), "Low-resolution face recognition: a review". *The Visual Computer*, 30, pp. 359-386
- [٦] Stan, Z. Li., Anil, K. Jain (2011), "*Handbook of Face Recognition*". Vol. 1, Springer London, UK. 1, pp. 7-9
- [٧] Moses, Y., Adini, Y., Ullman, S (1994), "Face Recognition: the Problem of Compensating for Changes in Illumination Direction". Vol. A, *Proceedings of the European Conference on Computer Vision*, pp. 286–296.
- [٨] Li, S. Z., Xin Wen Hou, Hong Jiang Zhang, and Qian Sheng Cheng. (2001). "Learning spatially localized, parts-based representation". *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001*, 1, 1-207-1–212.

- [۹] Yang, H., & Wang, X. A. (2016). "Cascade classifier for face detection". *Journal of Algorithms & Computational Technology*, 10(3), pp. 187–197.
- [۱۰] Ahonen, T., Hadid, A., and Pietikäinen, M. (2004). "Face Recognition with Local Binary Patterns". *Computer Vision - ECCV*, Vol. 3021, pp. 469–481.
- [۱۱] Hao, L., & Yu, W. (2019). "Face Image Set Recognition Based On Improved HOG-NMF and Convolutional Neural Networks". *Progress in Human Computer Interaction*, 2(1).
- [۱۲] Zhou, H., Yuan, Y., & Shi, C. (2009). "Object tracking using SIFT features and mean shift". *Computer Vision and Image Understanding*, 113(3), 345–352.
- [۱۳] Du, G., Su, F., and Cai, A. (2009), "Face recognition using SURF features". *MIPPR: Pattern Recognition and Computer Vision*, 7496, pp. 749628
- [۱۴] Turk, M. and Pentland, A. (1991), "Eigenfaces for Recognition". *Journal of Cognitive Neuroscience*, 3, 1, pp 71–86
- [۱۵] Simonyan, K., Parkhi, O., Vedaldi, A., & Zisserman, A. (2013). "Fisher Vector Faces in the Wild". *Proceedings of the British Machine Vision Conference*, 8.1-8.11.
- [۱۶] Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). "Face recognition by independent component analysis". *IEEE Transactions on Neural Networks*, 13(6), 1450–1464.
- [۱۷] Mika, S., Schölkopf, B., Smola, A., Müller, K., Scholz, M., Rätsch, G. (1999). "Kernel PCA and De-Noising in Feature Spaces". 11.
- [۱۸] بیگلری م. قویدل نیچران ج. میرزائی ف. سلیمانی ع. (۱۳۹۳)، "استخراج ویژگی کارا با استفاده از فیلتر گابور، LBP و LDA برای شناسایی چهره"، ششمین کنفرانس فناوری اطلاعات و دانش، شاهرود
- [۱۹] Moussa, M., Hmila, M., and Douik, A. (2018), "A Novel Face Recognition Approach Based on Genetic Algorithm Optimization". *Studies in Informatics and Control*, 27, 1

- [٢٠] Tuncer, T., Dogan, S., Abdar, M., & Pławiak, P. (2020). "A novel facial image recognition method based on perceptual hash using quintet triple binary pattern". *Multimedia Tools and Applications*.
- [٢١] Otto, C., Wang, D., & Jain, A. K. (2016). "Clustering Millions of Faces by Identity". ArXiv:1604.00989.
- [٢٢] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). "FaceNet: A unified embedding for face recognition and clustering". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823.
- [٢٣] Xi, M. (2015). "Local Binary Pattern Network: A Deep Learning Approach For Face Recognition". 89.
- [٢٤] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). "ImageNet classification with deep convolutional neural networks". *Communications of the ACM*, 60(6), 84–90.
- [٢٥] Kim, J., & Park, H. (2011). "Fast Nonnegative Matrix Factorization: An Active-Set-Like Method and Comparisons". *SIAM Journal on Scientific Computing*, 33(6), 3261–3281.
- [٢٦] Alahi, A., Ortiz, R., & Vandergheynst, P. (2012). "FREAK: Fast Retina Keypoint". *IEEE Conference on Computer Vision and Pattern Recognition*, 510–517.
- [٢٧] Phillips, P.J., Wechsler, H., Huang, J., Rauss, P. (1998), "The FERET database and evaluation procedure for face recognition algorithms". *Image and Vision Computing*, 16, 5, pp. 295–306
- [٢٨] Ojala T, Pietikäinen M, Harwood D. (1996), "A comparative study of texture measures with classification based on featured distributions". *Pattern Recognit*, 29, pp 51-59
- [٢٩] Lumaban, P., B, M. (2020). "CCTV-Based Surveillance System with Face Recognition Feature". *International Journal of Advanced Trends in Computer Science and Engineering*, 9, pp 349–355
- [٣٠] Huang, C. and Huang, J. (2017), "A Fast HOG Descriptor Using Lookup Table and Integral Image". arXiv:1703.06256

- [٣١] Lenc, L. and Král, P. (2015), "Automatic face recognition system based on the SIFT feature". *Computers & Electrical Engineering*, 46, pp. 256–272
- [٣٢] Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. (2008). "Speeded-Up Robust Features (SURF)". *Computer Vision and Image Understanding*, 110, pp, 346–359.
- [٣٣] Anand. B. (2016), "Face Recognition using SURF Features and SVM Classifier". *International Journal of Electronics Engineering Research*
- [٣٤] Nazari, A., Dehghan, H., Zehtabian, A., and Hassanpour, H. (2016), "Gender classification based on fuzzy clustering and principal component analysis". *IET Computer Vision*, 10, 3, pp 228–233
- [٣٥] Lu, J., Tan, Y.-P., & Wang, G. (2013). "Discriminative Multi-Manifold Analysis for Face Recognition from a Single Training Sample per Person". 8.
- [٣٦] Duong, V.-H., Bui, M.-Q., & Wang, J.-C. (2019). "Matrix Factorization on Complex Domain for Face Recognition". *Visual Object Tracking with Deep Neural Networks, IntechOpen*.
- [٣٧] Wild, S., Curry, J., & Dougherty, A. (2004). "Improving non-negative matrix factorizations through structured initialization". *Pattern Recognition*, 37(11), 2217–2232.
- [٣٨] Nikolaus, R. (2007). "Learning the parts of objects using non-negative matrix factorization (NMF)". *MMer Team*.
- [٣٩] Wang, Y., Jia, Y., Hu, C., & Turk, M. (2005). "Non-negative Matrix Factorization Framework For Face Recognition". *International Journal of Pattern Recognition and Artificial Intelligence*, 19(04), 495–511.
- [٤٠] Rezaei, M., & Derhami, V. (2018). "Improving LNMF Performance of Facial Expression Recognition via Significant Parts Extraction using Shapley Value". *Journal of AI and Data Mining*.

- [۴۱] Lee, D. D., and Seung, H. S., (2000). "Algorithms for non-negative matrix factorization". *Advances in Neural Information Processing Systems*, 13, pp.556–562.
- [۴۲] Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., and Plemmons, R. J. (2007). "Algorithms and applications for approximate nonnegative matrix factorization". *Computational Statistics & Data Analysis*, 52(1), 155–173.
- [۴۳] Lin, C.-J. (2007). "Projected Gradient Methods for Nonnegative Matrix Factorization". *Neural Computation*, 19(10), 2756–2779.
- [۴۴] نیکان ف، حسن پور ح، (۱۳۹۷)، پایان نامه ارشد: "استخراج ویژگی از تصاویر چهره در پایگاه داده حجیم"، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود،
- [۴۵] Rashid, S. J., Abdullah, A. I., & Shihab, M. A. (2020). "Face Recognition System Based on Gabor Wavelets Transform, Principal Component Analysis and Support Vector Machine". *International Journal on Advanced Science, Engineering and Information Technology*, 10(3), 959.
- [۴۶] بخشی م، حسن پور ح، فاتح م. (۱۳۹۷). "استخراج ویژگی های مکان-فرکانسی جهت بازیابی تصویر چهره از پایگاه داده حجیم تصاویر"، مجله مهندسی برق دانشگاه تبریز، دوره ۴۸، شماره ۲، صفحه ۵۰۹–۵۱۷
- [۴۷] نوریان خ، خسروی ح، (۱۳۹۶)، پایان نامه ارشد: "احراز هویت چهره با استفاده از ویژگی های مکان و فرکانس تصویر در مجموعه های بزرگ"، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود،
- [۴۸] Xin, M., Zhou, Y., and Yan, J. (2015). "Single sample face recognition using lgbp and locality preserving discriminant analysis". *Applied Mathematics and Information Sciences*, 9, 353–360.
- [۴۹] Zeng, J., Zhao, X., Gan, J., Mai, C., Zhai, Y., & Wang, F. (2018). "Deep Convolutional Neural Network Used in Single Sample per Person Face Recognition". *Computational Intelligence and Neuroscience*, 1–11.
- [۵۰] Min, R., Xu, S., & Cui, Z. (2019). "Single-Sample Face Recognition Based on Feature Expansion". *IEEE Access*, 7, 45219–45229.

-
- [٥١] Cover, T., & Hart, P. (1967). "Nearest neighbor pattern classification". *IEEE Transactions on Information Theory*, 13(1), 21–27.
- [٥٢] Viola, P., & Jones, M. (2001). "Rapid object detection using a boosted cascade of simple features". *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 1, I-511-I-518.
- [٥٣] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks". *IEEE Signal Processing Letters*, 23(10), 1499–1503.
- [٥٤] V, S., L, L., A, A., & Kumar N, P. (2019). "Histogram Equalization Techniques in Image Enhancement". *Journal of Information Technology & Software Engineering*, 09(01).
- [٥٥] Reza, A. M. (2004). "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement". *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, 38(1), 35–44.
- [٥٦] Jianbo Shi, & Tomasi. (1994). "Good features to track". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. CVPR-94*, 593–600.
- [٥٧] Harris, C. and Stephens, M. (1988). "A combined corner and edge detector". *In Proceedings of the Alvey Vision Conference*, pp. 23.1–23.6.
- [٥٨] Rosten, E., & Drummond, T. (2006). "Machine Learning for High-Speed Corner Detection". *Computer Vision – ECCV 2006*, Vol. 3951, pp. 430–443.
- [٥٩] Huiping Li, Doermann, D., and Kia, O. (2000). "Automatic text detection and tracking in digital video". *IEEE Transactions on Image Processing*, 9(1), 147–156.
- [٦٠] Wang, H., & Schmid, C. (2013). "Action Recognition with Improved Trajectories". *IEEE International Conference on Computer Vision*, 3551–3558.

- [٤١] Milborrow, S., Morkel, J., and Nicolls, F. (2010). "The MUCT Land-marked Face Database". *Pattern Recognition Association of South Africa*.
- [٤٢] Thomaz, C. E., & Giraldi, G. A. (2010). "A new ranking method for principal components analysis and its application to face image analysis". *Image and Vision Computing*, 28(6), 902–913.
- [٤٣] Abdullah-Al-Wadud, M., Kabir, M., Akber Dewan, M., & Chae, O. (2007). "A Dynamic Histogram Equalization for Image Contrast Enhancement". *IEEE Transactions on Consumer Electronics*, 53(2), 593–600.
- [٤٤] Solomon, C., & Breckon, T. (2011). "*Fundamentals of Digital Image Processing*". Wiley-Blackwell, pp, 62-63.
- [٤٥] Zhang, D., Chen, S., & Zhou, Z.-H. (2005). "A new face recognition method based on SVD perturbation for single example image per person". *Applied Mathematics and Computation*, 163(2), 895–907.
- [٤٦] Gao, Q., Zhang, L., & Zhang, D. (2008). "Face recognition using FLDA with single training image per person". *Applied Mathematics and Computation*, 205(2), 726–734.
- [٤٧] Liu, H.-H., Hsu, S.-C., & Huang, C.-L. (2014). "Single-sample-per-person-based face recognition using fast Discriminative Multi-manifold Analysis". *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 1–9.
- [٤٨] Changhui Hu, & Xiaobo Lu. (2015). "Single sample face recognition via lower-upper decomposition". *10th Asian Control Conference (ASCC)*, 1–6.
- [٤٩] Pan, J., Wang, X.-S., & Cheng, Y.-H. (2016). "Single-Sample Face Recognition Based on LPP Feature Transfer". *IEEE Access*, 4, 2873–2884.
- [٧٠] Perronnin, F., & Dugelay, J.-L. (2005). "Clustering face images with application to image retrieval in large databases". *Proceedings of SPIE - The International Society for Optical Engineering*, p. 256).

Abstract

Face recognition is one of the most common biometric methods of authentication and many research has been done to improve the performance of face recognition systems. This biometric method is more popular than other factors such as iris, fingerprint, and voice, due to the ease of use of the face and also no need for human interaction. However, there are many challenges that researchers are struggling with. One of these challenges is the large size of the datasets, which as the number of images increases, due to the inability to extract discriminative features and also the increase in the number of comparisons, the usual methods fail and system performance decreases.

The purpose of this article is to provide a robust face recognition method that performs well in large-scale datasets. The proposed method uses a simple and innovative clustering approach to hierarchically divide images into smaller clusters. Then, in each cluster, which the number of members is far less than the number of images in the entire dataset, face recognition is performed. Also, in the recognition phase, a combination of two Non-negative Matrix Factorization (NMF) features and Fast Retina Keypoint (FREAK) descriptors is used to be able to simultaneously use local and global characteristics of the image.

Some experiments were performed to evaluate the proposed method on the large-scale dataset FERET, as well as an extended image set that is collected from smaller standard datasets together. The proposed method achieved the accuracy of 98.36% and 94.54% on dataset FERET and the extended image set, respectively, which is about 2% better than the previous best work. The results of experiments show that the performance of the clustering approach is better than the conventional clustering methods. Also, it indicates the proposed method outperforms previous works in the face recognition task.

Keywords: Face Recognition, Non-negative Matrix Factorization, Large-scale Dataset, Clustering, FREAK descriptors, Nearest Neighbor



Faculty Of Computer & IT Engineering

MSc Thesis in Artificial Intelligent

**Face recognition in big databases using
hierarchical classifier**

By: Navid Abbaspoor Tasmaloo

Supervisor:

Hamid Hassanpour

Advisor:

Mohsen Biglari

September 2020