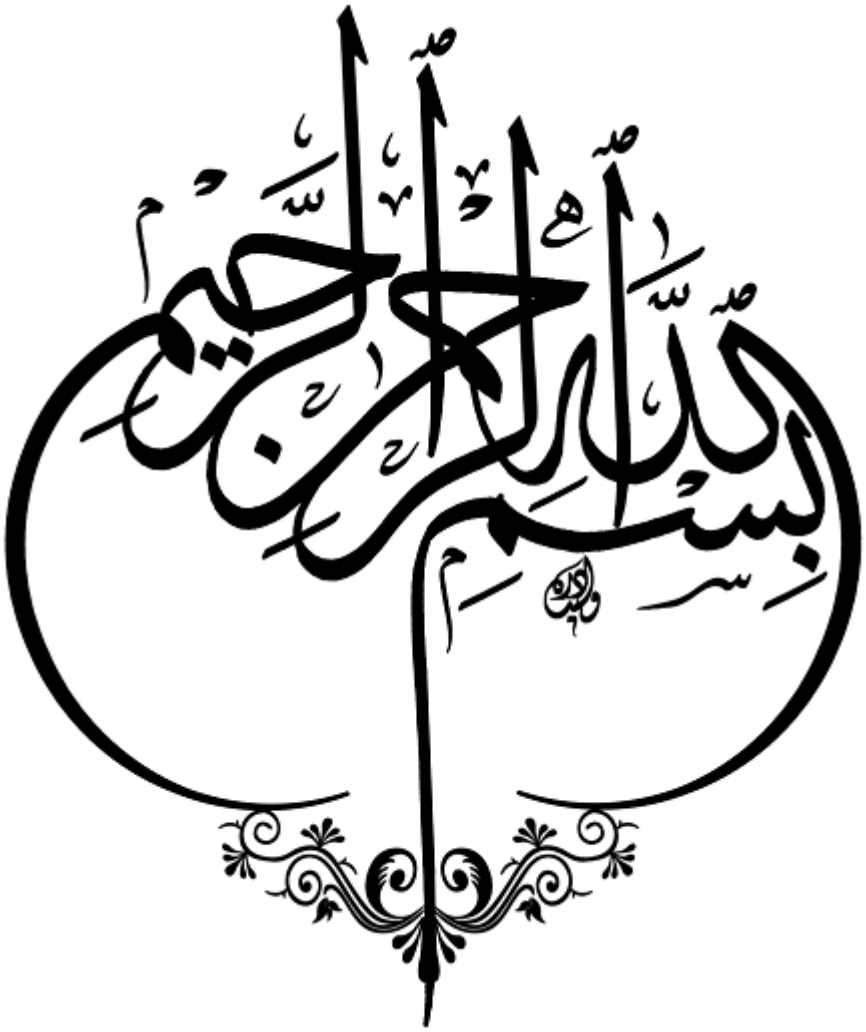

بسم الله الرحمن الرحيم

دانشکده مهندسی کامپیوتر

پایان‌نامه کارشناسی ارشد هوش مصنوعی

تشخیص چهره با استفاده از درهم‌سازی تصویر

نگارنده:

مهسا قاسمی

استاد راهنما

دکتر حمید حسن‌پور

تیر ماه ۱۴۰۰

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

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم **مهسا قاسمی** با شماره دانشجویی **۹۷۱۲۹۰۴** رشته **مهندسی کامپیوتر** گرایش **هوش مصنوعی** تحت عنوان **تشخیص چهره با استفاده از درهم‌سازی تصویر** که در تاریخ ۱۴۰۰/۰۴/۱۴ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار شد به شرح ذیل اعلام می گردد:

الف) درجه عالی: نمره ۲۰-۱۹ ☑ ب) درجه خیلی خوب: نمره ۱۸/۹۹ – ۱۸ ☐
ج) درجه خوب: نمره۱۷/۹۹– ۱۶ ☐ د) درجه متوسط: نمره ۱۵/۹۹-۱۴ ☐
ه) کمتر از ۱۴ غیر قابل قبول و نیاز به دفاع مجدد دارد ☐
نوع تحقیق: نظری ☐ عملی ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استاد راهنمای اول	دکتر حمید حسن‌پور	استاد	
۲- استاد راهنمای دوم			
۳- استاد مشاور			
۴- استاد داور اول	دکتر هدی مشایخی	استادیار	
۵- استاد داور دوم	دکتر حسین مرشدلو	استادیار	
۶- نماینده تحصیلات تکمیلی	مهندس محسن فرهادی	مربی	

نام و نام خانوادگی رئیس دانشکده: دکتر علیرضا الفی

تاریخ و امضاء و مهر دانشکده:

خدایا...[1]

به من زیستنی عطا کن که در لحظه مرگ بر بی ثمری لحظه ای که برای زیستن گذشته است حسرت نخورم ومردنی عطا کن که بر بیهودگیش سوگوار نباشم.

"چگونه زیستن" را تو به من بیاموز، "چگونه مردن" را خود خواهم آموخت.

به من توفیق تلاش در شکست، صبر در نومیدی، رفتن بی همراه، جهاد بی سلاح، کار بی پاداش، فداکاری در سکوت، دین بی دنیا، عظمت بی نام، خدمت بی نان، ایمان بی ریا، خوبی بی نمود، گستاخی بی خامی، مناعت بی غرور، عشق بی هوس، تنهایی در انبوه جمعیت ودوست داشتن بی آنکه دوست بداند، روزی کن.

[1] مناجاتی از دکتر علی شریعتی

تقدیم به پدر بزرگوار و مادر مهربانم

آن دو فرشته ای که از خواسته هایشان گذشتند، سختی ها را به جان خریدند و خود را سپر بلای مشکلات و ناملایمات کردند تا من به جایگاهی که اکنون در آن ایستاده ام برسم .

ای پدر از تو هرچه می گویم باز هم کم می آورم

خورشیدی شدی و از روشنایی ات جان گرفتم و لبریزم کردی از شوق

اکنون حاصل دستان خسته ات رمز موفقیتم شد

به خودم تبریک می گویم که تو را دارم و دنیا با همه بزرگیش مثل تو را ندارد .

و تو ای مادر، ای شوق زیبایی نفس کشیدن ،

عمری خستگی ها را به جان خریدی تا اکنون توانستی طعم خوش پیروزی را به من بچشانی

ره آوردی گران سنگ تر از این ارزان نداشتم تا به خاک پایتان نثار کنم ،

باشد که حاصل تلاشم نسیم گونه غبار خستگیتان را بزداید.

سپاس و ستایش مر خدای را جل و جلاله که آثار قدرت او بر چهره روز روشن، تابان است و انوار

حکمت او در دل شب تار، درفشان. آفریدگاری که خویشتن را به ما شناساند و درهای علم را بر ما گشود و

عمری و فرصتی عطا فرمود تا بدان، بنده ضعیف خویش را در طریق علم و معرفت بیازماید.

و با تقدیر و تشکر شایسته از استاد فرهیخته و فرزانه جناب آقای دکتر حمید حسن‌پور که با نکته های دلاویز و گفته

های بلند، صحیفه های سخن را علم پرور نمود و همواره راهنما و راه گشای نگارنده در اتمام و اکمال پایان نامه

بوده است.

تعهدنامه

اینجانب **مهسا قاسمی** دانشجوی دوره کارشناسی ارشد رشته **هوش مصنوعی** دانشکده **مهندسی کامپیوتر** دانشگاه صنعتی شاهرود، نویسنده پایان‌نامه با عنوان **شناسایی چهره با استفاده از درهم‌سازی تصویر**، تحت راهنمائی **دکتر حمید حسن‌پور** متعهد می‌شوم:

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

مهسا قاسمی

تیر ماه ۱۴۰۰

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم-افزارها 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 پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

سیستم شناسایی چهره کاربردهای فراوانی در حوزه‌های مختلف از جمله حفاظت از سیستم‌های کامپیوتری و منابع داده‌ای، شناسایی مجرمین و نظارت ویدئویی دارد. تاکنون روش‌های متنوعی برای شناسایی چهره ارائه شده است. پایگاه‌های داده مورد استفاده در این روش‌ها، معمولا شامل تعداد تصاویر کم هستند یا تحت شرایط خاص و کنترل شده، تهیه شده اند. به همین دلیل، افزایش حجم پایگاه داده یا وجود خرابی‌های موجود در تصاویر از جمله تغییرات روشنایی، نویز، انسداد و زاویه‌ی چهره عملکرد روش‌های موجود را به‌طور قابل توجهی کاهش می‌دهد.

در این پایان‌نامه با استفاده از مزایای درهم‌سازی تصویر، سیستمی برای بازشناسی چهره ارائه می‌شود که علاوه بر سرعت بالا، عملکرد مناسبی روی پایگاه‌های داده با تصاویر زیاد دارد و در صورت وجود برخی از شرایط کنترل نشده، دقت خود را تا حد زیادی حفظ می‌کند. روش پیشنهادی در سه گام پیش‌پردازش، استخراج ویژگی و تطابق چهره معرفی می‌شود. در اولین گام، کیفیت روشنایی تصاویر با استفاده از دو پیش‌پردازش تنظیم شدت روشنایی و برابرسازی هیستوگرام وفقی با محدودیت کنتراست بهبود می‌یابد. در گام استخراج ویژگی، دو دیدگاه بصری متفاوت مبتنی بر توابع درهم‌ساز، مورد استفاده قرار می‌گیرند. در دیدگاه جزئی‌نگر، ابتدا ویژگی‌های محلی از تصاویر استخراج می‌شوند، سپس این ویژگی‌ها با استفاده از الگوریتم درهم‌ساز حساس محلی به جداول درهم‌ساز نگاشت می‌شوند. در دیدگاه کلی‌نگر، ویژگی سراسری توسط روش درهم‌سازی تصویر مبتنی بر ضرایب تبدیل کسینوسی گسسته از تصاویر استخراج می‌شوند. در گام تطابق چهره، شناسایی تصویر چهره‌ی آزمون در دو مرحله انجام می‌شود. در مرحله‌ی اول، با استفاده از یک سیستم رأی‌گیری، شناسه‌هایی از مجموعه داده که از نظر دیدگاه بصری جزئی‌نگر، بیش‌ترین شباهت را با تصویر آزمون دارند، انتخاب می‌شوند. در مرحله‌ی دوم، از میان شناسه‌های منتخب مرحله‌ی اول، شبیه‌ترین شناسه به تصویر آزمون از لحاظ دیدگاه بصری کلی‌نگر، به عنوان شناسه‌ی نهایی انتخاب می‌شود.

به منظور ارزیابی روش پیشنهادی، پایگاه‌های داده‌ی ORL (۴۰۰ تصویر)، AR (۴۰۰۰ تصویر)، FERET (۴۰۱۷ تصویر) و یک مجموعه تصاویر جمع‌آوری‌شده (۳۳۶۸ تصویر) مورد آزمایش قرار گرفته‌اند و به ترتیب دقت‌های ۱۰۰، ۹۹، ۹۶/۴۷ و ۹۵/۹۰ درصد حاصل شده است که نشان‌دهنده عملکرد بسیار خوب سیستم پیشنهادی می‌باشد.

کلمات کلیدی: شناسایی چهره، درهم‌سازی تصویر، درهم‌ساز حساس محلی، پایگاه داده‌ی حجیم، سرعت شناسایی

لیست مقالات

Ghasemi M., Hassanpour H., “Improving Performance of Face Recognition in Large Datasets Using Image Hashing”, *IEEE Transactions on Image Processing* (Submitted on Jan 2021).

Ghasemi M., Hassanpour H., “A Three-stage Filtering Approach for Face Recognition Using Image Hashing”, *International Journal of Engineering* (Accepted 21 June 2021).

Ghasemi M., Hassanpour H., “Feature Extraction Using Image Hashing” (In progress).

فهرست مطالب

فهرست تصاویر

فهرست جداول

فصل ۱: سیستم های شناسایی چهره

۱-۱- مقدمه

تشخیص چهره یکی از مهمترین تکنیک‌های بیومتریک[۱] برای شناسایی افراد است. در این تکنیک با استفاده از تجزیه و تحلیل الگوهایی از چهره، شناسایی یا تأیید هویت انجام می‌گیرد. این فناوری در برابر سایر روش‌های بیومتریک مانند اثر انگشت، عنبیه چشم و اثر کف دست مزایای فراوانی دارد. که از آن جمله می‌توان به عدم نیاز به تماس مستقیم با فرد و امکان شناسایی افراد از فاصله دور بدون نیاز به عوامل انسانی اشاره کرد [۱,۲]. سیستم شناسایی چهره در اهداف و زمینه‌های مختلف از جمله امنیت اطلاعـات، کنترل و ثبت تردد در سیستم‌های حضور و غیاب، کنترل کارت‌های شناسایی، حفاظت از سیستم‌های کامپیوتری و منابع داده‌ای و تجارت الکترونیکی کاربرد دارد [۱]. به همین دلیل بسیاری از زمینه‌های هوش مصنوعی، شناسایی الگو و یادگیری ماشین به دنبال یافتن الگوریتم‌هایی برای بهبود این سیستم‌ها هستند. تاکنون سیستم‌های شناسایی چهره پیشرفت‌های قابل توجهی داشته‌اند. با این حال، عملکرد این سیستم‌ها با چالش‌هایی مواجه هستند [۳،۴]. در ادامه به مطالعه‌ی این عوامل می‌پردازیم. در این پژوهش، یک معماری جدید با استفاده از روش‌های درهم‌سازی مطرح می‌شود.

۱-۲- چالش‌های شناسایی چهره

در فناوری شناسایی چهره، چالش‌های متفاوتی وجود دارند که تاثیر مستقیمی بر عملکرد سیستم دارند. با مطالعه و تحلیل این مسائل، می‌توان تا حدودی این چالش‌ها را برطرف کرد. برخی از این چالش‌ها در جدول (۱-۱) نشان داده شده است. این عوامل را می‌توان در گروه‌های زیر طبقه‌بندی کرد [۵]:

[۱] Biometric

۱. **شرایط کنترل نشده:**

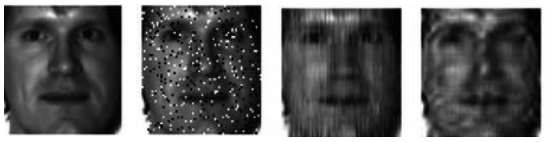
- تغییر زاویه‌ی چهره: تصاویر چهره بر اساس تغییرات زاویه‌ی دوربین نسبت به چهره، به‌صورت تمام‌رخ، سه‌رخ و نیم‌رخ تغییر می‌کنند. بنابراین توجه به این مساله در عملکرد سیستم شناسایی چهره، تاثیر قابل توجهی دارد. برخی از محققان روی این چالش تمرکز کردند و سیستم‌هایی مقاوم به تغییرات زاویه را ارائه دادند [۱،۶].
- عوامل پوششی: ویژگی‌هایی مانند عینک، ریش، ماسک در اندازه و شکل‌های مختلف، موجب انسداد برخی از نقاط چهره می‌شود. بنابراین این مساله نیز مورد توجه پژوهش‌گران قرار گرفته است [۷].
- شرایط تصویربرداری: عواملی از جمله تغییرات روشنایی و مشخصات دوربین، تاثیر زیادی بر ویژگی‌های مستخرج از تصویر می‌گذارد. لذا در ارتباط با این چالش نیز تحقیقات مهمی انجام شده است [۸،۹].
- حالت‌های چهره: حالت‌های چهره مانند اخم، لبخند، فریاد زدن و غیره بر ظاهر چهره تاثیر مستقیمی می‌گذارد [۴].

۲. **پایگاه‌های داده با حجم زیاد:** یکی از دلایل کاهش دقت برخی از سیستم‌های شناسایی چهره پایگاه داده‌ی حجیم است. به دلیل تعداد زیاد تصاویر در این نوع پایگاه‌های داده، قدرت تمایز با استفاده از ویژگی‌های مستخرج کاهش می‌یابد. به‌همین دلیل، سیستم‌هایی به منظور شناسایی چهره در پایگاه-های داده‌ی حجیم طراحی شده اند [۱۰،۱۱].

۳. **سرعت:** پیچیدگی زمانی، یک پارامتر مهم در سیستم شناسایی چهره است و در نظر گرفتن این مساله، عملکرد سیستم‌های شناسایی را بهبود می‌بخشد [۱۲].

علاوه بر موارد فوق، سیستم شناسایی چهره باید چالش‌های دیگری مانند وضوح چهره و تغییر سن را در نظر بگیرد و بتواند صرف نظر از شرایط محیطی و چالش‌های ذکرشده، عملکرد خوبی را ارائه دهد.

جدول ۱-۱- برخی از چالش‌های مربوط به سیستم‌های شناسایی چهره

چالش	مثال
گریم و آرایش چهره	
عوامل پوششی	
تغییر حالت چهره	
تغییر زاویه‌ی چهره	
افزایش سن	
وجود نویز در تصویر	

۱-۳- مولفه‌های سیستم شناسایی چهره

سیستم‌های شناسایی چهره اغلب از سه مرحله‌ی اصلی پیش‌پردازش، استخراج ویژگی و طبقه‌بندی بردار ویژگی تشکیل می‌شوند که در شکل (۱-۱) نمایش داده شده است [۱،۱۳]. در اولین مرحله، ابتدا اطلاعات ناحیه‌ی چهره مشخص می‌شود. در این راستا باید تا حد امکان از تاثیرات پس‌زمینه و سایر موارد، بر امر شناسایی جلوگیری شود. بنابراین استفاده از یک الگوریتم قابل اطمینان در این مرحله ضروری می‌باشد، زیرا تاثیر عمده‌ای روی عملکرد سیستم تشخیص چهره دارد. سپس به منظور بهبود کیفیت روشنایی تصاویر، از روش‌های تنظیم روشنایی استفاده می‌شود.

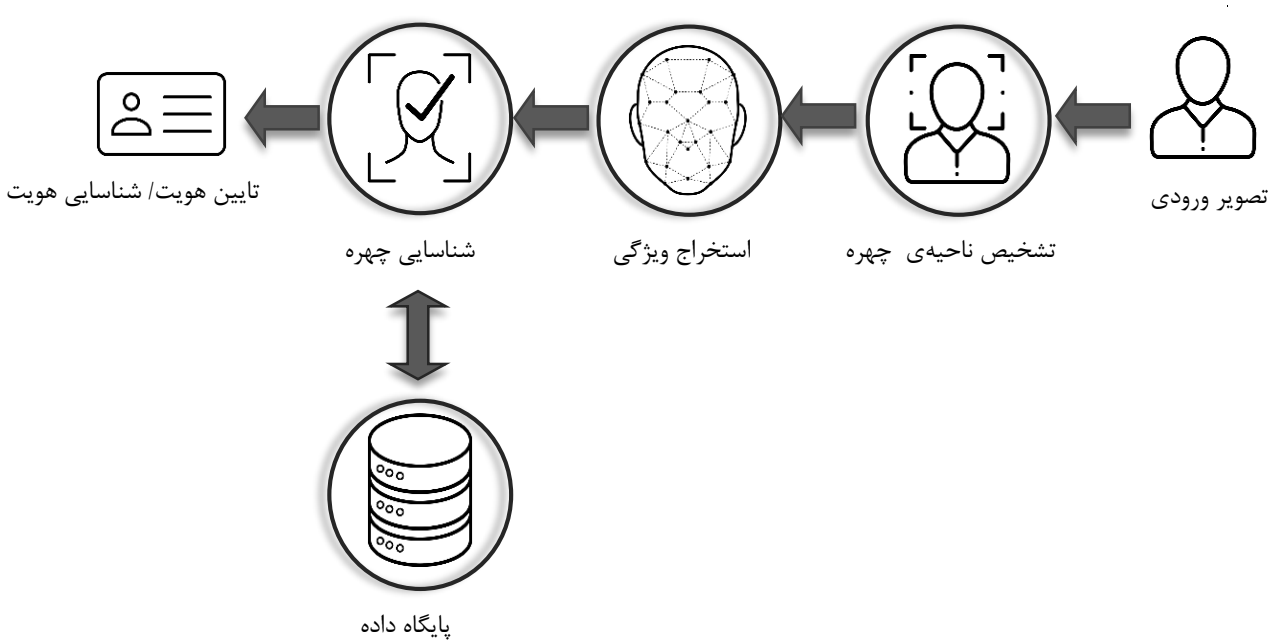
مرحله‌ی بعدی، استخراج ویژگی از تصاویر موجود در پایگاه‌داده می‌باشد. این مرحله یکی از مهم‌ترین و اساسی‌ترین بخش‌های سیستم شناسایی چهره است. بدلیل اینکه ابعاد فضای ویژگی زیاد است، استفاده مستقیم آن برای شناسایی چهره، دارای پیچیدگی محاسباتی زیادی می‌باشد. بنابراین هدف از این بخش، یافتن الگوهای منحصربه‌فرد موجود در هر چهره، به‌منظور کاهش ابعاد تصویر است. روش‌های استخراج ویژگی به سه دسته‌ی کلی ویژگی‌های سراسری، ویژگی‌های محلی و ویژگی‌های ترکیبی تقسیم می‌شوند [۵]. در روش‌های سراسری، کل تصویر چهره با استفاده از یک بردار توصیفگر ویژگی نمایش داده می‌شود که نشان‌دهنده‌ی اطلاعات مربوط به بافت و شکل تصویر چهره می‌باشد. اما روش‌های مبتنی بر ویژگی محلی، ویژگی‌های استخراجی از نقاطی مانند نوک بینی، چشم‌ها و دهان را استخراج می‌کنند. روش‌های ترکیبی از هر دو مزیت ویژگی‌های محلی و سراسری استفاده می‌کنند.

روش‌های درهم‌سازی تصویر[۱] دارای مزایای بیشتری نسبت به توصیفگرهای دیگر هستند و در پردازش داده‌های حجیم مورد توجه ویژه می‌باشند [۱۴،۱۵]. درهم‌سازی تصویر روشی کارآمد برای استخراج ویژگی-های ضروری یک تصویر و تبدیل آن به یک دنباله‌ی عدد کوتاه با اندازه‌ی ثابت است. این امر منجر به تولید ویژگی‌های فشرده‌تر، حافظه‌ی کم‌تر، هزینه‌ی محاسباتی پایین‌تر و عملکرد بهتر می‌شود. نکته حائز اهمیت

[۱] Image Hashing

در درهم‌سازی تصویر این است که یک تابع درهم‌ساز تصویر باید تغییرات در دامنه بصری را در نظر بگیرد و مقادیر درهم‌ساز را بر اساس ظاهر بصری تصویر تولید کند.

هنگامی که تصاویر در یک زیر فضا قرار گرفتند، به‌منظور تعیین هویت، بررسی شباهت میان الگوی چهره‌ی افراد توسط طبقه‌بندی کننده‌ها در مرحله‌ی سوم انجام می‌شود. انتخاب روش طبقه‌بندی بر عملکرد سیستم شناسایی چهره بسیار تاثیرگذار است، بنابراین باید از الگوریتمی با دقت بالا استفاده شود.



شکل ۱-۱- مراحل شناسایی چهره

۱-۴- اهداف پایان‌نامه

در این پایان‌نامه، یک روش موثر و کارآمد برای شناسایی چهره با استفاده از درهم‌سازی تصویر معرفی شده است. هدف این تحقیق، بهبود برخی از چالش‌های موجود در شناسایی چهره می‌باشد که در زیر به این موارد اشاره شده است:

۱- پایگاه‌های داده‌ی حجیم: با استفاده از روش‌های درهم‌سازی تصویر در این پژوهش، تصاویر به دنباله‌ای از اعداد با اندازه‌ی کوتاه و با قدرت تمایز بالا تبدیل می‌شوند. بنابراین ظرفیت ذخیره‌سازی افزایش می‌یابد و می‌توان تصاویر بیش‌تری را ذخیره و بررسی کرد.

۲- شرایط کنترل‌نشده: در این تحقیق، از دو نوع روش استخراج ویژگی محلی و سراسری مناسب استفاده شده است. به همین دلیل، روش پیشنهادی در صورت وجود برخی از شرایط کنترل نشده از جمله تغییرات زاویه‌ی چهره تا ۲۵± درجه، سن، حالت چهره و انسداد، عملکرد نسبتا خوبی دارد.

۳- کمبود تصاویر آموزشی به ازای هر فرد: روش پیشنهادی در حضور تنها یک تصویر به ازای هر فرد، عملیات شناسایی را با دقت بالایی انجام می‌دهد. این مهم، یکی از مزایای قابل توجه این روش می‌باشد، زیرا منجر به افزایش فضای ذخیره‌سازی می‌شود.

۴- پیچیدگی‌های زمانی و محاسباتی: یکی از مزایای مهم درهم‌سازی تصویر، محاسبه آسان و مقایسه سریع آن به دلیل طول بردارهای ویژگی کم می‌باشد.

همانطور که در مقدمه ذکر شد، هدف این پایان‌نامه، استفاده از مزایای درهم‌سازی تصویر در طراحی سیستمی برای شناسایی چهره است که علاوه بر سرعت بالا، عملکرد مناسبی روی پایگاه‌های داده با تصاویر زیاد دارد و در صورت وجود برخی از شرایط کنترل نشده، دقت خود را از دست ندهد.

۱-۵- روش حل

در این پژوهش، برای شناسایی چهره، یک معماری جدید با استفاده از توابع درهم‌سازی تصویر تعریف شده است. سیستم شناسایی چهره در این پژوهش شامل گام‌های پیش‌پردازش، استخراج ویژگی و تطابق چهره می‌باشد. در اولین گام، به‌منظور افزایش دقت سیستم، پیش‌پردازش‌هایی مانند تشخیص ناحیه‌ی چهره، تنظیم روشنایی و تغییر اندازه روی تصاویر انجام می‌شوند. در گام استخراج ویژگی، تصاویر مجموعه‌ی داده با دو دیدگاه بصری جزئی‌نگر و کلی‌نگر بررسی می‌شوند. در این راستا، ویژگی‌های محلی و سراسری تصاویر به کمک روش‌های درهم‌سازی استخراج می‌شوند و در جدول‌های درهم‌ساز ذخیره می‌شوند. در نهایت، تصاویر آزمون در گام تطابق چهره با استفاده از دو مرحله شناسایی می‌شوند. در مرحله‌ی اول با استفاده از یک سیستم رأی‌گیری، افرادی از مجموعه‌ی داده که بیش‌ترین شباهت را، از لحاظ ویژگی‌های محلی، به تصویر آزمون دارند، انتخاب می‌شوند و به مرحله‌ی دوم وارد می‌شوند. در مرحله‌ی دوم، از میان نامزدهای منتخب مرحله‌ی اول، فردی که بیش‌ترین شباهت را، از لحاظ ویژگی سراسر، به تصویر آزمون دارد به‌عنوان هویت نهایی تصویر آزمون انتخاب می‌شود.

۱-۶- پیش‌فرض‌های مسئله

در این پژوهش، پیش‌فرض‌هایی در نظر گرفته شده‌اند، که عبارتند از:

- تصاویر چهره در تمامی مراحل سیستم پیشنهادی، خاکستری[۱] هستند.
- اشخاص شامل دو جنسیت مرد و زن هستند و ملیت‌های متفاوتی دارند.
- تصاویر چهره ممکن است دارای سبیل، گوشواره و عینک باشد و در برخی موارد، چهره‌ی افراد دارای انسداد است.
- زاویه‌ی چهره در برخی موارد به‌صورت سه‌رخ می‌باشد.

[۱] Grayscale

- حداقل تعداد نمونه تصویر به‌ازای هر فرد، یک تصویر می‌باشد.
- تصاویر چهره در شرایط محیطی مختلفی تصویر برداری شده‌اند و کیفیت روشنایی در تصاویر متفاوت است.

۱-۷- ساختار پایان‌نامه

در ادامه، فصل دوم به مرور برخی از روش‌های شناسایی چهره می‌پردازد. فصل سوم، پیشینه و نحوه‌ی عملکرد الگوریتم‌های استفاده شده در این پژوهش را معرفی می‌کند. در فصل چهارم، روش پیشنهادی ارائه می‌شود. فصل پنجم به تشریح آزمایشات انجام شده و مقایسه روش پیشنهادی با روش‌های دیگر پرداخته است. در فصل ششم نیز، نتیجه‌گیری و کارهای پیشنهادی بیان می‌شود.

۱-۸- جمع‌بندی

در فصل اول، ابتدا به معرفی سیستم شناسایی چهره و کاربرد و اهمیت‌های آن پرداخته شد. سپس چالش-های موجود در این سیستم بررسی شد. در ادامه، موضوع و هدف از این پایان‌نامه ذکر شد و در نهایت ساختار پایان‌نامه بیان شد.

فصل ۲: مروری بر کارهای پیشین

۲-۱- مقدمه

در چند دهه‌ی اخیر، محققین زیادی در زمینه‌ی شناسایی چهره فعالیت کرده‌اند و روش‌های متفاوتی را با توجه به چالش‌های موجود در سیستم شناسایی چهره ارائه داده‌اند. همانطور که اشاره شد، استخراج ویژگی گام مهمی در شناسایی چهره است و باعث رقابت روش‌های مختلف می‌شود. این روش‌ها بر اساس ویژگی-های محلی، ویژگی‌های سراسری و ویژگی‌های ترکیبی به سه گروه تقسیم می‌شوند. در این فصل، ابتدا برخی از روش‌های مربوط به هر گروه بررسی می‌شوند. سپس چند نمونه روش درهم‌سازی تصویر معرفی می‌شود.

۲-۲- استخراج ویژگی مبتنی بر ویژگی‌های محلی

روش‌های مبتنی بر ویژگی محلی، تصاویر چهره را با استخراج توصیف‌گرهای ویژگی از نواحی مانند نوک بینی، چشم‌ها و دهان نشان می‌دهند. بردارهای ویژگی بر اساس روابط هندسی مانند مکان، زوایا و فاصله استخراج می‌شوند. مزیت این روش، استحکام در تغییراتی مانند روشنایی، چرخش و انسداد جزئی است. یعنی حتی اگر برخی از نقاط پوشیده یا تار باشند، می‌توان از نقاط دیگر برای شناسایی چهره استفاده کرد. اما این روش‌ها، از ظاهر کلی تصویر چهره غافل هستند و اغلب توضیحات مبهمی تولید می‌کنند. زیرا تنها چند پیکسل از تصویر برای محاسبه‌ی توصیف‌گر محلی در نظر گرفته می‌شود. روش تبدیل ویژگی مستقل از مقیاس[۱] (SIFT) در دسته‌ی روش‌های استخراج ویژگی محلی قرار دارد [۱۶]. این دسته از روش‌ها معمولاً زمانی استفاده می‌شوند که اطلاعات ساختار محلی مهم‌تر از اطلاعات کلی نظیر شدت روشنایی هستند.

[۱] Scale-Invariant Feature Transform

- توصیفگر SIFT برای آشکارسازی و توصیف ویژگی‌های کلیدی در تصاویر به کار می‌رود [۱۶]. ویژگی‌های استخراج شده با این روش در مقابل تغییراتی مانند چرخش، مقیاس، تاری، روشنایی و نویز مقاوم هستند. این روش، عموما در چهار مرحله‌ی شناسایی اکسترمم‌های فضای مقیاس[۱]، مکان‌یابی نقاط کلیدی، گرایش گماری و ایجاد توصیفگر پیاده‌سازی می‌شود. مرحله‌ی اول، سعی در شناسایی نقاطی را دارد که از دیدگاه‌های مختلف یک شیء قابل شناسایی باشند. در این راستا از یک فیلتر "فضای مقیاس[۲]" استفاده می‌شود. هدف از ایجاد این فیلتر، در نظر گرفتن مقیاس‌های مختلف به همراه ساختارهایی با سطوح تاری متفاوت از تصویر است. بدین منظور ابتدا تصویر با استفاده از روش درون‌یابی دوقطبی[۳] به نوع داده‌ی دوتایی[۴] تبدیل می‌شود. سپس، با استفاده از پیچش گوسی[۵] تعریف شده در رابطه (۲-۱)، تار می‌شود. این عمل در شکل (۲-۱) با پیکان نارنجی نشان داده شده است.

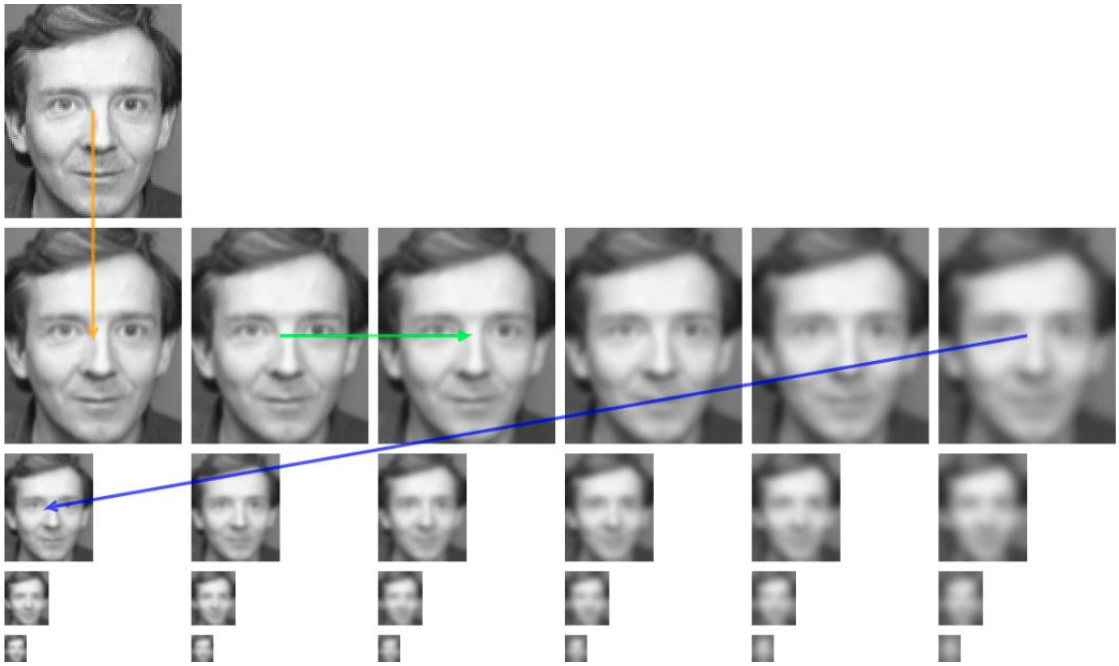
$$L(X,Y,\sigma) = G(X,Y,O\sigma) * I(X,Y) \qquad (۱-۲)$$

در این تابع، علامت * به معنای پیچش[۶]، G مقیاس متغیر گاوسی[۷]، I بیانگر تصویر ورودی و σ بیانگر اندازه‌ی مقیاس و O تصویر خروجی است. با توجه به شکل، هر تصویری که در سمت راست قرار دارد، نتیجه‌ی پیچش با همسایه‌ی سمت چپ خود است. با توجه به پیکان سبز، تصویر در هر مقیاس به تدریج تار می‌شود. همچنین، همانطور که پیکان آبی نشان می‌دهد، مقیاس تصویر در انتهای هر ردیف تغییر می‌کند و مرحله‌ی دیگری از پیچش آغاز می‌شود. این فرآیند تا زمانی که تصویر قابل پردازش باشد ادامه دارد. درنهایت، می‌توان از رابطه‌ی تفاوت گوسی[۸] (رابطه‌ی (۲-۲)) برای شناسایی نقاط کلیدی استفاده کرد.

$$\mathrm{D(X,Y,\sigma)} = \mathrm{L(X,Y,}O\sigma\mathrm{)} - \mathrm{L(X,Y,\sigma)} \qquad (۲-۲)$$

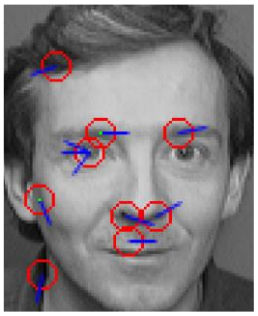
---

۱ Scale-Space Extrema Detection
۲ Scale-Space
۳ Bilinear Interpolation
۴ Double
۵ Gaussian Convolution
۶ Convolution
۷ Variable-scale Gaussian
۸ Difference of Gaussians



شکل ۲-۱- فیلتر فضای مقیاس

سپس اکسترمم‌های[۱] فضای مقیاس جستجو می‌شوند، زیرا این نقاط به‌عنوان نقاط کلیدی در نظر گرفته می‌شوند. شکل (۲-۲) این نقاط کلیدی را نشان می‌دهد. دایره قرمز بیانگر نقطه‌ی کلیدی به همراه همسایه‌های آن است و هر خط آبی درون این دایره نشان‌دهنده جهت نقطه‌ی کلیدی می‌باشد.



شکل ۲-۲- نمایش نقاط کلیدی

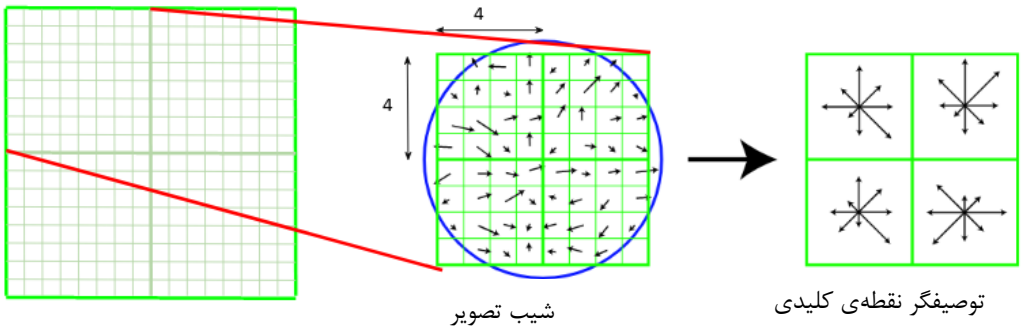
به‌منظور بررسی مقاومت نقاط کلیدی در مقابل نویز، در مرحله‌ی دوم، نقاطی که وضوح کمی دارند یا بسیار نزدیک به لبه قرار دارند، حذف می‌شوند. برای مقابله با نقاطی با وضوح کم، سری تایلور مرتبه دوم[۲] برای هر نقطه کلیدی محاسبه می‌شود، اگر مقدار حاصل، کمتر از حد آستانه (۰/۰۳) باشد، نقطه

[۱] Extremum

[۲] Second-order Taylor expansion

کلیدی حذف می‌شود. همچنین به‌منظور شناسایی نقاط کلیدی ضعیفی که به لبه‌ها نزدیک هستند، از ماتریس هسین مرتبه دوم[۱]استفاده می شود. پس از حذف نقاط کلیدی ناپایدار، برای هر نقطه کلیدی یک مقدار جهت‌یابی اختصاص داده می‌شود تا در مقابل چرخش مقاوم باشد.

در مرحله‌ی سوم، برای هر نقطه‌ی کلیدی، یک مقدار جهت‌یابی اختصاص داده می‌شود تا در مقابل چرخش مقاوم باشد. به‌منظور محاسبه‌ی بزرگی و جهت گرادیان در هر نقطه‌ی کلیدی، از همسایگی آن، یک هیستوگرام جهت‌یابی با ۳۶ قسمت در ۳۶۰ درجه، ایجاد می‌شود (هر قسمت ۱۰ درجه است). جهتی با حداکثر بزرگی، به‌عنوان جهت نقطه‌ی کلیدی در نظر گرفته می‌شود. در نهایت در مرحله‌ی چهارم، توصیفگر ویژگی محلی[۲]ایجاد می‌شود. در این مرحله، از جهت‌گیری‌ها و بزرگی پیکسل‌های همسایه استفاده می‌شود تا توصیف‌کننده‌ی منحصر به فردی برای نقطه‌کلیدی ایجاد شود. علاوه بر این، به دلیل اینکه از پیکسل‌های اطراف استفاده می‌شود، توصیف کننده ها تا حدی در تغییر روشنایی تصاویر مقاوم هستند. مطابق با شکل (۲-۳)، ابتدا یک پنجره ۱۶ × ۱۶ در اطراف نقطه‌ی کلیدی در نظر گرفته می‌شود. سپس این ناحیه به زیر پنجره‌های ۴ × ۴ تقسیم می شود. برای هر یک از این زیر پنجره‌ها، یک هیستوگرام جهت‌گیری ۸ بیتی تولید می‌شود. سپس تمام هیستوگرام‌ها در یک توصیف‌گر با مقدار ۱۲۸ جمع می‌شوند. این روش نقاط کلیدی را به‌خوبی شناسایی می‌کند، ولی به‌دلیل وجود محاسبات پیچیده، عملکرد سریعی در شناسایی نقاط کلیدی ندارد.



شکل ۲-۳- توصیفگر نهایی بدست آمده توسط روش SIFT [۱۶]

[۱] Second-order Hessian matrix

[۲] Keypoint Descriptor

زنگ[۱] و همکاران [۱۷]، یک سیستم شناسایی چهره با استفاده از توصیفگر SIFT ارائه دادند. در این روش ابتدا ویژگی‌های محلی SIFT از تصاویر چهره‌ی آموزش و آزمون استخراج می‌شوند. در نهایت، شناسه‌ای[۲] از مجموعه‌ی آموزش که ویژگی‌های محلی آن، بیشترین شباهت را با ویژگی‌های محلی تصویر آزمون داشته باشد، به‌عنوان شناسه‌ی چهره‌ی آزمون انتخاب می‌شود. به‌منظور افزایش سرعت در یافتن ویژگی‌های مشابه، از الگوریتم درهم‌ساز حساس محلی[۳] (LSH) استفاده کردند. توضیحات بیشتر در ارتباط با این الگوریتم در فصل سوم داده می‌شود. این روش، بر روی پایگاه‌داده‌های AR [۱۹] و ORL [۱۸] به ترتیب دقت‌های ۸۰/۰۸ و ۹۶/۱۴ را بدست آورده است [۲۰]. به‌دلیل استفاده از الگوریتم درهم‌ساز LSH، روش ارائه شده در مرجع [۱۷] می‌تواند تصاویر چهره را با سرعت زیادی شناسایی کند. اما همانطور که نتایج آزمایشات نشان می‌دهد، این روش در صورت وجود برخی از شرایط کنترل نشده مانند وجود انسداد، عملکرد خوبی ندارد.

- چن[۴] و همکاران [۲۱]، یک روش مبتنی بر درهم‌سازی ویژگی‌های محلی برای شناسایی چهره ارائه دادند. در این روش، یک تابع درهم‌ساز جدید با استفاده از بردارهای تفاضل پیکسل وصله‌دار[۵] (PDV) و روش لاگرانژ تقویت شده[۶] (ALM) تعریف می‌شود و با استفاده از آن، هر تصویر چهره به یک دنباله‌ی عدد باینری نگاشت داده می‌شود. در نهایت، خوشه‌بندی k-میانگین[۷] بر روی کدهای باینری اعمال می شود و آن‌ها را به خوشه‌های مختلفی تقسیم می‌کند. این روش با استفاده از پایگاه‌های داده‌ی FERET و LFW ارزیابی شد و به ترتیب دقت ۹۶/۲ و ۹۰/۲ درصد حاصل شد. همچنین زمان شناسایی به‌ازای یک چهره در این روش با مشخصات سیستم PC=2.70 GHz و RAM=24 G برابر با ۱۳۹میلی ثانیه است. دقت و سرعت شناسایی بالا در این روش قابل توجه است. اما این روش به تغییرات نوری حساس است [۲۱].

---

[۱] Zeng
[۲] Identity
[۳] Locality Sensitive Hashing
[۴] Chen
[۵] Patch-wise pixel Difference Vectors
[۶] Augmented Lagrange Method
[۷] k-means clustering

۲-۳- استخراج ویژگی بر اساس ویژگی‌های سراسری

در روش‌های مبتنی بر ویژگی سراسری، کل تصویر چهره با استفاده از یک بردار ویژگی نمایش داده می‌شود. این توصیفگر ویژگی، اطلاعات مربوط به بافت و شکل کلی تصویر چهره را استخراج می‌کند. یکی از مزایای روش‌های استخراج ویژگی‌های سراسری، کدگذاری آسان روی ظاهرکلی چهره است. به‌همین دلیل، توصیفگر سریع ایجاد می‌شود. به دلیل اینکه هر پیکسل از چهره در چگونگی ایجاد توصیفگر نقش دارد، اطلاعات زیادی در توصیفگر موجود می‌باشد. در این روش‌ها، معمولاً یک مرحله‌ی پیش‌پردازش به‌منظور بررسی روشنایی و وضعیت تصویر نیاز است. توصیف‌گرهای ویژگی سراسری مانند چهره ویژه[۱]و چهره فیشر[۲]در این دسته قرار دارند [۲۲]. در ادامه، برخی از این روش‌ها معرفی می‌شوند.

- تحلیل مولفه اساسی[۳](PCA) یک روش تبدیل خطی ساده و در عین حال محبوب و کارآمد برای شناسایی چهره است [۲۳]. PCA، ویژگی‌هایی که ارزش بیشتری دارند را استخراج می‌کند. در این روش فرض شده است که ماتریس X متشکل از بردارهای $\{x_1, x_2, x_3, \dots, x_N\}$ می‌باشد. بنابراین اگر ابعاد بردارها M باشد، ماتریس داده‌ها برابر با $M \times N$ می‌باشد. PCA میانگین و ماتریس کواریانس را محاسبه می‌کند و سپس فضای ویژگی را استخراج می‌کند. بردار x_i بیانگر یک تصویر چهره و N تعداد تصاویر را نشان می‌دهد. سپس میانگین و کواریانس با استفاده از رابطه (۲-۳) و (۲-۴) محاسبه می‌شود.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (۲-۳)$$

$$S = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T \qquad (۲-۴)$$

[۱] Eigenface
[۲] Fisherface
[۳] Principal Component Analysis

مقادیر ویژه λ_i و بردارهای ویژه v_i نیز با حل رابطه‌ی (۲-۵) بدست می‌آید.

$$Sv_i = \lambda_i v_i \qquad i = ۱,۲, \dots, M \tag{۵-۲}$$

در نهایت، بردارهای ویژه بر اساس مقدارهای ویژه به‌صورت نزولی مرتب می‌شود و در ماتریس $W_{M\times N}$ ذخیره می‌شود. لازم به ذکر است که هر ستون از این ماتریس نشان‌دهنده‌ی یک تصویر است که به این تصویر، چهره ویژه گفته می‌شود. سپس k بردار اول $W = \{v_۱, v_۲, v_۳, \dots, v_N\}$ انتخاب می‌شود. با استفاده از این روش، می‌توان هر تصویر از یک مجموعه تصویر را با استفاده از چهره‌های ویژه‌ی آن مجموعه نشان داد. بنابراین برای شناسایی چهره، ابتدا برای تمام نمونه‌های آموزش باید ماتریس W را محاسبه کرد. سپس با توجه به این ماتریس، تمام تصاویر را به ضرایب مربوطه تبدیل کرد. به منظور بازشناسی چهره نیز می‌توان از ضرایب چهره استفاده کرد. بدین صورت که اگر ضرایب دو چهره از نظر فاصله‌ی اقلیدسی[۱] به هم نزدیک باشند، احتمالا آن دو چهره به هم شباهت دارند. تورک[۲] و همکاران [۲۴]، این روش را روی یک پایگاه داده‌ی دست‌ساز[۳] حاوی ۲۵۰۰ تصویر از ۱۶ فرد آزمایش کردند. در این مجموعه به‌ازای هر فرد، سه حالت از تغییرات سر، وضوح و شرایط روشنایی وجود دارد. نتایج ارزیابی نشان می‌دهد که این روش در مقابل تغییرات روشنایی مقاوم است اما در حضور شرایط کنترل نشده‌ای که با تغییر زاویه‌ی چهره و اندازه همراه هستند، شناسایی قابل قبولی انجام نمی‌دهد [۲۴].

- گاوو[۴] و همکاران [۲۵] یک سیستم شناسایی چهره با استفاده از روش تجزیه و تحلیل خطی فیشر[۵] (FLDA) و تجزیه مقدار تکین[۶] (SVD) ارائه دادند. هدف اصلی این سیستم، حل چالش نمونه‌ی کم آموزشی (وجود تنها یک نمونه تصویر به‌ازای هر فرد) می‌باشد. فرض بر این است که تصویر چهره در ماتریسی به نام $A \epsilon R^{m\times n}$ ذخیره می‌شود $(m \geq n)$. همچنین SVD به‌صورت رابطه‌ی (۲-۷) بیان می‌شود.

---

۱ Euclidean
۲ Turk
۳ Handcraft
۴ Gao
۵ Fisher Linear Discrimination Analysis
۶ Singular Value Decomposition

$$A = \sum_{i=1}^{n} \sigma_i \cdot u_i \cdot v_i^T \qquad (۷-۲)$$

در رابطه‌ی فوق، بردارهای u_i و v_i، به ترتیب i امین ستون از $U \epsilon R^{m \times n}$ و $V \epsilon R^{m \times n}$ هستند. U و V بردارهای ویژه و σ_i مقدارهای تکین می‌باشد و به‌صورت نزولی مرتب شده‌اند. در تشکیل چهره، بزرگترین مقادیر تکین دارای اهمیت بیشتری هستند. بنابراین، می‌توان با استفاده از سه مقدار بزرگتر ماتریس تکین، تصویر اصلی را بازسازی کرد. با استفاده از تصویر اصلی و تصویر بازسازی شده، مجموعه‌ی جدید آموزش ایجاد می‌شود. بنابراین، در هر کلاس، دو نمونه تصویر از یک فرد وجود دارد. برای ارزیابی این روش، از پایگاه داده FERET [۲۶] استفاده شده است. مجموعه داده‌ی در نظر گرفته شده دارای ۴۰۰ تصویر از ۲۰۰ فرد می‌باشد و دقت شناسایی ۹۰/۵ درصد حاصل شده است. چالش وجود یک نمونه آموزشی به‌ازای هر فرد توسط این روش حل شده است. اما با وجود اینکه مجموعه داده‌ی در نظر گرفته شده برای ارزیابی این روش حجیم نمی‌باشد، دقت شناسایی خوبی را به‌دست نیاورده است.

- بخشی و همکاران [۲۷] روی چالش یک نمونه تصویر چهره به‌ازای هر فرد در پایگاه داده‌ی حجیم تمرکز کردند. این روش، پنجره‌های هم‌مرکز با ابعاد مختلف را روی تصویر چهره اعمال می‌کند و محتوای هر پنجره را به فضای فرکانسی منتقل می‌کند. سپس، از یک فیلتر مناسب استفاده می‌شود و فقط مولفه‌های فرکانسی با قابلیت جداسازی بالا بین تصاویر، حفظ می‌شوند. در نهایت، با بکارگیری فاصله‌ی اقلیدسی، فاصله‌ی بین بردار ویژگی تصاویر آموزش و آزمون را به‌دست می‌آورد. این روش با استفاده از پایگاه داده‌ی FERET [۲۶] ارزیابی و دقت ۹۰/۴ درصد حاصل شده است. زمان شناسایی چهره در این روش به ازای هر تصویر چهره روی سیستمی با مشخصات CPU= Intel Core i7, 2.0 GHz و Memory= 4 GB برابر با ۸۶۰ میلی ثانیه است. با وجود اینکه روش ارائه شده توانسته چالش کمبود نمونه‌های آموزشی را حل کند، اما به دلیل محاسبات پیچیده‌ی ویژگی‌های مکان-فرکانس، به زمان اجرای بالایی نیاز دارد.

- نیکان و حسن‌پور [۱۰] با بکارگیری روش تجزیه‌ی ماتریس نامنفی[۱](NMF)، سیستمی برای شناسایی چهره ارائه دادند. تمرکز این روش روی چالش نمونه‌های کم آموزشی (وجود فقط یک نمونه تصویر چهره به‌ازای هر فرد) و سرعت شناسایی در پایگاه داده‌ی حجیم است. در این سیستم، تصاویر چهره به دو قسمت به‌صورت عمودی تقسیم می‌شوند و دو ماتریس را تشکیل می‌دهند. سپس با استفاده از روش NMF، این ماتریس‌ها به ابعاد کمتری تجزیه می‌شوند و دو ماتریس ویژگی ایجاد می‌شوند. در نهایت، مقدار همبستگی[۲]بین قسمت‌های بالا و پایین محاسبه می‌شود و با هم ترکیب می‌شوند. شناسه‌ی تصویری از مجموعه‌ی داده که همبستگی بیشتری با تصویر آزمون دارد، به‌عنوان هویت تصویر آزمون انتخاب می‌شود. دقت این روش روی پایگاه‌داده‌ی FERET [۲۶] برابر با ۹۲/۷ درصد است. زمان شناسایی به‌ازای هر چهره روی سیستمی با مشخصات RAM=64 و CPU Core i7 برابر با ۶۴,۱ میلی ثانیه است. زمان شناسایی تصاویر چهره در این سیستم کاهش چشم‌گیری داشته است. اما یکی از پیش‌فرض‌های این روش این است که تمامی تصاویر چهره باید به‌صورت روبرو تصویر برداری شده باشند و بخشی از آن‌ها پوشیده نباشد [۱۰]. بنابراین، دقت این روش به برخی از شرایط محیطی کنترل نشده مانند تغییر زاویه‌ی چهره و وجود انسداد در تصویر محدود می‌باشد.

- شریف و همکاران [۲۸]، یک سیستم شناسایی چهره با استفاده از یک روش درهم‌سازی تصویر نوین مبتنی بر ضرایب DCT معرفی می‌کنند. در این روش، ابتدا تصویر نرمال می‌شود و به پنجره‌های غیر همپوشان تقسیم می‌شود. سپس در مرحله‌ی استخراج ویژگی، الگوریتم DCT روی پنجره‌های تصویر اعمال می‌شود و هر پنجره را به یک بردار ویژگی تبدیل می‌کند. از آنجا که ضرایب پایین الگوریتم DCT گویای ساختار کلی تصویر است، لذا تنها این ضرایب حفظ می‌شوند و باقی ضرایب حذف می‌شوند. در مرحله‌ی ایجاد درهم‌ساز، ابتدا میانگین بردارهای ویژگی بدست آمده از تمامی پنجره‌ها محاسبه می‌شود. سپس، خوشه‌بندی K-mean روی این بردار اعمال شده و چهار مرکز خوشه به‌عنوان

[۱] Non-negative Matrix Factorization

[۲] Correlation

مقدار نهایی درهم‌ساز تصویر در نظر گرفته می‌شود. روش ارائه شده بر روی پایگاه‌های داده‌ی ORL و Yale آزمایش شد و دقت‌های ۹۸/۲ و ۹۶/۶۶ درصد را بدست آورد. سرعت شناسایی زیاد در این روش بسیار قابل توجه است. اما این روش فقط با استفاده از پایگاه‌های داده‌ی کم حجم آزمایش شده است. با توجه به اینکه اندازه‌ی بردار ویژگی در روش ارائه شده در مرجع [۲۸] بسیار کوتاه است، می‌توان بیان کرد که قدرت تمایز با استفاده از این روش در پایگاه‌های داده‌ی حجیم کاهش می‌یابد [۱۰،۱۱].

۲-۴- روش‌های ترکیبی

همانطور که قبلا بیان شد، هر کدام از روش‌های استخراج ویژگی محلی و سراسری دارای مزایا و معایبی هستند. روش‌های استخراج ویژگی ترکیبی، با هدف حفظ مزایا و پوشاندن معایب این ویژگی‌ها، ویژگی-های سراسری و محلی را ترکیب می‌کنند [۲۹]. در ادامه برخی از این روش‌ها معرفی می‌شوند.

- کاسیو و همکاران [۳۰] با استفاده از رگرسیون حداقل مربعات جزئی[۱](PLSR) و الگوریتم LSH، روشی مبتنی بر درهم‌سازی تصویر در پایگاه‌های داده حجیم ارائه دادند. در مرحله‌ی آموزش این روش، یک مدل درهم‌ساز ایجاد می‌شود. در این مدل، تصاویر موجود در پایگاه داده با استفاده از یک توزیع برنولی به دو زیرگروه مثبت و منفی تقسیم می‌شوند. سپس با استفاده از الگوریتم PLSR، چهره‌های زیرگروه مثبت با هدف ۱+ از زیر گروه منفی با هدف ۱- متمایز می‌شوند. این مدل ایجاد شده به همراه شناسه‌ی چهره‌های زیر مجموعه‌ی مثبت ذخیره می‌شوند. این روال چندین بار تکرار می‌شود و در هر بار اجرا، اطلاعات بدست آمده ذخیره می‌شوند. در مرحله‌ی شناسایی چهره، ابتدا لیستی از آرا برابر با تعداد تصاویر چهره ایجاد می‌شود و تمامی آن‌ها با عدد صفر مقداردهی می‌شوند. سپس تصویر چهره‌ی آزمون به هر کدام از مدل‌های بدست آمده از مرحله‌ی آموزش ارائه می‌شود و یک مقدار رگرسیون[۲] بدست می‌آید. مقدار رای هر فرد با توجه به مقدار رگرسیون بدست آمده از مدل، افزایش می‌یابد. یعنی

[۱] Partial Least Squares Regression

[۲] Regression

هنگامی که چهره‌ی آزمون به اولین مدل ایجاد شده در اولین اجرا وارد می‌شود، اگر مقدار بدست آمده مثبت باشد، شناسه‌هایی که در اجرای اول مقدار مثبت داشتند، به اندازه‌ی میزان خروجی مقدار رگرسیون افزایش می‌یابند. درنهایت، چهره‌ای با بیشترین مقدار رگرسیون، به‌عنوان شناسه‌ی برتر انتخاب می‌شود. در این روش نیز، دقت شناسایی تصاویر چهره افزایش قابل توجهی داشته است. اما به‌دلیل استفاده از چندین نوع توصیفگر ویژگی، این روش به حافظه‌ی نسبتا زیادی نیاز دارد. عملکرد این روش با استفاده از پایگاه داده‌ی FERET [۲۶] شامل ۱۱۹۵ تصویر ارزیابی شد و دقت ۹۶/۴۷ درصد بدست آمد.

- عباس‌پور و همکاران [۱۱]، یک سیستم شناسایی چهره با استفاده از یک رویکرد خوشه‌بندی ساده ارائه دادند. در این روش، ابتدا تصاویر به دو قسمت بالا و پایین تقسیم می‌شوند. سپس روش NMF روی این قسمت‌ها اعمال می‌شود و اختلاف میاتگین دو بردار ویژگی محاسبه می‌شود. در نهایت برای استخراج ویژگی، تصاویر به دو خوشه تقسیم می‌شوند و داخل هر خوشه، از دو روش NMF و FREAK[۱] استفاده می‌شود. در نهایت، برای شناسایی تصویر از روش نزدیک‌ترین همسایه[۲](NN) استفاده شده است. این روش با استفاده از پایگاه داده‌ی FERET [۲۶] ارزیابی و دقت ۹۸/۳۶ درصد حاصل شد. زمان شناسایی چهره در این روش به ازای هر تصویر چهره روی سیستمی با مشخصات CPU= Intel Core i5 و Memory= 6 GB برابر با ۲۰۰۰ میلی ثانیه است. در مقایسه با روش نیکان [۱۰]، دقت شناسایی در این سیستم بهتر است. اما به‌دلیل استفاده از چند روش استخراج ویژگی، زمان شناسایی به‌صورت قابل توجهی افزایش یافته است.

- زنگنه و همکاران [۳۱] یک معماری جدید با استفاده از شبکه عصبی کانولوشن عمیق[۳](DCNNs)، برای شناسایی چهره با وضوح پایین پیشنهاد می‌کنند و به دقت شناسایی بالایی دست می‌یابند. معماری

[۱] Fast Retina Keypoint
[۲] Nearest Neighbor
[۳] Deep Convolutional Neural Network

پیشنهادی شامل دو بخش DCNN برای نگاشت تصاویر چهره با وضوح بالا و پایین به یک فضای مشترک با تبدیل غیرخطی می‌باشد. اولین بخش ۱۴ لایه دارد و نقش آن تبدیل تصاویر با وضوح بالا است. بخش دوم، تصاویر چهره با وضوح پایین را به فضای مشترک نگاشت می‌کند که شامل یک شبکه‌ی ۵ لایه با وضوح بالا است و به شبکه ۱۴ لایه متصل است. برای بروزرسانی وزن های DCNN از بهینه سازی مبتنی بر شیب استفاده شده است تا با انتشار مجدد خطا، فاصله بین جفت تصویرهای وضوح بالا و پایین را در فضای مشترک به حداقل برساند. این روش بر روی مجموعه داده‌های FERET و LFW ارزیابی شد و به‌ترتیب دقت‌های ۹۶/۷ و ۷۶/۳ درصد حاصل شد. نتایج نشان می‌دهد که این روش به طور قابل‌توجهی عملکرد شناسایی را برای تصاویر چهره آزمون با وضوح بسیار پایین بهبود می‌بخشد. اما روش‌ها یادگیری عمیق معمولا به تعداد نمونه آموزشی بالایی نیاز دارند. بنابراین در شرایطی که تنها یک نمونه به ازای هر فرد وجود دارد، نمی‌توان از روش‌های معمول یادگیری عمیق برای شناسایی استفاده کرد [۳۱].

- زنگ[۱] و همکاران [۳۲] با استفاده از روش یادگیری عمیق به شناسایی چهره پرداختند. تمرکز این روش روی حل چالش وجود تنها یک نمونه تصویر به‌ازای هر فرد با کمک روش نمونه‌گیری در حال گسترش[۲] است. در این روش، تصویر بدون حالت یک فرد از تصویر با حالت مختلف آن فرد کم می‌شود و سپس برای گسترش، با نمونه‌های دیگر جمع می‌شود. در نهایت، با استفاده از تصاویر بدست آمده، یک مدل شبکه عصبی کانولوشن[۳](CNN) را آموزش می‌دهد. این روش با استفاده از ۱۴۰۰ تصویر تحت حالت‌های مختلف مانند تغییرات زاویه و روشنایی از پایگاه داده‌ی FERET [۲۶] ارزیابی شد و دقت ۹۳/۹ درصد حاصل شد. این روش مشکل کمبود نمونه‌های آموزشی را به خوبی حل کرده است. اما با وجود اینکه تعداد تصاویر مجموعه داده کم است، دقت شناسایی آن از روش‌های دیگر مانند مراجع [۱۰،۱۱] کمتر است.

---

[۱] Zang
[۲] Expanding Sample Method
[۳] Convolutional Neural Network

۲-۵- روش‌های مبتنی بر درهم‌سازی تصویر

همانطور که در فصل اول بیان شد، با استفاده از توابع درهم‌ساز تصویر می‌توان محتوای بصری تصویر را به شیوه‌ای فشرده نشان داد. در ادامه چند روش درهم‌سازی تصویر بیان شده است. هدف تمامی این روش‌ها، ایجاد درهم‌سازهای مقاوم به تغییراتی مانند فشرده‌سازی، روشنایی و نویز است و می‌توانند در کاربردهای مختلف مانند شناسایی چهره و احراز هویت تصویر[۱] استفاده شوند.

- الگوی باینری محلی[۲](LBP) و مبتنی بر شیب: این روش شامل سه مرحله‌ی پیش‌پردازش، استخراج ویژگی و تولید درهم‌ساز می‌باشد [۳۳]. در مرحله‌ی پیش‌پردازش، ابتدا تصاویر رنگی به تصاویر خاکستری تبدیل می‌شوند. سپس اندازه‌ی تمامی تصاویر ورودی به یک استاندارد ثابت تغییر می‌کند. در مرحله‌ی دوم، تصویر ورودی به پنجره‌های غیر همپوشان تقسیم می‌شود و ویژگی الگوی باینری محلی متقارن مرکزی[۳] (CSLBP) [۳۴] برای هر پنجره از تصویر استخراج می‌شود. در مرحله‌ی تولید درهم‌ساز، ویژگی‌های استخراج شده به یک دنباله از اعداد حقیقی تبدیل می شوند.

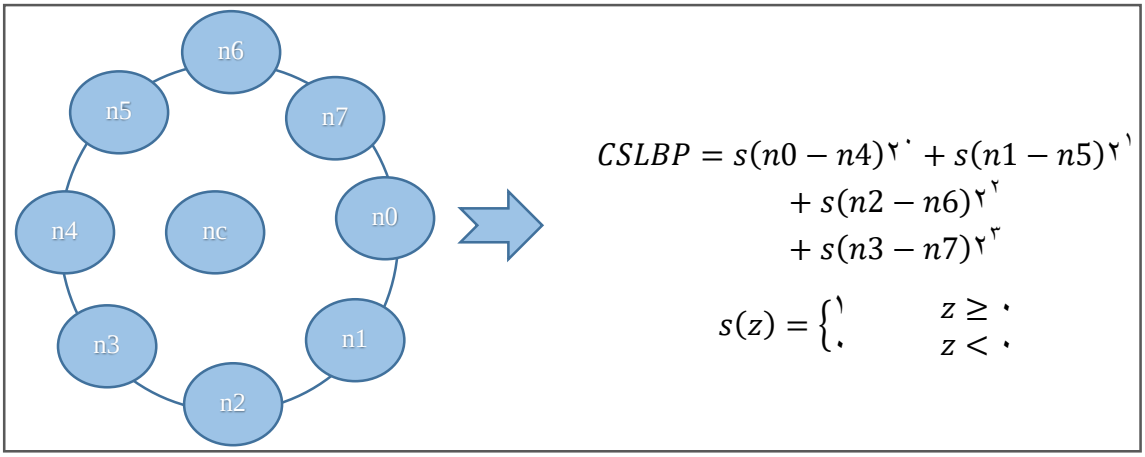
در مرجع فوق، یک توصیفگر ویژگی جدید ارائه می‌شود که ترکیبی از توصیفگرهای مبتنی بر CSLBP و مبتنی بر شیب می‌باشد. همانطور که در شکل (۲-۴) مشاهده می‌شود، روش CSLBP تفاضل بین جفت‌های متقارن مرکزی در پیکسل‌های مخالف یک همسایگی را محاسبه می‌کند. نکته‌ی حائز اهمیت این است که عملگر CSLBP، اطلاعات مربوط به ارزش مقدار تفاضل بین هر جفت پیکسل را از بین می‌برد و فقط اطلاعات علامت را استفاده می‌کند. بنابراین نیاز است از یک توصیفگر مبتنی بر شیب، برای بدست آوردن اندازه و جهت هر پیکسل تصویر استفاده کرد. در این صورت، اطلاعات مربوط به علامت و اندازه‌ی تفاضل بین هر دو جفت ذخیره می‌شود. هر بردار ویژگی دارای طول ۶۴ می‌باشد. در مرحله‌ی آخر برای ایجاد درهم‌ساز، ضرب داخلی بردار ویژگی و یک بردار وزن شبه تصادفی محاسبه

[۱] Image authentication

[۲] Local Binary Pattern

[۳] Center Symmetric Local Binary Pattern

می‌شود. این روش با مجموعه‌ای از ۱۰۰۰ تصویر رنگی از پایگاه داده‌ی PCGT [۳۵] آزمایش شد. استحکام بالا در برابر تغییر روشنایی، طول کم بردار درهم‌ساز و زمان اجرای کوتاه از مزایای این روش در نظر گرفته می‌شود.



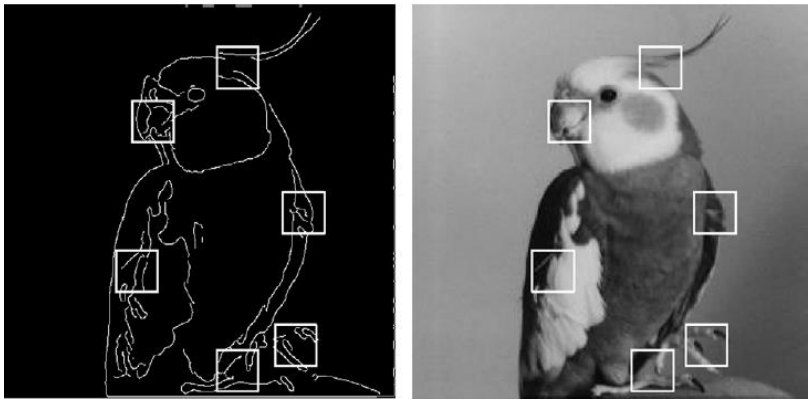
شکل ۲-۴- الگوریتم CSLBP

- در مرجع [۳۶]، یک روش درهم‌سازی تصویر قوی مبتنی بر ویژگی‌های ساختاری برجسته پیشنهاد شده است. در این روش، ابتدا در مرحله‌ی پیش‌پردازش، تصاویر در اندازه‌ای ثابت قرار می‌گیرند و با هدف کاهش نویز، با استفاده از گوسین پایین‌گذر فیلتر می‌شوند. در مرحله‌ی بعدی، ویژگی‌های ساختاری مبتنی بر لبه‌های برجسته‌ی تصویر، تشخیص داده می‌شوند. در این راستا، با استفاده از عملگر کَنی[۱] لبه‌های تصویر پیدا می‌شوند. بعد از آشکارسازی لبه، مقادیر ۰ و ۱ پیکسل‌ها در یک ماتریس به نام نقشه‌ی باینری قرار می‌گیرند. مقدار ۱ بیانگر وجود لبه و ۰ نشان‌دهنده‌ی عدم وجود لبه است. همانطور که در شکل (۲-۵) مشاهده می‌شود، نواحی که بیشترین تعداد پیکسل با مقدار یک را دارند، انتخاب می‌شوند. سپس، ضرایب تبدیل کسینوسی گسسته[۲](DCT) این نواحی و اطلاعات موقعیت مربوط به آنها ذخیره می‌شوند. در نهایت، بعد از کاهش ابعاد با روش PCA، درهم‌ساز نهایی ایجاد می‌شود. این روش با استفاده از پایگاه داده‌ی UCID [۳۷] ارزیابی شده است. این مجموعه داده شامل

[۱] Canny

[۲] Discrete Cosine Transform

۱۳۳۸ تصویر رنگی فشرده نشده است. نتایج نشان می‌دهند که الگوریتم پیشنهادی می‌تواند درهم-سازهایی با قدرت تمایز خوبی ایجاد کند.



شکل ۲-۵- انتخاب نواحی برجسته [۱۴]

- تبدیل فوریه گسسته و تبدیل لگاریتمی-قطبی: در مرجع [۳۸]، از تبدیل فوریه گسسته‌ی چهارگانه[۱] (QDFT) [۳۹] برای ساخت یک درهم‌ساز تصویر مقاوم به تغییرات روشنایی، نویز و چرخش استفاده می‌شود. در این راستا، ابتدا در مرحله‌ی پیش‌پردازش، تصویر ورودی به یک اندازه‌ی ثابت تبدیل می-شوند. سپس یک دایره روی مرکز تصویر در نظر گرفته و پیکسل‌های درون دایره انتخاب می‌شوند و باقی پیکسل‌ها حذف می‌شوند. در مرحله‌ی استخراج ویژگی، ابتدا تصویر به فضای لگاریتمی-قطبی تبدیل می‌شود و سپس ضرایب QDFT از تصویر استخراج می‌شوند. QDFT، یک روش برای مقابله‌ی همزمان با سه کانال تصاویر رنگی ارائه می‌دهد. ضرایب با فرکانس پایین شامل اطلاعات ساختاری و مهم تصویر هستند و ساختار اصلی تصویر را نشان می‌دهند. پس از انتخاب این ضرایب به‌عنوان ویژگی نهایی تصویر، از رابطه‌ی (۲-۶) برای تولید درهم‌ساز استفاده می‌شود.

$$F_G(i) = \begin{cases} ۱, & if\ z_{(i)} - z_{(i+۱)} \geq ۰, \\ ۰, & if\ z_{(i)} - z_{(i+۱)} < ۰, \end{cases} \quad i = ۱, \dots, p-۱. \tag{۲-۶}$$

[۱] Quaternion Discrete Fourier Transform

z بردار ویژگی مربوط به تصویر، p طول بردار ویژگی و F مقدار درهم‌ساز تصویر می‌باشد. این روش نیز با استفاده از ۱۰۰۰ تصویر از پایگاه داده‌ی UCID [۳۷] آزمایش شده است. هر تصویر با استفاده از چندین نوع عملیات از جمله نویز فلفل و نمک، نویز گوسی، تنظیم روشنایی و فیلترهای مختلف فشرده سازی JPEG، برش، مقیاس گذاری و چرخش دستکاری می‌شود، در نهایت ۴۹۹۵۰۰ تصویر ایجاد می‌شود. نتایج آزمایش، عملکرد خوب این روش را در ایجاد درهم‌سازهای مقاوم در مقابل نویز، فشردگی و تغییر مقیاس نشان می‌دهد.

۲-۶- جمع‌بندی

در این فصل، برخی از سیستم‌های شناسایی چهره و روش‌های درهم‌سازی تصویر معرفی شدند. ویژگی‌های استخراج شده از چهره را می‌توان به سه دسته‌ی محلی، سراسری و ترکیبی تقسیم کرد. خلاصه‌ی نقاط ضعف و قوت هر کدام از روش‌ها در جدول (۲-۱) گزارش شده است. بسیاری از روش‌های ذکر شده، آزمایشات را با پایگاه‌های داده‌ی کم‌حجم انجام داده‌اند. برخی از آنها نیز برای شناسایی چهره نیاز به چندین نمونه تصویر به ازای هر فرد دارند. همچنین، این روش‌ها شرایط کنترل نشده را به‌طور خاص در نظر نگرفته‌اند. در این پژوهش، روشی ارائه داده شده است که بتواند تصاویر چهره را با سرعت و دقت بالایی شناسایی کند. علاوه بر این، چالش‌های مربوط به پایگاه‌های داده‌ی حجیم، وجود تنها یک نمونه تصویر به‌ازای هر فرد و برخی شرایط کنترل نشده، در این پژوهش به‌طور خاص در نظر گرفته شدند. در فصل بعدی به معرفی الگوریتم‌های به‌کاررفته در این پژوهش پرداخته می‌شود.

جدول ۲-۱- نمونه روش‌های شناسایی چهره

نویسنده (سال)	روش	نقاط قوت	نقاط ضعف
چن و همکاران [۲۱] (۲۰۲۰)	(ALM+PDV) Hashing	دقت و سرعت شناسایی بالا	حساس به تغییرات نور
زنگنه و همکاران [۳۱] (۲۰۲۰)	DCNN	عدم نیاز به تصاویر با وضوح بالا	نیاز به نمونه آزمایشی زیاد
عباس‌پور و همکاران [۱۱] (۲۰۲۰)	NMF+FREAK	دقت شناسایی بالا	سرعت شناسایی پایین
نیکان و حسن‌پور [۱۰] (۲۰۲۰)	NMF	سرعت شناسایی بالا	حساس به تغییرات زاویه و پوشیدگی در چهره
بخشی و همکاران [۲۷] (۲۰۱۸)	FFT	دقت شناسایی بالا	محاسبات پیچیده و سرعت شناسایی پایین
زنگ و همکاران [۳۲] (۲۰۱۸)	CNN	عدم نیاز به نمونه آزمایشی زیاد	حساس به پایگاه داده‌ی حجیم
کاسیو و همکاران [۳۰] (۲۰۱۶)	(LSH+PLSR) Hashing	دقت شناسایی بالا	پیچیدگی حافظه‌ی
زنگ و همکاران [۱۷] (۲۰۰۹)	(SIFT+LSH) Hashing	سرعت شناسایی بالا	شرایط محیطی کنترل نشده
شریف و همکاران [۲۸] (۲۰۰۹)	DCT Hashing	دقت شناسایی بالا	حساس به پایگاه داده‌ی حجیم
گاوو و همکاران [۲۵] (۲۰۰۸)	FLDA+SVD	عدم نیاز به نمونه آزمایشی زیاد	حساس به پایگاه داده‌ی حجیم
تورک و همکاران [۲۴] (۱۹۹۱)	PCA	تغییرات روشنایی	حساس به تغییرات زاویه‌ی چهره

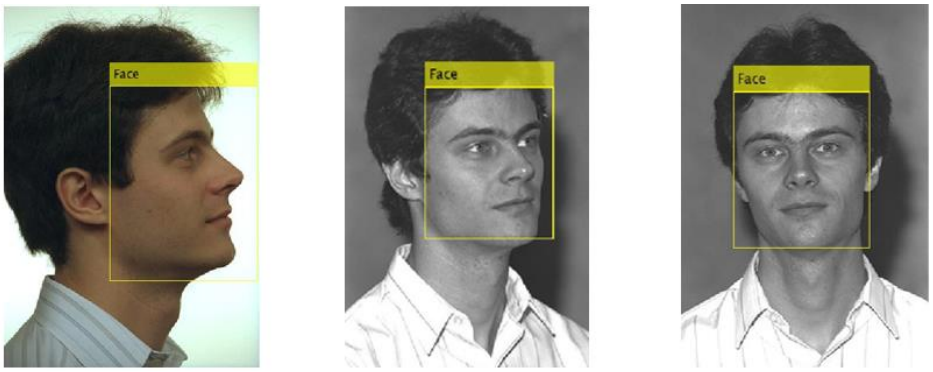
فصل ۳: پیشینه‌ی الگوریتم‌های به کار رفته

۳-۱- مقدمه:

همانطور که در فصل اول بیان شد، در این پژوهش، یک روش موثر و کارآمد برای شناسایی چهره با استفاده از روش‌های درهم‌سازی تصویر معرفی شده است. در این فصل، پیشینه‌ی الگوریتم‌های به کار رفته در این پژوهش به همراه عملکرد آنها توضیح داده می‌شوند.

۳-۲- روش تشخیص ناحیه‌ی چهره:

شبکه‌ی عصبی پیچشی چند منظوره‌ای[۱](MTCNN)، یک الگوریتم کارا و موثر برای تشخیص چهره است [۴۰]. در این الگوریتم، سه مرحله پیچش به‌صورت آبشاری روی تصویر چهره اعمال می‌شود. در مرحله‌ی اول، ابتدا پنجره‌های نامزد ناحیه‌ی چهره از طریق یک شبکه‌ی پیشنهاد[۲] (P-Net) تولید می شوند. سپس، نامزدهای همپوشان با استفاده از روش[۳]NMS ادغام می‌شوند. در مرحله‌ی دوم، تعداد زیادی از نامزدهای کاذب از طریق یک شبکه‌ی تصفیه[۴] (R-Net) حذف می‌شوند. مرحله‌ی سوم مشابه مرحله‌ی دوم است، اما هدف آن، شناسایی ناحیه‌ی چهره با نظارت و دقت بیشتری می‌باشد در نهایت، ناحیه‌ی چهره با استفاده از پنج نقطه مشخص می‌شود. این الگوریتم می‌تواند چهره‌های زاویه‌دار را با دقت بیشتری نسبت به الگوریتم معروف Viola-Jones [۱۲]، تشخیص دهد . نتایج حاصل از این روش در شکل (۳-۱) قابل مشاهده است.



شکل ۳-۱- نتیجه‌ی تشخیص ناحیه‌ی چهره در زوایای مختلف با استفاده از الگوریتم MTCNN

[۱] Multi-Task cascaded Convolutional Neural Networks

[۲] Proposal Network
[۳] Non-Maximum Suppression
[۴] Refine Network

۳-۳- روش‌های بررسی کیفیت روشنایی تصویر

پیش‌پردازش تصاویر چهره، یکی از مراحل مهم در سیستم‌های شناسایی چهره است. در این مرحله، تصاویر چهره از لحاظ روشنایی و کنتراست بهبود می‌یابند. برابرسازی هیستوگرام وفقی با محدودیت کنتراست (CLAH[۱])، یکی از روش‌های بهسازی تصویر است [۴۱]. در این روش، بهسازی تصویر به‌صورت غیر یکسان در تمامی نواحی تصویر انجام می‌شود. بدین صورت که تصویر به نواحی کاشی‌وار تقسیم می‌شود و در نواحی هموار بهسازی کمتری انجام می‌شود [۴۲]. این روش از تقویت نویز و اثر رنگ پریدگی جلوگیری می‌کند. شکل (۳-۲) نمونه‌ی از انجام این پیش‌پردازش را نشان می‌دهد.

الف) تصویر اصلی ب) تصویر پیش‌پردازش شده

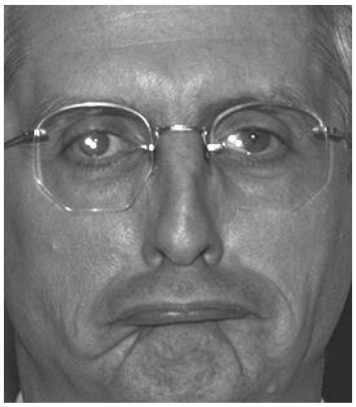
شکل ۳-۲- تصویر اصلی و تصویر با پیش‌پردازش CLAHE

همچنین، به منظور بهبود کیفیت روشنایی تصویر، می‌توان دامنه‌ی تغییرات روشنایی آن را تغییر داد [۴۱]. در این روش، با استفاده از رابطه‌ی (۳-۱) روشنایی تصویر تغییر می‌کند.

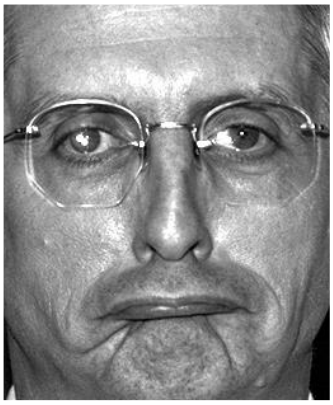
$$A(i,j) = I \times \left(\frac{max-min}{۲۵۵}\right) \qquad (۱-۳)$$

I نشان دهنده‌ی تصویر، max و min نیز به ترتیب بیشینه و کمینه‌ی مقدار آن می‌باشد. شکل (۳-۳) نمونه‌ی از انجام این پیش‌پردازش را نشان می‌دهد.

[۱] Contrast Limited Adaptive Histogram



الف) تصویر اصلی



ب) تصویر پیش‌پردازش شده

شکل ۳-۳- تصویر اصلی و تصویر با پیش‌پردازش تنظیم شدت روشنایی تصویر

۳-۴- روش استخراج ویژگی‌های محلی

هنگامی که می‌خواهیم ویژگی‌های محلی تصاویر چهره را با هم تطبیق دهیم، اغلب با چالش تغییرات مقیاس، زاویه و فاصله‌ی تصویر برداری از یک جسم روبرو می‌شویم. در این حالت، اگر یک ویژگی را از دو تصویر چهره، با بکارگیری از یک همسایگی با طول ثابت تطبیق دهیم، به دلیل داشتن تغییر مقیاس، شدت روشنایی آن‌ها با هم تطابق نخواهد داشت. در این راستا، داشتن یک ضریب مقیاس متناظر با هر یک از نقاط ویژگی به رفع این مشکل کمک می‌کند.

توصیفگر ویژگی‌های مقاوم تسریع داده شده[۱](SURF) [۴۳] توسط هربرت بای[۲]در سال ۲۰۰۶ پیشنهاد شده است و نسخه‌ی بهینه یافته‌ی الگوریتم SIFT، از لحاظ سرعت است. این روش در برابر تغییرات مقیاس، چرخش، تا حدودی تغییرات روشنایی و تغییرات ناشی از زاویه‌ی دید تصویر برداری مقاوم است. با استفاده از این الگوریتم، سعی در شناسایی نقاط کلیدی را داریم که از مقیاس‌های مختلف یک تصویر چهره قابل شناسایی هستند. برای یافتن این نقاط، مقیاس‌های مختلف تصویر به همراه ساختارهای سطوح تاری متفاوت، باید بررسی شوند. در این راستا، باید مشتقات یک تصویر را به کمک فیلترهای گوسی[۳]

[۱] Speeded Up Robust Features

[۲] Herbert Bay

[۳] Gaussian

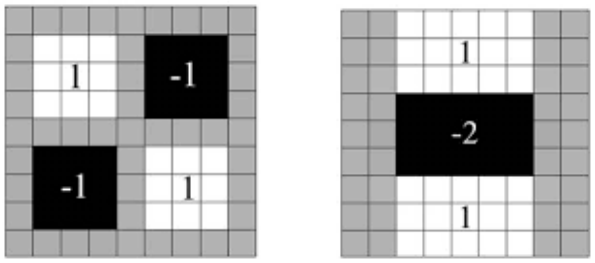
تخمین زد. این فیلترها از یک پارامتر به نام سیگما[۱]استفاده می‌کنند که بیانگر اندازه‌ی هسته است. این پارامتر با واریانس تابع گاوسی مورد استفاده برای ایجاد فیلتر، متناظر است و مقیاسی که در آن مشتقات تخمین زده می‌شوند را تعریف می‌کند. این تعریف به این معنی است که فیلتر با مقدار سیگمای بزرگ‌تر، جزئیات ظریف‌تری از تصویر را هموار می‌کند. اگر لاپلاسین[۲]یک نقطه از تصویر را با بکارگیری فیلتری‌های گاوسی در مقیاس‌های مختلف محاسبه کنیم، مقادیر مختلفی نیز بدست می‌آید. با توجه به تغییر پاسخ فیلتر در مقیاس‌های مختلف، می‌توان به یک منحنی رسید که در نهایت به یک مقدار بیشینه می‌رسد. اکنون اگر این مقدار بیشینه را برای دو تصویر از یک چهره که مقیاس‌های مختلفی دارند، استخراج کنیم، آنگاه نسبت این دو مقدار، سیگمای بیشینه متناظر با نسبت مقیاس‌هایی است که این دو تصویر چهره در آن تصویر برداری شده‌اند. در ادامه، نحوه‌ی پیاده‌سازی این الگوریتم توضیح داده می‌شود. برای یافتن ویژگی‌ها، ابتدا ماتریس هسین [۳]در هر پیکسل x و y از تصویر I توسط رابطه (۳-۲) محاسبه می‌شود.

$$E(x,y) = \begin{bmatrix} \frac{\delta^2 I}{\delta^2 x^2} & \frac{\delta^2 I}{\delta x \delta y} \\ \frac{\delta^2 I}{\delta x \delta y} & \frac{\delta^2 I}{\delta y^2} \end{bmatrix} \qquad (۳-۲)$$

دترمینان[۴]این ماتریس، قدرت این انحنا را معرفی می‌کند. به دلیل اینکه این ماتریس از مشتقات مرتبه‌ی دوم تشکیل شده است، می‌توان آن را به کمک هسته‌های گاوسی لاپلاسی با مقیاس‌های مختلف سیگما (σ) محاسبه کرد. در این صورت، ماتریس هسین تابعی از سه متغیر $E(x,y,\sigma)$ خواهد بود. در نهایت ویژگی‌ای مقاوم است که دترمینان ماتریس هسین آن هم در فضای مکانی و هم فضای مقیاس، یک بیشینه محلی باشد. محاسبه‌ی تمام این مشتقات در مقیاس‌های مختلف، پیچیدگی محاسباتی سنگینی دارد. لذا الگوریتم SURF این فرآیند را کارآمدتر می‌کند. بدین صورت که از هسته‌های گاوسی تخمینی که فقط با چند جمع صحیح درگیر هستند، استفاده می‌کند. شکل (۳-۴) ساختار این هسته‌ها را نشان می‌دهد.

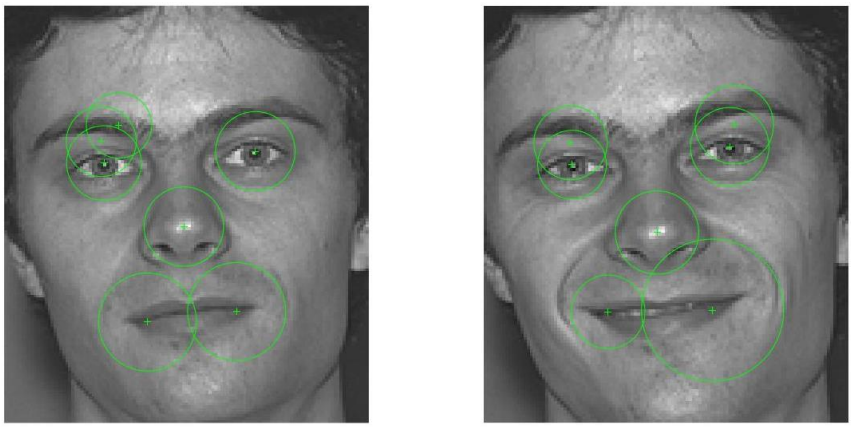
---

[۱] Sigma
[۲] Hessian
[۳] Laplacian
[۴] Determinant



شکل ۳-۴- ساختار هسته‌های گاوسی در توصیفگر SURF [۳۷]

از هسته‌ی سمت چپ به منظور تخمین مشتقات مرتبه‌ی دوم استفاده می‌شود. هسته‌ی سمت راست نیز مشتق مرتبه دوم را در جهت عمودی تخمین می‌زند. همچنین یک نسخه‌ی دورانی از هسته‌ی دوم، مشتق مرتبه‌ی دوم را در جهت افقی تخمین می‌زند. اندازه‌ی کوچکترین هسته برابر با ۹×۹ است و پیکسل متناظر سیگما ۰/۵ می‌باشد. درادامه، هسته‌هایی با اندازه‌های بزرگتر به‌صورت پیاپی اعمال می‌شوند. با یافتن بیشینه‌ی محلی، مکان دقیق نقاط ویژگی از طریق درون‌یابی در فضای تصویر و مقیاس به دست می‌آیند. بنابراین، این الگوریتم، یک موقعیت و یک مقیاس برای هر یک از ویژگی‌های تشخیص داده شده تعریف می‌کند. ضریب مقیاس، اندازه‌ی یک پنجره در اطراف نقطه ویژگی را تعریف می‌کند. بنابراین، صرف نظر از اینکه تصویری که ویژگی به آن تعلق دارد، در چه مقیاسی تصویر برداری شده است، اطلاعات بصری یکسانی دارد. طول بردار توصیفگر هر ویژگی محلی از تصویر برابر با ۶۴ می‌باشد و تعداد ویژگی‌های محلی به ازای هر تصویر چهره متفاوت است. می‌توان تمامی توصیفگرهای مربوط به یک تصویر چهره را به همراه شناسه‌ی مربوط به فرد، در یک ماتریس ذخیره کرد. شکل (۳-۵)، ویژگی‌های محلی استخراج‌شده از چند نمونه تصویر را نشان می‌دهد. هر دایره‌ی سبز در تصاویر، بیانگر یک ویژگی محلی می‌باشد. همانطور که در شکل مشاهده می‌شود، ویژگی‌های استخراج شده از تصاویر متعلق به یک فرد، با هم اشتراک دارند. یعنی نقاط کلیدی وجود دارند که به ازای هر حالت و مقیاسی از چهره ثابت هستند. به‌عنوان مثال، در مورد تصاویر نمونه‌ی (الف) در شکل شماره (۳-۵)، نقاط کلیدی در هر دو نمونه تقریبا برابر هستند. یعنی ویژگی‌های محلی‌شان نزدیک به هم هستند. مورد (ب) نیز، با وجود اینکه زاویه‌ی چهره‌ی فرد تغییر کرده است، اما ویژگی‌های محلی در اکثر موارد مشترک هستند.



الف) دو نمونه تصویر از فرد شماره یک

ب) دو نمونه تصویر از فرد شماره دو

شکل ۳-۵- نمایش ویژگی‌های توصیفگر محلی SURF

۳-۵- درهم‌سازی تصویر

درهم‌سازی تصویر روشی کارآمد برای استخراج ویژگی‌های ضروری یک تصویر و نگاشت آن‌ها به یک دنباله‌ی کوتاه از اعداد با اندازه‌ی ثابت است. بدین‌صورت که تابع درهم‌ساز به داده‌ی دلخواه اعمال می‌شود و داده‌ای با اندازه‌ی کمتر و به اصطلاح درهم‌ساز تولید می‌کند. درهم‌سازهای حاصل از تصاویر با حجم ذخیره‌سازی کمتر و ویژگی‌های قوی‌تر در جدول ذخیره می‌شوند. به‌همین دلیل، عملیات احراز هویت با توجه به درهم‌ساز هر تصویر، با سرعت، دقت و قدرت بالایی انجام می‌شود. LSH، یک روش درهم‌سازی است که توابع آن، مقادیر درهم‌ساز را به‌صورت تصادفی ایجاد می‌کنند [۴۴،۴۵]. در ادامه، ابتدا به مفهوم

جدول درهم‌ساز پرداخته می‌شود، سپس مفهوم LSH بیان می‌شود. در نهایت، یک نسخه‌ی توسعه یافته‌ی الگوریتم LSH به نام درهم‌سازی حساس به مکان بر اساس توزیع‌های پایدار[۱] (P-stable LSH) معرفی می‌شود.

استفاده از الگوریتم LSH مزایای زیر را به دنبال دارد:

۱- به دلیل نداشتن محاسبات پیچیده، زمان پاسخگویی بسیار کم است.

۲- در این الگوریتم، اقدامات مربوط به هر مرحله مشخص است و برای آن توجیهی وجود دارد.

۳- سیستم به ترتیب داده‌های ورودی وابسته نیست و در هر مرحله فقط روی داده‌ی جاری تمرکز می‌کند.

۴- تشخیص نویز به‌صورت خودکار انجام می‌شود، زیرا درهم‌سازهای پراکنده حذف می‌شوند.

۵- سیستم، تنها با محاسبه داده‌ی جدید به روزرسانی می‌شود و نیاز به محاسبه دوباره‌ی کل داده‌ها نیست.

۶- این روش برای جریان داده[۲] مناسب است و عملکرد خوبی دارد.

۳-۵-۱- جدول درهم‌ساز

یک جدول درهم‌ساز، ساختمان داده‌ای است که به ما اجازه می‌دهد به سرعت از یک نماد به یک مقدار برسیم. با ورود یک نماد، می‌توان با استفاده از تابع درهم‌ساز، یک مقدار صحیح بدست آورد که به عنوان اندیس جدول درهم‌ساز مورد استفاده قرار می‌گیرد. بنابراین، یک نماد با تعداد زیادی از بیت و حروف به یک اندیس از جدول درهم‌ساز نگاشت می‌شود. درصورتی که دو نماد به یک اندیس از جدول درهم‌ساز نگاشت شوند، برخورد اتفاق می‌افتاد. یک تابع درهم‌ساز مناسب، دو نماد که به هم نزدیک هستند را در یک مکان[۳] با اندیس یکسان از جدول درهم‌ساز نگاشت می‌دهد.

---

[۱] P-stable distribution LSH
[۲] Streaming data
[۳] Bucket

۳-۵-۲- درهم‌ساز حساس به مکان

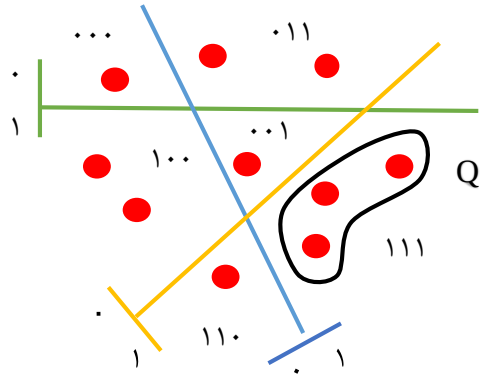
این الگوریتم، مبتنی بر یک ایده‌ی ساده است که اگر دو نقطه به هم نزدیک باشند، آنگاه پس از نگاشته شدن توسط یک تابع درهم‌ساز، این نقاط همچنان به هم نزدیک می‌مانند. بررسی نزدیکی بر اساس یک معیار فاصله که روی اعضا اعمال می‌شود، انجام می‌پذیرد. اعضایی که بر اساس این معیار فاصله به یکدیگر نزدیک‌ترند، با احتمال بالا به یک اندیس از جدول نگاشت داده می‌شوند. می‌توان از این روش برای خوشه-بندی داده‌ها و جستجوی نزدیکترین همسایه استفاده کرد. برای یک دامنه‌ی S از مجموعه نقاط با فضای متری Z، یک خانواده‌ی LSH مانند زیر تعریف می‌شود:

تعریف (۱-۱) : یک خانواده‌ی $H = \{h: S \to U\}$ را $(R, \mathrm{CR}, \mathrm{P}_1, \mathrm{P}_2)$ حساس برای Z گوییم، که برای $(\mathrm{R} < \mathrm{CR}, \mathrm{P}_2 < \mathrm{P}_1)$، اگر برای هر دو عضو $p, q \in S$

- اگر $Z(p,q) \leq R$ آنگاه $Pr_H[\mathrm{h(q)} = \mathrm{h(p)}] \geq \mathrm{P}_1$
- اگر $Z(p,q) \geq CR$ آنگاه $Pr_H[h(q) = h(p)] \leq P_2$

تعریف (۱-۱) بیان می‌کند، اگر یک تابع درهم‌ساز h را از خانواده LSH انتخاب کنیم، در این صورت هر دو نقطه که در یک همسایگی نزدیک در فاصله‌ی کمتر از R قرار دارند، با احتمال بیشتر از P_1 به یک اندیس یکسان از جدول درهم‌ساز نگاشت می‌شوند. همچنین، نقاطی که در فاصله‌ی بیشتر از CR نسبت بهم قرار دارند، با احتمال کمتر از P_2 نگاشت یکسانی دارند. یک درهم‌ساز حساس به مکان استاندارد، نگاشت‌ها را به‌صورت تصادفی انجام می‌دهد. به‌عنوان مثال، اگر تعدادی خط در فضای دو بعدی رسم کنیم که به‌صورت تصادفی، تعدادی از نقاط در یک سمت آن قرار بگیرند و باقی در سمت دیگر خط قرار بگیرند. برای دو نقطه که به‌هم نزدیک هستند، احتمال اینکه این خط تصادفی طوری رسم شود که نقاط در دو سمت مخالف قرار بگیرند، خیلی کم است. نمونه‌ای از نگاشت دودویی دو بعدی الگوریتم LSH در شکل (۳-۶) نمایش داده شده است. اگر یک خط تصادفی بنفش رنگ را در این فضا عبور دهیم، نقاطی که در سمت پایین این خط قرار دارند، مقدار یک و نقاطی که در سمت بالای این خط قرار دارند، مقدار صفر

می‌گیرند. به همین ترتیب، دو خط تصادفی دیگر در این فضا رسم می‌شود و به ازای هر خط، مقدار صفر یا یک به نقاط اختصاص داده می‌شود. همانطور که مشاهده می‌شود، فضا به زیربخش‌هایی تقسیم می‌شود. نقاطی که در یک زیر بخش قرار گرفتند، به یک اندیس از جدول درهم‌ساز نگاشت می‌شوند. به‌عنوان مثال، نقاطی که در زیربخش Q قرار گرفتند، به اندیس ۱۱۱ از جدول درهم‌ساز نگاشت می‌شوند. با توجه به این مثال، می‌توان ایده‌ی الگوریتم LSH را به‌خوبی درک کرد. یعنی، نقاطی که بهم نزدیک باشند، احتمال اینکه خطی آنها را به دو قسمت مختلف تقسیم کند، خیلی کم است. بنابراین پس از نگاشته‌شدن، این نقاط همچنان به‌هم نزدیک هستند.

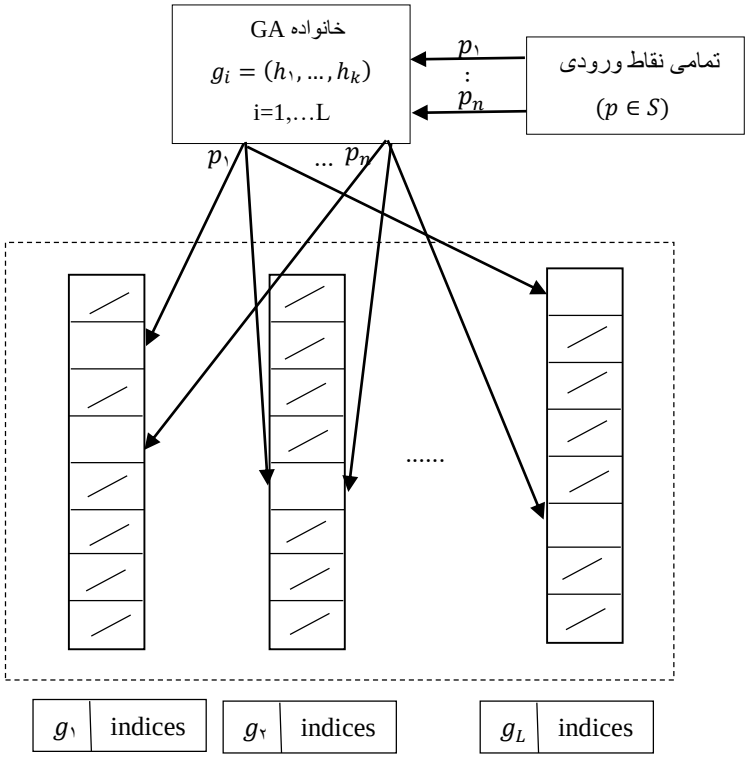


شکل ۳-۶- مثال مربوط به تعریف LSH

همانطور که پیشتر بیان شد، هر دو نقطه‌ی نزدیک، با احتمال بیشتر از P_1 به یک اندیس یکسان از جدول درهم‌ساز نگاشت می‌شوند. همچنین، نقاطی که در فاصله‌ی دورتری نسبت بهم قرار دارند، با احتمال کمتر از P_2 نگاشت یکسانی دارند. زمانی الگوریتم LSH به خوبی عمل می‌کند که اختلاف بین P_1 و P_2 به اندازه‌ی کافی بزرگ باشد. می‌توان این فاصله را با اعمال چندین تابع درهم‌ساز افزایش داد. با اعمال هر تابع درهم‌ساز، یک بُعد به سیستم کدگذاری اضافه خواهد شد. بر این اساس، با اضافه کردن چندین تابع درهم‌ساز، اختلاف بین نقاط غیر همسایه افزایش خواهد یافت. در این راستا، یک خانواده‌ی $GA = \{g: S \to U^k\}$ تعریف می‌شود. به‌طوریکه $g(v) = (h_1(p), \dots, h_k(p))$ و $h_i \in H$ می‌باشد. تعداد L تابع $g_1, g_2, \dots, g_L$ به‌صورت مستقل و یکنواخت از GA انتخاب می‌شوند. در مورد L و k بعدا توضیحات بیشتری داده می‌شود.

با توجه به توضیحات فوق، فرض کنید دو مجموعه نقاط $p_{1,2,\dots,n}$ و $q_{1,2,\dots,k}$ داشته باشیم. ابتدا با استفاده از توابع درهم‌سازی، مجموعه نقاط p را در مکانی از حافظه ذخیره می‌کنیم (عملیات آموزش). همانطور که در شکل (۳-۷) نشان داده شده است، در زمان ذخیره‌سازی، هر $p \in S$ (نقاط مجموعه داده) در مکان $g_j(p)$ برای $j = ۱, \dots L$ نگاشت داده می‌شوند. پس از ذخیره‌سازی، می‌توان نقاط q را بازیابی کرد (عملیات آزمون). اگر دو نقطه‌ی p و q نزدیک باشند، با احتمال بسیار زیادی به یک اندیس از جدول درهم‌ساز نگاشت می‌شوند. بنابراین، مکان بازیابی آن‌ها یکسان خواهد بود.

با استفاده از این ایده می‌توان همسایه‌های نزدیک به یک نقطه‌ی ورودی را یافت. در این راستا، اگر بخواهیم همسایه‌های نزدیک به نقطه‌ی ورودی q را بیابیم، ابتدا تمامی نقطه‌های p که در اندیس‌های $g_۱(q), g_۲(q), \dots, g_L(q)$ نگاشت شده اند را بازیابی می‌کنیم. سپس از میان این نقطه‌ها، نزدیک‌ترین همسایه‌ها را جستجو می‌کنیم. همچنین می‌توان فقط L' نقطه را بازیابی کرد و برای جستجو تنها از این نقاط استفاده کرد. نشان داده شده است که اگر L' سه برابر L باشد، به خوبی می‌تواند مساله‌ی همسایه‌ها نزدیک را حل کند [۴۵]. مقدار J به‌صورت $J = n^{\rho}$ در نظر گرفته می‌شود، که در آن $\rho = \frac{\ln 1/p_1}{\ln 1/p_2}$ است. مقدار K به‌صورت $K = log^{n}_{1/P_2}$ بدست می‌آید. بنابراین LSH در زمان $O(n^{\rho})$ اجرا می‌شود. یعنی زمانی که $P_1 > P_2$ است، نسبت به n زیر خطی می‌باشد. حافظه‌ی مورد نیاز این الگوریتم، $O(dn + n^{1+\rho})$ می‌باشد.



شکل ۳-۷- نمایش خانواده‌ی GA از LSH

نسخه‌های مختلفی از الگوریتم LSH ارائه شده است که در آن توابع درهم‌ساز و معیارهای اندازه‌گیری نزدیکی بین نقاط، متفاوت است. توضیحات کامل در مورد برخی از این روش‌ها در مرجع [۲۰] بیان شده است. در ادامه‌ی این گزارش، یک نسخه‌ی الگوریتم P-stable LSH معرفی می‌شود.

۳-۵-۳- درهم‌سازی بر اساس توزیع‌های پایدار

الگوریتم LSH استاندارد برای فضای متری همینگ طراحی شده است. به‌همین دلیل، برای استفاده در یک فضای متری دیگر، به عملیات جایسازی[۱] نیاز دارد که این امر، منجر به بار محاسباتی می‌شود. P-stable LSH، یک خانواده‌ی LSH بر اساس توزیع‌های پایدار p ارائه می‌دهد که روی $p\epsilon(۰,۲]$ قابل اجرا است [۴۶]. از مزایای این الگوریتم می‌توان به این موارد اشاره کرد: پیاده‌سازی سریع و آسانی دارد، می‌تواند

[۱] Embedding

مستقیما روی داده‌های موجود در فضای اقلیدسی اجرا شود، دارای زمان پاسخ‌گویی سریعتری است و دقت بهتری نسبت به الگوریتم LSH دارد. در ادامه، ابتدا مفهوم توزیع‌های پایدار بیان می‌شود. سپس چگونگی ایجاد توابع درهم‌ساز در الگوریتم P- stable LSH بیان می‌شود.

۳-۵-۳-۱- مفهوم توزیع‌های پایدار

در نظریه‌ی آمار، توزیع T را پایدار گویند اگر برای هر مجموعه از اعداد حقیقی $v_1, v_2, \dots, v_n$ و هر مجموعه از متغیرهای تصادفی $Y_1, Y_2, \dots, Y_n$ مستقل با توزیع یکسان T، متغیر تصادفی $\sum_i v_i Y_i$ توزیع یکسانی با توزیع متغیر تصادفی $(\sum_i |v_i|^b)^{1/b} Y$ داشته باشد. که در آن Y یک متغیر تصادفی با توزیع T است و $b \in (0,2]$ می‌باشد.

هدف استفاده از توزیع پایدار، تولید یک بردار تصادفی x با بُعد d می‌باشد، به‌طوریکه مقادیر آن به‌طور مستقل توسط توزیع پایدار انتخاب شود. اگر یک بردار ورودی v با بُعد d داشته باشیم، ضرب داخلی $a \cdot v$ یک متغیر تصادفی است که به صورت $(\sum_i |v_i|^b)^{1/b} Y (i.e., ||v||_b Y)$ توزیع می‌شود. a برداری است که از توزیع نرمال پیروی می‌کند و ابعادی برابر با بردار ویژگی ورودی دارد.

۳-۵-۳-۲- چگونگی ایجاد تابع درهم‌ساز

P-stable LSH برای نگاشت ورودی از ضرب داخلی $a \cdot v$ استفاده می‌کند. برای دو ورودی با بردارهای (v_1, v_2)، فاصله‌ی بین نگاشت‌های آنها $(a.v_1 - a.v_2)$ به صورت $||v_1 - v_2||_b X$ توزیع می‌شود که در آن X توزیع پایدار b است. تابع درهم‌سازی که نتیجه می‌شود در رابطه‌ی (۳-۳) تعریف شده است.

$$h_j(v) = \left\lfloor \frac{a \cdot v + bb}{r} \right\rfloor \qquad (۳-۳)$$

پارامتر r نشان دهنده‌ی عرض هر خانه از جدول درهم‌ساز می‌باشد، bb یک متغیر تصادفی یکنواخت در بازه $[0, r]$ است. افزایش عرض r باعث افزایش تعداد نقاطی می‌شود که به هر اندیس نگاشت می‌شوند. برای یافتن همسایه‌های نزدیک، باید یک جستجوی خطی روی تمامی نقاطی که در یک اندیس یکسان با

نقطه‌ی ورودی قرار می‌گیرند انجام داده شود. بنابراین تغییر r، یک موازنه بین جدول بزرگتر با جستجوی خطی کوچکتر یا جدول کوچکتر با تعداد نقاط بیشتر ایجاد می‌کند.

برای کاهش تاثیر انتخاب مقدار نادرست r در هر نگاشت، L نگاشت مستقل تشکیل می‌شود و همسایه‌ها با کمک همه‌ی نگاشت‌ها بدست می‌آیند. در این حالت، احتمال اینکه نقاط نزدیک در تمامی نگاشت‌ها در اندیس نادرستی قرار بگیرد، بسیار کم است. احتمال اینکه دو نقطه به یک اندیس از جدول درهم‌ساز نسبت داده شوند توسط رابطه‌ی (۳-۴) بدست می‌آید.

$$t(c) = Pr_{a,bb}\left[h_{a,bb}(v_1) = h_{a,bb}(v_2)\right] = \int_0^r \frac{1}{c} f_p\left(\frac{t}{c}\right)\left(1 - \frac{t}{r}\right) dt \qquad (۳-۴)$$

که در آن $f_p(t)$ تابع چگالی توزیع پایدار است و $c = ||v_1 - v_2||_b$ می‌باشد. برای هر خانه از جدول درهم‌ساز با عرض r، احتمال برخورد با افزایش فاصله‌ی c کاهش می‌یابد.

۳-۶- درهم‌سازی بر اساس تبدیل کسینوسی گسسته

همانطور که بیان شد، درهم‌ساز تصویر، یک امضای منحصر به فرد است که محتوای بصری تصویر را به شیوه‌ای مخصوص، به‌صورت فشرده نشان می‌دهد. روش‌های درهم‌سازی تصویر، ویژگی‌های ضروری تصویر را استخراج می‌کنند و به یک دنباله از اعداد با طول کوتاه، نگاشت می‌دهند. یک روش درهم‌سازی تصویر، باید قابلیت تمایز خوبی برای تصاویر غیر مشابه داشته باشد. همچنین، درهم‌ساز تصویر باید نسبت به تغییراتی مانند فشرده‌سازی JPEG، تاری، نویز و برخی عملیات مشابه دیگر ثابت باشد. ضرایب DCT می‌توانند محتوای بصری تصویر را به خوبی بیان کنند. همچنین، ضرایب DCT با فرکانس پایین ضروری تر از ضرایب فرکانس بالا هستند، زیرا مقادیر بیشتری دارند و حاوی بخش اصلی انرژی سیگنال هستند [۴۷]. به‌همین دلیل، روش‌های زیادی برای تولید درهم‌ساز، از ضرایب DCT استفاده می‌کنند. در ادامه، یکی از روش درهم‌سازی تصویر مبتنی بر ضرایب DCT معرفی می‌شود.

در الگوریتم مرجع [۴۸]، برای ایجاد درهم‌ساز، فرکانس‌های کم اهمیت DCT حذف می‌شوند و فقط مهمترین فرکانس‌های تصویر حفظ می‌شوند. در روش مورد نظر، درهم‌ساز هر تصویر با استفاده از سه مرحله ایجاد می‌شود. در مرحله‌ی اول، پیش‌پردازش‌هایی از جمله تنظیم روشنایی، تغییر اندازه‌ی تصویر و رفع نویز در تصویر انجام می‌شود. برای تنظیم روشنایی تصویر، مقدار تمام پیکسل‌ها با یک مقدار ثابت بهبود می‌یابد. در ادامه، با استفاده از درون‌یابی دو خطی، اندازه‌ی تصویر به $M \times M$ تغییر می‌یابد. همچنین، عواملی مانند خرابی حاصل از فشرده سازی JPEG و نویز، توسط فیلتر گوسین کاهش می‌یابد. در مرحله‌ی دوم، ابتدا تصویر به پنجره‌های غیر همپوشان تقسیم می‌شود، سپس از آن‌ها برای ایجاد ماتریس ویژگی‌های شامل ضرایب DCT استفاده می‌شود. سرانجام، با فشرده‌سازی ماتریس‌های ویژگی، یک دنباله‌ی عدد باینری تولید می‌شود.

همانطور که در بالا توضیح داده شد، در مرحله‌ی دوم، تصویر به پنجره‌های غیر همپوشان $m \times m$ تقسیم می‌شود. تعداد کل پنجره‌های تصویر ورودی $N = (\frac{M}{m})^۲$ در نظر گرفته می‌شود. فرض می‌شود B_i ، i امین پنجره است. که از بالا به پایین و از چپ به راست نمایه می شود. ($۱ \le i \le N$). DCT به هر پنجره اعمال می شود. ضریب DCT در ردیف j و ستون k مانند رابطه (۳-۵) بدست می‌آید.

$$t(c) = Pr_{a,b}\left[h_{a,b}(v_۱) = h_{a,b}(v_۲)\right] = \int_۰^r \frac{۱}{c} f_p\left(\frac{t}{c}\right)\left(۱ - \frac{t}{r}\right) dt \quad C_i(u,v) = \qquad (۵-۳)$$
$$a(u)a(v)\sum_{j=۰}^{m-۱}\sum_{k=۰}^{m-۱} B_i\,(j,k) cos\left[\frac{(k+۱)u\pi}{k}\right] cos\left[\frac{(k+۱)v\pi}{k}\right]$$

$u, v = ۰, ۱, \dots, m-۱$

$a(u)$ نیز مانند رابطه‌ی (۳-۶) بدست می‌آید.

$$a(u)=\begin{cases} \sqrt{۱/m}, & if\ u = ۰ \\ \sqrt{۲/m} & Otherwise \end{cases} \qquad (۶-۳)$$

به همین روال ضرایب DCT در ردیف اول و ستون اول مانند رابطه‌ی (۳-۷) و (۳-۸) بدست می‌آید.

$$C_i(۱, v) = \frac{\sqrt{۲}}{m}\sum_{j=۰}^{m-۱}\sum_{k=۰}^{m-۱} B_i(j,k) cos\left[\frac{(k+۱)v\pi}{k}\right] \qquad (۷-۳)$$

$v = ۰, \dots, m-۱$

$$C_i(u,۱) = \frac{\sqrt{۲}}{m}\sum_{j=۰}^{m-۱}\sum_{k=۰}^{m-۱} B_i(j,k)cos\left[\frac{(k+۱)u\pi}{k}\right] \quad (۸-۳)$$

$$u = ۰, ۱, \dots, m-۱$$

ضرایب در ستون و ردیف اول تا n امین ضریب $(n \leq m)$ در دو ماتریس ویژگی ذخیره می‌شوند. در ادامه، ابتدا داده‌ها با استفاده از میانگین و انحراف معیار هر ماتریس ویژگی نرمال می‌شوند. سپس، فاصله‌ی اقلیدسی از ماتریس‌های ویژگی هر پنجره محاسبه می‌شود. برای تولید درهم‌ساز نهایی، فاصله‌ی اقلیدسی حاصل از تمامی پنجره‌ها به‌صورت $d=\{d_۱, d_۲, \dots, d_N\}$ در کنار هم قرار می‌گیرند و با استفاده از رابطه‌ی (۹-۳) به یک درهم‌ساز باینری نگاشت می‌شوند.

$$s_i = \begin{cases} ۰ & if\ d_i < T \\ ۱ & Otherwise \end{cases} \quad (۹-۳)$$

که در آن T مقدار متوسط مقدار مرتب شده از توالی تمامی فاصله‌ها است. در نتیجه، درهم‌ساز تصویر به‌صورت $HA = [s_۱, s_۲, \dots, s_N]$ به دست می‌آید. این الگوریتم برای نقاطی که از نظر بصری شبیه هستند، درهم‌سازهای مشابه به هم تولید می‌کند. به‌طوریکه فاصله‌ی همینگ بین درهم‌ساز های حاصل از دو تصویر مشابه، بسیار کم است.

۳-۷- نتیجه‌گیری

در این فصل، پیشینه و نحوه‌ی عملکرد الگوریتم‌های مورد استفاده در این تحقیق، بیان شد. ابتدا روش آشکارسازی چهره به‌نام MTCNN توضیح داده شد. سپس، روش‌های بررسی کیفیت روشنایی CLAH و تنظیم شدت روشنایی معرفی شدند. سپس نحوه‌ی عملکرد چندین الگوریتم درهم‌سازی تصویر به‌نام‌های LSH، P-stable LSH و درهم‌سازی تصویر مبتنی بر ضرایب DCT توضیح داده شد. در فصل بعدی، با استفاده از این الگوریتم‌ها، سیستم شناسایی چهره‌ی پیشنهادی در این پژوهش را معرفی می‌کنیم.

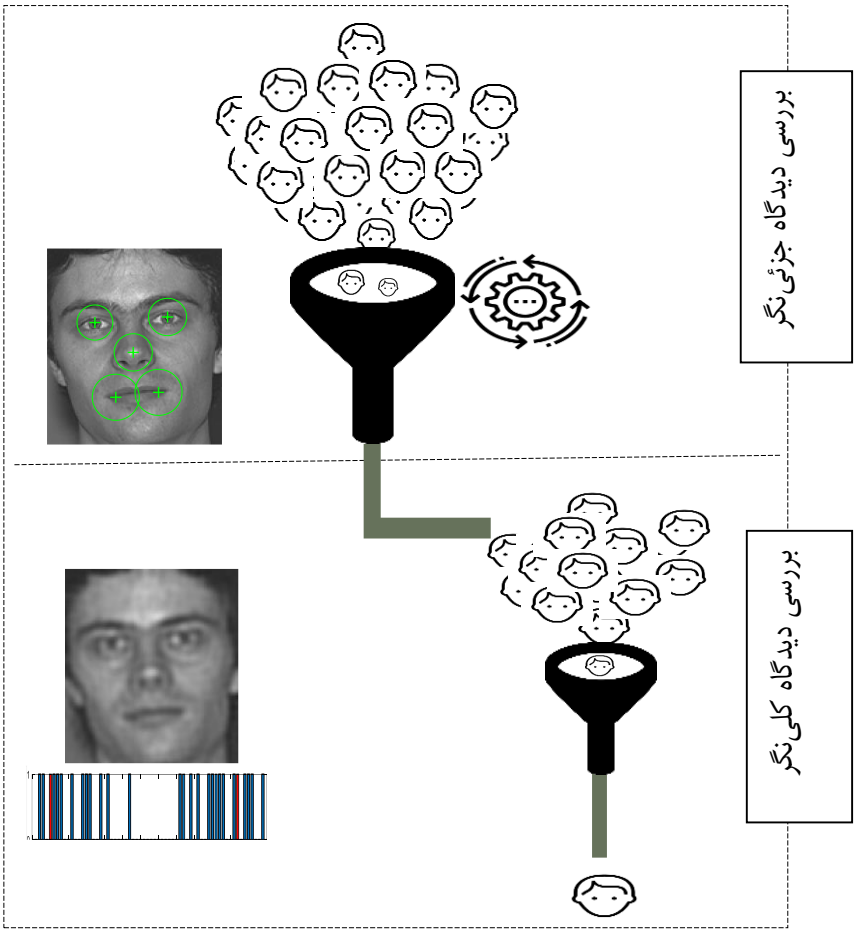
فصل ٤: روش پیشنهادی

۴-۱- مقدمه

تمرکز این پایان‌نامه، روی چالش‌های پایگاه داده‌ی حجیم، شرایط محیطی کنترل نشده، تعداد کم نمونه‌های آموزشی (وجود تنها یک تصویر به ازای هر نفر) و پیچیدگی‌های زمانی می‌باشد. هدف اصلی این تحقیق، ارائه‌ی یک سیستم شناسایی چهره با استفاده از روش‌های درهم‌سازی تصویر است، به‌طوریکه بتواند شخص مورد نظر را در یک پایگاه داده‌ی حجیم و به سرعت شناسایی کند. علاوه بر این، از عهده‌ی مشکلاتی مانند برخی شرایط محیطی کنترل نشده و تعداد کم نمونه‌های آموزشی برآید.

همانطور که در فصل اول بیان شد، سیستم بینایی انسان، پردازش‌های مختلفی را در طی چندین مرحله روی تصویر چهره اعمال می‌کند و سپس آن را با دقت بالایی شناسایی می‌کند. نکته‌ی حائز اهمیت این است که تمامی فرآیندهای پردازشی، تنها در کسری از ثانیه اتفاق می‌افتند. امروزه، بشر سعی در ارائه‌ی سیستم‌های شناسایی چهره‌ای دارد که مشابه سیستم بینایی انسان عمل کند. بنابراین، این پایان‌نامه سعی دارد روشی ارائه دهد تا با مفاهیم شهودی تطابق داشته باشد. همانطور که در شکل (۴-۱) مشاهده می‌شود، تصاویر چهره با استفاده از دیدگاه‌های بصری متفاوت، مورد بررسی قرار می‌گیرند. هر دیدگاه بصری، در یک مرحله‌ی متفاوت از الگوریتم پیشنهادی پردازش می‌شود. علاوه بر این، به منظور افزایش سرعت ارزیابی و ظرفیت تصاویر موجود در پایگاه‌داده، در هر مرحله، از توابع درهم‌سازی تصویر استفاده می‌شود.

سیستم شناسایی چهره در این پژوهش شامل گام‌های پیش‌پردازش، استخراج ویژگی و تطابق چهره می‌باشد. در اولین گام، به‌منظور افزایش دقت سیستم، پیش‌پردازش‌هایی مانند تشخیص ناحیه‌ی چهره، تنظیم روشنایی و تغییر اندازه روی تصاویر انجام می‌شوند. در گام استخراج ویژگی، تصاویر مجموعه‌ی داده با دو دیدگاه بصری جزئی‌نگر و کلی‌نگر بررسی می‌شوند. در این راستا، ویژگی‌های محلی و سراسری تصاویر به کمک روش‌های درهم‌سازی استخراج می‌شوند و در جدول‌های درهم‌ساز ذخیره می‌شوند.



شکل ۴-۱- مراحل کلی روش پیشنهادی

در نهایت، تصاویر آزمون در گام تطابق چهره شناسایی می‌شوند. گام پیش‌پردازش در بخش‌های (۴-۲) و (۴-۳) معرفی شده است. گام استخراج ویژگی به‌طور کامل در بخش (۴-۴) توضیح داده شده است و گام تطابق چهره در بخش (۴-۵) معرفی شده است.

۴-۲- تشخیص ناحیه‌ی چهره

همانطور که در مقدمه ذکر شد، یکی از چالش‌های مورد بررسی در این تحقیق، توانایی شناسایی چهره در شرایط کنترل نشده است. وجود زوایای مختلف، تغییرات نور و انسداد بخشی از تصویر چهره در محیط‌های کنترل نشده، شناسایی چهره را به امری چالش برانگیز تبدیل می‌کند. بنابراین، آشکارسازی چهره یکی از

بخش‌های اصلی سیستم محسوب می‌شود. روش‌های زیادی برای آشکارسازی چهره ارائه شده است، اما بسیاری از این روش‌ها توانایی آشکارسازی چهره در زوایای مختلف را ندارند.

مطالعات اخیر نشان می‌دهند که روش‌های یادگیری عمیق، می‌توانند عملکرد چشمگیری را در آشکارسازی چهره به دست آورند. بنابراین در این تحقیق، از روش MTCNN [۴۰] برای تشخیص ناحیه‌ی چهره‌ی مربوط به هر تصویر استفاده می‌کنیم. همانطور که در بخش (۳-۲) توضیح داده شد، این الگوریتم، نواحی چهره در زوایای مختلف را با دقت بهتری نسبت به روش معروف Viola-Jones [۱۲]، آشکار می‌کند.

۴-۳- پیش‌پردازش

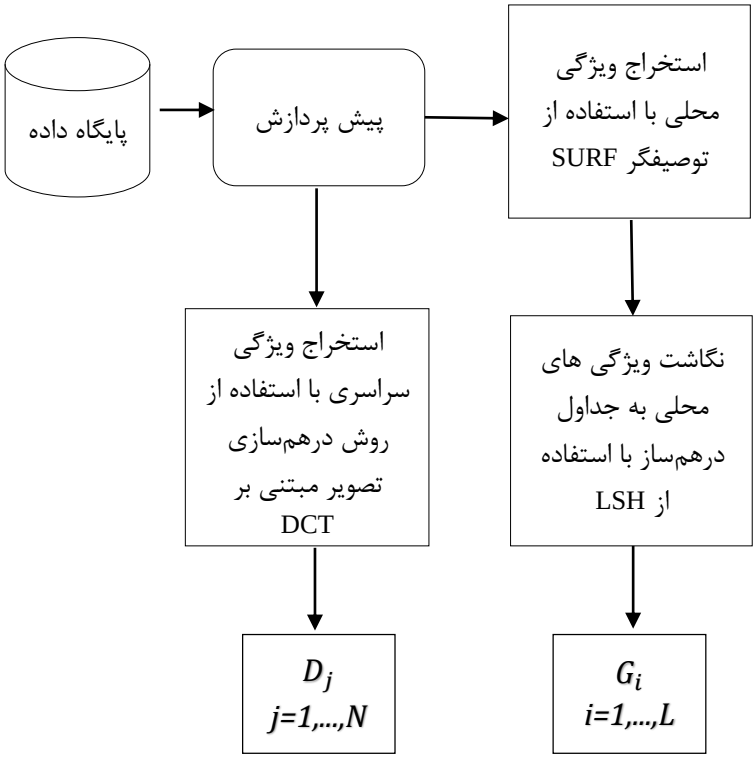
استفاده از اطلاعات دقیق‌تر و صحیح‌تر از تصاویر چهره، منجر به افزایش دقت سیستم شناسایی چهره می‌شود. در این تحقیق، از دو روش برابرسازی هیستوگرام وفقی با محدودیت کنتراست و تنظیم شدت روشنایی تصویر استفاده می‌کنیم و روشنایی و وضوح تصاویر را بهبود می‌بخشیم [۴۱،۴۲]. نحوه‌ی عملکرد این روش‌ها، به طور کامل در بخش (۳-۳) توضیح داده شده است.

۴-۴- گام استخراج ویژگی

همانطور که در مقدمه بیان شد، روش پیشنهادی در سه گام پیش‌پردازش، استخراج ویژگی و تطابق چهره معرفی می‌شود. در گام استخراج ویژگی، قصد داریم تصاویر موجود در پایگاه داده را با استفاده از دو دیدگاه بصری جزئی‌نگر و کلی‌نگر ارزیابی و بررسی کنیم. همچنین، به دلیل اهمیت ظرفیت تعداد تصاویر موجود در پایگاه داده و سرعت شناسایی تصاویر، از توابع درهم‌سازی تصویر استفاده می‌کنیم. همانطور که در فصل سوم بیان شد، درهم‌سازهای حاصل از تصاویر با حجم ذخیره‌سازی کمتر و ویژگی‌های قوی‌تر ذخیره می‌شوند. علاوه بر این، عملیات احراز هویت با توجه به درهم‌ساز هر تصویر، با سرعت، دقت و قدرت بالایی انجام می‌شود. در دیدگاه بصری جزئی‌نگر، ویژگی‌های محلی تصاویر مورد بررسی قرار می‌گیرند. در این

راستا، ابتدا ویژگی‌های محلی توسط توصیفگر SURF از تصاویر موجود در مجموعه داده استخراج می‌شوند. توضیحات مربوط به این روش در فصل سوم بیان شده است. در ادامه، از روش p-stable LSH استفاده می‌شود و ویژگی‌های محلی به مکان‌های جدول درهم‌ساز نگاشت داده می‌شوند [۴۶]. همانطور که بخش (۳-۴) توضیح داده شد، ویژگی‌های محلی مشابه به یک مکان از جدول درهم‌ساز نگاشت داده می‌شوند.

در ادامه، به‌منظور بررسی دیدگاه بصری کلی‌نگر، ویژگی سراسری توسط یک روش درهم‌سازی تصویر مبتنی بر ضرایب DCT، که در بخش (۳-۴) ارائه شد، ایجاد می‌شود. در این روش، هر تصویر به یک دنباله‌ی عدد باینری با اندازه‌ی کوتاه تبدیل و ذخیره می‌شود. بنابراین، در این گام، تصاویر با دو نوع دیدگاه بصری مختلف تحلیل می‌شوند و ویژگی‌های محلی و سراسری حاصل از توابع درهم‌ساز مختلف، ذخیره می‌شوند. از درهم-سازهای بدست آمده از گام استخراج ویژگی، می‌توان برای شناسایی چهره در گام بعدی استفاده کرد. شکل (۴-۲) فرآیند گام استخراج ویژگی از مدل پیشنهادی را نشان می‌دهد. ویژگی های محلی در جداول درهم‌ساز $G_i,\ i=1,...,L$ نگاشت و ذخیره می‌شوند. همچنین، ویژگی‌های سراسری در جدول درهم‌ساز $D_j,\ j=1,...,N$ ذخیره می‌شوند. L بیانگر تعداد جداول درهم‌ساز و N تعداد تصاویر مجموعه داده را نشان می‌دهد.



شکل ۴-۲- گام استخراج ویژگی در مدل پیشنهادی

۴-۵- گام تطابق چهره

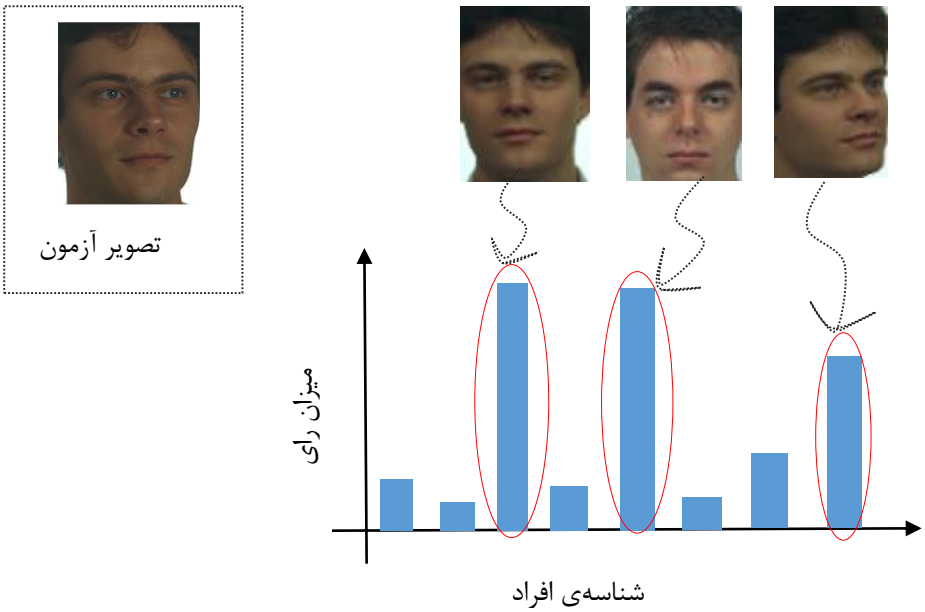
سومین گام از مدل پیشنهادی، گام تطابق چهره می‌باشد که در این مرحله تصاویر آزمون شناسایی می‌شوند. برای تعیین هویت تصویر چهره‌ی آزمون، در ابتدا با هدف بهبود روشنایی، پیش‌پردازش روی تصویر چهره انجام می‌شود. سپس با استفاده از دو مرحله، هویت تصویر آزمون شناسایی می‌شود. در ادامه، این دو مرحله به تفصیل توضیح داده می‌شوند.

۴-۵-۱- انتخاب شناسه‌های برتر

در اولین مرحله، شناسه‌هایی که از لحاظ دیدگاه بصری جزئی‌نگر به تصویر آزمون شبیه هستند انتخاب می‌شوند. منظور از شناسه، هویت فرد در مجموعه داده است. بدین منظور، تنها شناسه‌هایی که ویژگی‌های محلی‌شان، بیشترین شباهت را با ویژگی‌های محلی تصویر آزمون دارد، انتخاب می‌شوند و باقی شناسه‌ها

حذف می‌شوند. در این راستا، ابتدا بردارهای ویژگی محلی تصویر آزمون با استفاده از توصیفگر SURF استخراج می‌شوند. سپس با استفاده از روش p-stable LSH، ویژگی‌های محلی استخراج شده از تصویر آزمون را به مکان‌های جدول‌های درهم‌ساز $G_i, i=1,...,L$ نگاشت می‌دهیم. در بخش (۳-۵)، چگونگی یافتن ویژگی‌های محلی مشابه توسط الگوریتم LSH بیان شده است. به همین روش، به ازای هر ویژگی محلی موجود در چهره‌ی آزمون، تعدادی از نزدیک‌ترین ویژگی‌های محلی از میان ویژگی‌های محلی موجود در مجموعه داده، جستجو می‌شود.

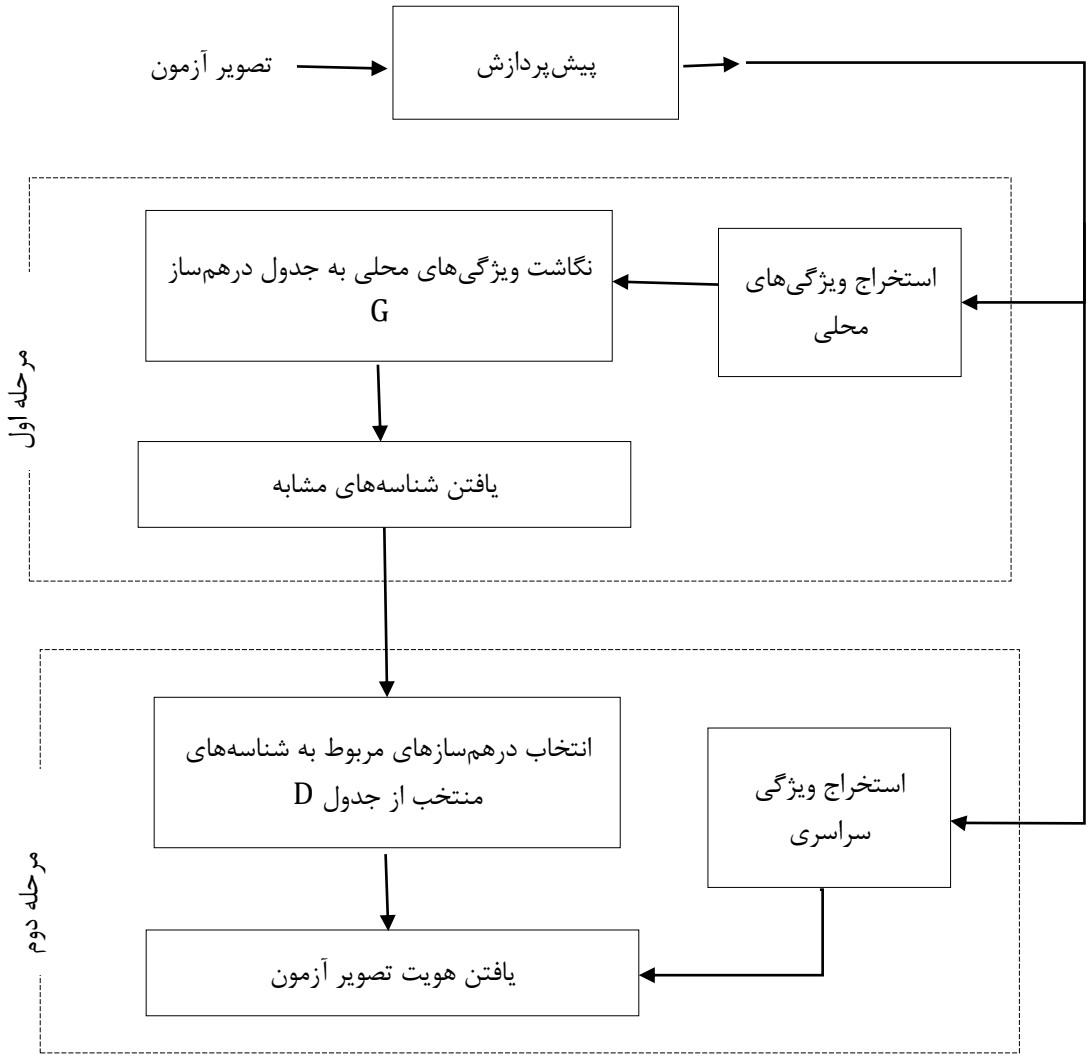
هر یک از ویژگی‌های محلی یافت شده، یک رای به شناسه‌ی چهره‌ای که به آن تعلق دارد، به حساب می‌آید و معکوس فاصله‌ی اقلیدسی تا ویژگی مورد نظر در تصویر آزمون، وزن آن رای را نشان می‌دهد. در ادامه، تمامی آرای مربوط به هر شناسه، با هم جمع می‌شوند. مانند شکل (۴-۳)، شناسه‌هایی که بالاترین میزان رای را کسب کردند، به‌عنوان شبیه‌ترین شناسه‌ها انتخاب می‌شوند. در حقیقت، شناسه‌هایی که تعداد رای کمی بدست آورده‌اند، حذف می‌شوند، زیرا به ندرت با تصویر آزمون مطابقت دارند. در مرحله‌ی دوم، شناسه‌های منتخب، از لحاظ دیدگاه بصری کلی‌نگر مورد بررسی قرار می‌گیرند. نکته‌ی مهم این است که اگر اختلاف بین تعداد آرای یک شناسه بسیار بیشتر از سایر شناسه‌ها باشد، نیازی به مرحله‌ی دوم نیست. در فصل پنجم، توضیحات بیشتری در ارتباط با حد آستانه‌ی در نظر گرفته شده در این پژوهش داده می‌شود.



شکل ۴-۳- شبیه‌ترین شناسه‌ها به تصویر آزمون

۴-۵-۲- شناسایی هویت تصویر آزمون

در مرحله‌ی دوم، ابتدا ویژگی سراسری تصویر آزمون را به کمک درهم‌ساز مبتنی بر ضرایب DCT استخراج می‌شود. سپس، مقایسه‌ای بین فاصله‌ی همینگ درهم‌ساز تصویر آزمون و درهم‌سازهای تصاویر شناسه‌های منتخب مرحله‌ی اول (که در داخل $D_j, j=1,...,N$ ذخیره شده اند) انجام می‌شود. در نهایت، شناسه‌ی نزدیک‌ترین تصویر به‌عنوان هویت تصویر آزمون انتخاب می‌شود. شکل (۴-۴)، گام تطابق چهره را نمایش می‌دهد.



شکل ۴-۴- روال کلی گام تطابق چهره

۴-۶- جمع‌بندی

در این فصل، الگوریتم پیشنهادی برای شناسایی چهره در سه گام پیش‌پردازش، استخراج ویژگی و تطابق چهره معرفی شد. در اولین گام، تصاویر چهره از نظر کیفیت روشنایی بررسی می‌شوند و نواحی چهره در تصاویر مشخص می‌شوند. در گام استخراج ویژگی، تصاویر با دو دیدگاه بصری جزئی‌نگر و کلی‌نگر، با استفاده از روش‌های درهم‌سازی تصویر مورد بررسی قرار می‌گیرند. در این راستا، ویژگی‌های محلی تصاویر استخراج

می‌شوند و در جداول درهم‌ساز نگاشت می‌شوند. ویژگی سراسری تصاویر نیز به کمک یک روش درهم‌سازی تصویر محاسبه و ذخیره می‌شود. در گام تطابق چهره، ابتدا ویژگی‌های محلی تصویر آزمون استخراج می‌شوند. سپس این ویژگی‌ها در جدول درهم‌ساز حاصل از گام استخراج ویژگی، نگاشت می‌شوند. با توجه به این جدول درهم‌ساز، شبیه‌ترین شناسه‌ها از لحاظ دیدگاه بصری جزئی‌نگر، به تصویر آزمون انتخاب می‌شوند. سپس از میان شناسه‌های منتخب، شبیه‌ترین شناسه به تصویر آزمون، از لحاظ دیدگاه بصری کلی‌نگر، به‌عنوان هویت تصویر آزمون انتخاب می‌شود. در فصل بعدی، آزمایشات و نتایج ارزیابی شرح داده می‌شود.

فصل ۵: نتایج و ارزیابی

۵-۱- مقدمه

همانطور که در فصل چهارم بیان شد، مدل پیشنهادی دارای سه گام پیش‌پردازش، استخراج ویژگی و تطابق چهره می‌باشد. در این فصل، ابتدا پایگاه‌های داده‌ی مورد استفاده در این پژوهش بررسی می‌شوند. سپس به بررسی آزمایشات و نتایج ارزیابی مدل پیشنهادی پرداخته می‌شود.

روال کلی الگوریتم ارائه شده در این پایان‌نامه به صورت زیر است:

۱. تشخیص نواحی چهره در تصاویر با استفاده از الگوریتم MTCNN.
۲. اعمال پیش‌پردازش روی تصاویر.
۳. اجرای الگوریتم SURF و ذخیره‌ی ویژگی‌های محلی هر تصویر به همراه شناسه‌ی مربوط به آن‌ها.
۴. اجرای الگوریتم LSH و نگاشت ویژگی‌های محلی به مکان‌های جدول درهم‌ساز.
۵. اجرای الگوریتم درهم‌سازی تصویر مبتنی بر ضرایب DCT و ایجاد ویژگی سراسری برای هر تصویر.
۶. اجرای مرحله‌ی ۱ تا ۵ برای تصاویر آموزش و آزمون.
۷. انتخاب شناسه‌هایی از مجموعه داده که ویژگی‌های محلی‌شان بیشترین شباهت را با ویژگی‌های محلی تصویر آزمون دارند.
۸. شناسایی هویت نهایی تصویر آزمون با مقایسه‌ی ویژگی سراسری شناسه‌های منتخب با ویژگی سراسری تصویر آزمون.

۵-۲- پایگاه داده

در این تحقیق از سه پایگاه داده‌ی معمول به نام‌های ORL [۱۸]، AR [۱۹]، FERET [۲۶] و یک مجموعه تصاویر جمع آوری شده، استفاده شده است. در مورد این مجموعه تصاویر توضیحات بیشتری داده می‌شود. در ادامه این پایگاه داده‌ها معرفی شده اند.

۵-۲-۱- پایگاه داده‌ی FERET

FERET یک بانک اطلاعاتی بزرگ از تصاویر چهره است که در سال‌های ۱۹۹۳ تا ۱۹۹۶، جمع‌آوری شده است [۲۶]. این مجموعه، حاوی ۱۴۰۵۱ تصویر چهره می‌باشد که در یک محیط نیمه کنترل شده ضبط شده اند. تصاویر، شامل ۱۳ موقعیت چهره در زوایای مختلف می‌باشند و تغییراتی مانند روشنایی، حالت چهره، داشتن و نداشتن عینک، ریش، زاویه‌ی چهره و سن را شامل می‌شوند. لازم به ذکر است، تعداد تصاویر چهره به ازای هر فرد موجود در پایگاه داده متفاوت می‌باشد. شکل (۵-۱) نمایش مختصری از تصاویر این پایگاه داده را نشان می‌دهد.

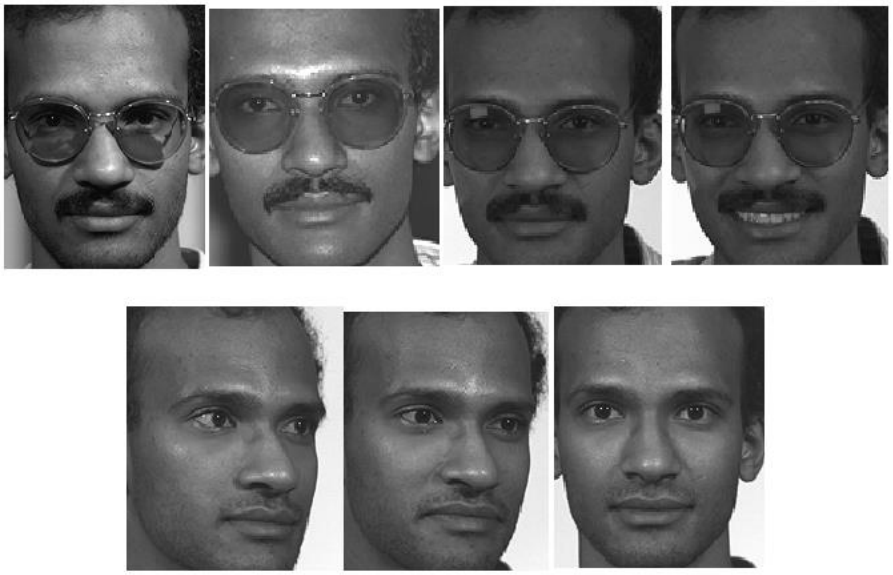
شکل ۵-۱- نمونه تصاویر پایگاه داده FERET [۲۶]

این پایگاه داده از زیرمجموعه تصاویر مختلف با چالش‌های متفاوت تشکیل شده است. در این تحقیق، دو زیرمجموعه تصاویر از پایگاه داده‌ی FERET در نظر گرفته شده است. اولین زیر مجموعه شامل ۱۹۸۰ تصویر چهره از ۹۹۰ فرد می‌باشد. تصاویر چهره در این زیرمجموعه، همگی به‌صورت رو به جلو هستند و فقط حالات چهره‌ی مربوط به هر فرد متفاوت است. دلیل در نظر گرفتن این زیر مجموعه، ارزیابی مدل

پیشنهادی هنگام وجود تنها یک نمونه تصویر به ازای هر فرد در پایگاه داده‌ی حجیم می‌باشد. نمونه‌ای از تصاویر این زیرمجموعه در شکل (۵-۲) قابل مشاهده می‌باشد.

شکل ۵-۲- نمونه تصاویر اولین زیرمجموعه‌ی پایگاه داده FERET [۲۶]

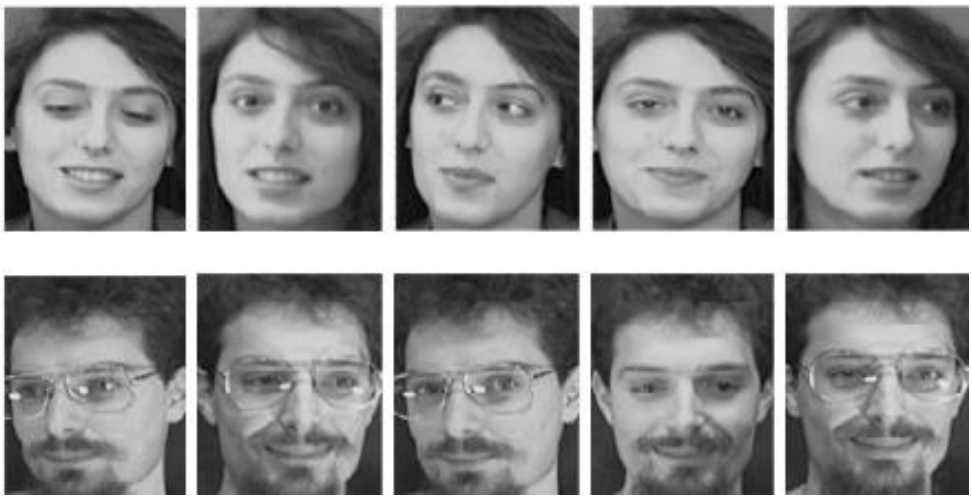
زیر مجموعه‌ی دیگری از این پایگاه داده با تعداد ۴۰۱۷ تصویر چهره از ۹۹۴ فرد مختلف در نظر گرفته شده است. دلیل در نظر گرفتن این زیر مجموعه، ارزیابی مدل پیشنهادی با تعداد بیشتری تصویر است. لازم به ذکر است، تعداد تصاویر چهره‌ی در نظر گرفته شده به ازای هر فرد متفاوت است. یعنی پراکندگی نمونه‌های مربوط به هر کلاس متفاوت است. همچنین، تصاویر در این زیر مجموعه دارای چالش‌های بیشتری مانند افزایش سن، شرایط روشنایی متفاوت، تغییر حالت مو و ریش و از همه مهمتر، تغییر زاویه‌ی چهره تا مقدار ±۲۵ درجه می‌باشد. نمونه‌ای از تصاویر این زیر مجموعه در شکل (۵-۳) مشاهده می‌شود.



شکل ۵-۳- نمونه تصاویر دومین زیرمجموعه‌ی پایگاه داده FERET [۲۶]

۵-۲-۲- پایگاه داده‌ی ORL

این پایگاه داده شامل ۴۰۰ تصویر از ۴۰ نفر در ۱۰ حالت مختلف چهره می‌باشد [۱۸]. تصاویر در این پایگاه داده شامل تغییراتی مانند خندیدن یا نخندیدن، عینک زدن یا نزدن، چشمان بسته یا باز و تا حدودی زاویه‌ی چهره هستند. همچنین این تصاویر در شرایط نوری متفاوت ضبط شده‌اند. نمونه‌ای از تصاویر این پایگاه داده در شکل (۵-۴) مشاهده می‌شود.



شکل ۵-۴- نمونه تصاویر پایگاه داده ORL [۱۸]

۵-۲-۳- پایگاه داده‌ی AR

این پایگاه داده شامل ۴۰۰۰ تصویر چهره از ۱۲۶ فرد مختلف می‌باشد [۱۹]. تصاویر در روزهای مختلف با شرایط نوری مختلف، تغییرات چهره‌ی متفاوت مانند فریاد زدن، خندیدن، عصبانیت و وجود انسداد ضبط شده اند. دلیل انتخاب این پایگاه داده، بررسی مدل پیشنهادی هنگام وجود انسداد در تصاویر می‌باشد. نمونه‌ای از تصاویر این پایگاه داده در شکل (۵-۵) قابل مشاهده است.

شکل ۵-۵- نمونه تصاویر پایگاه داده AR [۱۹]

۵-۲-۴- مجموعه تصاویر جمع آوری شده

عباس‌پور و همکاران [۱۱] برای بررسی ظرفیت سیستم شناسایی چهره با تعداد افراد بیشتر، تصاویر چندین پایگاه داده را جمع‌آوری کرده‌اند. تصاویر از پایگاه داده‌های FERET [۲۶]، MUCT [۴۹]، PICS [۵۰]، FEI [۵۱] و Face94 [۵۲] انتخاب شده‌اند. این تصاویر، اغلب به‌صورت روبرو هستند و حالات چهره و روشنایی در تصاویر متفاوت است. تعداد کل اشخاص در این مجموعه برابر با ۱۶۸۴ نفر می‌باشد و از هر فرد تنها دو نمونه تصویر وجود دارد، یعنی تعداد کل تصاویر برابر با ۳۳۶۸ می‌باشد. نمونه‌ای از تصاویر این پایگاه داده در شکل ۵-۶ قابل مشاهده است.

شکل ۵-۶- نمونه تصاویر پایگاه داده‌ی جمع‌آوری‌شده [۴۹-۵۲]

۵-۳- پارامترهای مدل پیشنهادی

برخی از روش‌های مورد استفاده در این تحقیق نیاز به تنظیم پارامتر دارند. جدول (۵-۱) مقادیر پارامترهای در نظر گرفته شده‌ی این روش‌ها را نشان می‌دهد. LSH در اولین ردیف، شامل پارامترهای l، r، K و n می‌باشد. پارامتر L، تعداد جداول درهم‌ساز را نشان می‌دهد. پارامتر K بیانگر طول درهم‌ساز هر خانه است و پارامتر r، عرض هر خانه‌ی درهم‌ساز را نشان می‌دهد. در فصل سوم به توضیح این پارامترها پرداخته شده است. همچنین، پارامتر n، تعداد ویژگی‌های مشابه به‌ازای هر ویژگی محلی ورودی را نشان می‌دهد.. در روش درهم‌سازی تصویر مبتنی بر DCT، فیلتر گوسین ۳×۳ در نظر گرفته شده است. اندازه‌ی تصویر در این روش، ۶۴×۶۴ و اندازه‌ی پنجره ۸×۸ در نظر گرفته شده است. همانطور که در فصل چهارم بیان کردیم، در مرحله‌ی اول از گام تطابق چهره، شناسه‌های غیر مشابه به تصویر آزمون از نظر دیدگاه بصری جزئی‌نگر، حذف می‌شوند. در این راستا، شناسه‌هایی که مقدار رای آن‌ها از B درصد بالاترین مقدار رای، بیشتر باشد، برای مرحله‌ی نهایی از تطابق چهره انتخاب می‌شوند.

جدول ۵-۱- مقادیر پارامترهای استفاده شده در مدل پیشنهادی

الگوریتم مورد استفاده	پارامتر	مقدار
الگوریتم p-Stable LSH	K،L	۵
	r	۲
	n	۱۰
درهم‌سازی مبتنی بر ضرایب DCT	گوسین	۳×۳
	M	۶۴
	m	۸
مرحله اول از گام تطابق چهره	B	٪۶۵

۵-۴- نتایج روش پیشنهادی

در این تحقیق، برای اجرای برنامه از زبان برنامه‌نویسی متلب[۱]استفاده شده است. لازم به ذکر است که مشخصات سیستم CPU: Intel Core i5-4300M, 2.60GHz و Memory:4 GB می‌باشد. در ادامه، نتایج الگوریتم پیشنهادی در آزمایشات مربوط به پایگاه‌های داده‌ی مختلف بیان شده است. همانطور که در فصل اول بیان شد، در این پژوهش قصد داریم با استفاده از مزایای درهم‌سازی تصویر، سیستمی برای بازشناسی چهره ارائه کنیم که علاوه بر سرعت بالا، عملکرد مناسبی روی پایگاه‌های داده با تصاویر زیاد دارد و در صورت وجود برخی از شرایط کنترل نشده، دقت خود را تا حد زیادی حفظ می‌کند. علاوه بر این، در حضور تنها یک تصویر به ازای هر فرد، عملیات شناسایی را با دقت بالایی انجام دهد. تصاویر در پایگاه داده‌ی AR شامل چالش‌های زیادی از جمله انسداد، تغییرات شدید روشنایی و حالت چهره هستند. همچنین تصاویر پایگاه داده‌ی FERET شامل چالش‌های بیشتری مانند تغییرات زاویه‌ی چهره تا ۲۵± درجه، سن، حالت چهره و انسداد هستند. برای ارزیابی دقت روش پیشنهادی در شرایط کنترل نشده از پایگاه‌داده‌های فوق استفاده شده است. تاثیر حجم پایگاه داده در عملکرد الگوریتم پیشنهادی با استفاده از پایگاه داده‌های FERET با تعداد بیش از ۴۰۰۰ تصویر و پایگاه‌داده‌ی جمع‌آوری شده با تعداد بیش از ۳۰۰۰ تصویر مورد ارزیابی قرار گرفته است. علاوه بر این، برای ارزیابی روش پیشنهادی در حضور تنها یک نمونه تصویر به‌ازای هر فرد، آزمایشاتی روی پایگاه داده‌های فوق در شرایط محدود شده انجام شده است.

۵-۴-۱- دقت مدل پیشنهادی

جدول(۵-۲)، نتایج دقت مدل پیشنهادی را با استفاده از پایگاه‌های داده‌ی ORL، AR، FERET و مجموعه تصاویر جمع‌آوری‌شده نشان می‌دهد. معیار ارزیابی در این پژوهش از طریق رابطه‌ی (۵-۱) بدست آمده

[۱] MATLAB

است. I بیانگر تعداد کل تشخیص‌های درست در سیستم شناسایی چهره و U نشان‌دهنده‌ی تعداد کل تصاویر آزمون می‌باشد.

$$\boldsymbol{Accuracy = \frac{I}{U}} \qquad (۱-۵)$$

با در نظر گرفتن آزمایشات مختلف می‌توان مدل شناسایی چهره را با تعداد مختلفی از تصاویر آموزش داد که در این حالت، نتیجه‌ی ارزیابی قابل اعتمادتر است. افراد به‌صورت تصادفی از بین تمامی افراد مجموعه داده انتخاب می‌شوند. سپس ۲۰ بار مراحل روش پیشنهادی تکرار می‌شوند و در نهایت، میانگین دقت‌های بدست آمده ثبت می‌شود.

در اولین آزمایش پایگاه‌داده‌ی ORL، به ازای هر فرد، ۶۰ درصد از تصاویر برای گام استخراج ویژگی مورد استفاده قرار گرفتند و باقی به‌عنوان تصاویر آزمون تعریف شدند. در آزمایش دوم مربوط به همین پایگاه داده، ۵۰ درصد از تصاویر در گام استخراج ویژگی مورد استفاده قرار گرفتند و باقی به‌عنوان تصاویر آزمون در نظر گرفته شدند. لازم به ذکر است که این تصاویر به‌صورت تصادفی انتخاب شده‌اند. همانطور که از نتایج مشخص است، مدل پیشنهادی، دقت شناسایی ۱۰۰ درصد را در هر دو آزمایش بدست آورده است.

پایگاه داده‌ی AR به‌منظور ارزیابی دقیق‌تر مدل پیشنهادی در شرایط نیمه‌کنترل شده‌ای که شامل انسداد نیز می‌باشد، استفاده شده است. این پایگاه داده نیز با دو آزمایش مختلف ارزیابی شده است. در آزمایش اول، ۶۰ درصد از تصاویر هر فرد موجود در پایگاه داده به صورت تصادفی برای گام استخراج ویژگی در نظرگرفته شده است. باقی تصاویر نیز به‌عنوان تصاویر آزمون در نظر گرفته شده‌اند. در آزمایش دوم، ۵۰ درصد از تصاویر به صورت تصادفی برای گام استخراج ویژگی و باقی برای تصاویر آزمون در نظر گرفته شده‌اند. نتایج نشان می‌دهد که الگوریتم پیشنهادی در مقابل چالش وجود انسداد در تصاویر نیز به خوبی عمل می‌کند.

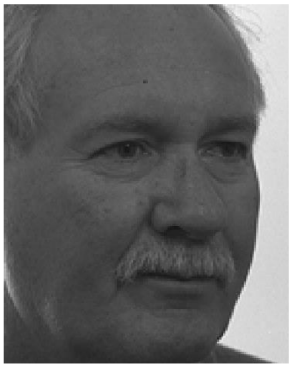
پایگاه داده‌ی FERET یکی از معروف‌ترین پایگاه‌های داده برای ارزیابی مدل، در مقابل چالش حجم بالای پایگاه داده است. در آزمایش اول این پایگاه داده، تعداد ۱۹۸۰ تصویر از ۹۹۰ فرد در نظر گرفته شده است. یعنی به ازای هر فرد، تنها دو نمونه تصویر موجود می‌باشد. در این مورد، به ازای هر فرد، یک تصویر برای

گام استخراج ویژگی و یک تصویر به‌عنوان آزمون در نظر گرفته شده است. در آزمایش دوم این پایگاه داده، تعداد ۴۰۱۷ تصویر از ۹۹۴ فرد در نظر گرفته شده است. تعداد تصاویر به ازای هر فرد متفاوت است. همانطور که در بخش معرفی پایگاه داده بیان شد، دلیل در نظر گرفتن این زیر مجموعه، ارزیابی مدل پیشنهادی با تعداد تصاویر و چالش‌های بیشتر است. در آزمایش دوم پایگاه داده‌ی FERET، ۷۰ درصد از تصاویر هر فرد برای گام استخراج ویژگی استفاده شده است. در مجموعه تصاویر جمع‌آوری‌شده نیز تعداد ۱۶۸۴ فرد وجود دارد و از هر فرد تنها یک نمونه تصویر برای گام استخراج ویژگی استفاده شده است.

جدول ۵-۲- نتایج دقت مدل پیشنهادی (درصد)

پایگاه داده	آزمایش اول	آزمایش دوم
ORL	۱۰۰	۱۰۰
AR	۹۹	۹۷
FERET	۹۶٫۴۷	۹۲٫۸۶
تصاویر جمع‌آوری‌شده	۹۵٫۹۰	-

در ادامه، مواردی از تشخیص‌های نادرست توسط مدل پیشنهادی را در شکل (۵-۷)، نشان دادیم. همانطور که مشاهده می‌شود، حتی انسان‌ها ممکن است چهره‌های موجود در هر ردیف را شبیه به هم بدانند.




الف) تصویر ورودی از مجموعه آزمون ب) تشخیص اشتباه


الف) تصویر ورودی از مجموعه آزمون ب) تشخیص اشتباه


الف) تصویر ورودی از مجموعه آزمون ب) تشخیص اشتباه

شکل ۵-۷: شناسایی نادرست تصویر چهره در روش ارائه شده

۵-۵- بررسی نتایج

در بخش قبلی، نتایج حاصل از عملکرد مدل پیشنهادی با استفاده از سه نوع پایگاه داده‌ی مختلف بررسی شد. در بخش (۵-۵-۱)، این نتایج با برخی از روش‌های شناسایی چهره، مقایسه می‌شوند. همچنین، همانطور که بیان شد، حجم تصاویر پایگاه داده نقش مهمی را در عملکرد مدل پیشنهادی ایفا می‌کند. به‌همین دلیل، در بخش (۵-۵-۲)، تاثیر حجم پایگاه داده در عملکرد الگوریتم پیشنهادی بررسی می‌شود. همچنین، تاثیر طول درهم‌سازهای حاصل از تصاویر، بر عملکرد سیستم شناسایی چهره در بخش (۵-۵-۳) تحلیل می‌شود. در انتها، پیچیدگی زمانی این الگوریتم در بخش (۵-۵-۴) مورد بررسی قرار می‌گیرد.

۵-۵-۱- مقایسه با روش‌های دیگر شناسایی چهره

در این بخش، دقت مدل پیشنهادی با برخی از روش‌های شناسایی چهره مقایسه می‌شود. برای مقایسات، از هر سه پایگاه داده‌ی FERET، ORL و AR استفاده شده است. جدول (۵-۳)، نتایج مقایسات بین دقت مدل پیشنهادی و سایر روش‌ها را با استفاده از پایگاه داده‌ی ORL نشان می‌دهد. همانطور که در بخش (۵-۴-۱) بیان شد، در آزمایش اول، ۶۰ درصد از مجموعه‌ی داده برای تصاویر آزمون استفاده شده است. در آزمایش دوم، ۵۰ درصد از مجموعه داده به عنوان تصاویر آزمون در نظر گرفته شده است. همانطور که نتایج نشان می‌دهد، دقت مدل پیشنهادی در هر دو آزمایش برابر با ۱۰۰ درصد می‌باشد و عملکرد بهتری را نسبت به باقی روش‌ها دارد. در ادامه، مقایسه‌ای بین دقت مدل پیشنهادی با دیگر روش‌ها با استفاده از پایگاه داده‌ی AR در جدول (۵-۴) انجام شده است. مدل پیشنهادی با استفاده از این پایگاه داده، عملکرد بهتری نسبت به باقی روش‌ها دارد. در انتها، نتایج مقایسات بین مدل پیشنهادی و سایر روش‌ها با استفاده از پایگاه داده‌ی FERET در جدول (۵-۵) بررسی شده است. مقایسات با استفاده از ۵ حالت مختلف مجموعه انجام شده است. در هر حالت، تعداد تصاویر مجموعه داده و تعداد افراد متفاوت می‌باشد.

مرجع [۲۰] صرفا از روش درهم‌ساز LSH برای شناسایی چهره استفاده کرده است. این روش عملکرد خوبی را در پایگاه داده‌ی ORL دارد. اما به‌دلیل اینکه فقط به ویژگی‌های محلی توجه می‌کند و از ویژگی‌های سراسری تصاویر غافل است، در صورت وجود برخی از شرایط کنترل نشده مانند وجود انسداد، عملکرد خوبی ندارد. همچنین، در مرجع [۳۰] علاوه بر در نظر گرفتن روش درهم‌ساز LSH، از ترکیب چندین توصیفگر ویژگی استفاده شده است. ولی به‌دلیل افزایش محاسبات، این روش به فضای ذخیره‌سازی زیادی نیاز دارد و دقت آن از روش پیشنهادی در این پژوهش کمتر است. مرجع [۲۱] از دقت شناسایی بالایی برخوردار است. اگرچه که اختلاف کمی در دقت مشاهده می‌شود. اما روش پیشنهادی در مقایسه با مرجع [۲۱]از سرعت شناسایی بالاتری برخوردار است. همچنین، دقت مدل پیشنهادی با برخی از روش‌های استخراج ویژگی مبتنی بر الگوهای باینری محلی مانند مراجع [۵۳،۶۲]، مقایسه شده است. نتایج نشان می‌دهد که دقت مدل پیشنهادی از این روش‌ها بیشتر می‌باشد. از طرفی دیگر، یکی از روش‌های شناسایی چهره‌ی مطرح که عملکرد خوبی دارد، روش مبتنی بر شبکه‌های عصبی می‌باشد. در این تحقیق، دقت مدل پیشنهادی را با چند روش جدید مبتنی بر شبکه‌های عصبی موجود در مراجع [۵۵،۵۶]، مقایسه کرده‌ایم. علاوه بر اینکه دقت این روش‌ها از روش پیشنهادی کمتر است، این روش‌ها نیاز به تعداد نمونه‌ی بیشتری در مرحله‌ی آموزش خود دارند. روش‌های یادگیری عمیق مانند مرجع [۳۱] نیز به تعداد نمونه‌های آموزشی بیشتری نیاز دارند و قادر به حل مشکل یک نمونه به‌ازای هر فرد نیستند [۵۷]. البته همانطور که در فصل دوم بیان شد، روش‌های یادگیری عمیق برای حل مشکل کمبود داده‌های آموزشی ارائه شدند. اما این روش‌ها، عملکرد خوبی روی پایگاه‌داده‌های حجیم ندارند. به‌عنوان مثال، مرجع [۳۲]، یک مدل با استفاده از ترکیب روش‌های سنتی و یادگیری عمیق ارائه کرده است. اما با وجود اینکه تعداد تصاویر مجموعه داده کم است، دقت شناسایی حاصل در این روش کمتر از روش پیشنهادی است. مرجع [۶۶]، با استفاده از یک شبکه‌ی عمیق، نمونه‌های تصویر جدید ایجاد می‌کند و به شناسایی چهره می‌پردازد. روش‌های دیگر نیز برای حل چالش نمونه‌ی کم تصاویر آموزشی، ارائه شده‌اند. اما این روش‌ها نیز فقط در پایگاه داده‌هایی با حجم کم عملکرد خوبی دارند. مراجع [۱۰،۱۱] با هدف رفع چالش نمونه‌های کم آموزشی در پایگاه داده‌ی

حجیم، روش‌هایی مبتنی بر NMF ارائه داده است. همانطور که در جدول (۵-۵) نشان داده شده است، دقت روش [۱۰] با افزایش حجم پایگاه داده به تعداد ۱۶۸۴ فرد، به‌طور چشمگیری کاهش می‌یابد. همچنین روش پیشنهادی در مقایسه با مرجع [۱۱] از سرعت شناسایی بالاتری برخوردار است. توضیحات بیشتر در ارتباط با زمان شناسایی در بخش (۵-۵-۴) داده می‌شود. علاوه بر این، با افزایش حجم داده به تعداد ۱۶۸۴ فرد، دقت مرجع [۱۱] کاهش بیشتری را در مقایسه با روش پیشنهادی دارد. در سال‌های اخیر، برخی روش‌ها با استفاده از الگوریتم‌های استخراج ویژگی مطرح، روش‌هایی را برای حل چالش شرایط کنترل نشده در سیستم شناسایی چهره ارائه دادند [۵۸،۵۹،۶۰] . برخی از این روش‌ها، نتایج خوبی روی پایگاه داده‌های کوچک دارند ولی در پایگاه داده‌های حجیم با افت مواجه می‌شوند. در نهایت، با انجام مقایسات بین برخی از روش‌های پیشین و روش پیشنهادی، مشاهده می‌شود که روش پیشنهادی در چالش‌های مطرح شده عملکرد بهتری بدست آورده است. در این پژوهش، چالش‌های پایگاه داده‌ی حجیم و وجود تنها یک نمونه تصویر به‌ازای هر فرد حل شده‌اند. این روش عملکرد خوبی در شرایط تقریبا کنترل نشده از جمله تغییرات زاویه‌ی چهره تا ±۲۵ درجه، سن و انسداد دارد. سرعت شناسایی در روش پیشنهادی نیز افزایش چشمگیری داشته است.

جدول ۵-۳- مقایسه دقت مدل پیشنهادی با روش‌های دیگر در پایگاه داده ORL

روش	آزمایش اول (٪)	آزمایش دوم (٪)
SVM [۶۱]	۷۸٫۴۷	۷۵٫۷۷
LBP(4×4) [۶۲]	۸۸	۸۳
LGP+BAW-GSA [۵۳]	۹۲٫۵	۹۰٫۷
IKLDA + PNN [۵۵]	۹۳٫۹۵	۹۱٫۴۴
PLSH [۲۰]	۹۴٫۱۶	۹۲٫۳۳
ESGK [۵۸]	۹۷٫۵	۹۶
روش پیشنهادی	**۱۰۰**	**۱۰۰**

جدول ۵-۴- مقایسه دقت مدل پیشنهادی با روش‌های دیگر در پایگاه داده AR

روش	آزمایش اول (٪)	آزمایش دوم (٪)
PLSH [۲۰]	۸۰	۷۶
SVM [۶۱]	۸۷٫۳۱	۸۵٫۶۸
SRC [۶۳]	۹۶	۹۵
IKLDA + PNN [۵۵]	۹۷٫۴۶	۹۵٫۸۶
روش پیشنهادی	**۹۹**	**۹۷**

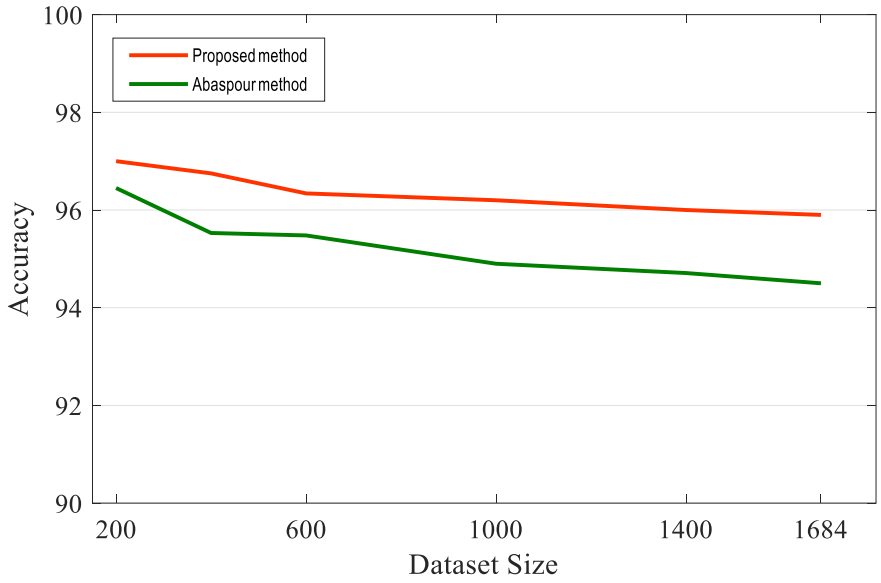
جدول ۵-۵- مقایسه دقت مدل پیشنهادی با روش‌های دیگر در پایگاه داده FERET

روش	تعداد تصاویر در پایگاه داده	تعداد افراد در پایگاه داده	دقت (٪)
FLDA-SVD [۲۵]	۴۰۰	۲۰۰	۹۰٫۵
DMMA [۵۶]	۴۰۰	۲۰۰	۹۳
TDL [۳۲]	۴۰۰	۲۰۰	۹۳٫۹
KCFT [۶۶]	۴۰۰	۲۰۰	۹۳٫۱۶
روش پیشنهادی	۴۰۰	۲۰۰	**۹۷٫۳۸**
FFT [۲۷]	۱۹۸۰	۹۹۰	۹۰٫۴
NMF [۱۰]	۱۹۸۰	۹۹۰	۹۲٫۷۳
ANMF [۱۱]	۱۹۸۰	۹۹۰	۹۸٫۳۶
LFH [۲۱]	۱۹۸۰	۹۹۰	۹۶٫۶
روش پیشنهادی	۱۹۸۰	۹۹۰	**۹۶٫۴۷**
NMF [۱۰]	۳۳۶۸	۱۶۸۴	۸۹٫۸۹
ANMF [۱۱]	۳۳۶۸	۱۶۸۴	۹۴٫۵۴
روش پیشنهادی	۳۳۶۸	۱۶۸۴	**۹۵٫۹۰**

ESRC [٦٤]	۱۴۰۰	۲۰۰	۵۱٫۶۰
LBP [٦٢]	۱۴۰۰	۲۰۰	۷۴٫۹۰
RRNN [٥٤]	۱۴۰۰	۲۰۰	۷۷٫۹۰
PCA [٦٥]	۱۴۰۰	۲۰۰	۸۳٫۱۶
ESGK [٥٨]	۱۴۰۰	۲۰۰	۸۴٫۵
PCA&LDA [٥٩]	۱۴۰۰	۲۰۰	۸۶٫۱۰
SESRC&LDF [٦٠]	۱۴۰۰	۲۰۰	۹۳٫۷۵
روش پیشنهادی	۱۴۰۰	۲۰۰	**۹۴٫۹۴**
روش پیشنهادی	۴۰۱۷	۹۹۴	**۹۲٫۸۶**

۵-۵-۲- تاثیر حجم پایگاه داده بر دقت الگوریتم پیشنهادی

همانطور که در فصل دوم بیان شد، تاکنون روش‌های خوبی برای شناسایی چهره ارائه شدند. اما بسیاری از این روش‌ها، عملکرد خوبی را در پایگاه داده‌های حجیم نشان نمی‌دهند. یعنی زمانی که حجم پایگاه داده افزایش یابد، دقت این روش‌ها با افت قابل توجهی مواجه می‌شود. با افزایش تعداد افراد در مجموعه داده، قدرت تمایز بین شناسه‌ها کاهش می‌یابد. در این بخش، روند تغییر دقت مدل پیشنهادی با افزایش تعداد داده بررسی می‌شود. شکل (۵-۸) روند تغییر عملکرد روش پیشنهادی و روش عباسپور [۱۱] را با افزایش حجم داده در مجموعه تصاویر جمع‌آوری شده نشان می‌دهد. در هر مرحله، افراد به‌صورت تصادفی اضافه می‌شوند و درنهایت کل پایگاه داده با تعداد ۱۶۸۴ فرد مورد ارزیابی قرار می‌گیرد. همانطور که مشاهده می‌شود، دقت روش پیشنهادی، افت کمتری در مقابل افزایش حجم داده دارد. این امر، نشان‌دهنده‌ی کارایی مناسب روش پیشنهادی در حجم داده‌ی بالا می‌باشد. روش نیکان [۱۰] نیز دقت شناسایی ۸۹/۸۹ درصد را با ۱۶۸۴ فرد بدست آورد.



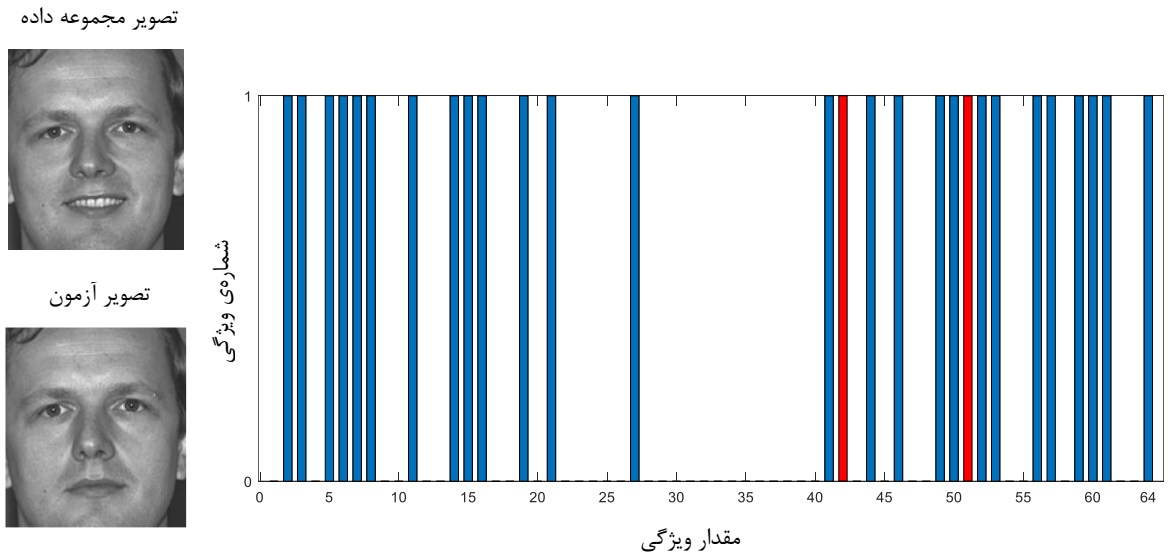
شکل ۵-۸- مقایسه دقت شناسایی روش پیشنهادی با روش عباس‌پور [۱۱]

۵-۵-۳- تاثیر طول درهم‌ساز بر دقت الگوریتم پیشنهادی

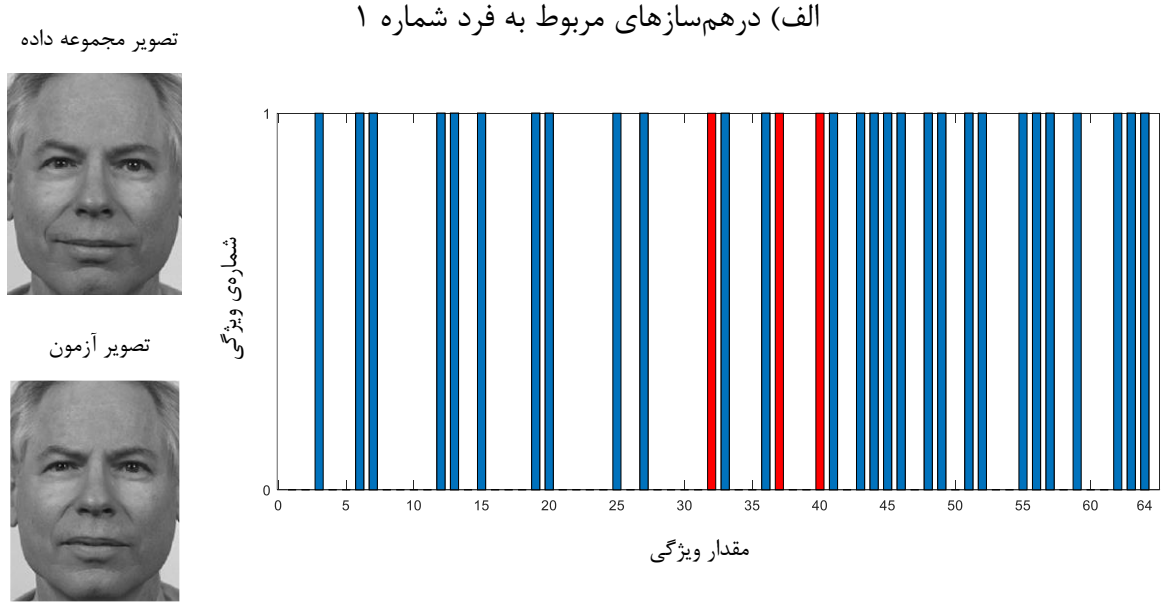
همانطور که در فصل چهارم بیان شد، یکی از روش‌های درهم‌سازی تصویر در روش پیشنهادی، روش درهم‌سازی تصویر مبتنی بر ضرایب DCT است. همانطور که در فصل سوم بیان شد، در این روش، تصویر به پنجره‌هایی تقسیم می‌شود و از هر پنجره، ویژگی استخراج می‌شود. در نهایت ویژگی‌ها بهم می‌پیوندند و به اندازه‌ی کوتاه‌تری نگاشت داده می‌شوند. در این پژوهش، مقدار M=64 و m=8 به‌عنوان بهترین جواب بدست آمده است. بنابراین با توجه به توضیحات فصل سوم، به ازای هر تصویر، یک توالی از اعداد باینری با طول ۶۴ بدست می‌آید. شکل (۵-۹) درهم‌سازهای حاصل از چند نمونه تصویر پایگاه داده‌ی FERET را نشان می‌دهد. در این شکل، مشاهده می‌شود که درهم‌سازهای بدست آمده از تصاویر (الف) و (ب) به هم نزدیک هستند. در واقع، فاصله‌ی همینگ[۱] بین آن‌ها بسیار کم می‌باشد. مقدار درهم‌ساز دو تصویر در ستون‌های قرمز با هم متفاوت است. نکته‌ی قابل اهمیت، اندازه پنجره تصاویر است. باید تعادلی بین طول درهم‌ساز و قدرت تمایز آن به وجود آورد. زیرا اندازه‌ی پنجره بزرگ به معنی ویژگی‌های کمتر است که به ناچار قابلیت تمایز را کاهش می‌دهد. هنگامی که اندازه‌ی پنجره کاهش می‌یابد، تمایز می‌تواند بهبود یابد،

[۱] Hamming

اما دقت به آسانی تحت تاثیر اصلاحات جزئی قرار می‌گیرد. علاوه بر این، اندازه‌ی پنجره کوچکتر، طول درهم‌ساز را افزایش می‌دهد. در این بخش، تاثیر اندازه‌ی درهم‌ساز بر عملکرد مدل پیشنهادی بررسی می‌شود. شکل (۵-۱۰)، تاثیر اندازه‌ی درهم‌ساز را بر دقت مدل پیشنهادی با استفاده از پایگاه داده‌ی FERET نشان می‌دهد. همانطور که مشاهده می‌شود، زمانی که طول درهم‌ساز برابر با ۶۴ می‌باشد، دقت شناسایی برابر با ۹۶/۴۷ است. پس از آن با شیب ملایمی افزایش و سپس کاهش می‌یابد. بنابراین، درهم‌ساز با اندازه‌ی ۶۴ از لحاظ طول و دقت، بهتر از باقی اندازه‌ها است.

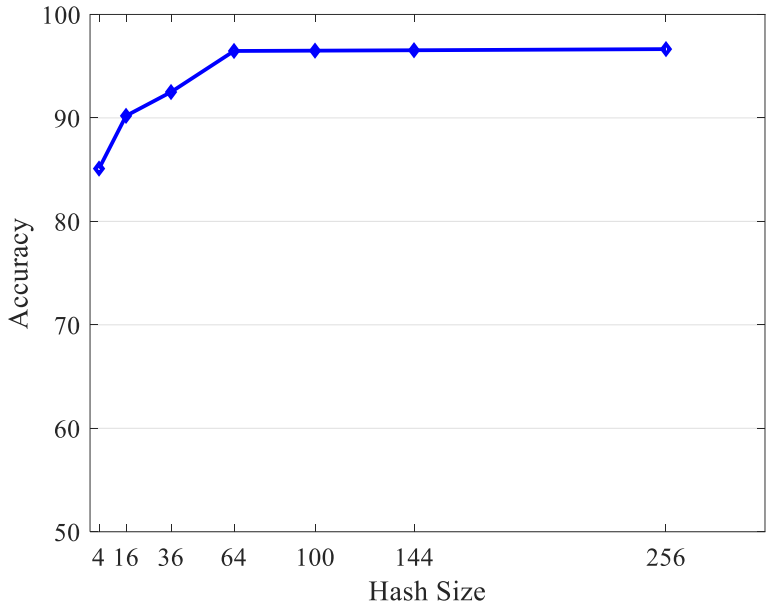


الف) درهم‌سازهای مربوط به فرد شماره ۱



ب) درهم‌سازهای مربوط به فرد شماره ۲

شکل ۵-۹- نمایش درهم‌سازهای ایجاد شده از چند نمونه تصویر با استفاده از ضرایب DCT



شکل ۵-۱۰- تاثیر اندازه‌ی درهم‌ساز بر عملکرد مدل پیشنهادی

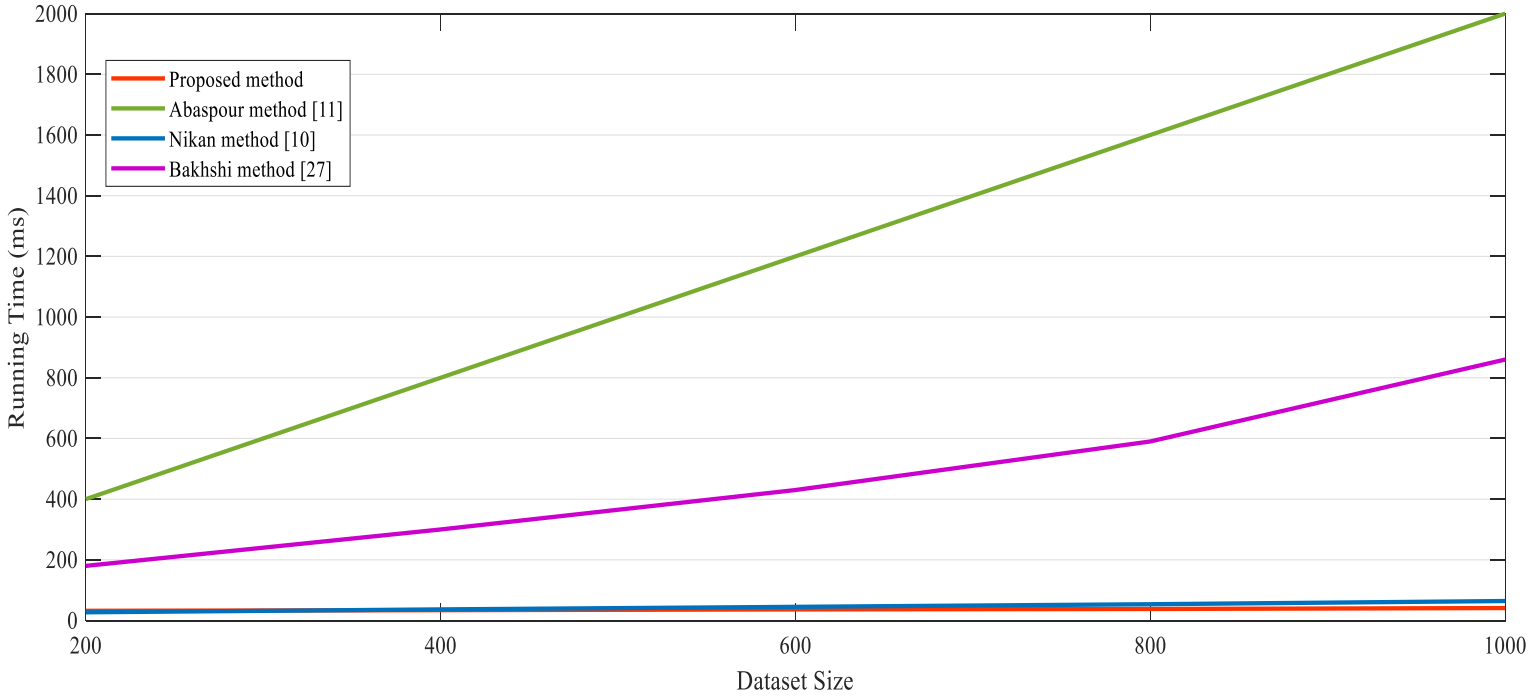
۵-۵-۴- پیچیدگی زمانی

برای پیاده‌سازی روش پیشنهادی، از سیستمی با مشخصات CPU: Intel Core i5-4300M, 2.60GHz و Memory:4 GB استفاده شده است. در این بخش، سرعت شناسایی تصویر در روش پیشنهادی محاسبه می‌شود. در این راستا، از زیر مجموعه‌ی اول پایگاه داده‌ی FERET (تعداد ۱۹۸۰ تصویر چهره و تمامی تصاویر به‌صورت روبرو هستند) استفاده می‌شود. سرعت شناسایی، حاصل میانگین ۱۰ تصویر آزمون می‌باشد. جدول (۵-۶)، زمان اجرا را به ازای یک تصویر آزمون در مراحل مختلف روش پیشنهادی نشان می‌دهد. همانطور که قبلا بیان شد، در برخی موارد، نیاز به اجرای مرحله‌ی دوم گام تطابق چهره نمی‌باشد. یعنی نیازی به استخراج درهم‌ساز تصویر مبتنی بر DCT و همچنین مقایسه‌ی نهایی نیست. بنابراین سرعت شناسایی تصویر بیشتر خواهد شد.

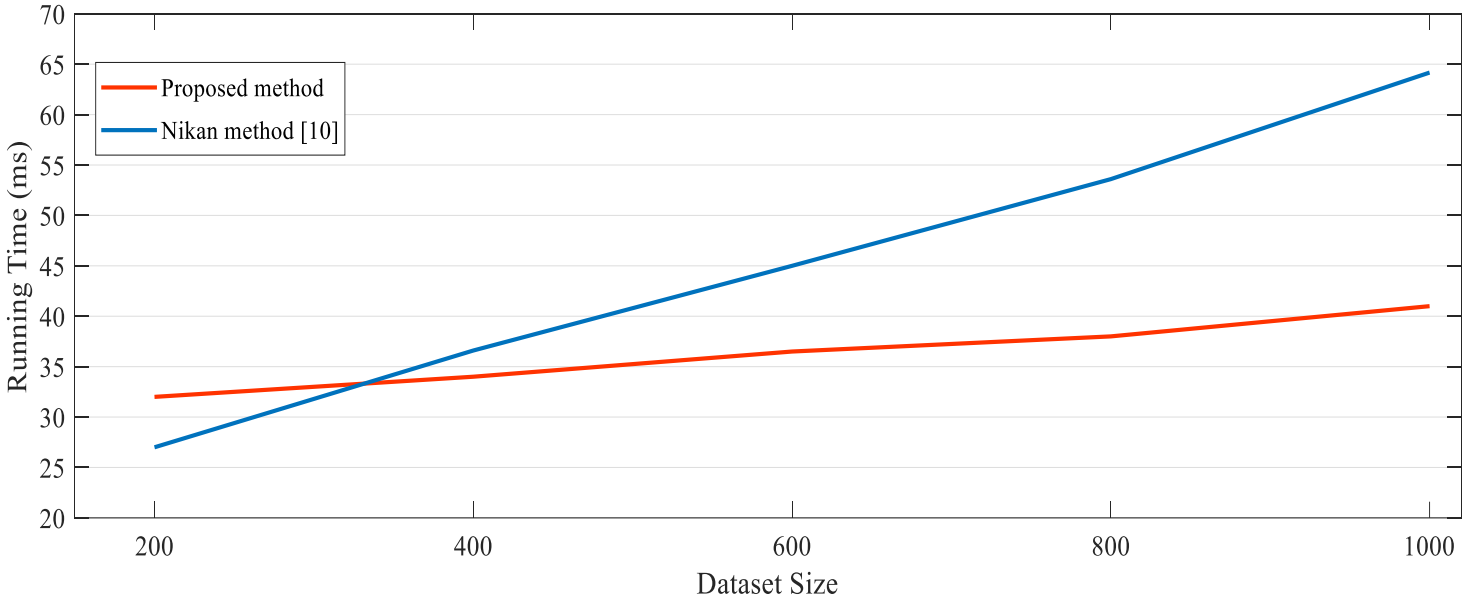
جدول ۵-۶- مدت زمان اجرای هر مرحله از مدل پیشنهادی

بخش الگوریتم	مدت زمان (میلی ثانیه)
پیش پردازش	۴
استخراج ویژگی‌های محلی	۷
نگاشت ویژگی‌های محلی و انتخاب شناسه‌های منتخب	۹
استخراج ویژگی سراسری	۱۰
تعیین هویت تصویر آزمون	۷

در ادامه، سرعت شناسایی با استفاده از حجم‌های مختلف داده نیز محاسبه شده است. منظور از افزایش حجم پایگاه داده، افزایش تعداد افراد در پایگاه داده می‌باشد. همانطور که در شکل (۵-۱۱) مشاهده می‌شود، مقایسه‌ای بین سرعت شناسایی روش پیشنهادی و برخی از روش‌های مرتبط (با استفاده از زیر مجموعه‌ی اول پایگاه داده‌ی FERET) انجام شده است. لازم به ذکر است مشخصات سیستم در این روش‌ها متفاوت است که در فصل دوم بیان شده است. به‌دلیل اینکه اختلاف زیادی بین زمان شناسایی روش پیشنهادی و روش نیکان [۱۰] با روش عباس‌پور [۱۱] و روش بخشی [۳۱] وجود دارد، بخشی از نمودار در شکل (۵-۱۲) بزرگنمایی شده است. همانطور که مشاهده می‌شود، با افزایش حجم پایگاه داده، زمان شناسایی روش پیشنهادی با شیب کمتری افزایش می‌یابد. مرجع نیکان [۱۰]، یک روش شناسایی چهره در پایگاه داده‌ی حجیم ارائه داده است و سرعت شناسایی این روش برابر با ۶۴,۱ میلی ثانیه می‌باشد. لازم به ذکر است که این نتیجه روی سیستمی با RAM برابر با ۶۴ و CPU Core i7 اجرا شده است. اما سرعت شناسایی روش پیشنهادی برابر با ۴۱ میلی ثانیه می‌باشد و با مشخصات سیستمی ضعیفتر گزارش شده است. از این مشاهده می‌توان نتیجه گرفت که استفاده از درهم‌سازی تصویر در روش پیشنهادی، سرعت شناسایی را افزایش می‌دهد.



شکل ۵-۱۱- مقایسه زمان شناسایی روش پیشنهادی با روش‌های مرتبط



شکل ۵-۱۲- مقایسه زمان شناسایی روش پیشنهادی با روش نیکان [۱۰]

۵-۶- جمع‌بندی

در این فصل، ابتدا پایگاه‌های داده‌ی مورد استفاده در این تحقیق معرفی شدند. سپس مقادیر پارامترهای مورد استفاده در الگوریتم پیشنهادی بیان شدند. همچنین، الگوریتم پیشنهادی ارزیابی شد و نتایج حاصل، مورد بررسی قرار گرفت و با روش‌های مرتبط دیگر مقایسه شد. در انتها، زمان اجرای روش پیشنهادی نیز مورد بررسی قرار گرفت. در این پژوهش، به دلیل استفاده از روش‌های درهم‌سازی تصویر به منظور بررسی دو دیدگاه بصری متفاوت از تصویر، مدل شناسایی چهره‌ی پیشنهادی به عملکرد بهتری نسبت به سایر روش‌های موجود از لحاظ سرعت و دقت رسیده است.

فصل ٦: نتیجه‌گیری

۶-۱- جمع‌بندی از پایان‌نامه

هدف این پایان نامه، استفاده از روش‌های درهم‌سازی تصویر برای شناسایی چهره در پایگاه داده‌ی حجیم است. به‌طوریکه عملکرد خوبی روی پایگاه‌های داده با شرایط تقریبا کنترل نشده داشته باشد و بتواند با دقت مناسب و سرعت خوبی تصاویر چهره را شناسایی کند.

در این تحقیق، تصاویر چهره با دو دیدگاه بصری جزئی‌نگر و کلی‌نگر مورد بررسی قرار می‌گیرند. همچنین برای تسریع بررسی این ویژگی‌ها، از روش‌های درهم‌سازی تصویر استفاده شده است. روش پیشنهادی در سه گام پیش‌پردازش، استخراج ویژگی و تطابق چهره معرفی شده است. در اولین گام، تصاویر چهره از نظر کیفیت روشنایی بررسی می‌شوند و نواحی چهره در تصاویر مشخص می‌شوند. در گام استخراج ویژگی، ابتدا ویژگی‌های محلی از تصاویر چهره‌ی موجود در پایگاه داده استخراج می‌شوند. سپس این ویژگی‌های محلی به مکان‌های جدول درهم‌ساز نگاشت می‌شوند. همچنین به‌منظور ارزیابی تصاویر از یک دیدگاه دیگر با استفاده از ویژگی‌های سراسری، از یک روش درهم‌سازی تصویر دیگر نیز استفاده می‌شود و درهم‌ساز مربوط به هر تصویر ذخیره می‌شود.

در گام تطابق چهره، ابتدا ویژگی‌های محلی تصویر آزمون استخراج می‌شود و به جدول درهم‌ساز حاصل از گام استخراج ویژگی، نگاشت می‌شوند. سپس با استفاده از یک سیستم رای‌دهی، شناسه‌هایی از مجموعه داده که بیشترین شباهت را از لحاظ دیدگاه بصری جزئی‌نگر با تصویر آزمون دارند، انتخاب می‌شوند. در نهایت، به‌منظور شناسایی هویت تصویر آزمون، ویژگی‌های سراسری تصویر آزمون با استفاده از یک روش درهم‌سازی تصویر، استخراج می‌شود و با ویژگی‌های سراسری شناسه‌های منتخب مرحله‌ی اول مقایسه می‌شود.

در این پایان‌نامه، از مجموعه پایگاه‌های FERET، AR، ORL و ترکیبی از پایگاه‌های داده‌ی MUCT، PICS، FEI و Face94 استفاده شده است. دو زیر مجموعه از پایگاه داده‌ی FERET در نظر گرفته شده است. زیر مجموعه‌ی اول شامل ۱۹۸۰ تصویر چهره از ۹۹۰ فرد می‌باشد. زیر مجموعه‌ی دوم شامل ۴۰۱۷

تصویر چهره از ۹۹۴ فرد می‌باشد. پایگاه داده‌ی AR شامل ۴۰۰۰ تصویر چهره از ۱۰۰ فرد است. پایگاه داده‌ی ORL نیز شامل ۴۰۰ تصویر از ۴۰ فرد می‌باشد. مجموعه تصاویر جمع‌آوری‌شده نیز شامل ۳۳۶۸ تصویر از ۱۶۸۴ فرد است. دقت شناسایی چهره در پایگاه داده‌ی ORL، AR، FERET و مجموعه تصاویر جمع‌آوری‌شده به ترتیب برابر ۱۰۰، ۹۹، ۹۶/۴۷ و ۹۵/۹۰ درصد است. همچنین سرعت شناسایی تصویر با استفاده از زیر مجموعه‌ی اول پایگاه داده‌ی FERET، ۴۱ میلی ثانیه بدست آمد. همانطور که در نمودار شکل (۵-۱۱) در فصل پنجم نشان داده شد، با افزایش حجم پایگاه داده، سرعت شناسایی تصویر با شیب اندکی کاهش می‌یابد. این مساله بیانگر عملکرد مناسب روش پیشنهادی در پایگاه‌های داده‌ی حجیم است. از مزیت‌های روش پیشنهادی می‌توان به چندین مورد اشاره کرد: به دلیل استفاده از روش‌های درهم‌سازی تصویر، ظرفیت روش پیشنهادی بالا است، یعنی عملکرد خوبی در حجم زیادی از تصاویر دارد. علاوه بر این، بکارگیری توابع درهم‌سازی تصویر منجر به افزایش سرعت شناسایی چهره می‌شود. در حالیکه پیچیدگی زمانی در بسیاری از روش‌ها، با افزایش تعداد شناسه‌های موجود در پایگاه داده به‌طور چشمگیری افزایش می‌یابد. یکی دیگر از مزیت‌های مهم مدل پیشنهادی، توانایی شناسایی چهره در شرایط تقریبا کنترل نشده است. تصاویر چهره در این شرایط شامل انسداد، تغییر زاویه‌ی چهره، نویز، افزایش سن، تغییرات روشنایی می‌باشند. به‌همین دلیل این روش می‌تواند در کاربردهایی مانند دوربین‌های نظارتی و امنیتی، به خوبی عمل کند. همچنین با توجه به نتیجه‌ی عملکرد مدل پیشنهادی روی پایگاه داده‌ی AR، می‌توان بیان کرد که روش پیشنهادی عملکرد خوبی را در صورت وجود شرایطی که افراد مجبور به پوشیدن ماسک باشند، ارائه می‌دهد. علاوه بر این، طول بردارهای ویژگی در روش پیشنهادی کم است و نیازی به روش‌های کاهش ابعاد ندارد.

۶-۲- پیشنهادهایی برای کارهای آینده

- اعمال پیش‌پردازش‌های مناسب‌تر مانند فیلتر هومومورفیک می‌تواند در عملکرد بهتر سیستم شناسایی تاثیر گذار باشد.
- روش پیشنهادی به تغییرات روشنایی زیاد، حساس است. بنابراین، مقاوم‌سازی روش در مقابل این چالش، منجر به عملکرد بهتر شناسایی چهره خواهد شد.
- می‌توان به جای پنجره‌گذاری، تصاویر را به صورت مفهومی بخش‌بندی کنیم و سپس درهم‌ساز را ایجاد کنیم. به عنوان مثال، از اجزای چهره مانند بینی، لب و چشم‌ها استفاده کرد و ضرایب DCT را تنها از این نواحی استخراج کرد. این امر باعث کوتاه شدن طول درهم‌ساز نیز خواهد شد.
- در این تحقیق، حد آستانه‌ی انتخاب شبیه‌ترین شناسه‌ها در مرحله‌ی اول از گام تطابق چهره، به‌صورت سعی و خطا مشخص شده است. انتخاب این مقدار، نیاز به مطالعه‌ی عمیق‌تری دارد. اگر بتوان مقدار دقیق‌تری را مشخص کرد، می‌تواند در بهبود نرخ شناسایی موثر باشد.
- در مرحله‌ی اول از گام تطابق چهره، می‌توان به هر یک از شناسه‌های منتخب، امتیاز وزن‌دار داد. در این صورت، شناسایی در مرحله‌ی دوم، با کمک امتیازات وزن‌دار حاصل از مرحله‌ی اول، با دقت بیشتری انجام می‌شود.
- می‌توان یک دیدگاه بصری دیگر به مدل پیشنهادی اضافه کرد و چهره را با دقت بیشتری بررسی و شناسایی کرد.

منابع

[1] Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A., (2003), "Face recognition: A literature survey", *ACM Computing Surveys*, Vol. 35, No. 4, 399–458.

[2] Pentland, A., (2000), "Looking at people: Sensing for ubiquitous and wearable computing", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No.1, 107–110.

[3] Best-Rowden, L., Han, H., Otto, C., Klare, B. F., and Jain, A. K., (2014), "Unconstrained face recognition: Identifying a person of interest from a media collection", *IEEE Transactions on Information Forensics and Security*, Vol. 9, No. 12, 2144–2157.

[4] Ding, C. and Tao, D., (2017), "Pose-invariant face recognition with homography-based normalization", *Pattern Recognition*, Vol. 66, No. July, 144–152.

[5] Stan, Z. Li., Anil, K. Jain., (2011), "Local Representation of Facial Features", *Handbook of Face Recognition.*

[6] I. Masi et al., "Learning Pose-Aware Models for Pose-Invariant Face Recognition in the Wild", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 379-393, 1 Feb. 2019, doi: 10.1109/TPAMI.2018.2792452.

[7] Weihua O., Xinge Y., Dacheng T., (2014), "Robust face recognition via occlusion dictionary learning", *Pattern Recognition*, Volume 47, no 4, Pages 1559-1572,

[8] 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.

[9] 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. 359386.

[10] Nikan, F. and Hassanpour, H., (2020), "Face recognition using non-negative matrix factorization with a single sample per person in a large database", *Multimedia Tools and Applications*, 28265–28276.

[۱۱] عباسپور طسمالو ن، (۱۳۹۹)، پایان‌نامه ارشد:"استفاده از دسته‌بند سلسله‌مراتبی در شناسایی چهره در پایگاه داده حجیم"، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود.

[12] Viola, P., Jones, M., (2001)., "Rapid object detection using a boosted cascade of simple features", In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1.

[13] Li, S. Z., Xin Wen Hou, Hong Jiang Zhang, and Qian Sheng Cheng. (2001)." Learning spatially localized, partsbased representation". *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR 2001, 207–212.

[14] Vadlamudi, L., Vaddella, R., (2017), "A Review of Robust Hashing Methods for Content Based Image Authentication based on DWT", *IEEE Transactions on Information Forensics and Security*, Vol. 3, No. 4.

[15] Shiyuan, H., Bokun, W., (2020), "Bidirectional Discrete Matrix Factorization Hashing for Image Search", *IEEE Transactions on Cybernetics*, Vol. 50, No. 9, 4157–4168.

[16] Lowe, D. G., (2004), "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, Vol. DOI: 60, No. 2, 91–110.

[17] Zeng, Z., Fang, T., Shah, S., and Kakadiaris, I. A., (2009), "Local feature Hashing for face recognition", *IEEE 3rd International Conference on Biometrics*, 1-8.

[18] Ferdinando, S., Andy, H., (1994), "Parameterisation of a Stochastic Model for Human Face Identification", *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*.

[19] Martinez, A., Benavente, R., (1998), "The AR face database", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1090–1104.

[20] Dehghani, M., Moeini, A., and Kamandi, A., (2019), "Experimental Evaluation of Local Sensitive Hashing Functions for Face Recognition", *5th IEEE International Conference on Web Research*, 184–195.

[21] Chen, J., and Zu. Y., (2020), "Local Feature Hashing with Binary Auto-Encoder for Face Recognition," *in IEEE Access*, Vol. 8, 37526-37540.

[22] Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J., (1996), "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *Lecture Notes in Computer Science*, Vol. 1064, No. 7, 45–58.

[23] Zafaruddin, G. M. and Fadewar, H. S., (2018), "Face recognition using eigenfaces", *Advances in Intelligent Systems and Computing*, Vol. 810, 855–864.

[24] Turk, M. and Pentland, A., (1991), "Eigenfaces for Face Detection / Recognition - RESUME", *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1–11.

[25] Xue Gao, Q., Zhang, L., and Zhang, D., (2008), "Face recognition using FLDA with single training image per person", *Applied Mathematics and Computation*, Vol. 205, No. 2, 726–734.

[26] Philips, P. J., Moon, P. J., and Rizvi, H., (2000), "The ferret evaluation methodology for face-recognition algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1090–1104.

[۲۷] 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

[28] Sharif, M., Ayub, K., Sattar, D., Raza, M., (2012), "Enhanced and Fast Face Recognition by Hashing Algorithm" *Journal of Applied Research and Technology*, Vol. 10.

[29] Schwartz, W. R., Guo, H., Choi, J., and Davis L., (2012), “Face identification using large feature sets,” *IEEE Transactions on Image Processing*, Vol. 21, No. 4, 2245-2255.

[30] Dos Santos, C. E., Kijak, E., Gravier, G., and Schwartz, W. R., (2016), “Partial least squares for face hashing”, *Neurocomputing*, Vol. 213, 34–47.

[31] Zangeneh, E., Rahmati, M., Mohsenzadeh, Y., (2020), “Low resolution face recognition using a two-branch deep convolutional neural network architecture”, *Expert Systems with Applications*, Vol. 139, 0957-4174.

[32] Zeng, J., Zhao, X., Gan, J., Mai, C., Zhai, Y., and Wang, F., (2018), “Deep Convolutional Neural Network Used in Single Sample per Person Face Recognition”, *Computational Intelligence and Neuroscience*.

[33] Patil, V. and Sarode, T., (2019), “Compressed Image Hashing using Minimum Magnitude CSLBP”, *Journal of AI and Data Mining*, Vol. 7, No. 2, 287–297.

[34] Xiao, J. and Wu, G. (2011). “A robust and compact descriptor based on center-symmetric LBP”, *IEEE 6th International Conference on Image and Graphics (ICIG),* Vol.7, 45219–45229.

[35] Ng, T., Chang, S., Hsu, Y. & Pepeljugoski, M. (2005), Columbia Photographic Images and Photorealistic Computer Graphics Dataset.

[36] Qin, C., Chen, X., Dong, J., and Zhang, X., (2016), “Perceptual image hashing with selective sampling for salient structure features”, *Displays*, Vol. 45, 26–37.

[37] Schaefer, G., Michal, S., (2004), “UCID: an uncompressed color image database”, *Storage and Retrieval Methods and Applications for Multimedia*, Vol. 5307.

[38] Pei, S.C., Domg, J.J., Chang, J.H., (2001), “Efficient implementation of quaternion Fouriertransform, convolution, and correlation by 2-D complex FFT”, *IEEE Trans. Signal Process*. Vol. 49, 2783–2797.

[39] Ouyang, J., Coatrieux, G., and Shu, H., (2015), “Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform*”, Digital Signal Processing*, Vol. 41, 98–109.

[40] Zhang, K., Zhang, Z., Li, Z., Member, S., Qiao, Y., and Member, S., (2016), "(MTCNN) Multi-task Cascaded Convolutional Networks", *IEEE Signal Processing Letters*, Vol. 23, No. 10, 1499–1503

[41] Gonzalez, E., Woods, R., (2006), “Digital Image Processing Using MATLAB,” *Prentice Hall*.

[۴۲] حسنپور ح. و اسدی امیری س. (۱۳۹۴)، "مفاهیم جامع پردازش تصویر دیجیتال به همراه پیاده سازی الگوریتمها با MATLAB"، چاپ اول، انتشارات دانشگاه صنعتی شاهرود.

[43] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L., (2008), “Speeded-Up Robust Features (SURF)”, *Computer Vision and Image Understanding*, Vol. 110, No. 3, 346–359.

[44] Gionis, A, Indyk P, and Motwani, R, (1999), “Similarity Search in High Dimensions via Hashing”, *25th International Conference on Very Large Data Bases*, 518–529.

[45] Indyk, P. and Motwani, R., (1998), “Approximate nearest neighbors”, *Proceedings of the thirtieth annual ACM*, 604–613.

[46] Datar, M., Immorlica, N., and Indyk, P., (2004), “Locality-Sensitive Hashing Scheme Based on p-Stable Distributions”, *Proceedings of the twentieth annual symposium on Computational geometry*, 253–262.

[47] Fridrich, G. M., (2000) “Robust hash functions for digital watermarking”, *in International Conference on Information Technology: Coding and Computing*, 27–29.

[48] Tang, Z., Yang, F., Huang, L., and Zhang, X., “Robust image hashing with dominant DCT coefficients”, *Optik*, Vol. 125, No. 18, (2014), 5102–5107.

[49] Milborrow, S., Morkel, J., (2010), “The MUCT Landmarked Face Database”, *Proceedings Pattern Recognition Association of South Africa*.

[50] “Psychological Image Collection at Stirling”, (2004), *School of Natural Sciences University of Stirling*.

[51] Thomaz, C.E., (2012), “FEI Face Database”, *Image Processing Laboratory, FEI University Center.*

[52] Libor, S., (2009), Facial Images: Faces94, *Computer Vision Science Research Projects*.

[53] Chakraborti, T. and Chatterjee, A., (2014), “Engineering Applications of Artificial Intelligence A novel binary adaptive weight GSA based feature selection for face recognition using local gradient patterns, modified census transform, and local binary patterns”, *Engineering Applications of Artificial Intelligence*, Vol. 33, 80–90.

[54] Li, Y., Zheng, Cui, W., Z., and Zhang, T., (2018), “Face recognition based on recurrent regression neural network”, *Neurocomputing*, Vol. 297, 50–58.

[55] Ouyang, A., Liu, Y., Pei, S., Peng, X., He, M., and Wang, Q., (2020), “A hybrid improved kernel LDA and PNN algorithm for efficient face recognition”, *Neurocomputing*, Vol. 393, 214–222.

[56] Lu, J., Tan, Y., and Wang, G., (2011), “Discriminative Multi-Manifold Analysis for Face Recognition from a Single Training Sample per Person”, *IEEE International Conference on Computer Vision*, 1943–1950.

[57] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017).” ImageNet classification with deep convolutional neural networks”. *Communications of the ACM*, 60(6), 84–90.

[58] Dora, L., Agrawal, S., Panda, R., and Abraham, A., (2017), “An evolutionary single Gabor kernel based filter approach to face recognition”, *Engineering Applications of Artificial Intelligence*, Vol. 62.

[59] Cureton, E. E. and Agostino, R. B., (2019), "Face Recognition by Independent Component Analysis Marian", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 13, No. 6, 296–338.

[60] Liao, M. and Gu, X., (2019), "Face recognition approach by subspace extended sparse representation and discriminative feature learning" *Neurocomputing*.

[61] Cortes, C., Vapnik, (1995), "V. Support-vector networks", *Machine Learning*, Vol.297, 273–297.

[62] Ojala, D., Pietik¨ainen, T., Harwood, M., (1996), "A comparative study of texture measures with classification based on feature distributions", *Pattern Recognition*, 51–59.

[63] Wang, J., Lu, C., Wang, M., Li, P., Yan, S., and Hu, X., (2014), "Robust face recognition via adaptive sparse representation", *IEEE Transactions on Cybernetics*, Vol. 44, No. 12, 2368–2378.

[64] Deng, W., Hu, J., and Guo, J., (2012), "Extended SRC: Undersampled face recognition via intraclass variant dictionary", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 9, 1864–1870.

[65] Martinez, A. M. and Kak, A. C., (2001), "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, 228–233.

[66] Min, R., Xu, S., & Cui, Z, (2019)." Single-Sample Face Recognition Based on Feature Expansion". *IEEE Access*, 7, 45219–45229.

واژه نامه

Image Hashing	درهم‌سازی تصویر
Locality Sensitive Hashing	درهم‌ساز حساس محلی
Speed Up Robust Features	ویژگی‌های قوی تسریع داده شده
Discrete Cosine Transform	تبدیل کسینوسی گسسته
Scale- Invariant Feature Transform	تبدیل ویژگی مستقل از مقیاس
Identity	شناسه
Scale-Space Extrema Detection	تشخیص اکسترمم‌های فضای مقیاس
Scale-space	فضای مقیاس
Bilinear interpolation	درون‌یابی دوقطبی
Gaussian convolution	کانولوشن گوسی
Variable-scale Gaussian	متغیر گوسی
Keypoint Localization	محلی‌سازی نقاط کلیدی
Second-order Taylor expansion	سری دوم تایلر
Orientation Assignment	گرایش گماری
Principal Component Analysis	تحلیل مولفه اساسی
Center Symmetric Local Binary Pattern	الگوی باینری محلی متقارن مرکزی
Canny	کنی
Streaming data	جریان داده
Indexing	نمایه‌گذاری
Partial Least Squares Regresion	رگرسیون حداقل مربعات جزئی
K nearest neighbor	K نزدیکترین همسایه
Occlusion	انسداد
Cascade	آبشاری
Contrast Limited Adaptive Histogram	برابرسازی هیستوگرام وفقی با محدودیت کنتراست
Illumination variation	تغییرات روشنایی

Pose variation	تغییرات زاویه
Fisherface	چهره فیشر
Eigenface	چهره ویژه
Biometric	زیست سنجی
Convolution Neural Network	شبکه عصبی کانولوشن
Image Intensity	شدت روشنایی تصویر
Classifiers	طبقه کننده‌ها
factor	عامل
Convolve	پیچش
Local Binary Pattern	الگوی دودویی محلی
Nearest Neighbor	نزدیکترین همسایه
Down Sampling	نمونه برداری کم
Resolution	وضوح
Resolution Low	وضوح پایین
Holistic Feature	ویژگی سراسری
Local Feature	ویژگی محلی
Feature Haar	ویژگی هار
Histogram Orientation Gradian	هیستوگرام گرادیان جهت‌دار
hash	درهم‌ساز

جدول اختصارات

KNN	K Nearest Neighbor	K نزدیک‌ترین همسایه
Clahe	Contrast Limited Adaptive Histogram	برابرسازی هیستوگرام وفقی با محدودیت کنتراست
SVD	Singular Value Decomposition	تجزیه مقدار تکین
FLDA	Fisher Linear Discrimination Analysis	حلیل خطی فیشر
PCA	Principal Component Analysis	تحلیل مؤلفه اصلی
CNN	Convolution Neural Network	شبکه عصبی کانولوشن
LBP	Local Binary Pattern	الگوی محلی دودویی
SIFT	Scale-Invariant Feature Transform	تبدیل ویژگی مستقل از مقیاس
SURF	Speeded Up Robust Features	ویژگی‌های قوی تسریع یافته
LSH	Locality Sensitive Hashing	درهم‌سازی حساس محلی
DCT	Discrete Cosine Transform	تبدیل کسینوسی گسسته
CSLBP	Center Symmetric Local Binary Pattern	الگوی باینری محلی متقارن مرکزی
QDFT	Quaternion Discrete Fourier Transform	تبدیل فوریه گسسته
PLS	Partial Least Squares	حداقل مربعات جزئی
MTCNN	Multi-Task cascaded Convolutional Neural Networks	شبکه‌ی عصبی کانواوشن آبشاری چند وظیفه‌ای

Abstract

Face recognition plays a crucial role in video surveillance, access control systems, and forensic. Although numerous techniques exist for face recognition, they are limited to both a controlled environment and plenty of training data, especially for recognition in large datasets. In addition, increasing the number of identities in the dataset hardens the computational complexity of recognition. To address these problems, a framework is proposed in this research that employs hashing functions for feature extraction in face recognition. The framework leverages a cascaded architecture with two stages of analyzing different visual information based on image hashing. Firstly, we produce candidates by rejecting a large number of dissimilar identities in terms of local visual information. Similar identities are found quickly through the random independent hash functions inspired by Locality-Sensitive Hashing (LSH). Secondly, we further refine candidates and recognize the most similar identity according to global visual information. The global feature is obtained by hashing each face into a high-quality binary feature space using Discrete Cosine Transform (DCT) coefficients. We evaluated the performance of the proposed method on the FERET, ORL, and AR datasets. It is witnessed that our method improves the recognition rate and significantly reduces the computational complexity of face recognition.

Keywords: Face recognition, Image hashing, LSH, DCT based hashing algorithm, SURF, Large datasets.

Shahrood University of Technology

Faculty of Computer Engineering and Information Technology

M.Sc. Thesis in Artificial Intelligence Engineering

Face Recognition Using Image Hashing

By:

Mahsa Ghasemi

Supervisor:

Prof. Hamid Hassanpour

July 2021