

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی کامپیوتر و فناوری اطلاعات  
پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

## استخراج ویژگی از تصاویر چهره در پایگاه داده حجیم

نگارنده: فاطمه نیکان

استاد راهنما

دکتر حمید حسن پور

بهمن ماه ۱۳۹۷

ای سزاور شنای خویش

ای سگرکننده‌ی عطای خویش

ای شیرین‌نمانده‌ی بلا‌ی خویش

ره‌ی بذات خود از شنای تو عاجز

و به عقل خود از شناخت منت تو عاجز

و به توان خود از سزای تو عاجز

«خواجہ عبداللہ انصاری»

ماحصل آموخته‌هایم را تقدیم می‌کنم

به آنان که مهر آسمانی‌شان آرام بخش آلام زمینی‌ام است

به استوارترین تکیه‌گاهم، دستان پر مهر پدرم

به سبزترین نگاه زندگیم، چشمان سبز مادرم

که هرچه آموختم در مکتب عشق شما آموختم و هرچه بگوختم قطره‌ای از دریای سیکران مهربانیان را سپاس  
توانم بگویم.

امروز، مستی‌ام به امید شامت و فردا کلید باغ بهشتم رضای شما

ره آوردی گران‌سنگ‌تر از این ارزان نداشتم تا به خاک پایتان نثار کنم،

باشد که حاصل تلاشم نسیم کوزه‌خوار خشکی‌تان را برزید.

بوسه بر دستان پر مهرتان

همچنین تقدیم به برادرانم که وجودشان مایه‌ی دلگرمی من است.

پاس خدای را که هر چه دارم از اوست، به امید آن که توفیق یابم جز خدمت به خلق او نکوشم.

از پدر و مادر عزیزم... این دو معلم بزرگوارم... که همواره بر کوتاهی و درستی من، قلم عفو کشیده و گریمانه

از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی یار و یاور بی چشم داشت برای من بوده اند؛

از استاد دانشمند و فرزانه ام جناب آقای دکتر حسن پور که در کمال سع صدر، با حسن خلق و فروتنی از بیچ

کلی در این عرصه بر من دریغ نمودند و زحمات راهبانی این پایان نامه را بر عهده گرفتند؛ کمال تشکر و قدردانی

را دارم.

همچنین لازم می دانم از هم کلاسی و دوست عزیز و مهربانم خانم مهندس خدا بخشی و دانشجویان آزمایشگاه

پردازش تصویر و داده کاوی برای مساعدت ها و همراهی هایشان قدردانی نمایم.

باشد که این خردترین، بخشی از زحمات آنان را پاس گوید.

# تعهد نامه

اینجانب فاطمه نیکان دانشجوی دوره کارشناسی ارشد رشته مهندسی کامپیوتر هوش مصنوعی دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی شاهرود نویسنده پایان نامه استخراج ویژگی از تصاویر چهره در پایگاه داده حجیم تحت راهنمایی دکتر حمید حسن پور متعهد می شوم:

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

تاریخ ۹۷/۱۱/۷

## امضای دانشجو

### مالکیت نتایج و حق نشر

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

## چکیده

سیستم شناسایی چهره یکی از حوزه‌های تحقیقاتی مهمی است که برای شناسایی هویت افراد بکار می‌رود. این سیستم، در زمینه‌های مختلفی از جمله امنیتی، پردازش تصویر، نظارت ویدئویی و شناسایی مجرمین کاربرد دارد. همچنین با پیشرفت روز افزون دوربین‌های دیجیتالی، توسعه سیستم شناسایی چهره نیز افزایش یافته است.

سیستم شناسایی چهره با چالش‌هایی نظیر پوشیده بودن بخشی از چهره، زاویه سر، تغییرات روشنایی، حالت‌های چهره، حجم بالای پایگاه داده و تعداد کم نمونه‌های آموزشی روبه‌رو است. هر کدام از روش‌های موجود در این حوزه، به بررسی و حل یک یا چند چالش پرداخته‌اند. یکی از اشکال‌هایی که می‌توان به اغلب روش‌های موجود گرفت، این است که؛ باید چندین چهره از یک شخص وجود داشته باشد، تا سیستم بتواند به خوبی عمل کند. معمولاً روش‌های شناسایی چهره، با استفاده از پایگاه داده‌های کم حجم، توانسته‌اند به دقت بالایی دست یابند. در حالی که اگر حجم پایگاه داده افزایش یابد، عملکرد این روش‌ها به‌طور قابل توجهی کاهش می‌یابد.

تمرکز اصلی این پایان‌نامه، روی سه چالش حجم پایگاه داده، تعداد کم نمونه‌های آموزش (به ازای هر شخص، یک تصویر) و سرعت سیستم است. روش ارائه شده، به کمک تجزیه ماتریس نامنفی، ویژگی‌های اساسی در ساختار چهره را استخراج می‌کند. در این روش، براساس مجموعه تصاویر چهره موجود در پایگاه داده مورد بررسی، ماتریس پایه و ماتریس وزن‌ها، حاصل می‌شود. ماتریس پایه شامل چندین نسخه از عناصر چهره مانند دهان، بینی و دیگر اجزای صورت، در مکان‌ها یا شکل‌های مختلف است. با ترکیب این عناصر و ماتریس وزن از یک چهره، می‌توان تغییرات متفاوت از آن چهره را تولید کرد. بر این اساس، برای جستجوی یک چهره در پایگاه داده، صرفاً جستجو روی بردارهای ویژگی تصاویر چهره موجود در پایگاه داده انجام می‌شود. در این تحقیق، به منظور تشکیل دقیق‌تر عناصر ساختاری، ماتریس پایه‌ی جداگانه‌ای برای نیمه بالا و پایین مجموعه تصاویر چهره تشکیل می‌گردد. همچنین برای بهبود کیفیت روشنایی تصاویر از سه پیش‌پردازش، یکسان‌سازی هیستوگرام، برابرسازی هیستوگرام وفقی با محدودیت کنتراست و تنظیم شدت استفاده شده است. روش ارائه شده با استفاده از پایگاه داده FERET، که حاوی یک تصویر چهره از ۹۹۰ فرد مختلف می‌باشد، مورد ارزیابی قرار گرفته است. روش پیشنهادی، با سرعت بازیابی مناسب و نرخ بازیابی نزدیک به ۹۳ درصد تصاویر چهره را از پایگاه داده بازیابی می‌کند.

## کلمات کلیدی

استخراج ویژگی، پایگاه داده حجیم، شناسایی چهره، سرعت بازیابی، وجود تنها یک تصویر به ازای هر فرد، تجزیه ماتریس نامنفی

# لیست مقالات مستخرج از پایان نامه

• نیکان ف، حسن پور ح، "بازیابی سریع تصویر چهره در پایگاه داده حجیم"، بیست و چهارمین کنفرانس سالانه ملی انجمن کامپیوتر ایران، دانشگاه صنعتی شریف، تهران ۱۳۹۷ (ارسال شده)

- Nikan F., Hassanpour H., "Face Recognition Using Non-negative Matrix Factorization with a Single Sample per Person from in a Large Database", *Computer vision and image understanding* (in progress).
- Nikan F., Hassanpour H., "Low-resolution Face Recognition with Single Sample per Person by using NMF", *Pattern Recognition and Image Analysis* (in progress).



## فهرست مطالب

فهرست اشکال ..... ل

فهرست جداول ..... ن

### فصل اول: سیستم‌های شناسایی چهره

۱-۱- مقدمه ..... ۲

۱-۲- چالش‌های سیستم شناسایی چهره ..... ۲

۱-۳- روش‌های شناسایی چهره ..... ۳

۱-۴- هدف و نوآوری پایان‌نامه ..... ۵

۱-۵- پیش فرض‌های تحقیق ..... ۶

۱-۶- ساختار پایان‌نامه ..... ۷

۱-۷- جمع‌بندی ..... ۸

### فصل دوم: مروری بر کارهای پیشین

۲-۱- مقدمه ..... ۱۰

۲-۲- روش‌های استخراج ویژگی بر اساس ویژگی‌های سراسری ..... ۱۰

۲-۲-۱- چهره ویژه ..... ۱۰

۲-۳- روش‌های استخراج ویژگی بر اساس ویژگی‌های محلی ..... ۱۲

۲-۳-۱- الگوی دودویی محلی (LBP) ..... ۱۲

۲-۴- روش‌های استخراج ویژگی ترکیبی ..... ۱۴

۲-۴-۱- ترکیب روش GABOR WAVELET با روش تجزیه ماتریس نامنفی (NMF) .. ۱۴

۲-۵- سایر روش‌ها ..... ۱۵

۲-۵-۱- روش ویژگی‌های مکان-فرکانسی ..... ۱۵

۲-۵-۲- ترکیبی از روش هیستوگرام گرادیان جهت‌دار (HOG) و تبدیل فوریه زمان

کوتاه (STFT) ..... ۱۵

- ۳-۵-۲- شناسایی چهره با استفاده از روش تجزیه و تحلیل خطی فیشر (FLDA) ..... ۱۷
- ۴-۵-۲- شبکه عصبی عمیق مورد استفاده با یک نمونه برای شناسایی چهره ..... ۱۸
- ۵-۲- جمع بندی ..... ۱۹

### فصل سوم: روش پیشنهادی

- ۱-۳- مقدمه ..... ۲۲
- ۲-۲- آشکارسازی چهره ..... ۲۳
- ۱-۲-۳- تصویر انتگرال ..... ۲۳
- ۲-۲-۳- طبقه بندی آبخاری مبتنی بر الگوریتم آدابوست ..... ۲۴
- ۳-۳- پیش پردازش ..... ۲۵
- ۱-۳-۳- یکسان سازی هیستوگرام ..... ۲۶
- ۲-۳-۳- برابر سازی هیستوگرام وفقی با محدودیت کنتراست ..... ۲۸
- ۳-۳-۳- تنظیم شدت روشنایی تصویر ..... ۲۸
- ۴-۳- استخراج ویژگی ..... ۲۹
- ۱-۴-۳- تجزیه ماتریس نامنفی ..... ۲۹
- ۵-۳- معیار شباهت ..... ۳۷
- ۱-۵-۳- ضریب همبستگی ..... ۳۷
- ۲-۵-۳- معیار فاصله اقلیدسی ..... ۳۸
- ۳-۵-۳- فاصله کسینوسی ..... ۳۸
- ۶-۳- جمع بندی ..... ۴۲

### فصل چهارم: ارزیابی روش پیشنهادی و بررسی نتایج

- ۱-۴- مقدمه ..... ۴۴
- ۲-۴- پایگاه داده ..... ۴۵
- ۳-۴- نتایج آشکارسازی چهره ..... ۴۶
- ۴-۴- نتایج روش پیشنهادی ..... ۴۶

۴۷	.....	۱-۴-۴ به دست آوردن مقدار مناسب مرتبه‌ی (R) روش NMF
۵۵	.....	۲-۴-۴ آزمایش با تصاویر وضوح پایین
۵۷	.....	۵-۴- بررسی نتایج
۶۰	.....	۶-۴- جمع‌بندی

### فصل پنجم: نتیجه‌گیری

۶۴	.....	۱-۵- خلاصه و جمع‌بندی
۶۵	.....	۲-۵- پیشنهاد ادامه کار
۶۷	.....	مراجع
۷۱	.....	واژه‌نامه
۷۳	.....	جدول اختصارات

## فهرست اشکال

- شکل (۱-۱): فرآیند کلی الگوریتم ارائه شده برای شناسایی چهره ..... ۶
- شکل (۲-۱): نمونه تصاویر پایگاه داده FERET ..... ۷
- شکل (۱-۲): عملیات الگوریتم LBP اصلی از مرجع [۱۳] ..... ۱۳
- شکل (۲-۲): عملیات الگوریتم LBP دایره‌ای [۱۳] ..... ۱۳
- شکل (۳-۲): (a) تصویر اصلی (b) تصویر  $\mathcal{E}$  ..... ۱۸
- شکل (۴-۲): طرح اصلی روش نمونه‌گیری گسترش یافته ..... ۱۹
- شکل (۱-۳): فرآیند کلی الگوریتم ارائه شده ..... ۲۲
- شکل (۲-۳): نمونه‌هایی از مستطیل‌های استفاده شده در الگوریتم Viola-Jones ..... ۲۳
- شکل (۳-۳): محاسبه مجموع مقادیر پیکسلی یک ویژگی مستطیلی، با استفاده از چهار آرایه در تصویر انتگرال در [۵] ..... ۲۴
- شکل (۴-۳): برخی از ویژگی‌های انتخاب شده توسط آداپوست در [۵] ..... ۲۵
- شکل (۵-۳): نمونه‌ای از استخراج ناحیه چهره از تصویر ورودی با استفاده از الگوریتم Viola-Jones ..... ۲۵
- شکل (۶-۳): نمونه‌ای از تصاویر پایگاه داده ..... ۲۶
- شکل (۷-۳): تصویر اصلی و تصویر با پیش‌پردازش یکسان‌سازی هیستوگرام ..... ۲۷
- شکل (۸-۳): تصویر اصلی و تصویر با پیش‌پردازش clahe ..... ۲۸
- شکل (۹-۳): تصویر اصلی و تصویر با پیش‌پردازش تنظیم شدت روشنایی تصویر ..... ۲۹
- شکل (۱۰-۳): نمونه‌ای از بیان معادله (۳-۵)؛ [آمین تصویر در ماتریس پایگاه داده، که با مجموع حاصلضرب ماتریس  $H$  و  $W$  تخمین زده شده [۳۵] ..... ۳۱
- شکل (۱۱-۳): نمایش تجزیه یک چهره با استفاده از روش NMF [۳۳] ..... ۳۲
- شکل (۱۲-۳): الگوریتم قواعد به‌روزرسانی ضربی [۳۹] ..... ۳۳
- شکل (۱۳-۳): الگوریتم کاهش گرادیان [۳۹] ..... ۳۳
- شکل (۱۴-۳): الگوریتم کمترین مربعات متناوب [۳۹] ..... ۳۴
- شکل (۱۵-۳): مقایسه بردارهای ویژگی تصاویر مختلف در روش ارائه شده ..... ۳۶
- شکل (۱۶-۳): بازیابی صحیح تصویر ۳۱۰ از مجموعه آزمون با بیشترین مقدار همبستگی نشان داده شده با تصویر ۳۱۰ از پایگاه‌داده در روش ارائه شده ..... ۴۰
- شکل (۱۷-۳): نمونه‌ای از مقایسه تصویر آزمون با سه تصویر از مجموعه تصاویر پایگاه داده با بیشترین شباهت ..... ۴۱
- شکل (۱-۴): نمونه‌هایی از مجموعه‌های پایگاه داده FERET ..... ۴۶
- شکل (۲-۴): استخراج ناحیه چهره پایگاه داده FERET با استفاده از الگوریتم Viola-Jones ..... ۴۶
- شکل (۳-۴): نمونه تصاویر از مجموعه آموزش و آزمون مورد بررسی. سطر اول مجموعه آموزش و سطر دوم مجموعه آزمون ..... ۴۷
- شکل (۴-۴): میانگین دقت و بیشینه دقت در ۱۰ بار اجرا با اندازه‌های مختلف تصویر در روش پیشنهادی قبل از پیش‌پردازش ..... ۴۹
- شکل (۵-۴): نمونه‌ای از ماتریس پایه ..... ۴۹

- شکل (۴-۶): مقایسه تصویر به دست آمده با استفاده از ماتریس دیکشنری پایگاه داده و تصویر آزمون در روش پیشنهادی  
 ۵۱.....
- شکل (۴-۷): تأثیر پیش پردازش در نرخ بازیابی .....  
 ۵۱.....
- شکل (۴-۸): شناسایی نادرست تصویر آزمون با یکی از تصاویر مجموعه آموزش در روش ارائه شده .....  
 ۵۲.....
- شکل (۴-۹): شناسایی درست تصویر آزمون با یکی از تصاویر مجموعه آموزش در روش ارائه شده .....  
 ۵۲.....
- شکل (۴-۱۰): روند تغییرات نرخ بازیابی با افزایش حجم پایگاه داده در الگوریتم ارائه شده .....  
 ۵۴.....
- شکل (۴-۱۱): عملکرد روش ارائه شده با اعمال پیش پردازش .....  
 ۵۵.....
- شکل (۴-۱۲): مقایسه روش پیشنهادی و روش [۱۵] زمانی که تصاویر آزمون وضوح پایینی دارند. ....  
 ۵۶.....
- شکل (۴-۱۳): شناسایی نادرست تصویر چهره در روش ارائه شده .....  
 ۵۷.....
- شکل (۴-۱۴): مقایسه زمان بازیابی یک تصویر آزمون در حجم‌های مختلف پایگاه داده روش پیشنهادی با روش [۱۵]  
 ۵۸.....

## فهرست جداول

- جدول (۱-۲): عملکرد الگوریتم پیشنهادی [۱۶] روی پایگاه داده FERET ..... ۱۶
- جدول (۱-۳): مقایسه روش‌های NMF ..... ۳۵
- جدول (۱-۴): مجموعه تصاویر پایگاه داده FERET ..... ۴۵
- جدول (۲-۴): به دست آوردن مرتبه‌ی (۲) روش NMF و سایز مناسب تصویر و بدون اعمال پیش‌پردازش ..... ۴۸
- جدول (۳-۴): مقایسه معیارهای شباهت در روش پیشنهادی ..... ۵۱
- جدول (۴-۴): نرخ و زمان بازیابی الگوریتم ارائه شده در حجم‌های متفاوت پایگاه داده ..... ۵۳
- جدول (۵-۴): بررسی زمانی روش پیشنهادی ..... ۵۸
- جدول (۶-۴): مقایسه روش پیشنهادی با روش‌های دیگر ..... ۵۹

# فصل اول

## سیستم‌های شناسایی چهره

## ۱-۱- مقدمه

در دهه‌های اخیر، شناسایی چهره به یکی از حوزه‌های تحقیقاتی مهم تبدیل شده است. سیستم بینایی انسان به صورتی است که می‌تواند در جمعیتی انبوه، چهره شخص دیده شده را به سرعت شناسایی کند. لذا نه تنها محققان علوم کامپیوتر، بلکه دانشمندان علوم اعصاب و روانشناسان نیز به این حوزه، علاقه‌مندند. زیرا با تحقیق در این حوزه، می‌توانند به چگونگی عملکرد مغز انسان پی ببرند. هر خصوصیت فیزیولوژیکی یا رفتاری انسان می‌تواند به عنوان یک ویژگی زیست‌سنجی<sup>۱</sup> استفاده شود [۱]. شناسایی چهره، اثر انگشت، عنبیه چشم، اثر کف دست همگی یک سیستم زیست‌سنجی هستند. اما تشخیص هویت افراد با استفاده از تصاویر چهره، در مقایسه با سایر مؤلفه‌های زیست‌سنجی دارای مزیت است. زیرا برای تشخیص هویت، تنها از تصویر چهره افراد استفاده می‌شود و تماسی با خود فرد ندارد [۲،۳]. شناسایی چهره در کاربردهای مختلف امنیتی، پردازش تصویر، شناسایی اتوماتیک سریع و بدون دخالت شخص، مورد استفاده قرار می‌گیرد. همچنین در زمینه‌هایی مانند تأیید یا تشخیص هویت افراد، سیستم‌های امنیتی نظیر بانک‌ها، نظارت ویدئویی، ردگیری و شناسایی مجرمین، ثبت حضور و غیاب، ارتباطات اینترنتی، استفاده در بازی‌های کامپیوتری و غیره نیز استفاده می‌شود. این نکته نیز قابل ذکر است که توسعه دوربین‌های دیجیتالی پیشرفته، در بالا رفتن طرفداران این حوزه مؤثر بوده است.

با وجود همه پیشرفت‌هایی که در این زمینه به دست آمده، باید به این نکته توجه داشت که عملکرد رضایت‌بخش سیستم‌های شناسایی چهره فقط در سناریوهای کنترل شده قابل دستیابی است و در مواجهه با سناریوهای دنیای واقعی، به طور قابل توجهی کاهش می‌یابد [۴].

## ۱-۲- چالش‌های سیستم شناسایی چهره

سیستم شناسایی چهره، با چالش‌هایی روبه‌رو است، که در زیر به چند مورد از آن‌ها اشاره می‌کنیم.

---

<sup>۱</sup> Biometric



- انسداد<sup>۱</sup>: پوشیده بودن بخشی از چهره، مانند عینک، ریش، ماسک؛ که باعث می‌شود شناسایی چهره با مشکل مواجه شود.
  - وجود تنها یک تصویر به ازای هر فرد: اگر به ازای هر شخص، تنها یک تصویر موجود باشد؛ آموزش سیستم شناسایی چهره به سختی انجام می‌شود. زیرا در شرایط مختلف، ظاهر هر شخص ممکن است تغییر کند، معمولاً روش‌های موجود نیاز دارند تا چندین چهره از یک شخص را داشته باشند.
  - وضوح پایین تصویر<sup>۲</sup>: به علت هزینه و فضای ذخیره‌سازی، کیفیت تصاویر عکس‌برداری شده معمولاً با وضوح پایین صورت می‌گیرد.
  - حجم بالای پایگاه داده: در سیستم شناسایی چهره با حجم‌های بالا، احتمال دارد که دو شخص متفاوت را مشابه یک‌دیگر اعلام کند.
  - تغییرات روشنایی<sup>۳</sup>: با کم و یا زیاد بودن نور بر روی چهره، ممکن است استخراج ویژگی‌های چهره با مشکل روبه‌رو شود.
  - سرعت: سرعت بازیابی یکی از پارامترهای حائز اهمیت در سیستم‌های امنیتی است [۵].
  - زاویه چهره<sup>۴</sup>: یکی از عوامل مؤثر در نرخ شناسایی، زاویه چهره در تصویر است.
- همچنین در این سیستم، چالش‌های دیگری از جمله، تغییر سن، تغییر حالت‌های چهره، تعیین جنسیت و غیره نیز وجود دارد.

### ۱-۳- روش‌های شناسایی چهره

با توجه به اهمیت سیستم شناسایی چهره و چالش‌های موجود در این حوزه، روش‌های مختلفی ارائه شده است. سیستم شناسایی چهره شامل سه مرحله آشکارسازی چهره، استخراج ویژگی و مقایسه است،

<sup>1</sup> Occlusion

<sup>2</sup> Low Resolution

<sup>3</sup> Illumination variation

<sup>4</sup> Pose variation

که اکثر روش‌ها به مرحله استخراج ویژگی پرداخته‌اند. در شناسایی چهره، ویژگی‌هایی که از چهره استخراج می‌شوند، جنبه حیاتی دارند. زیرا اگر این مرحله خوب عمل نکند، حتی بهترین طبقه‌بندکننده‌ها<sup>۱</sup> هم نمی‌توانند نتایج مناسب را تولید کنند [۶]. عمده این روش‌ها به سه دسته تقسیم می‌شوند: ویژگی‌های سراسری<sup>۲</sup> [۷،۸]، ویژگی‌های محلی<sup>۳</sup> [۹،۱۰] و ترکیبی.

## ۱. ویژگی‌های سراسری:

روش‌هایی که بر اساس ویژگی‌های سراسری، استخراج ویژگی را انجام می‌دهند. این روش‌ها، تمام تصویر چهره را به عنوان ورودی پردازش کرده و آن را به یک بردار ویژگی که معمولاً دارای ابعاد بالایی است، تبدیل می‌کنند.

## ۲. ویژگی‌های محلی:

روش‌هایی که بر اساس ویژگی‌های محلی هستند، ابتدا الگوی ساختار هر ناحیه محلی از چهره را توصیف می‌کنند. سپس تمام این الگوهای به دست آمده را به یک بردار ویژگی تبدیل می‌نمایند [۱۱].

## ۳. ویژگی‌های ترکیبی:

روش‌های ترکیبی به روش‌هایی می‌گویند که هم از ویژگی‌های محلی و هم از ویژگی‌های سراسری استفاده می‌کنند. این که ویژگی‌ها به چه شکلی با یکدیگر ترکیب شوند، فاکتور مهمی است. زیرا باید ترکیب آن‌ها به گونه‌ای باشد که بتوانند هم مزیت یکدیگر را حفظ کنند و هم ضعف یکدیگر را پوشش دهند.

روش‌های شناسایی چهره در فصل دوم، مورد بررسی قرار خواهند گرفت.

---

<sup>1</sup> Classifiers

<sup>2</sup> Holistic Feature

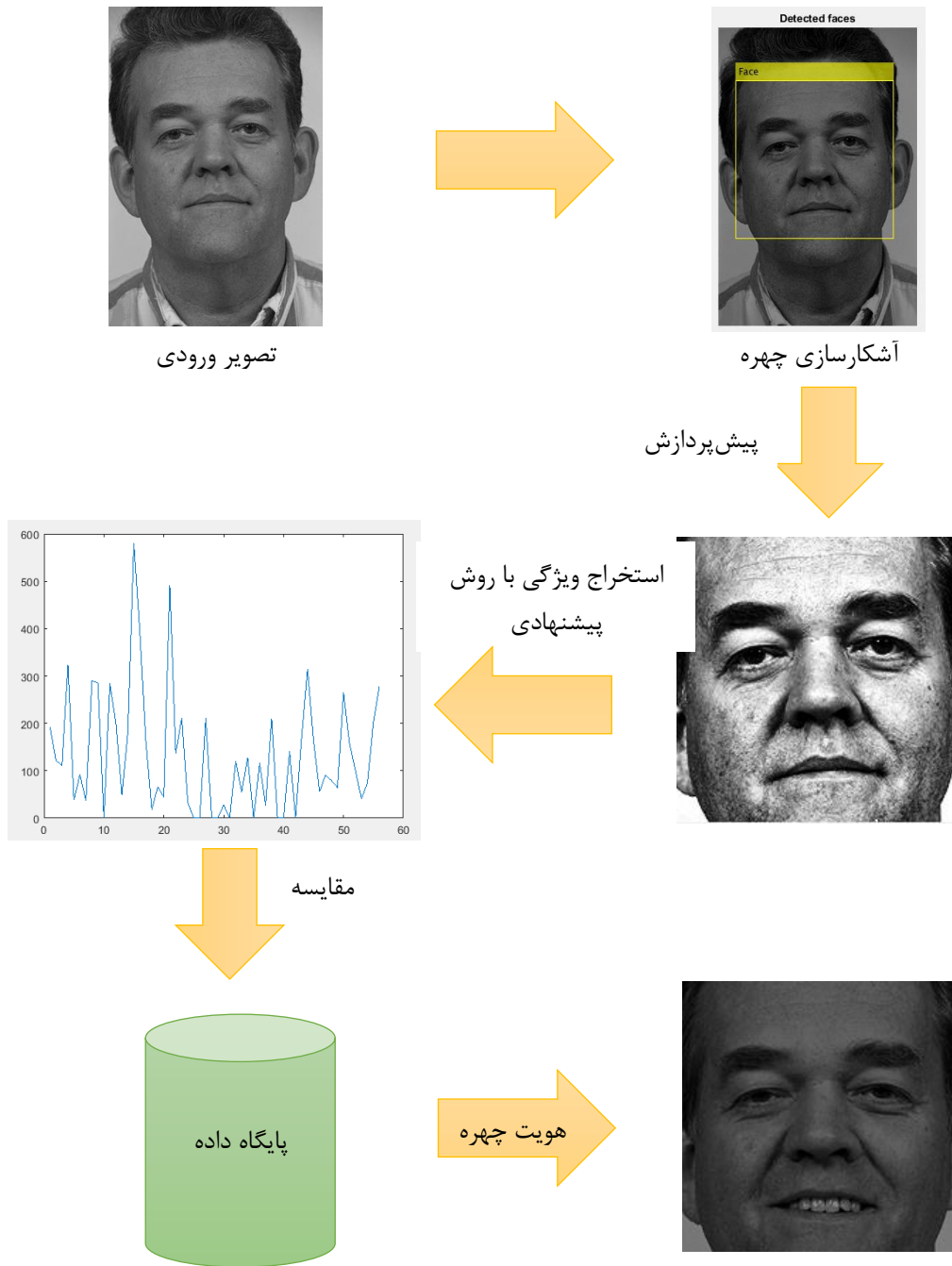
<sup>3</sup> Local Feature

## ۴-۱- هدف و نوآوری پایان نامه

عنوانی که برای این پایان نامه انتخاب شده "استخراج ویژگی از تصاویر چهره در پایگاه داده حجیم" است. در این عنوان دو عبارت کلیدی (استخراج ویژگی و پایگاه داده حجیم) وجود دارد که جهت تمرکز این تحقیق را نشان می‌دهد. علاوه بر این، در این تحقیق، از هر فرد تنها یک تصویر در پایگاه داده موجود است. در واقع مسأله‌ای که در این پایان نامه مطرح شده، با هدف بهبود برخی چالش‌ها در شناسایی چهره است. از جمله بهبودهای حاصل می‌توان به موردهای زیر اشاره کرد:

- **استفاده از روش پیشنهادی در پایگاه داده حجیم:** روش پیشنهادی، ویژگی‌هایی را استخراج می‌کند که فاصله‌ی بردارهای ویژگی بین افراد مختلف حداکثر شود، که این حاکی از آن است که ظرفیت روش بالا است. یعنی اگر تعداد تصاویر اضافه شود، خللی در عملکرد روش پیشنهادی به وجود نمی‌آید.
- **وجود تنها یک تصویر به ازای هر فرد:** روش‌های دیگر جهت آموزش نیاز به چند تصویر چهره در پایگاه داده دارند. اما در این پایان نامه، روش پیشنهادی با استفاده از تنها یک تصویر به ازای هر فرد، عملیات شناسایی را می‌تواند انجام دهد. از این رو فضای ذخیره‌سازی کاهش یافته و می‌توان به سادگی از آن استفاده کرد.
- **دستیابی به سرعت مناسب:** همان‌طور که در فصل چهارم بیشتر توضیح خواهیم داد، استخراج ویژگی‌های مناسب و طول بردار ویژگی کم در به دست آوردن سرعت مناسب، مؤثر بوده است.

با توجه به مطالب بیان شده، هدف این تحقیق، استخراج ویژگی‌هایی از تصاویر چهره است که قابلیت تفکیک بین تصاویر با تعداد زیاد را داشته باشد. علاوه بر این، بتواند در پایگاه داده حجیم، تصویر مورد نظر را با سرعتی مناسب بازیابی کند. روال کلی پایان نامه در شکل (۱-۱) نشان داده شده است.



شکل (۱-۱): فرآیند کلی الگوریتم ارائه شده برای شناسایی چهره

### ۱-۵- پیش فرض‌های تحقیق

ما در این تحقیق، پیش فرض‌هایی را در نظر گرفته‌ایم؛ که در زیر آمده‌اند.

- تصاویر چهره آموزش و آزمون، خاکستری فرض شده و همگی تمام رخ و کامل هستند.
- روش ارائه شده تنها با یک تصویر به ازای هر فرد در پایگاه داده، آموزش داده می‌شود.

- یکی از چالش‌های مورد بررسی در این پایان‌نامه، حجیم بودن پایگاه داده است.
- از پایگاه داده FERET، که حاوی یک تصویر چهره از ۹۹۰ افراد مختلف می‌باشد، استفاده شده است. در فصل چهارم به معرفی این پایگاه داده خواهیم پرداخت.



شکل (۱-۲): نمونه تصاویر پایگاه داده FERET

## ۱-۶- ساختار پایان‌نامه

این پایان‌نامه در پنج فصل تدوین شده است. در ادامه این پایان‌نامه، در فصل دوم، به بررسی اجمالی برخی از کارهای پیشین که در حوزه شناسایی چهره انجام شده، پرداخته می‌شود. در فصل سوم، روش پیشنهادی این تحقیق، به منظور استخراج ویژگی از تصاویر چهره ارائه می‌شود. فصل چهارم، آزمایشات انجام شده با روش پیشنهادی، مورد بررسی قرار می‌گیرد و با روش‌های پیشین از نظر سرعت و دقت

بازیابی تصویر مقایسه می‌شوند. در نهایت، در فصل پنجم، نتیجه‌گیری و پیشنهادهایی برای ادامه کارهای آینده نیز مطرح می‌شود.

## ۱-۷- جمع‌بندی

در این فصل، ابتدا اهمیت و کاربرد سیستم شناسایی چهره را معرفی و بعد به چالش‌ها و مشکلاتی پرداختیم، که این سیستم با آن مواجه است. سپس موضوع و هدف پایان‌نامه شرح داده شد؛ پیش‌فرض‌های روش پیشنهادی بیان شدند و در نهایت ساختار اصلی پایان‌نامه معرفی گردید.

# فصل دوم

## مروری بر کارهای پیشین

## ۲-۱- مقدمه

با توجه به اهمیت شناسایی چهره، محققین زیادی در این زمینه فعالیت کرده و بسته به چالش‌های موجود، روش‌های مختلفی را ارائه داده‌اند. بیشتر روش‌های موجود، همگی به نوعی از چهره، ویژگی استخراج می‌کنند. همان‌طور که اشاره شد، این روش‌ها بر اساس ویژگی‌های سراسری، ویژگی‌های محلی و ویژگی‌های ترکیبی به سه دسته تقسیم می‌گردند. در این فصل ابتدا، روش‌های مربوط به هر دسته را به صورت اجمالی معرفی می‌کنیم. سپس در ادامه آن، به بررسی چند روش شناسایی چهره می‌پردازیم.

## ۲-۲- روش‌های استخراج ویژگی بر اساس ویژگی‌های سراسری

در این نوع روش‌ها، کل ناحیه چهره به عنوان ورودی به سیستم شناسایی چهره داده می‌شود. سیستم با تبدیل آن به یک بردار ویژگی، عمل تشخیص هویت را انجام می‌دهد. مزیتی که این نوع روش‌ها دارند این است که، می‌توانند تمام اطلاعات بافت و شکل دقیق را که برای شناسایی چهره مفید هستند، حفظ کنند.

الگوریتم‌هایی مانند چهره‌های ویژه<sup>۱</sup> (Eigenface) [۸]، چهره فیشر<sup>۲</sup> (Fisherface) [۱۲] از این دسته هستند.

## ۲-۲-۱- چهره ویژه

تجزیه و تحلیل مؤلفه‌های اصلی<sup>۳</sup> PCA یکی از ابتدایی‌ترین روش‌های مورد استفاده برای شناسایی چهره است [۱۱]. بر اساس تعریف ارائه شده از PCA، فرض بر این است که ماتریس  $X$  از بردارهای  $\{x_1, x_2, \dots, x_n\}$  تشکیل شده است که هر کدام به صورت ستونی در ماتریس قرار داده شده‌اند. بنابراین با توجه به ابعاد بردارها ( $M$ ) ماتریس داده‌ها به صورت  $M \times N$  است. این روش با محاسبه میانگین و ماتریس کواریانس تصاویر، فضای ویژگی را از تصاویر استخراج می‌کند.

---

<sup>۱</sup> Eigenface

<sup>۲</sup> Fisherface

<sup>۳</sup> Principal Component Analysis



جزئیات الگوریتم PCA به شرح زیر است:

هر تصویر چهره به یک بردار تبدیل می‌شود. به این صورت که  $X_{M \times N} = \{x_1, x_2, \dots, x_N\}$  هر کدام از  $x_i$ ها یک تصویر چهره و  $N$  تعداد تصاویر می‌باشد.

۱. با استفاده از رابطه (۱-۲) میانگین  $\mu$  محاسبه می‌شود.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (1-2)$$

۲. با استفاده از رابطه (۲-۲) ماتریس کواریانس  $S_{M \times M}$  به دست می‌آید.

$$S = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \quad (2-2)$$

۳. با حل رابطه (۳-۲) مقادیر ویژه  $\lambda_i$  و بردارهای ویژه  $v_i$  محاسبه می‌شوند. طول هر یک از بردارهای ویژه  $M \times 1$  است.

$$Sv_i = \lambda_i v_i \quad i = 1, 2, \dots, M \quad (3-2)$$

۴. بر اساس مقدار ویژه‌های متناظر، بردارهای ویژه به صورت نزولی مرتب شده و در ماتریس  $W_{M \times M}$  ذخیره می‌شوند. سپس  $k$  بردار اول  $W = \{v_1, v_2, \dots, v_k\}$  برداشته می‌شود. هر یک از ستون‌های ماتریس  $W$ ، یک تصویر را نشان می‌دهد که به آن چهره ویژه (Eigenface) می‌گویند.

به این ترتیب می‌توان هر تصویر از یک مجموعه تصویر را بر اساس ضرایبی از چهره ویژه‌های آن مجموعه نوشت. برای شناسایی چهره کافی است به جای هر چهره با ضرایب آن کار کنیم. اگر ضرایب دو چهره از نظر معیار فاصله اقلیدسی به هم نزدیک باشند، احتمالاً آن دو چهره شبیه به هم هستند. بنابراین الگوریتم شناسایی چهره، در این حالت ابتدا ماتریس  $W$  را برای تمامی نمونه‌های آموزش محاسبه می‌کند و سپس تمام تصاویر را با استفاده از این ماتریس به ضرایب مربوطه تبدیل می‌کند. ضرایب تصویر ورودی آزمون با ضرایب تصاویر آموزش مقایسه می‌شوند و تصویر نزدیک‌تر در پایگاه داده، شناسایی می‌شود.

این روش روی پایگاه داده‌های دست‌ساز<sup>۱</sup> که حاوی ۲۵۰۰ تصویر از ۱۶ نفر بود؛ که به ازای هر شخص ترکیبات سه حالت از تغییرات زاویه‌ی سر، اندازه، شرایط روشنایی و وضوح<sup>۲</sup> آزمایش شد. نرخ شناسایی برای تغییرات روشنایی ۰.۹۶٪، برای تغییرات زاویه ۰.۸۵٪ و برای تغییرات اندازه به ۰.۶۴٪ رسید. روش ارائه شده، در برابر تغییرات روشنایی مقاوم است، اما به تغییرات اندازه یا مقیاس حساس است [۸].

## ۳-۲- روش‌های استخراج ویژگی بر اساس ویژگی‌های محلی

در این نوع روش‌ها، ویژگی‌های مشخص و معنایی مانند چشم‌ها، بینی یا دهان استخراج می‌شوند و با توجه به رابطه‌ای که بین این اعضا وجود دارد، شناسایی صورت می‌گیرد. این روش‌ها، ابتدا الگوی ساختار هر عنصر محلی را توصیف می‌کنند. سپس تمام این الگوهای به‌دست آمده را به یک بردار ویژگی تبدیل می‌کنند. در نهایت هویت فرد را با توجه به بردار ویژگی به‌دست آمده تشخیص می‌دهند.

### ۳-۲-۱- الگوی دودویی محلی<sup>۳</sup> (LBP)

ایده اصلی روش LBP، به دست آوردن اطلاعات بافت تصویر است [۱۳]. در این روش ابتدا تصویر را به چند ناحیه تقسیم می‌کند تا بتواند ویژگی‌های LBP را استخراج کند. همچنین هیستوگرام هر ناحیه محاسبه می‌شود. از الحاق هیستوگرام‌های به دست آمده از تمام ناحیه‌های تصویر، بردار ویژگی ساخته می‌شود، تا تصویر چهره را توصیف کند.

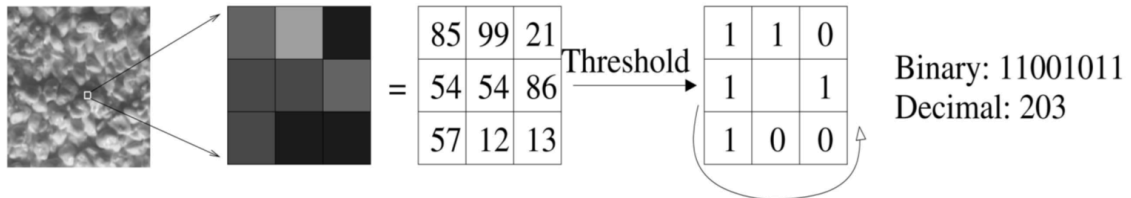
همان‌طور که در شکل (۱-۲) مشاهده می‌کنید، تمام پیکسل‌های تصویر، برچسب‌گذاری شده است. با تعریف یک مربع  $3 \times 3$ ، که پیکسل مورد نظر در مرکز آن قرار دارد؛ مقایسه بین پیکسل مورد نظر و پیکسل‌های همسایه‌ی آن، بر اساس حد آستانه تعیین شده، صورت می‌گیرد. اگر اختلاف مقدار پیکسل همسایه بزرگ‌تر یا مساوی پیکسل مرکز باشد، مقدار یک و در غیر این صورت مقدار صفر را می‌گیرد.

<sup>1</sup> handcrafted

<sup>2</sup> Resolution

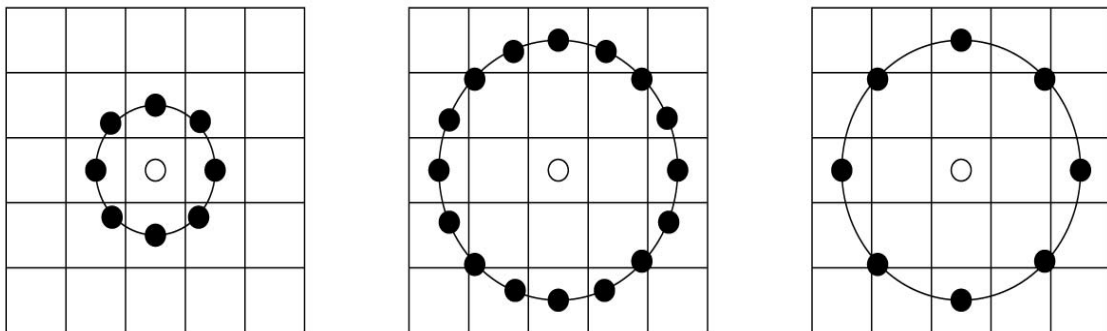
<sup>3</sup> Local Binary Pattern

نتیجه‌ی به دست آمده به باینری تبدیل می‌شود. با انجام این عمل برای تمام پیکسل‌های تصویر، تصویری که حاصل می‌شود، به آن تصویر LBP می‌گوییم.



شکل (۲-۱): عملیات الگوریتم LBP اصلی از مرجع [۱۳]

الگوریتم‌های LBP بر اساس نوع همسایگی، نحوه قرارگیری همسایه‌ها و تعداد آن‌ها دسته‌بندی می‌شوند. به عنوان مثال در شکل (۲-۲) نحوه قرارگیری همسایه‌ها به صورت دایره‌ای شکل نشان داده است.



شکل (۲-۲): عملیات الگوریتم LBP دایره‌ای [۱۳]

تصویر LBP به دست آمده به  $m$  ناحیه تقسیم می‌شود. از هر ناحیه هیستوگرام محاسبه و با کنار هم قرار دادن آن‌ها، بردار ویژگی ساخته می‌شود. در کل روش LBP بردار ویژگی‌های تصاویر موجود در پایگاه داده و بردار ویژگی تصویر آزمون را به دست می‌آورد. سپس با استفاده از کلاسه‌بند نزدیک‌ترین همسایگی<sup>۱</sup> (NN) و با معیار فاصله کای دو<sup>۲</sup> مقایسه بین بردار ویژگی آزمون با کل بردار ویژگی‌های تصاویر پایگاه داده انجام می‌شود. این روش با استفاده از پایگاه داده FERET مورد ارزیابی قرار گرفته و به نرخ بازیابی ۹۷ درصد دست یافته است.

<sup>۱</sup> Nearest Neighbor

<sup>۲</sup> Chi Square

روش ارائه شده در مرجع [۱۳] علاوه بر این که توانسته به نرخ بازیابی خوبی دست یابد، اما با پیمایش تک تک پیکسل‌های تصویر و عملیات الگوریتم خود، دارای حجم محاسباتی زیادی است.

## ۲-۴- روش‌های استخراج ویژگی ترکیبی

این نوع روش‌ها از ترکیب ویژگی‌های سراسری و ویژگی‌های محلی استفاده می‌کنند. هدف از ترکیب این ویژگی‌ها، حفظ مزیت‌ها و پوشاندن ضعف‌های یکدیگر است. در ادامه روشی که از ویژگی ترکیبی استفاده کرده است، معرفی خواهد شد.

### ۲-۴-۱- ترکیب روش Gabor Wavelet با روش تجزیه ماتریس نامنفی<sup>۱</sup> (NMF)

پورنومو و همکاران از ترکیب دو روش NMF و Gabor Wavelet برای استخراج ویژگی از چهره استفاده می‌کنند [۱۴]. روش Gabor Wavelet دارای یک تابع موجک دو بعدی است. با استفاده از هسته<sup>۲</sup> گابور استخراج ویژگی از چهره انجام می‌شود. تمام تصاویر پایگاه داده در یک ماتریس به نام  $V$  ذخیره شده و با هسته گابور پیچش<sup>۳</sup> می‌شوند. سپس تمام ویژگی‌های به دست آمده در ماتریس  $G$  ذخیره شده، و برای کاهش ابعاد از روش NMF استفاده می‌شود.

در نهایت این روش با استفاده از  $K$  نزدیک‌ترین همسایه<sup>۴</sup> (KNN) مقایسه انجام داده و شناسایی می‌کند. این مقاله از پایگاه داده ORL که شامل ۴۰۰ تصویر از ۴۰ نفر است، استفاده کرده و به دقت ۹۵ درصد دست‌یافته است.

---

<sup>1</sup> Non-negative Matrix Factorization

<sup>2</sup> Kernel

<sup>3</sup> Convolve

<sup>4</sup> K Nearest Neighbor

## ۲-۵- سایر روش ها

### ۲-۵-۱- روش ویژگی‌های مکان-فرکانسی

بخشی و همکاران به بررسی چالش پایگاه داده حجیم که از هر نمونه فقط یک تصویر وجود داشته باشد، پرداخته است [۱۵]. روش ارائه شده، با استفاده از ویژگی‌های مکان-فرکانسی و با اعمال پنجره‌های هم‌مرکز با ابعاد مختلف روی تصویر چهره، محتوای هر پنجره را به فضای فرکانسی منتقل می‌کند. تغییر مؤلفه‌های فرکانسی در پنجره‌های مختلف، فضای ویژگی تصویر را تشکیل می‌دهند. در نهایت با استفاده از معیار فاصله اقلیدسی، فاصله‌ی بین بردار ویژگی تصویر آزمون را بردارهای ویژگی تصاویر پایگاه داده به دست می‌آورد. در این روش از پیش‌پردازشی استفاده نشده است. نرخ بازیابی الگوریتم ارائه شده، با استفاده از پایگاه داده FERET، ۹۰٫۴ درصد است. زمان بازیابی این روش به ازای هر تصویر آزمون، روی سیستمی با مشخصات CPU: Intel Core i7, 2.0 GHz و Memory: 4 GB برابر با ۸۶۰ میلی ثانیه است.

روش ارائه شده در [۱۵]، با این‌که توانسته روشی نوینی را پیشنهاد دهد، اما به علت استفاده از ویژگی‌های مکان-فرکانسی، نیاز به زمان بیشتری دارد تا ویژگی‌ها را از چهره استخراج کند.

### ۲-۵-۲- ترکیبی از روش هیستوگرام گرادیان جهت‌دار<sup>۱</sup> (HOG) و تبدیل فوریه زمان

#### کوتاه<sup>۲</sup> (STFT)

نوریان و همکاران روشی مبتنی بر استخراج ویژگی در پایگاه داده حجیم را ارائه داده‌اند [۱۶]. در روش ارائه شده، به منظور پیش‌پردازش تصاویر، از الگوریتم‌های یکسان‌سازی هیستوگرام، اصلاح گاما و برابرسازی هیستوگرام وفقی با محدودیت کنتراست استفاده کرده است. اندازه تصاویر در نظر گرفته شده برابر با ۳۶۰×۳۶۰ است.

<sup>1</sup> Histogram Orientation Gradient (HOG)

<sup>2</sup> Short Time Fourier Transform (STFT)

روش ارائه شده در این مرجع، تصویر ورودی را بلوک‌بندی کرده و از هر بلوک با استفاده از الگوریتم HOG، اطلاعات گرادیان را استخراج می‌کند. سپس با چسباندن بردارهای به دست آمده به هم، بردار ویژگی ساخته می‌شود.

روش ارائه شده، با استفاده از الگوریتم STFT اطلاعات فرکانسی تصویر را به دست می‌آورد. به این ترتیب که تصویر ورودی را بلوک‌بندی کرده و از هر بلوک تبدیل فوریه می‌گیرد. سپس با چسباندن بردارهای بلوک‌های مختلف به هم، بردار ویژگی نهایی ساخته می‌شود. در نهایت با ترکیب بردارهای ویژگی به دست آمده از این دو روش، عمل شناسایی چهره را انجام می‌دهد. بر اساس نرخ بازشناسی، تعداد بلوک‌های استفاده شده در الگوریتم HOG را، ۶۴ و ۱۴۴ در نظر گرفته است. همچنین تعداد بلوک‌های STFT، ۲۵ و تعداد ویژگی‌های فرکانسی به ازای هر بلوک را ۸۱ انتخاب کرده است. در آخر با استفاده از الگوریتم K نزدیک‌ترین همسایه، نزدیک‌ترین تصویر در پایگاه داده را به تصویر آزمون پیدا می‌کند.

نرخ بازیابی این روش به ۹۵/۹۵ درصد رسیده است. نتایج ارزیابی روش ارائه شده در جدول زیر آمده است.

جدول (۱-۲): عملکرد الگوریتم پیشنهادی [۱۶] روی پایگاه داده FERET

پایگاه داده	ویژگی‌های ترکیبی	طول بردار ویژگی	نرخ بازیابی
FERET	stft81+hog64	۲۶۰۱	۹۵/۹۵ %

روش پیشنهادی در این مرجع، توانسته به نرخ بازیابی مناسبی دست یابد. اما همان‌طور که اشاره شد، اندازه تصاویر ۳۶۰×۳۶۰ در نظر گرفته شده است، و این حاکی از آن است که این روش، نیاز به تصاویر با وضوح بالا دارد.

## ۲-۵-۳- شناسایی چهره با استفاده از روش تجزیه و تحلیل خطی فیشر<sup>۱</sup> (FLDA)

از روش FLDA نمی‌توان برای چالش وجود تنها یک تصویر به ازای هر فرد استفاده کرد [۱۷]. زیرا در هر کلاس تنها یک نمونه آموزشی وجود دارد و این روش به دنبال ویژگی‌هایی است که بتواند واریانس نمونه‌های درون کلاسی را کاهش دهد. گاوو<sup>۲</sup> و همکاران با استفاده از تجزیه مقدار تکین<sup>۳</sup> (SVD)، این مشکل را حل می‌کنند.

فرض کنید تصویر چهره در ماتریسی به نام  $A \in R^{m \times n}$  ذخیره شده است که  $m \geq n$  است. SVD به صورت زیر بیان می‌شود [۱۸]:

$$A = \sum_{i=1}^n \sigma_i \cdot u_i \cdot v_i^T \quad (۴-۲)$$

که  $u_i$  و  $v_i$ ، به ترتیب نامین ستون از  $U \in R^{m \times m}$  و  $V \in R^{n \times n}$  هستند.  $U$  و  $V$  بردارهای ویژه و  $\sigma_i$  مقدارهای تکین است که به صورت نزولی مرتب شده هستند. بزرگترین مقدارهای تکین اهمیت بیشتری در تشکیل چهره دارند. لذا با استفاده از سه مقدار بزرگتر ماتریس تکین، تصویر اصلی بازسازی می‌شود. همچنین اختلاف تصویر اصلی را با تصویر بازسازی شده به دست می‌آورد.

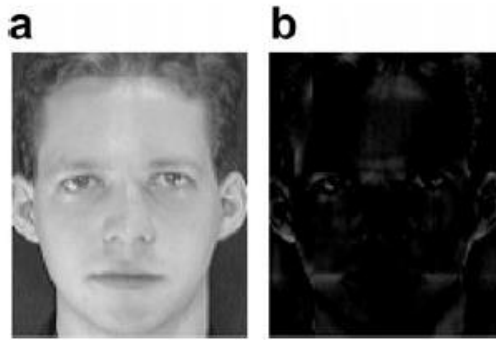
$$\varepsilon = A - \hat{A} \quad (۵-۲)$$

که  $\hat{A} = \sum_{i=1}^3 \sigma_i \cdot u_i \cdot v_i^T$  است. شکل (۳-۲)، تصویر اصلی از یک فرد و تصویر  $\varepsilon$  آن را نشان می‌دهد. با استفاده از تصاویر به دست آمده، مجموعه جدید آموزش را می‌سازد. به این ترتیب که هر کلاس دارای دو نمونه تصویر از یک فرد (یک تصویر اصلی و یک تصویر بازسازی شده) است. از این دو نمونه تصویر برای ارزیابی ماتریس پراکندگی بین کلاس‌ها استفاده می‌شود. همچنین با استفاده از تصویر  $\varepsilon$  رابطه (۵-۲) ماتریس پراکندگی درون کلاس محاسبه می‌گردد.

<sup>1</sup> Fisher Linear Discriminant Analysis

<sup>2</sup> Gao

<sup>3</sup> Singular Value Decomposition



شکل (۲-۳): (a) تصویر اصلی (b) تصویر  $\varepsilon$

در این مقاله، از چندین پایگاه داده استفاده شده، که یکی از آن‌ها پایگاه داده FERET است. با استفاده از ۴۰۰ تصویر از ۲۰۰ فرد از این پایگاه داده، روش ارائه شده مورد ارزیابی قرار گرفته است. نرخ بازیابی این روش ۹۰٫۵ درصد می‌باشد.

روش ارائه شده در این مرجع [۱۷]، توانسته است که چالش یک نمونه آموزشی به ازای هر فرد را حل کند. اما با توجه به تعداد تصاویر استفاده شده در مقایسه با روش‌های موجود، نرخ بازیابی خوبی را به دست نیآورده است.

## ۲-۵-۴- شبکه عصبی عمیق مورد استفاده با یک نمونه برای شناسایی چهره

زنگ<sup>۱</sup> و همکاران به چالش "تنها یک تصویر به ازای هر فرد" پرداخته‌اند [۱۹]. برای حل این چالش، چون روش یادگیری عمیق برای آموزش نیاز به تعداد زیادی نمونه دارد، ابتدا با استفاده از روش نمونه‌گیری در حال گسترش<sup>۲</sup> نمونه‌های آموزش را افزایش می‌دهد. به این منظور، روش نمونه‌گیری در حال گسترش، از تصاویر حالت‌های مختلف هر شخص استفاده می‌کند. فرض کنید  $m$  شخص، با  $n - 1$  حالت تصویر متفاوت از آن فرد وجود دارد. تصویر بدون حالت فرد، از تمام تصاویر حالت‌های مختلف آن فرد، کم شده و سپس برای گسترش، با نمونه‌های شخص دیگر، جمع می‌شود (شکل (۲-۴)).

<sup>1</sup> Zeng

<sup>2</sup> Expanding Sample Method





شکل (۲-۴): طرح اصلی روش نمونه‌گیری گسترش یافته

در این مقاله از یادگیری انتقال<sup>۱</sup> استفاده شده است. یادگیری انتقال از داده‌های کمکی برای یادگیری یک مدل بهره می‌گیرد، و سپس با استفاده از مدل به‌دست آمده، سناریوی دیگر را انجام می‌دهد. در این‌جا نیز به کمک شبکه عصبی کانولوشن<sup>۲</sup> (CNN) [۲۰] آموزش دیده با پایگاه داده CASIA-WebFace، مدل را به‌دست می‌آورد. به منظور دقیق‌تر تنظیم شدن شبکه، از نمونه‌های در حال گسترش استفاده شده تا مدل دقیق‌تری برای اجرای آزمایش با دیتاست اصلی به دست آید. روش ارائه شده با استفاده از پایگاه داده FERET ارزیابی می‌شود. در این آزمایش، از ۲۰۰ نفر با ۱۴۰۰ تصویر که تحت شرایط حالت‌های مختلف، زاویه‌های متفاوت و تغییرات روشنایی بوده، استفاده شده است. روش پیشنهادی به نرخ بازیابی ۹۳٫۹ درصد دست یافته است.

روشی که زنگ و همکاران ارائه داده‌اند، مشکل تک نمونه آموزشی را توانستند در یادگیری عمیق حل کنند. اما باید توجه داشت که اگر هیچ تصویری از شخص وجود نداشته باشد، روش نمونه‌گیری در حال گسترش دیگر استفاده‌ای نخواهد داشت و نمی‌توان نمونه‌های آموزشی را افزایش داد.

## ۲-۵- جمع‌بندی

در این فصل به بررسی روش‌های موجود در زمینه سیستم شناسایی چهره پرداخته شد. ویژگی‌های استخراج شده از چهره به سه دسته، سراسری، محلی و ترکیبی تقسیم شدند. چند روش از روش‌های مرسوم استخراج ویژگی در این فصل تشریح شد.

<sup>1</sup> Transfer Learning

<sup>2</sup> Convolutional Neural Network

همان‌طور که روش‌ها معرفی شدند، هرکدام مزایا و معایبی داشتند. ویژگی‌های استخراج شده از چهره، به علت ابعاد بالا، به روش‌های کاهش بعد مانند PCA و غیره نیاز دارند. الگوریتم الگوی باینری محلی، با توجه به عملیات الگوریتم خود، دارای حجم محاسباتی زیادی است. در اکثر روش‌های موجود، آزمایش را با تعداد افراد کمتر از ۲۰۰ نفر انجام داده‌اند، و یا این‌که از هر شخص چندین تصویر موجود بوده، که همین امر باعث می‌شود سیستم عملکرد بهتری را داشته باشد.

با توجه به مطالب ارائه شده، ما به دنبال روشی هستیم که بتواند ویژگی‌های مناسبی را استخراج کند، که در پایگاه داده‌های حجیم که به ازای هر فرد تنها یک تصویر موجود است، بتواند به نرخ بازیابی خوبی دست یابد. به عبارت دیگر، بتوان فاصله‌ی بردارهای ویژگی بین افراد مختلف را حداکثر کرد. علاوه بر این اگر بردار ویژگی استخراج شده نیز کم حجم باشد، در سرعت بازیابی می‌تواند تأثیر بسزایی نیز داشته باشد. در فصل بعد به معرفی روش پیشنهادی، استخراج ویژگی و معیار شباهت می‌پردازیم.

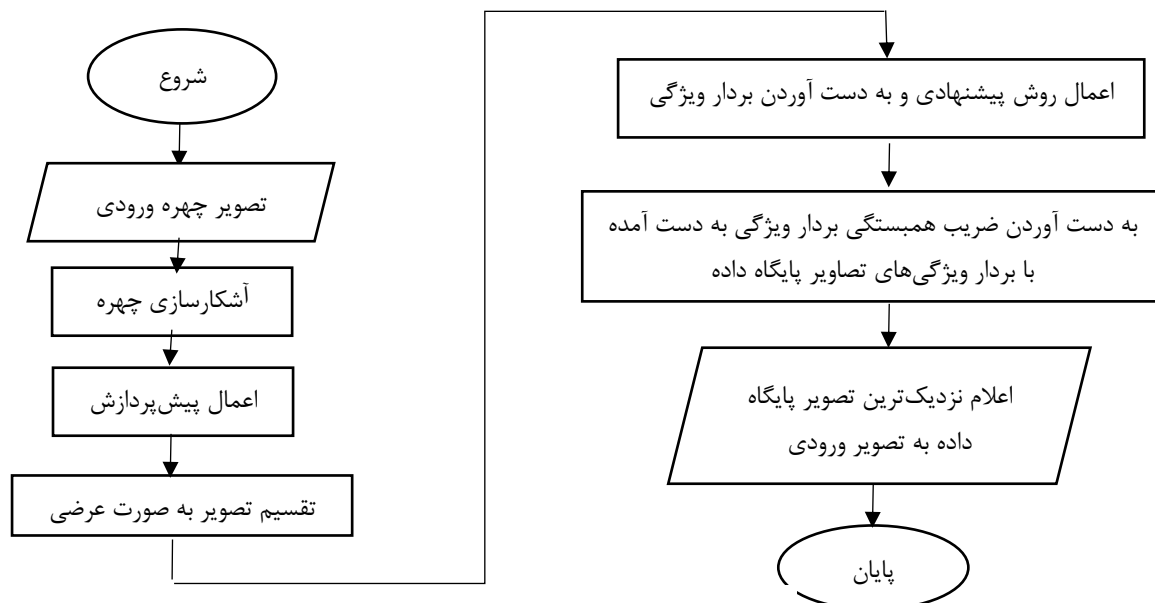
فصل سوم

روش پیمایشی

### ۳-۱- مقدمه

همان‌طور که اشاره شد، تمرکز اصلی این پایان‌نامه، روی سه چالش حجیم بودن پایگاه داده تصاویر چهره، تعداد کم نمونه‌های آموزش (موجود بودن تنها یک تصویر به ازای هر نفر) و سرعت بازیابی تصویر است. هدف اصلی این تحقیق، ارائه روشی بر مبنای استخراج ویژگی است، که بتواند از عهده مشکلات حجیم بودن پایگاه داده، تعداد نمونه‌های کم آموزش و در عین حال بالا بودن سرعت بازیابی سیستم برآید.

همان‌طور که در فصل یک اشاره شد، سیستم بینایی انسان، به‌گونه‌ای است که می‌تواند در جمعیتی انبوه، شخص مورد نظر را به سرعت بازیابی کند. لذا بشر سعی دارد که سیستم‌های شناسایی چهره‌ای را ارائه دهد، که به سیستم بینایی انسان نزدیک باشند. روش پیشنهادی در این پایان‌نامه نیز، تلاش می‌کند تا ویژگی‌هایی را که با مفاهیم شهودی بهتر مطابقت دارد؛ استخراج کند. در این تحقیق، روش پیشنهادی بردارهای ویژگی هر تصویر چهره را استخراج می‌کند. سپس با استفاده از معیار شباهت مناسب، بردار ویژگی تصویر ورودی با بردارهای ویژگی تصاویر پایگاه داده مقایسه شده و هویت شخص شناسایی می‌شود. شکل زیر، فلوچارت فرآیند کلی الگوریتم ارائه شده را نشان می‌دهد. در ادامه روش پیشنهادی را به‌طور مفصل توضیح خواهیم داد.



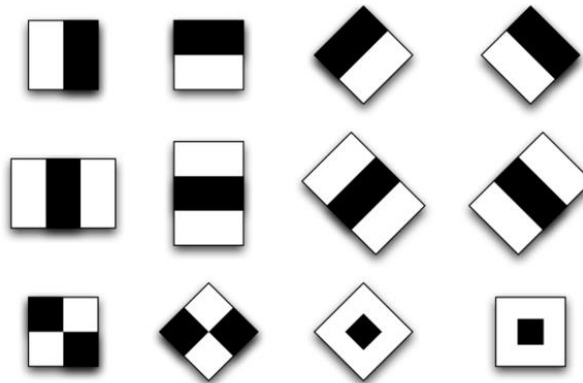
شکل (۳-۱): فرآیند کلی الگوریتم ارائه شده

## ۲-۳- آشکارسازی چهره

برای آشکارسازی چهره از الگوریتم معروف Viola-Jones استفاده شده است. این الگوریتم به طور مفصل در مرجع [۵] بیان شده است. در این بخش تنها به طور خلاصه، این روش مرور خواهد شد. این الگوریتم حاوی دو بخش تصویر انتگرال و طبقه‌بندی آبخاری<sup>۱</sup> مبتنی بر الگوریتم آدابوست<sup>۲</sup> است.

### ۱-۲-۳ تصویر انتگرال

از تصویر ورودی، ویژگی‌های هار<sup>۳</sup> استخراج می‌شوند. این ویژگی‌ها در واقع مستطیل‌هایی هستند که نشان‌دهنده بخش‌های مختلف صورت است و روی تصویر با روند تکراری قرار می‌گیرند (شکل ۲-۳). ویژگی‌های مستطیلی را می‌توان بسیار سریع محاسبه کرد، که آن را تصویر انتگرال می‌نامند [۵].



شکل (۲-۳): نمونه‌هایی از مستطیل‌های استفاده شده در الگوریتم Viola-Jones

ساخت تصویر انتگرال در معادله (۱-۳) آمده است.

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \quad (1-3)$$

<sup>1</sup> Cascade

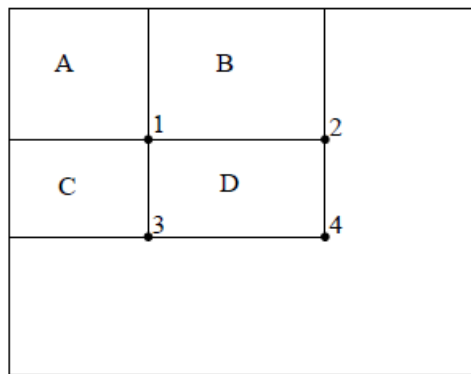
<sup>2</sup> Ada Boost

<sup>3</sup> Haar Feature

تصویر انتگرال  $ii(x,y)$  و تصویر اصلی  $i(x,y)$  است. به عنوان مثال شکل (۳-۳)، مجموع مقادیر مستطیل ABCD، به کمک تصویر انتگرال با معادله (۲-۳) محاسبه می‌شود. در ادامه با یک مثال، رابطه زیر توضیح داده می‌شود.

$$\sum_{(x,y) \in ABCD} i(x,y) = ii(D) + ii(A) - ii(B) - ii(C) \quad (2-3)$$

محاسبه ویژگی‌های هار، به صورت تفاوت روشنایی بین دو یا چهار مستطیل انجام می‌شود.



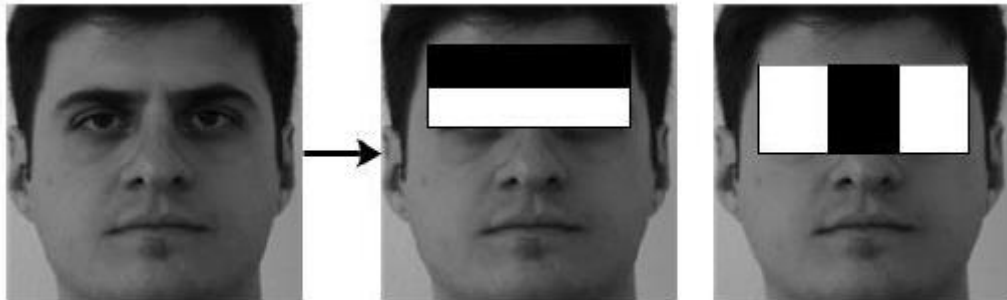
شکل (۳-۳): محاسبه مجموع مقادیر پیکسلی یک ویژگی مستطیلی، با استفاده از چهار آرایه در تصویر انتگرال در [۵]  
 شکل (۳-۳) مجموع مقادیر پیکسلی یک ویژگی مستطیلی را با استفاده از چهار آرایه در تصویر انتگرال محاسبه می‌کند. مقدار تصویر انتگرال در موقعیت ۱ برابر با مجموع مقادیر پیکسلی ناحیه A می‌باشد. این مقدار در موقعیت ۲ برابر با  $A + B$  است و در موقعیت ۳ و ۴ به ترتیب برابر با  $A + C$  و  $A + B + C + D$  می‌باشد. مجموع مقادیر پیکسلی در ناحیه D می‌تواند این‌گونه محاسبه شود:  $4 + 1 - (2 + 3)$

این الگوریتم از یک پنجره  $n \times n$  برای طی مراحل قرار دادن این مستطیل‌ها روی صورت استفاده می‌کند. با در نظر گرفتن تعداد و سایزهای مختلف این مستطیل‌ها، به ازای هر تصویر، نیاز به محاسبات و زمان زیادی است. از این‌رو از الگوریتم آدابوست استفاده می‌کند.

### ۲-۲-۳ طبقه‌بندی آبخاری مبتنی بر الگوریتم آدابوست

آدابوست یک الگوریتم یادگیری ماشین است و هدف آن، انتخاب مجموعه‌ای از بهترین طبقه‌بندهای ضعیف (که عملکرد خوبی ندارند) و تشکیل یک طبقه‌بند قوی است. ویژگی‌های انتخاب شده توسط

آدابوست، معنی‌دار و قابل تفسیر هستند. شکل (۳-۴) چند مورد از ویژگی‌های انتخاب شده توسط این الگوریتم را نشان می‌دهد. یکی از ویژگی‌های انتخاب شده، اختلاف ناحیه چشم‌ها و پیشانی و ویژگی دوم فاصله بین چشم‌ها است.



شکل (۳-۴) : برخی از ویژگی‌های انتخاب شده توسط آدابوست در [۵]

در نهایت مجموعه‌ای از طبقه‌بندی‌های قوی، سیستم نهایی را تشکیل می‌دهند.



(ب) ناحیه آشکارسازی چهره

(الف) تصویر اصلی

شکل (۳-۵): نمونه‌ای از استخراج ناحیه چهره از تصویر ورودی با استفاده از الگوریتم Viola-Jones

### ۳-۳- پیش پردازش

به منظور بالا بردن نرخ بازیابی، می‌توان در صورت نیاز، عمل پیش‌پردازش را روی تصاویر اعمال کرد. در این جا نیز برای بالا بردن کیفیت و بهبود روشنایی تصاویر چهره از چند پیش‌پردازش استفاده شده است.

در ادامه به معرفی چند پیش پردازش استفاده شده در این پایان نامه می پردازیم. سپس روش پیشنهادی استخراج ویژگی را توضیح خواهیم داد و در نهایت معیار شباهت مناسب را معرفی خواهیم کرد. قبل از استخراج ویژگی، ممکن است که تصاویر نیاز به پیش پردازش داشته باشند. در شکل زیر چند نمونه از تصاویر پایگاه داده FERET آورده شده است. ما می خواهیم با اعمال پیش پردازش، به کیفیت روشنایی مطلوبی برسیم.



شکل (۳-۶): نمونه‌ای از تصاویر پایگاه داده

با اعمال پیش پردازش روی تصاویر، این امکان به وجود می آید که اطلاعات از تصاویر، دقیق تر و صحیح تر استخراج شوند و همین امر در بهبود نرخ بازیابی مؤثر است. از پیش پردازش‌های تصویر می توان به اصلاح گاما [۲۲]، یکسان سازی هیستوگرام<sup>۱</sup>، تنظیم شدت<sup>۲</sup> روشنایی، برابری هیستوگرام<sup>۳</sup> و محدودیت کنتراست<sup>۳</sup> (clahe) اشاره کرد. در این تحقیق، به منظور بهبود روشنایی از روش یکسان سازی هیستوگرام، برابری هیستوگرام<sup>۳</sup> و محدودیت کنتراست (clahe) و تنظیم شدت تصویر استفاده شده است. با اعمال این سه روش، عملکرد بازشناسی چهره بهبود یافت. نتایج اعمال پیش پردازش در فصل چهارم توضیح داده خواهد شد.

### ۳-۳-۱ یکسان سازی هیستوگرام

یکسان سازی هیستوگرام یکی از روش‌های بهسازی تصویر است، که در آن بازه پویای تصویر گسترش می یابد [۲۳]. بازه پویای تصویر، اختلاف روشن ترین سطح خاکستری یک نقطه از تصویر با تاریک ترین

<sup>1</sup> Histogram equalization

<sup>2</sup> Image Intensity

<sup>3</sup> Contrast Limited Adaptive Histogram



نقطه در تصویر است. این روش با استفاده از یک تابع تبدیل صعودی، مقادیر پیکسل‌های تصویر را به گونه‌ای تغییر می‌دهد که هیستوگرام تصویر به دست آمده، تا حد امکان یکنواخت باشد.

از تابع توزیع تجمعی<sup>۱</sup> مطابق با رابطه (۳-۳) در یکسان‌سازی هیستوگرام استفاده می‌شود:

$$p_k = T(q_k) = (L - 1) \sum_{i=0}^k d_q(q_i) = \frac{(L-1)}{N} \sum_{i=0}^k n_i \quad k = 0, 1, 2, \dots, L - 1 \quad (3-3)$$

در این رابطه،  $q$  تصویر ورودی،  $p$  تصویر حاصل،  $d_q(q_i)$  مقدار نام هیستوگرام نرمال شده تصویر  $q$ ،  $n_i$  تعداد پیکسل‌های با سطح خاکستری  $i$ ،  $N$  تعداد کل پیکسل‌های تصویر،  $L$  بیشینه سطح خاکستری و  $k$  مقدار سطح خاکستری در بازه  $[0, L-1]$  است. در انتها، مقادیر  $p$  باید به مقادیر صحیح در بازه  $[0, L-1]$  گرد<sup>۲</sup> شوند [۲۳].

روش یکسان‌سازی هیستوگرام، یک روش سراسری است، لذا تمام نقاط تصویر را به یک نسبت بهبود می‌دهد. از این رو ممکن است باعث واضح‌تر شدن نویز در نقاط نویزی تصویر شود [۲۴]. همچنین ممکن است باعث رنگ پریدگی<sup>۳</sup> در بعضی قسمت‌های تصویر شود [۲۵]. شکل (۷-۳) نمونه‌ای از نتیجه یکسان‌سازی هیستوگرام را نشان می‌دهد. همان‌طور که مشاهده می‌کنید، تصویر (ب) نسبت به تصویر (الف) از کیفیت بهتری برخوردار شده است.



(ب) تصویر پیش‌پردازش شده

(الف) تصویر اصلی

شکل (۷-۳) : تصویر اصلی و تصویر با پیش‌پردازش یکسان‌سازی هیستوگرام

<sup>1</sup> Cumulative density function

<sup>2</sup> Round

<sup>3</sup> Wash out

### ۲-۳-۳ برابر سازی هیستوگرام وفقی با محدودیت کنتراست

این روش تصویر را به ناحیه‌هایی کاشی‌وار تقسیم‌بندی می‌کند. روش قبل، همه نواحی تصویر را به یک میزان بهسازی می‌کند، اما در این روش به صورت تطبیقی روی نواحی تصویر انجام می‌شود [۲۶]. در این پیش‌پردازش، بهسازی کمتری در نواحی هموار انجام می‌شود، از این‌رو مانع از تقویت نویز و اثر رنگ‌پریدگی در تصویر می‌گردد [۲۳]. در شکل (۳-۸) نمونه‌ای از انجام این پیش‌پردازش را مشاهده می‌کنید.



ب) تصویر پیش‌پردازش شده



الف) تصویر اصلی

شکل (۳-۸): تصویر اصلی و تصویر با پیش‌پردازش clahe

### ۳-۳-۳ تنظیم شدت روشنایی تصویر

با استفاده از این پیش‌پردازش می‌توان دامنه تغییرات روشنایی یک تصویر را تغییر داد [۲۷]. می‌توان از معادله تبدیل (۳-۴) برای این پیش‌پردازش استفاده کرد.

$$A(i,j) = I * ((max - min)/255) \quad (۳-۴)$$

$I$  تصویر و  $max$  و  $min$  به ترتیب بیشینه مقدار پیکسل‌ها و کمینه‌ی آن‌ها است. شکل (۳-۹) نمونه‌ای از انجام این پیش‌پردازش را نشان می‌دهد. همان‌طور که مشاهده می‌کنید، تصویر (ب) با کیفیت مطلوب‌تری نسبت به تصویر ورودی (الف) است.



شکل (۳-۹): تصویر اصلی و تصویر با پیش پردازش تنظیم شدت روشنایی تصویر

### ۳-۴- استخراج ویژگی

استخراج ویژگی یکی از مهم ترین مراحل بازشناسی چهره محسوب می شود، و تمرکز اصلی این پایان نامه معطوف به این قسمت است. یکی از روش های استخراج ویژگی، تجزیه ماتریس است، که می توان به روش هایی مانند تحلیل مؤلفه اصلی (PCA)، کوانتیزاسیون برداری<sup>۱</sup> (VQ)، تحلیل مؤلفه مستقل<sup>۲</sup> (ICA) و تجزیه نامنفی ماتریس (NMF) اشاره کرد [۲۱].

### ۳-۴-۱ تجزیه ماتریس نامنفی

داده های اطلاعاتی به صورت خام، معمولاً بعد بالایی دارند، و همین اغلب باعث افزایش هزینه محاسباتی و کاهش کارایی سیستم ها می شود. ما به دنبال این هستیم که داده ها را بتوان به یک فضای جدید با ابعاد کمتر، و در عین حال بار اطلاعاتی غنی تر نگاشت کنیم [۲۸].

روش تجزیه ماتریس نامنفی به گروهی از الگوریتم ها در شاخه جبر خطی گفته می شود که یک ماتریس را به دو ماتریس به نام ماتریس های عامل<sup>۳</sup> تجزیه می کنند، با این شرط که هر سه ماتریس فاقد مقادیر منفی باشد. این نامنفی بودن باعث می شود نتایج کار قابل تفسیرتر شود. حفظ نامنفی بودن در تحلیل

<sup>1</sup> Vector Quantization

<sup>2</sup> Independent Component Analysis

<sup>3</sup> Factor

اطلاعات، خواص اصلی داده‌ها را حفظ می‌کند. به عنوان مثال، نامنفی بودن عناصر ماتریس  $W$ ، باعث می‌شود که ستون‌های پایه قابل تجسم باشند [۲۹ و ۳۰].

از ساده‌ترین اصولی که در ادراک انسان وجود دارد، درک ساختمان اشیاء بر اساس اجزای آن است و بینایی انسان این قابلیت را دارد که وجود یا عدم وجود یک شی را به سادگی تشخیص می‌دهد [۲۸]. از این‌رو روش NMF، با فرض این‌که انسان اشیاء را بر اساس جزئیاتشان یاد می‌گیرد، الهام گرفته شده است و به همین علت کاربردهای بسیاری در حوزه‌ی پردازش تصویر دارد [۳۱ و ۳۲].

فرض کنید، تصاویر پایگاه داده در یک ماتریس  $m \times n$  به نام  $V$  ذخیره می‌شوند. ستون‌های این ماتریس شامل  $m$  عنصر نامنفی از هر تصویر است. ماتریس  $V$  به دو ماتریس  $W$  و  $H$  تجزیه می‌شود. معادله (۵-۳) نشان می‌دهد که هر ستون از داده‌های ماتریس  $V$  را می‌توان به صورت ترکیب خطی از  $r$  ستون ماتریس  $W$  نوشت:

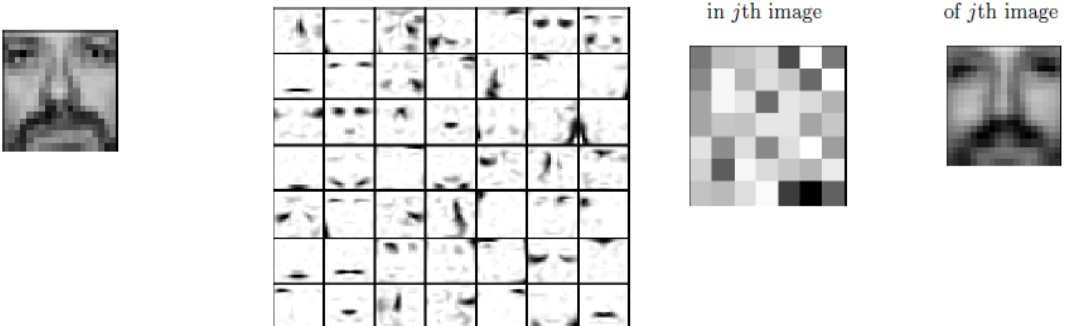
$$V_{mn} \cong (WH)_{mn} = \sum_{a=1}^r W_{ma} H_{an} \quad (۵-۳)$$

ضرایب این ترکیب خطی مؤلفه‌های بردار  $H$  هستند. لذا به ماتریس  $W$ ، ماتریس پایه و به ستون‌های آن، بردارهای پایه‌ای می‌گویند. همچنین ماتریس  $H$ ، ماتریس وزن خواهد بود؛ که رابطه یک به یک با هر یک از چهره‌ها در ماتریس  $V$  دارد [۳۳ و ۳۴]. ضرب ماتریس  $WH$  یک تقریب برای تجزیه ماتریس  $V$  است.  $r$  مرتبه‌ی روش NMF است. مرتبه‌ی یک ماتریس، برابر با حداکثر تعداد بردارهای مستقل خطی در یک ماتریس است. مجموعه‌ای از بردارها زمانی دارای وابستگی خطی خواهند بود، که بتوان حداقل یکی از بردارها را به صورت ترکیب خطی باقی بردارهای مجموعه بیان کرد. از معادله زیر برای انتخاب مقدار مرتبه  $r$  استفاده می‌شود [۳۳]. تعداد ویژگی‌ها را با مرتبه ( $r$ ) می‌توان تعیین نمود.

$$(n + m)r < nm \quad (۶-۳)$$

$m$  تعداد پیکسل‌های کل تصاویر و  $n$  تعداد تصاویر پایگاه داده را نشان می‌دهد. مقدار انتخابی  $r$  بسیار حائز اهمیت است. زیرا اگر این مقدار، مناسب انتخاب نشود، ممکن است که دو چهره متفاوت را یکسان شناسایی کند. لذا با توجه به معادله (۶-۳)، می‌توان گفت که مقدار  $r$  وابسته به تعداد تصاویر است.

پارامتر  $r$  به طور کلی با آزمون و خطا تنظیم می‌شود؛ که اگر ماتریس  $V$  بسیار بزرگ باشد، این فرآیند می‌تواند وقت‌گیر باشد [۳۳]. شکل (۳-۱۰)، معادله (۳-۵) را بیان می‌کند.

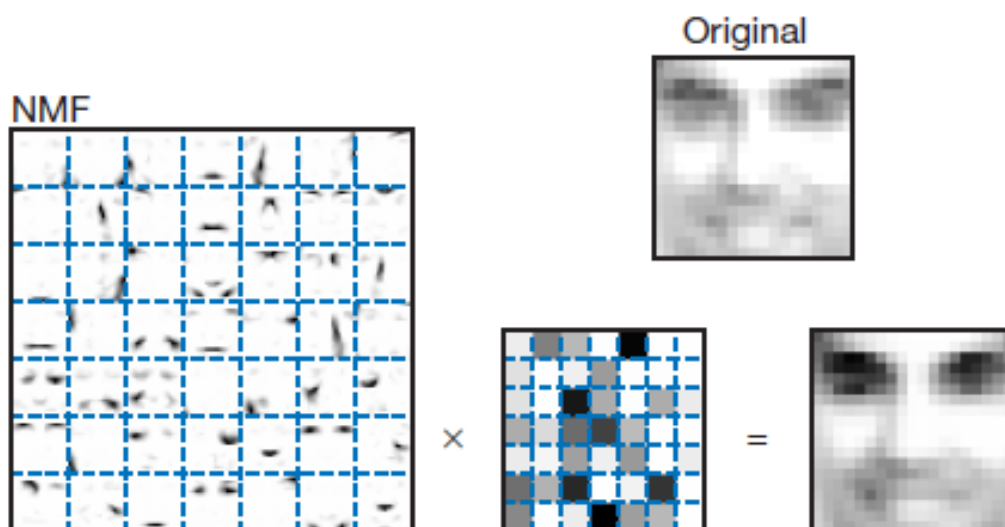
$$\underbrace{V(:,j)}_{j\text{th facial image}} \approx \sum_{k=1}^r \underbrace{W(:,k)}_{\text{facial features}} \underbrace{H(k,j)}_{\text{importance of features in } j\text{th image}} = \underbrace{WH(:,j)}_{\text{approximation of } j\text{th image}}$$


شکل (۳-۱۰): نمونه‌ای از بیان معادله (۳-۵)؛ آزمین تصویر در ماتریس پایگاه داده، که با مجموع حاصلضرب ماتریس  $W$  و  $H$  تخمین زده شده [۳۵]

این نکته قابل ذکر است که تمام عناصر دو ماتریس  $W$  و  $H$  غیرمنفی هستند. لذا محدودیت‌های نامنفی پذیری این روش، با مفهوم شهودی (با ترکیب قطعات به دست آمده، یک شکل کلی ساخته می‌شود)، سازگاری دارد. این بدان معنا است که روش NMF ویژگی‌های مبتنی بر جزء را یاد می‌گیرد [۳۳]. همان‌طور که در شکل (۳-۱۱) مشاهده می‌کنید، تصاویر پایه دارای چندین نسخه از دهان، بینی و دیگر قسمت‌های صورت هستند که نسخه‌های مختلف آن‌ها در مکان و یا شکل‌های متفاوت قرار دارند. با ترکیب این قطعه‌ها از یک چهره، تغییرهای مختلف از آن چهره تولید می‌شود. البته این نکته قابل ذکر است که هر چهره از تمام قطعه‌های موجود استفاده نمی‌کند [۳۳].

تابع هدف در روش تجزیه ماتریس نامنفی، نرْم Frobenius است که در رابطه زیر بیان شده است [۳۶].

$$\min f(W,H) = \frac{1}{2} \|V - WH\|_F^2 \quad (۳-۷)$$



شکل (۳-۱۱): نمایش تجزیه یک چهره با استفاده از روش NMF [۳۳]

تابع هدف (۳-۷)، هم‌زمان نسبت به دو عامل ماتریسی  $W$  و  $H$  نامحدب است. لذا کمینه‌ی سراسری را نمی‌توان یافت. در بهترین حالت، تنها رسیدن به کمینه محلی انتظار می‌رود [۳۷]. از این‌رو همگرایی یکی از هدف‌های روش NMF است. لذا الگوریتم‌های پیشنهادی برای NMF را در سه گروه الگوریتم قواعد به‌روزرسانی ضربی<sup>۱</sup>، الگوریتم کاهش گرادیان<sup>۲</sup> و الگوریتم بر پایه روش کمترین مربعات متناوب<sup>۳</sup> طبقه‌بندی کرد که به اختصار بررسی می‌شوند [۳۸ و ۳۹].

الگوریتم به‌روزرسانی ضربی توسط لی و سانگ معرفی شد [۳۷]، که تابع هدف (۳-۷) را با قواعد ضربی به‌روزرسانی می‌کند شکل (۳-۱۲).

<sup>۱</sup> Multiplicative Update rules

<sup>۲</sup> Gradient Descent

<sup>۳</sup> Alternating Least Square

```

MULTIPLICATIVE UPDATE ALGORITHM FOR NMF
W = rand(m, k); % initialize W as random dense matrix
H = rand(k, n); % initialize H as random dense matrix
for i = 1 : maxiter
(MU) H = H .* (WTA)./(WTWH + 10-9);
(MU) W = W .* (AHT)./(WHHT + 10-9);
end

```

شکل (۳-۱۲): الگوریتم قواعد به‌روزرسانی ضربی [۳۹]

سرعت همگرایی این روش پایین است، زیرا برای همگرایی نیاز به تعداد حلقه‌های تکرار و محاسبات زیاد در هر حلقه دارد.

شکل (۳-۱۳)، الگوریتم کاهش گرادیان را نشان می‌دهد که نحوه به‌روزرسانی ماتریس‌ها را بیان می‌کند.  $\mathcal{E}_W$  و  $\mathcal{E}_H$  پارامترهای اندازه گام<sup>۱</sup> هستند. انتخاب مقدار این پارامترها در همگرایی تابع هدف، اهمیت دارد. روش‌هایی مانند قواعد به‌روزرسانی ضربی که پیش‌تر بیان شد، این پارامترها را به‌طور هوشمندانه انتخاب می‌کنند، اما همان‌طور که بیان شد، نیاز به محاسبات زیادی در حلقه دارد که سرعت همگرایی را کاهش می‌دهد.

```

BASIC GRADIENT DESCENT ALGORITHM FOR NMF
W = rand(m, k); % initialize W
H = rand(k, n); % initialize H
for i = 1 : maxiter
H = H - εH ∂f / ∂H
W = W - εW ∂f / ∂W
end

```

شکل (۳-۱۳): الگوریتم کاهش گرادیان [۳۹]

در شکل (۳-۱۴)، الگوریتم کمترین مربعات متناوب را نشان می‌دهد. در پیش‌تر اشاره شد که به علت هم‌زمان در نظر گرفتن ماتریس  $W$  و  $H$ ، بهینه‌سازی نامحدب است. لذا در این الگوریتم، یک‌بار به

<sup>1</sup> Step size

عنوان مثال ماتریس  $W$  را ثابت در نظر می‌گیرد و ماتریس  $H$  را با استفاده از روش محاسبه کمترین مربعات به دست می‌آورد. بار دیگر ماتریس  $H$  را ثابت و ماتریس  $W$  را با این روش به دست می‌آورد.

```

BASIC ALS ALGORITHM FOR NMF
W = rand(m, k); % initialize W as random dense matrix or use another

for i = 1: maxiter
    (LS) Solve for H in matrix equation WTWH = WTA.
    (NONNEG) Set all negative elements in H to 0.
    (LS) Solve for W in matrix equation HHTWT = HAT.
    (NONNEG) Set all negative elements in W to 0.
end
    
```

شکل (۳-۱۴): الگوریتم کمترین مربعات متناوب [۳۹]

در هر مرحله الگوریتم، مقادیر منفی حاصل از جواب کمترین مربعات را با صفر جایگزین می‌کند، به این دلیل الگوریتم کمترین مربعات نامنفی متناوب<sup>۱</sup> (ANLS) نامیده می‌شود. این الگوریتم، ضمن برقراری شرط نامنفی بودن ماتریس‌های تولید شده، به خلوت<sup>۲</sup> (تنکی) آن‌ها نیز کمک می‌کند. در مورد همگرایی روش‌های کمترین مربعات متناوب، اثبات شده است که هر نقطه حدی به دست آمده  $W$  و  $H$ ، یک نقطه ایستا<sup>۳</sup> است. نقطه ایستا، نقطه حدی است که مشتق تابع در آن به صفر میل می‌کند [۳۹].

در این تحقیق از الگوریتم چرخش بلوک اصلی<sup>۴</sup> (BPP) برای حل مسأله کمترین مربعات نامنفی متناوب استفاده شده است [۳۶]. این الگوریتم هم از جنبه‌ی نرخ بازیابی و هم از جنبه‌ی زمان بازیابی، از سایرین بهتر عمل کرده است.

یک تقریب اولیه خوب، می‌تواند به کاهش تعداد تکرارها در الگوریتم‌های NMF کمک کند. از آن جایی که هیچ چیزی درباره توزیع مینیمم مشخص نیست، در این پایان‌نامه، مقداردهی اولیه تصادفی انجام

<sup>1</sup> Alternating Non-negative Least Squares

<sup>2</sup> Sparse

<sup>3</sup> Stationary Point

<sup>4</sup> Block Principal Pivoting



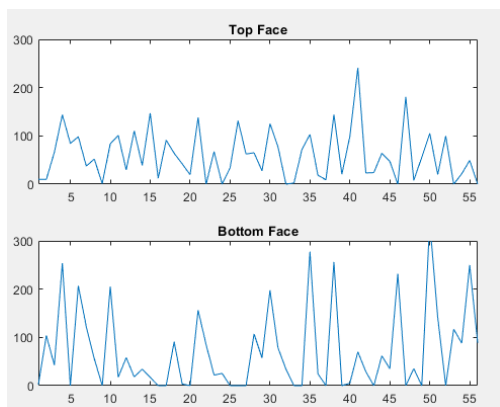
می‌شود. در واقع در بسیاری از کاربردهای موجود، مقداردهی اولیه تصادفی، به خاطر سادگی آن به کار می‌رود.

جدول (۳-۱): مقایسه روش‌های NMF

نرخ بازیابی (درصد)	روش
۸۹٫۲۹	روش NMF با قواعد به‌روز رسانی ضربی
۸۶٫۵۶	روش NMF با الگوریتم کاهش گرادیان [۴۰]
۸۵٫۷۵	روش NMF با ALS پایه
۹۲٫۷۳	روش NMF با ANLS-BBP (روش پیشنهادی)

یک تقریب اولیه خوب، می‌تواند به کاهش تعداد تکرارها در الگوریتم‌های NMF کمک کند. از آن جایی که هیچ چیزی درباره توزیع مینیمم مشخص نیست، در این پایان‌نامه، مقداردهی اولیه تصادفی انجام می‌شود. در واقع در بسیاری از کاربردهای موجود، مقداردهی اولیه تصادفی، به خاطر سادگی آن به کار می‌رود.

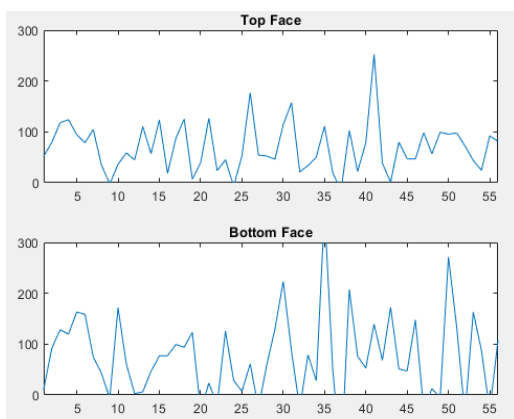
نوآوری که در این تحقیق انجام شده این بوده است که، تصاویر چهره به صورت عرضی نصف می‌شوند تا اجزای پایه هر ناحیه دقیق‌تر تشکیل شود. سپس روی این دو قسمت به‌دست آمده، روش NMF به صورت جداگانه اعمال می‌گردد و دو ماتریس پایه و وزن را حاصل می‌دهد. ماتریس وزن همان ماتریس فضای ویژگی است، یعنی هر ستون از آن، بردار ویژگی متناظر با یک تصویر از پایگاه داده مورد بررسی می‌باشد. شکل (۳-۱۵) بردارهای ویژگی متناظر با نیمه بالا و نیمه پایین تصویر مورد نظر با دو تصویر نمونه از پایگاه داده مورد بررسی را نشان می‌دهد. در این شکل، طول بردار ویژگی (یا مقدار  $r$ ) هر کدام از ناحیه‌ها (نیمه‌ی بالا و نیمه‌ی پایین) برابر با ۵۶ است. در فصل بعد در مورد مقدار  $r$  بیشتر توضیح خواهیم داد.



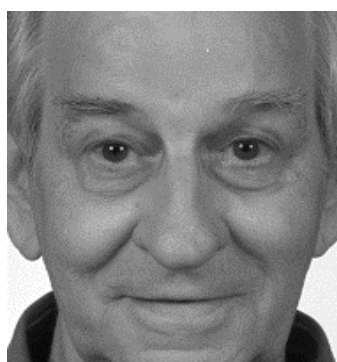
بردار ویژگی مناظر



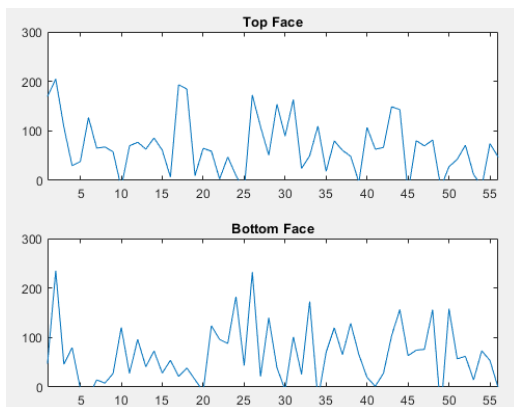
الف) نمونه‌ای از تصویر پایگاه داده



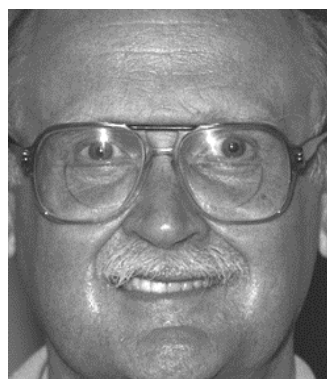
بردار ویژگی مناظر



ب) تصویر مورد بررسی



بردار ویژگی مناظر



ج) نمونه دیگری از تصویر پایگاه داده

شکل (۳-۱۵): مقایسه بردارهای ویژگی تصاویر مختلف در روش ارائه شده

### ۳-۵- معیار شباهت

پس از به دست آوردن بردار ویژگی تصویر آزمون، آن را با بردار ویژگی تصاویر موجود در پایگاه داده مقایسه می‌کنیم. در این تحقیق، برای این منظور از مفهومی به نام ضریب همبستگی<sup>۱</sup> استفاده شده که در ادامه بیشتر توضیح داده خواهد شد.

### ۳-۵-۱ ضریب همبستگی

ضریب همبستگی میزان ارتباط و وابستگی دو متغیر را نشان می‌دهد [۴۱]. این ضریب شدت ارتباط خطی و میزان همبستگی دو متغیر را تعیین می‌کند. این ضریب برای دو متغیر تصادفی  $x$  و  $y$  به صورت زیر تعریف می‌شود [۳۸]:

$$r_{xy} = \frac{cov(x,y)}{S_x S_y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (۳-۸)$$

در رابطه بالا، کواریانس با  $cov$ ، انحراف معیار با  $S$  و  $\bar{x}$  و  $\bar{y}$  به ترتیب میانگین متغیرهای تصادفی  $x$  و  $y$  است.

در صورت کسر، تغییر  $x$  و  $y$  را نسبت به هم نشان می‌دهد، اما در مخرج، تغییرات هر متغیر به صورت جداگانه اندازه‌گیری می‌شود. کواریانس به تنهایی به مقیاس اندازه‌گیری متغیرها وابسته است. از این رو معیار مناسبی برای نشان دادن همبستگی نیست. اما با تقسیم به مقیاس داده‌ها این همبستگی از بین رفته و معیار مناسبی برای همبستگی می‌شود.

ضریب همبستگی مقداری مطلق و در بازه  $[-۱$  و  $+۱]$  است؛ که اگر دو متغیر مستقل از هم باشند، صفر و هرچه همبستگی بین دو متغیر بیشتر باشد به  $+۱$  و  $-۱$  نزدیک‌تر است. در صورت مثبت بودن ضریب همبستگی، نشان‌دهنده آن است که با افزایش (کاهش) یک متغیر، متغیر دیگر نیز افزایش (کاهش)

<sup>۱</sup> Correlation Coefficient

می‌یابد. اما در صورت منفی بودن آن، افزایش (کاهش) یک متغیر با کاهش (افزایش) متغیر دیگر همراه است [۴۲].

### ۳-۵-۲ معیار فاصله اقلیدسی

معیار دیگری که در این تحقیق در نظر گرفته شده، فاصله اقلیدسی است. فاصله اقلیدسی در واقع همان فاصله مکانی بین دو متغیر  $x$  و  $y$  است که توسط قضیه فیثاغورث به دست می‌آید. رابطه (۳-۹) بیان‌گر فاصله اقلیدسی بین متغیر  $x$  و  $y$  است.

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (۳-۹)$$

### ۳-۵-۳ فاصله کسینوسی

یکی از معیارهای یافتن شباهت بین دو بردار، محاسبه زاویه بین آن دو بردار است. با استفاده از معادله (۳-۱۰)، کسینوس زاویه بین دو بردار محاسبه می‌شود.

$$\cos(\theta) = \frac{d_1 \cdot d_2}{|d_1||d_2|} \quad (۳-۱۰)$$

که در آن  $d_1$  و  $d_2$  بردارهای مورد بررسی هستند. اگر دو بردار هم جهت باشد، مقدار فاصله کسینوسی برابر است با ۱، در غیر این صورت برابر با صفر است.

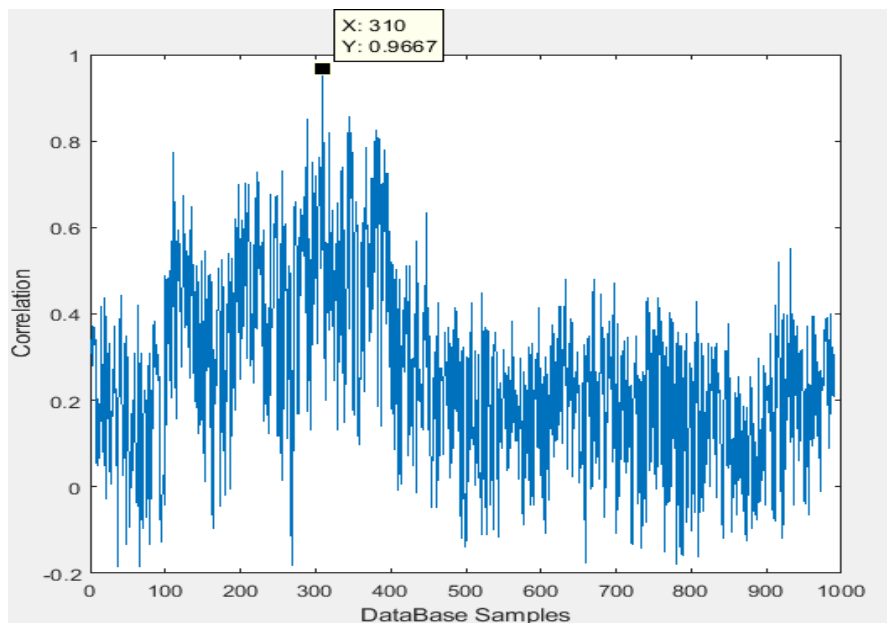
در شکل (۳-۱۵)، بردار هر ناحیه از تصویر آزمون متناظر با بردار هر ناحیه از تصویر پایگاه داده مقایسه می‌شود. همان‌طور که مشاهده می‌کنید، بردار ویژگی قسمت (ب)، با بردار ویژگی قسمت (الف) شباهت زیادی دارد؛ در حالی که بردار ویژگی قسمت (ب) با بردار ویژگی قسمت (ج)، این شباهت دیده نمی‌شود. همین امر را می‌توان با به دست آوردن مقدار ضریب همبستگی نشان داد. ضریب همبستگی بین بردار ویژگی قسمت بالای چهره (الف) و (ب) شکل (۳-۱۵)، برابر با ۰,۶۶۱۸ و ضریب همبستگی بین بردار ویژگی قسمت (ب) و (ج)، برابر با ۰,۴۰۷۲ است. لذا می‌توان گفت که ضریب همبستگی برای نشان دادن شباهت بین دو تصویر مناسب است.

ضریب همبستگی نیمه بالای تصویر آزمون با نیمه بالای کل تصاویر پایگاه داده؛ و همچنین ضریب همبستگی نیمه پایین تصویر آزمون با نیمه پایین کل تصاویر پایگاه داده، محاسبه می‌شود. سپس دو مقدار ضریب همبستگی به دست آمده را با ضریبی جمع کرده و شناسایی بر اساس آن انجام می‌شود (رابطه (۱۱-۳)). دو تصویر در صورتی مشابه اعلام می‌شوند که مقدار ضریب همبستگی‌شان، بیشترین مقدار در بین کل تصاویر پایگاه داده باشد. مقادیر بهینه  $\alpha$  و  $\beta$  با استفاده از مجموعه داده‌های اعتبارسنجی به دست آمده است.

$$A = \alpha \cdot correlation(Up) + \beta \cdot correlation(Down) \quad (11-3)$$

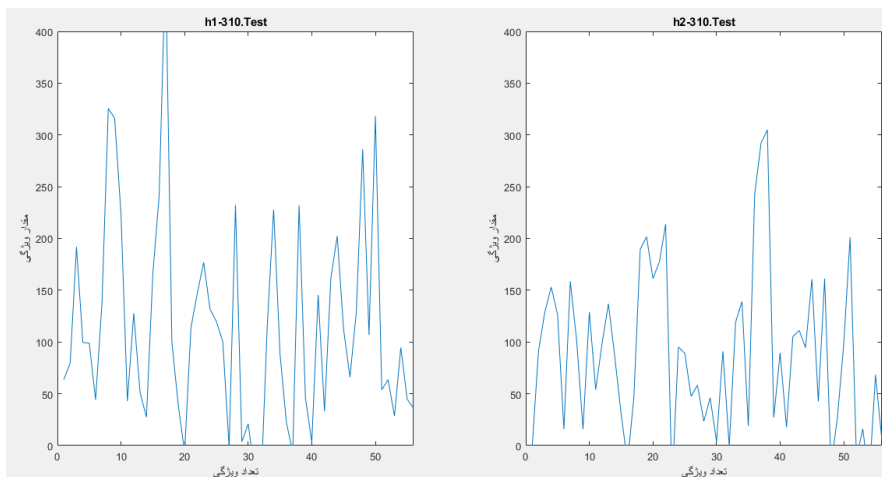
شکل (۱۶-۳) نمونه‌ای از ضریب همبستگی، بردار ویژگی یک تصویر آزمون با بردارهای ویژگی کل تصاویر پایگاه داده را نشان می‌دهد. با استفاده از معیار ضریب همبستگی، شباهت تصویر آزمون، با تمام تصاویر پایگاه داده محاسبه می‌شود. با توجه به مقادیر به دست آمده، بیشترین مقدار ضریب همبستگی متعلق به تصویر ۳۱۰ از مجموعه تصاویر پایگاه داده است. این به این معنی است که نزدیک‌ترین تصویر چهره به تصویر ورودی (تصویر آزمون)، تصویر ۳۱۰ می‌باشد. تصویر ۳۱۰ از مجموعه تصاویر پایگاه داده و از مجموعه آزمون همراه با بردار ویژگی‌های متناظر آن، در شکل (۱۷-۳) آمده است.

در شکل (۱۶-۳)، دو تصویری که بیشترین مقدار ضریب همبستگی بعد از تصویر ۳۱۰ دارند، متعلق به تصویر ۲۸۷ و ۳۴۵ از مجموعه تصاویر پایگاه داده است. ضریب همبستگی تصویر آزمون با تصویر ۲۸۷ برابر با ۰,۸۷۹۸ و با تصویر ۳۴۵ برابر با ۰,۸۸۰۳ است. شکل (۱۷-۳) بردار ویژگی‌های این دو تصویر را نیز نشان می‌دهد.



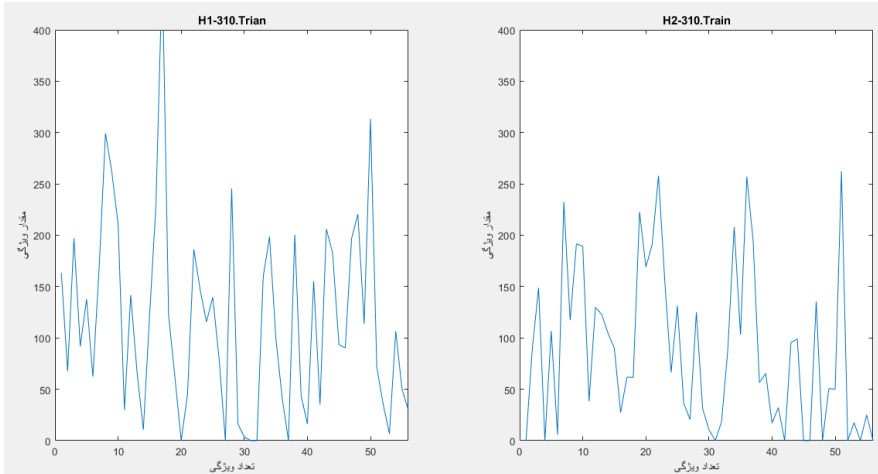
شکل (۳-۱۶): بازیابی صحیح تصویر ۳۱۰ از مجموعه آزمون با بیشترین مقدار همبستگی نشان داده شده با تصویر ۳۱۰ از پایگاه داده در روش ارائه شده

در این تحقیق، از معیارهای فاصله اقلیدسی و فاصله کسینوسی نیز استفاده شده و با یکدیگر مقایسه شده‌اند. با توجه به نتایج مقایسه، معیار ضریب همبستگی، بهتر از سایرین عمل کرده است. نتایج مقایسه معیارها در فصل چهارم توضیح داده خواهد شد.



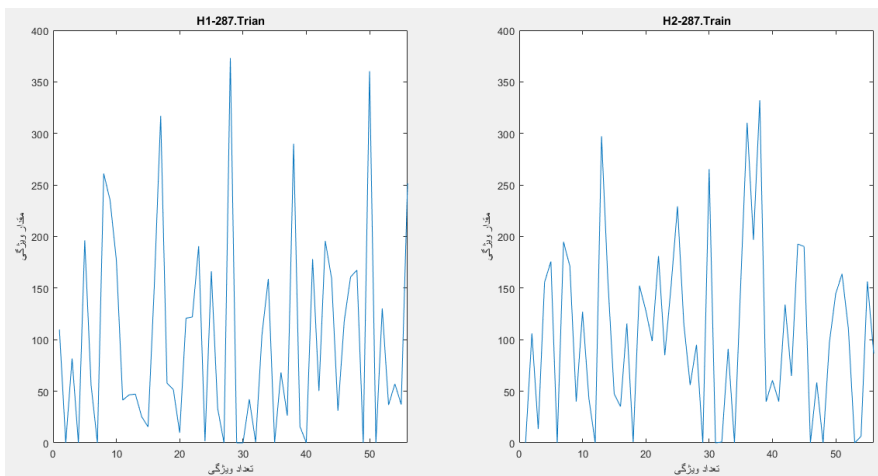
الف) تصویر ۳۱۰ از مجموعه آزمون

بردار ویژگی متناظر



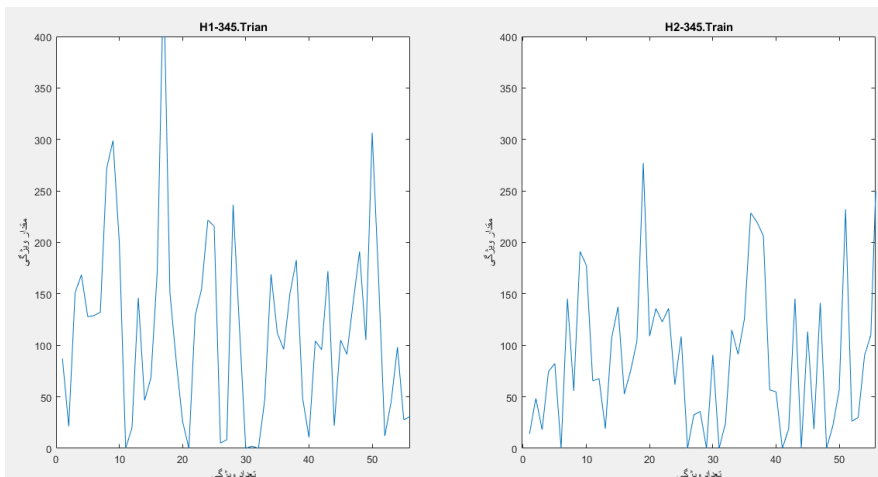
ب) تصویر ۳۱۰ از مجموعه تصاویر پایگاه داده

بردار ویژگی متناظر



ج) تصویر ۲۸۷ از مجموعه تصاویر پایگاه داده

بردار ویژگی متناظر



د) تصویر ۳۴۵ از مجموعه تصاویر پایگاه داده

بردار ویژگی متناظر

شکل (۳-۱۷): نمونه‌ای از مقایسه تصویر آزمون با سه تصویر از مجموعه تصاویر پایگاه داده با بیشترین شباهت

### ۳-۶- جمع‌بندی

در این فصل، الگوریتم پیشنهادی برای شناسایی چهره ارائه شد. ابتدا به کمک الگوریتم Viola-Jones ناحیه چهره، از تصویر استخراج می‌گردد. ویژگی‌های اساسی در ساختار چهره با استفاده از الگوریتم تجزیه ماتریس نامنفی استخراج می‌شوند. با تجزیه ماتریس تصاویر چهره موجود در پایگاه داده، ماتریس دیکشنری که دارای ویژگی‌های پایه چهره است، تشکیل می‌شود. علاوه بر ماتریس دیکشنری، ماتریس وزن نیز به دست می‌آید، که هر چهره با ترکیب عناصر اولیه موجود در دیکشنری و وزن متناظرش، توصیف می‌گردد. ماتریس وزن، شامل بردارهای ویژگی تصاویر چهره هستند؛ از این‌رو فضای ویژگی مسأله محسوب می‌شود. در مرحله آخر نیز، با استفاده از معیار شباهت مناسب، در فضای ویژگی، مقایسه بین بردار ویژگی تصویر آزمون و بردارهای ویژگی کل تصاویر پایگاه داده انجام شده تا تصویر بازیابی نهایی، به دست آید. در فصل بعد، آزمایش‌های انجام شده و نتایج ارزیابی ارائه خواهد شد.



# فصل چهارم

## ارزیابی روش پیشنهادی و بررسی نتایج

## ۴-۱- مقدمه

همان‌طور که بیان شد، در این پایان‌نامه تمرکز بر مرحله استخراج ویژگی است که روش تجزیه ماتریس نامنفی پیشنهاد شد. همچنین معیارهای مقایسه بین تصویر آزمون و تصاویر پایگاه داده نیز معرفی شدند. در این فصل، ابتدا پایگاه داده مورد بررسی، را معرفی می‌کنیم. سپس نتایج مرحله آشکارسازی چهره را نشان داده و بعد به ارزیابی و بررسی نتایج استخراج ویژگی می‌پردازیم. در بخش استخراج ویژگی، ابتدا در مورد چگونگی انتخاب اندازه تصویر و مقدار مرتبه‌ی روش ارائه شده بحث کرده، سپس اعمال پیش‌پردازش و تأثیر آن روی بهبود دقت را نشان می‌دهیم. بعد از آن، چگونگی پیاده‌سازی روش پیشنهادی و مقایسه بین معیارهای شباهت انجام شده را بیان می‌کنیم. در نهایت نتایج به‌دست آمده از جنبه سرعت و دقت را با چند روش دیگر مقایسه می‌کنیم. در آخر فصل نیز به جمع‌بندی نتایج و آزمایش‌های انجام شده، می‌پردازیم. روال کلی الگوریتم ارائه شده در این تحقیق به صورت زیر است:

۱. آشکارسازی چهره از تصاویر پایگاه داده با استفاده از الگوریتم Viola-Jones
۲. تبدیل تصاویر پایگاه داده به بردار و ذخیره آن‌ها در یک ماتریس
۳. اعمال پیش‌پردازش
۴. اجرای روش NMF و به‌دست آوردن ماتریس  $W$  (ماتریس پایه) و ماتریس  $H$  (ماتریس ویژگی)
۵. اجرای مرحله ۱ تا ۳ برای تصویر آزمون
۶. با استفاده از ماتریس  $W$  (که از تجزیه ماتریس تصاویر پایگاه داده حاصل شد) و بردار تصویر آزمون، بردار ویژگی تصویر آزمون به‌دست می‌آید.
۷. با استفاده از معیار ضریب همبستگی، مقایسه بین بردار ویژگی تصویر آزمون و بردارهای ویژگی تصاویر پایگاه داده انجام می‌شود.

۸. تصویری که بیشترین مقدار ضریب همبستگی را دارد، نزدیکترین تصویر به تصویر آزمون است.

#### ۴-۲- پایگاه داده

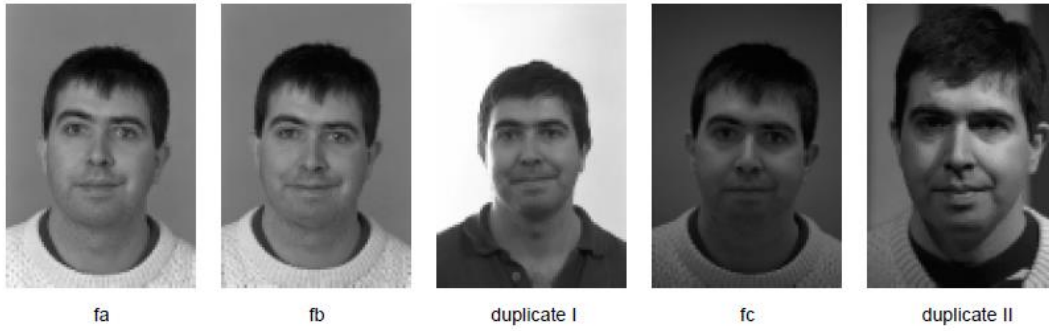
پایگاه داده FERET حاوی ۱۴۰۵۱ تصویر هشت بیتی سیاه و سفید با ۱۳ موقعیت چهره است. در جدول (۴-۱)، مجموعه تصاویر در این پایگاه داده را بیان می‌کند. در این تحقیق، تنها از تصاویر روبه‌رو استفاده شده است. این تصاویر شامل تغییراتی مانند روشنایی، اندازه، حالت چهره، سن، داشتن و یا نداشتن عینک و یا ریش می‌باشد.

جدول (۴-۱): مجموعه تصاویر پایگاه داده FERET

نام مجموعه	زاویه چهره	توصیف چهره	تعداد تصویر
fa	رو به رو	حالت عادی چهره	۹۹۰
fb	رو به رو	تغییر حالت چهره، بعد از گرفته شدن تصویر fa از هر شخص، با کمی تغییر حالت چهره، این تصاویر تهیه شدند.	۹۹۰
fc	رو به رو	با یک دوربین دیگر و در شرایط روشنایی متفاوت عکس‌برداری شده	۱۹۴
Dup1	چپ و راست	افزایش سن، داشتن یا نداشتن عینک، تغییر اندازه، روشنایی، حالت مو	۷۳۴
Dup2	چپ و راست	زیرمجموعه Dup1 و تصویر گرفته شده ۵۴۰ روز با تصویر fa فاصله دارد.	۲۲۸

در شکل (۴-۱)، از هر مجموعه پایگاه داده FERET، یک تصویر را نشان می‌دهد. تصویر dup1 در طول یک سال از تصویر fa گرفته شده و تصویر dup2 حداقل یک سال بعد از تصویر fa گرفته شده است

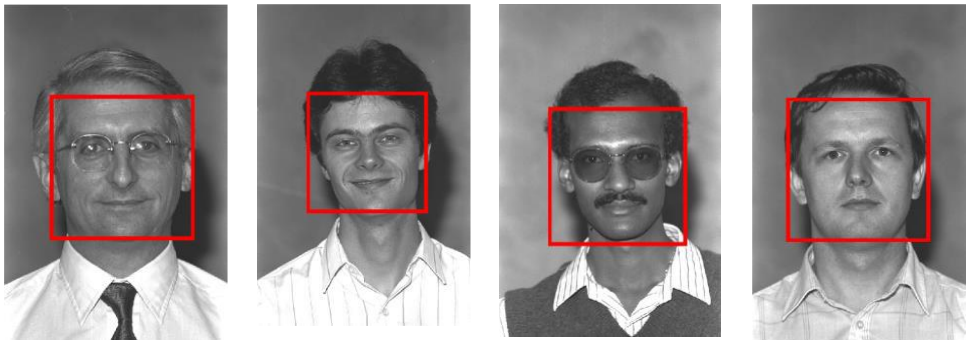
[۴۳].



شکل (۴-۱): نمونه‌هایی از مجموعه‌های پایگاه داده FERET

### ۴-۳- نتایج آشکارسازی چهره

برای آشکارسازی ناحیه چهره، از الگوریتم Viola-Jones و توابع از پیش تعریف شده در نرم‌افزار متلب استفاده شده است [۵]. نتایج حاصل از الگوریتم Viola-Jones روی چند نمونه از پایگاه داده FERET در شکل زیر نشان داده شده است.



شکل (۴-۲): استخراج ناحیه چهره پایگاه داده FERET با استفاده از الگوریتم Viola-Jones

### ۴-۴- نتایج روش پیشنهادی

در این پایان‌نامه، مجموعه fa، مجموعه آموزش و مجموعه fb، مجموعه آزمون قرار داده شد. نمونه‌ای از تصاویر مجموعه fa و fb در شکل (۴-۳) نشان داده شده‌اند. این نکته قابل ذکر است که افراد مجموعه fb همان افراد مجموعه fa هستند، که کمی تغییر حالت داده‌اند.



شکل (۴-۳): نمونه تصاویر از مجموعه آموزش و آزمون مورد بررسی. سطر اول مجموعه آموزش و سطر دوم مجموعه آزمون

#### ۴-۴-۱ به دست آوردن مقدار مناسب مرتبه‌ی (r) روش NMF

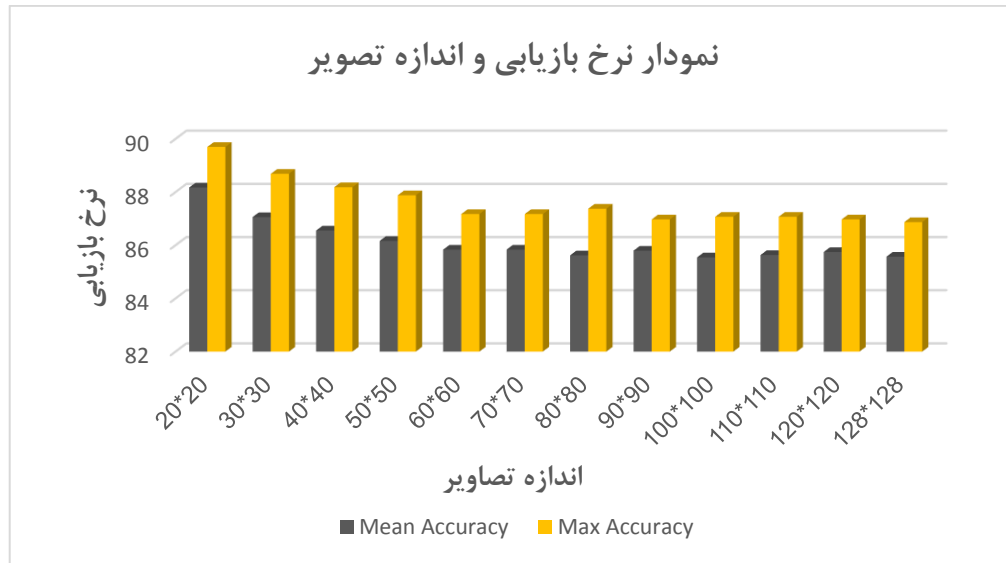
به منظور دقیق‌تر تشکیل شدن اجزای ماتریس پایه و ماتریس وزن در روش NMF، تصاویر به صورت عرضی نصف شده‌اند. همچنین همان‌طور که در بخش ۳-۳-۲ اشاره شد، برای در نظر گرفتن مقدار مرتبه روش NMF، برنامه را چندین بار اجرا نموده‌ایم تا با سعی و خطا بتوانیم، به مقدار r مناسب دست یابیم. لذا هم برای به دست آوردن اندازه تصویر مناسب و هم مقدار r، روش ارائه شده را ۱۰ بار و هر بار با مرتبه‌های (r) مختلف مورد آزمایش قرار داده‌ایم. در نهایت با توجه به میانگین صحت روش، r و اندازه تصویر مناسب انتخاب شدند. آزمایش‌های انجام شده در جدول (۴-۲) آمده است. قابل ذکر است که مقدار صحت، از به دست آمدن تعداد شناسایی‌های درست به کل تصاویر مجموعه آزمون است.

همان‌طور که در جدول (۴-۲) مشاهده می‌کنید، هنگامی که اندازه تصویر به  $20 \times 20$  تغییر می‌یابد، میانگین صحت روش برابر با ۸۸,۱۶ درصد است. همچنین بیشترین مقدار صحت به دست آمده با اجرای این روش در ۱۰ بار تکرار، برابر با ۸۹,۶۹ درصد می‌باشد. صحت حاصل شده با این اندازه، نسبت به صحت سایر اندازه‌های تصویر آورده شده در جدول (۴-۲) بیشتر بوده است. بنابراین برای ادامه‌ی کار، اندازه تصویر را برابر با  $20 \times 20$  و rank آن را ۵۶ در نظر می‌گیریم.

جدول (۴-۲): به دست آوردن مرتبه‌ی (۲) روش NMF و سایز مناسب تصویر و بدون اعمال پیش‌پردازش

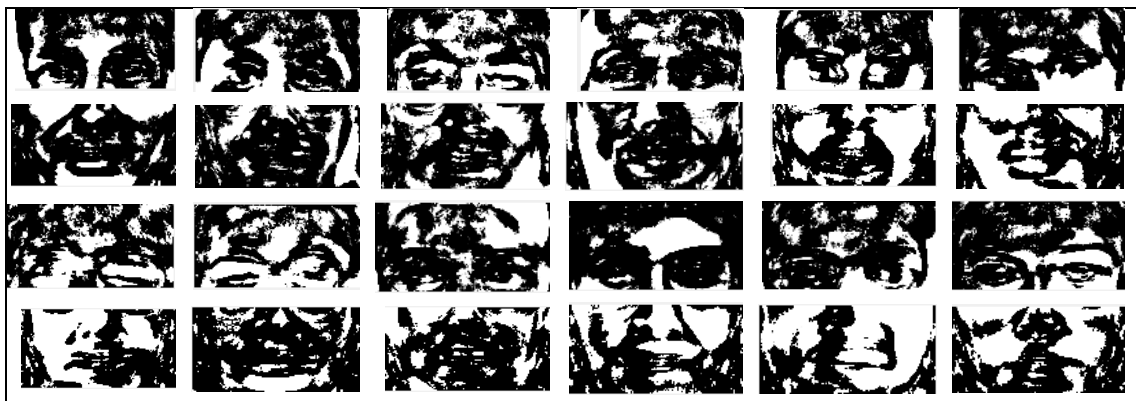
اندازه تصاویر	مرتبه	میانگین صحت	انحراف معیار	بیشینه
۸×۸	۲۰	۸۳,۹۳	۰,۶۷	۸۶,۲۶
۱۰×۱۰	۳۰	۸۵,۳۶	۰,۶۳	۸۶,۴۶
۲۰×۲۰	۵۶	۸۸,۱۶	۰,۲۰	۸۹,۶۹
۳۰×۳۰	۶۳	۸۷,۰۵	۰,۶۴	۸۸,۶۸
۴۰×۴۰	۶۹	۸۶,۵۵	۰,۲۴	۸۸,۱۸
۵۰×۵۰	۸۸	۸۶,۱۶	۰,۲۲	۸۷,۸۷
۶۰×۶۰	۱۱۰	۸۵,۸۳	۰,۱۵	۸۷,۱۷
۷۰×۷۰	۸۹	۸۵,۸۳	۰,۲۵	۸۷,۱۷
۸۰×۸۰	۱۰۹	۸۵,۶۱	۰,۴۸	۸۷,۳۷
۹۰×۹۰	۶۵	۸۵,۷۹	۰,۱۱	۸۶,۹۶
۱۰۰×۱۰۰	۷۶	۸۵,۵۴	۰,۵۹	۸۷,۰۷
۱۱۰×۱۱۰	۶۱	۸۵,۶۳	۰,۳۲	۸۷,۰۷
۱۲۰×۱۲۰	۶۷	۸۵,۷۴	۰,۳۷	۸۶,۹۶
۱۲۸×۱۲۸	۶۰	۸۵,۵۶	۰,۴۷	۸۶,۸۶

با توجه به نمودار شکل (۴-۴) بهترین عملکرد روش ارائه شده، هنگامی است که اندازه تصویر  $20 \times 20$  و مقدار مرتبه‌ی آن (۲) برابر با ۵۶ باشد.



شکل (۴-۴): میانگین دقت و بیشینه دقت در ۱۰ بار اجرا با اندازه‌های مختلف تصویر در روش پیشنهادی قبل از پیش‌پردازش

همان‌طور که توضیح داده شد، با اعمال روش NMF، روی ماتریس  $V$ ، آن را به دو ماتریس  $H$  و  $W$  تجزیه می‌کند. ماتریس  $W$  شامل ویژگی‌های پایه برای تشکیل چهره مانند چشم، بینی، ابرو و غیره است. این‌طور می‌توان گفت که ماتریس  $W$ ، ماتریس دیکشنری<sup>۱</sup> و ماتریس  $H$  فضای ویژگی را تشکیل می‌دهند، و همچنین مرتبه‌ی  $(r)$  روش NMF تعداد ویژگی‌های چهره را نشان می‌دهد.



شکل (۴-۵): نمونه‌ای از ماتریس پایه

در این روش، ابتدا تصاویر پایگاه داده به اندازه  $20 \times 20$  تغییر می‌یابد. سپس به منظور تشکیل دقیق‌تر عناصر ساختاری ماتریس پایه، هر کدام از تصاویر به صورت عرضی نصف می‌شوند. سپس نیمه‌ی بالا و

<sup>۱</sup> Dictionary

نیمه‌ی پایین تصاویر چهره، به دو بردار مجزا تبدیل شده و به ترتیب در ماتریس‌های  $V_1$  و  $V_2$  ذخیره می‌شوند. در نهایت، اندازه‌ی هر کدام از ماتریس‌های  $V_1$  و  $V_2$  برابر با  $200 \times n$  خواهد شد. به این ترتیب ستون‌های ماتریس، برابر با تعداد تصاویر و سطرهای آن برابر با تعداد پیکسل‌های تصویر می‌باشد.

$$V_{i200 \times 990} \cong W_{Train_i 200 \times r} H_{Train_i r \times 990} \quad i = 1, 2 \quad (1-4)$$

اگر داده به اندازه کافی موجود باشد و بخواهیم چهره‌ی یک فرد را با کمک ماتریس  $W_{Train}$ ، (به عنوان مثال با عینک یا بدون عینک) بسازیم، به ماتریس  $H$  (فضای ویژگی) مربوط می‌شود. لذا بهتر است که در فضای  $H$  شناسایی‌ها را انجام شوند.

به این نکته نیز باید توجه داشت که با داشتن ماتریس  $W_{Train}$  و یک ماتریس ویژگی ( $\bar{H}$ ) جدید، می‌توان چهره‌های جدیدی را خلق کرد. بنابراین با داشتن تصویر آزمون  $Y$  و با استفاده از ماتریس  $W_{Train}$  می‌توان  $H_{Test} = \bar{H}$  را به کمک تقسیم به دست آورد. از این‌رو می‌توان چهره‌های جدیدی از همان افراد در مجموعه آموزش را به وجود آورد. به عنوان مثال در شکل (۴-۶)، تصویر قسمت ج، تصویر ساخته شده از آنچه بالا گفته شده، می‌باشد.

$$H_{Test_i} = W_{Train_i}^{-1} \times Y_i \quad (2-4)$$

در این جا سایز ماتریس  $V_i$ ، برابر با  $200 \times 990$  است، که با توجه به رابطه (۳-۶) حداکثر مقدار rank را می‌توان ۱۶۶ در نظر گرفت.

$$(990 + 200) \times r < 990 \times 200 \quad r < 166$$





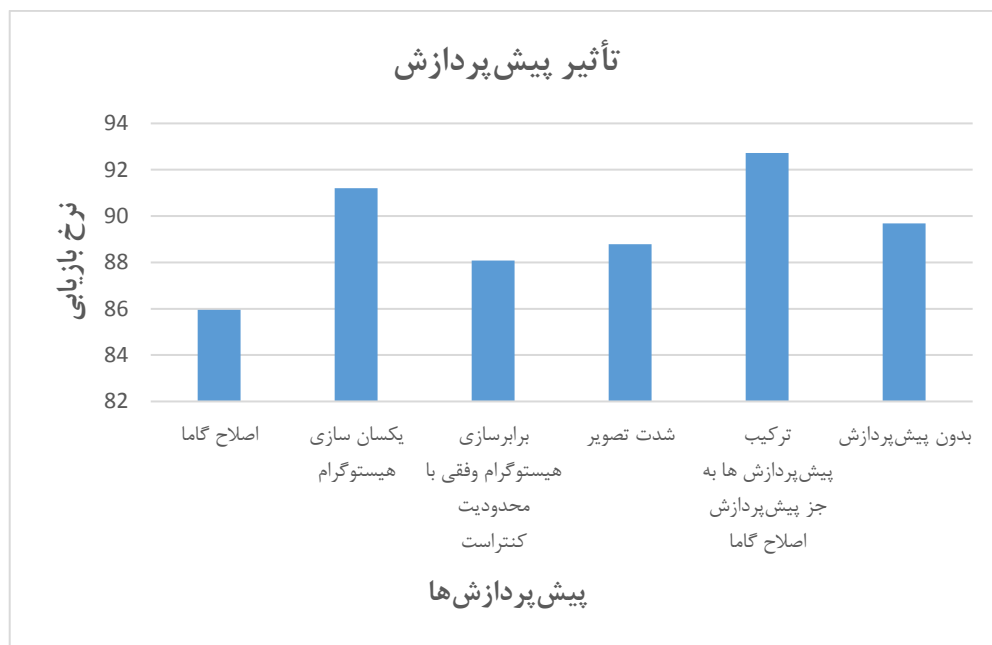
الف) تصویر آموزش

ب) تصویر آزمون

ج) تصویر ساخته شده

شکل (۴-۶): مقایسه تصویر به دست آمده با استفاده از ماتریس دیکشنری پایگاه داده و تصویر آزمون در روش پیشنهادی

با توجه به آن چه در بخش ۳-۳-۱ بیان شد، با اعمال پیش پردازش اطلاعات تصاویر، دقیق تر و صحیح تر استخراج می شوند و همین امر در بهبود نرخ بازیابی مؤثر است. در شکل زیر، تأثیر هر کدام از پیش پردازش های انجام شده را نشان می دهیم. با توجه به شکل (۴-۷)، اعمال ترکیب سه پیش پردازش، نرخ بازیابی تصویر را به ۹۲٫۷۳ درصد بهبود داده است.



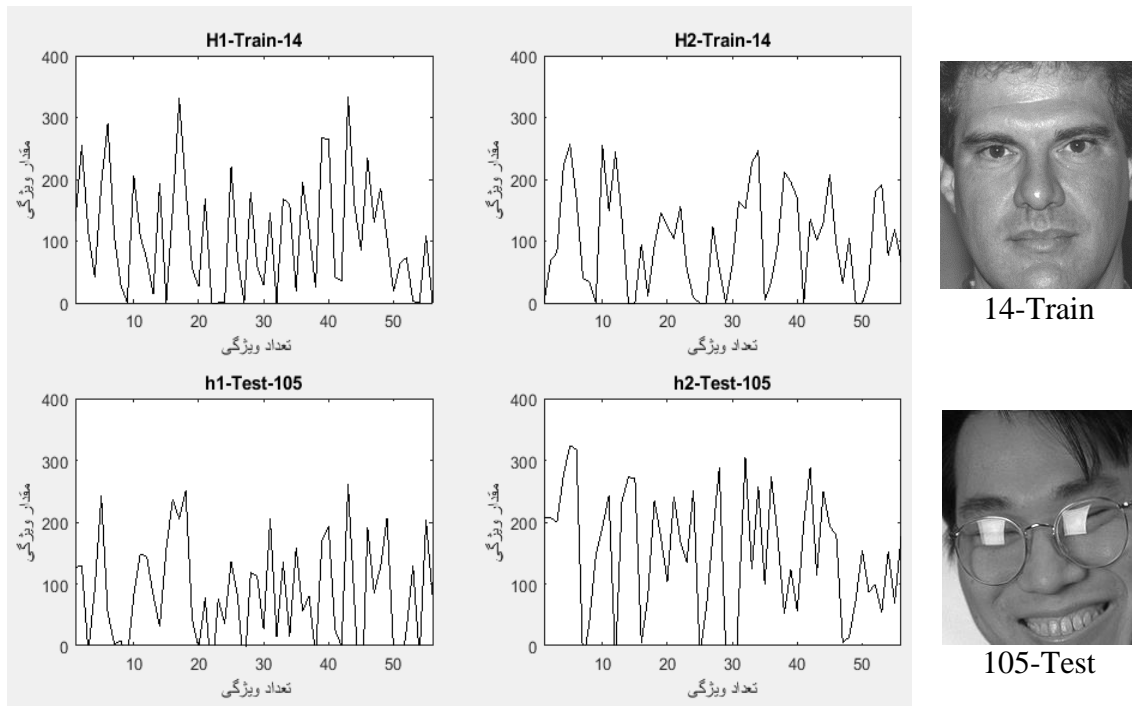
شکل (۴-۷): تأثیر پیش پردازش در نرخ بازیابی

به منظور به دست آوردن معیار مناسب، از سه معیاری که در بخش ۳-۴ معرفی شده است، استفاده کرده و نتایج در جدول (۴-۳) آمده اند.

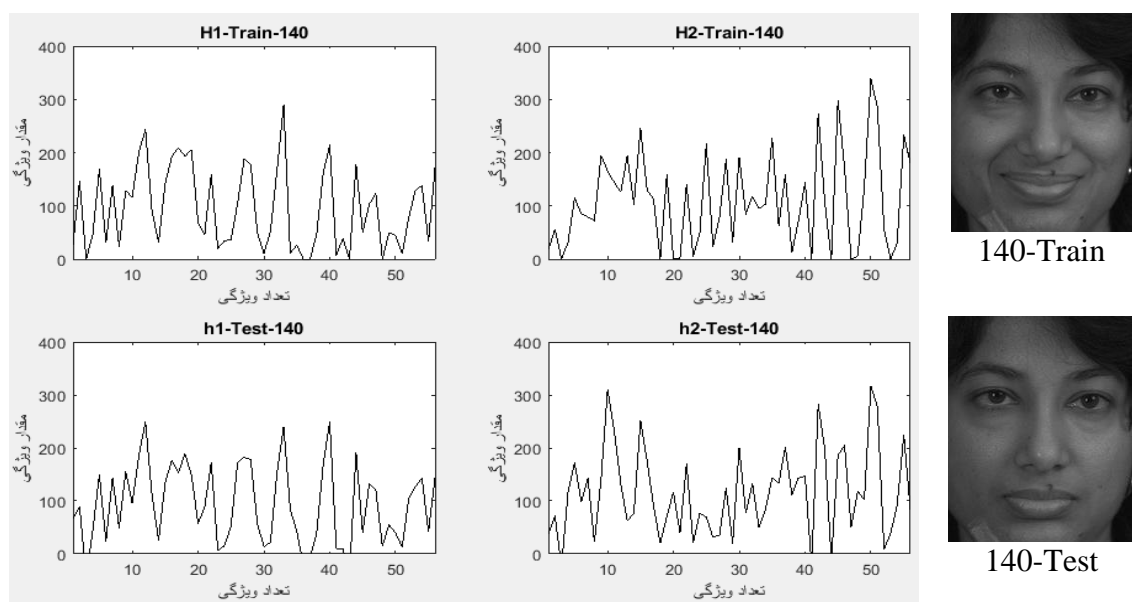
جدول (۴-۳): مقایسه معیارهای شباهت در روش پیشنهادی

معیار	ضریب همبستگی	فاصله اقلیدسی	فاصله کسینوسی
دقت	۹۲٫۷۳	۹۰٫۸	۹۰٫۷

با توجه به جدول (۳-۴) و همچنین شکل (۳-۱۶)، می‌توان گفت که معیار مناسب برای این تحقیق، معیار ضریب همبستگی است؛ زیرا این معیار توانسته که مشابه تصویر چهره‌ی آزمون را در تصاویر پایگاه داده با مقدار ضریب همبستگی بالایی بیابد. در ادامه چند تصویر و بردار ویژگی‌های متناظرشان را که روش پیشنهادی، آن‌ها را درست و یا نادرست شناسایی کرده، آمده است.



شکل (۴-۸): شناسایی نادرست تصویر آزمون با یکی از تصاویر مجموعه آموزش در روش ارائه شده



شکل (۴-۹): شناسایی درست تصویر آزمون با یکی از تصاویر مجموعه آموزش در روش ارائه شده

در این تحقیق، برنامه را با زبان برنامه‌نویسی متلب، با استفاده از سیستمی که مشخصات آن، CPU: Intel Core (TM) i7-5820k, 3.30 GHZ و Memory: 64 GB می‌باشد؛ اجرا گرفته‌ایم. در جدول (۴-۴) نتایج الگوریتم پیشنهادی، در حجم‌های متفاوت پایگاه داده آورده شده است. همچنین سرعت بازیابی، حاصل محاسبه میانگین از ۱۰ تصویر آزمون می‌باشد. در این جا سرعت بازیابی به ازای یک تصویر آزمون، محاسبه شده است. همچنین در این بررسی، به صورت تصادفی چندین تصویر آزمون انتخاب شده و میانگین سرعت‌های آن‌ها، در جدول (۴-۴) آمده است.

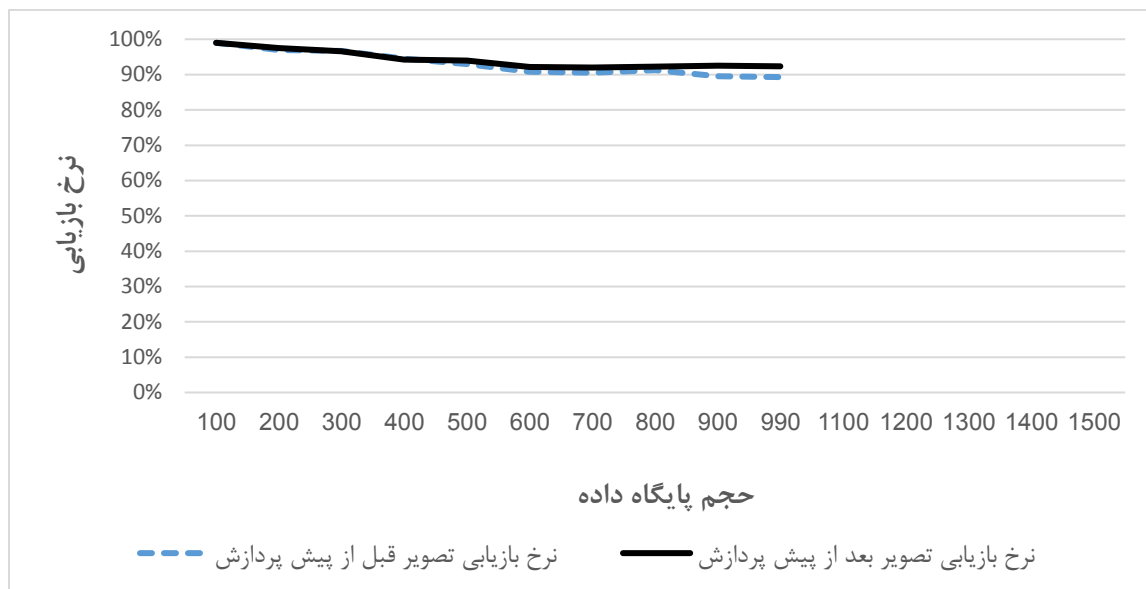
جدول (۴-۴): نرخ و زمان بازیابی الگوریتم ارائه شده در حجم‌های متفاوت پایگاه داده

مشخصات	حجم پایگاه داده	زمان بازیابی (ms) قبل از پیش پردازش	نرخ بازیابی قبل از پیش پردازش	زمان بازیابی (ms) پس از پیش پردازش	نرخ بازیابی پس از پیش پردازش
CPU: Intel Core (TM) i7-5820k, 3.30 GHZ Memory: 64 GB	۱۰۰	۱۴,۳	۹۹٪	۲۰,۷	۹۹٪
	۲۰۰	۱۹,۳	۹۷٪	۲۷	۹۷,۵٪
	۳۰۰	۲۳,۲	۹۶,۶۷٪	۳۲,۲	۹۶,۶٪
	۴۰۰	۲۶,۶	۹۴,۵٪	۳۶,۶	۹۴,۲۵٪
	۵۰۰	۳۳,۶	۹۳٪	۳۸,۸	۹۴٪
	۶۰۰	۳۶,۵	۹۰,۸۳٪	۴۵	۹۲,۱۶٪
	۷۰۰	۴۱	۹۰,۵۷٪	۴۸,۱	۹۲٪
	۸۰۰	۴۵,۷	۹۱,۲۵٪	۵۳,۶	۹۲,۲۵٪
	۹۰۰	۵۰,۵	۸۹,۵۵٪	۵۹,۵	۹۲,۵۵٪
	۹۹۰	۵۴,۲	۸۹,۲۹٪	۶۴,۱	۹۲,۳۲٪

همان‌طور که در شکل (۴-۱۰) مشاهده می‌شود، با افزایش حجم پایگاه داده، نرخ بازیابی با شیب اندکی کاهش می‌یابد. این نتیجه حاکی از آن است که روش پیشنهادی در حجم بالا، از کارایی مناسبی برخوردار است. در واقع نشان می‌دهد که ظرفیت این روش بالا است. زیرا با این روش توانسته‌ایم، فاصله‌ی بردارهای ویژگی بین افراد مختلف را حداکثر کنیم.

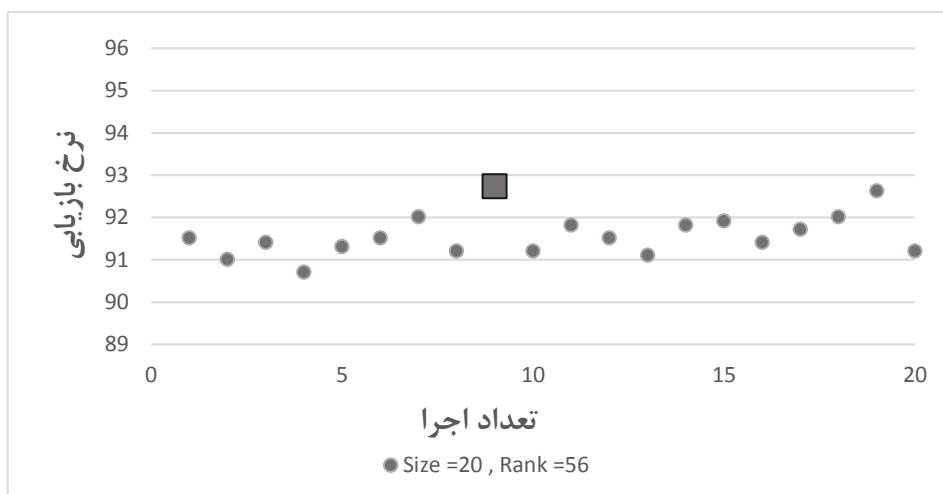
منظور از افزایش حجم پایگاه داده، افزایش تعداد افراد در پایگاه داده است، که به ازای هر شخص، تنها یک تصویر موجود باشد.

همچنین با توجه به کاهش شیب نرخ بازیابی روش پیشنهادی، پس از اعمال پیش‌پردازش، نسبت به قبل از اعمال آن، نشان‌دهنده‌ی تأثیر اعمال آن بر نرخ بازیابی است. زیرا زمانی که حجم پایگاه داده افزایش پیدا کرد، اعمال پیش‌پردازش باعث استخراج ویژگی‌های دقیق‌تری شد و روش پیشنهادی توانست به نرخ بازیابی بهتری دست یابد.



شکل (۴-۱۰): روند تغییرات نرخ بازیابی با افزایش حجم پایگاه داده در الگوریتم ارائه شده

همان‌طور که در شکل (۴-۱۱) مشاهده می‌کنید؛ بهترین عملکرد روش ارائه شده با اعمال پیش‌پردازش با مرتبه‌ی ۵۶، در ۲۰ بار اجرا، برابر با ۹۲٫۷۳ درصد و میانگین آن برابر با  $91.61 \pm 0.25$  درصد است.



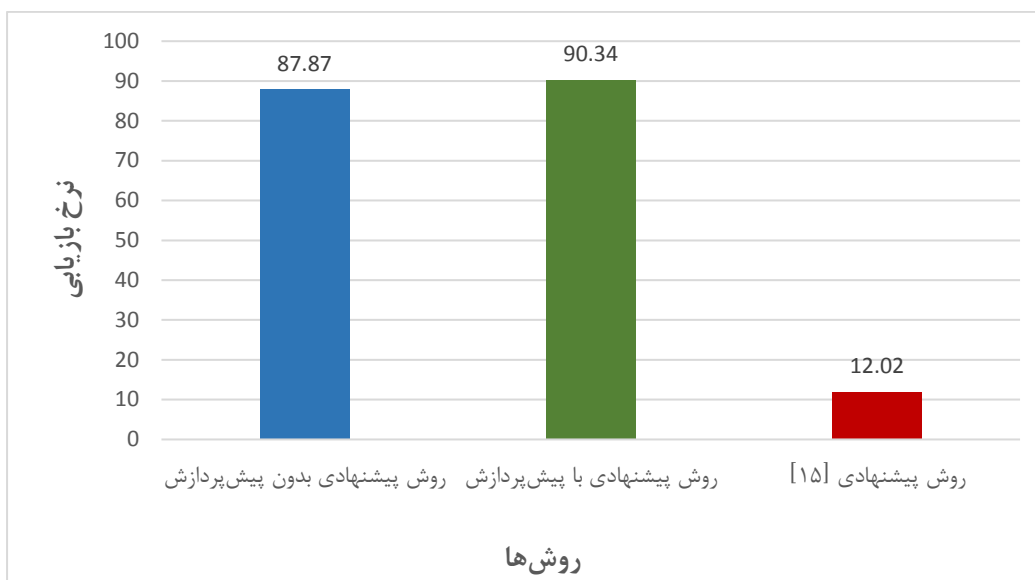
شکل (۴-۱۱): عملکرد روش ارائه شده با اعمال پیش پردازش

#### ۴-۴-۲ آزمایش با تصاویر وضوح پایین<sup>۱</sup>

آزمایش دیگری به منظور عملکرد روش پیشنهادی در مواقعی که تصاویر آزمون از وضوح پایینی برخوردار باشند، انجام شده است. به این صورت که، تصاویر مجموعه آزمون را با استفاده از نمونه برداری کم<sup>۲</sup> به اندازه  $10 \times 10$  و دوباره به اندازه  $20 \times 20$  تغییر داده و روش را آزمایش کرده ایم. حال با معیار شباهت، بردارهای ویژگی را با یکدیگر مقایسه می کنیم. شکل (۴-۱۲) مقایسه بین روش پیشنهادی و روش ارائه شده در [۱۵] را زمانی که تصاویر آزمون از وضوح پایینی برخوردار هستند؛ نشان می دهد. همان طور که مشاهده می کنید، روش پیشنهادی عملکرد بسیار بهتری نسبت به روش [۱۵] دارد. این حاکی از آن است که روش ارائه شده در این پایان نامه به وضوح تصویر حساس نیست. همین امر سبب می شود تا در کاربردهایی نظیر دوربین های نظارتی و امنیتی، که وضوح تصویر پایین است، بهتر عمل کند.

<sup>1</sup> Low Resolution

<sup>2</sup> Down Sampling



شکل (۴-۱۲): مقایسه روش پیشنهادی و روش [۱۵] زمانی که تصاویر آزمون وضوح پایینی دارند.

به منظور مقایسه‌ی دو روش پیشنهادی این پایان‌نامه و روش پیشنهاد شده در [۱۵]، بعضی از تصاویر را ۱۰ درجه چرخش داده و آزمایش انجام شد. روش [۱۵] توانست با نرخ بازیابی ۹۹ درصد به خوبی جواب دهد، اما روش پیشنهادی در این پایان‌نامه نسبت به چرخش تصاویر بسیار حساس است و لذا نرخ بازیابی مناسبی را نتیجه نداد.

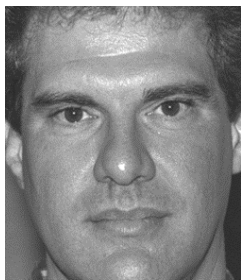
بنابراین این‌طور نتیجه می‌شود که روش پیشنهادی این پایان‌نامه نسبت به وضوح تصاویر حساس نیست، اما به چرخش تصویر بسیار حساس است. اما روش پیشنهاد شده در [۱۵] نسبت به وضوح تصاویر بسیار حساس است، در حالی که به چرخش تصویر حساس نیست.

در این روش فرض ما بر این بوده که تصاویر چهره افراد از روبه‌رو باشند و چهره‌ها زاویه‌ای نداشته باشند. اما در این روش، برخی از تصاویری را که نتوانسته به درستی شناسایی کند، به علت زاویه‌دار بودن چهره بوده است. به عنوان مثال در شکل (۴-۱۳)، سطر اول و دوم، نشان می‌دهد که روش ارائه شده نتوانسته شناسایی را به درستی انجام دهد. همچنین در سطر سوم این شکل نیز، به دلیل شباهت بسیار زیاد این دو فرد به یک‌دیگر، سیستم دچار شناسایی اشتباه شده است.

لذا همان‌طور که در بالا بیان شد، روش ارائه شده در این تحقیق، در زمانی که تصاویر چرخش داشته باشند، نمی‌تواند به نرخ بازیابی مناسبی دست یابد.

#### ۴-۵- بررسی نتایج

در این تحقیق، الگوریتم پیشنهادی را با زبان برنامه‌نویسی متلب و با استفاده از سیستمی که مشخصات آن در بالا اشاره شد؛ اجرا گرفته‌ایم. زمان‌های هر بخش از الگوریتم پیشنهادی، محاسبه شده و نتایج آن در جدول (۴-۵) آمده است.



(ج) تشخیص اشتباه



(ب) تصویر مورد انتظار از پایگاه داده



(الف) تصویر ورودی از مجموعه آزمون



(ج) تشخیص اشتباه



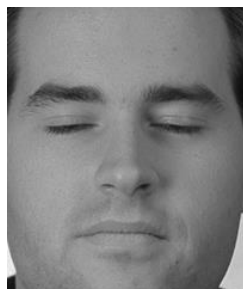
(ب) تصویر مورد انتظار از پایگاه داده



(الف) تصویر ورودی از مجموعه آزمون



(ج) تشخیص اشتباه



(ب) تصویر مورد انتظار از پایگاه داده



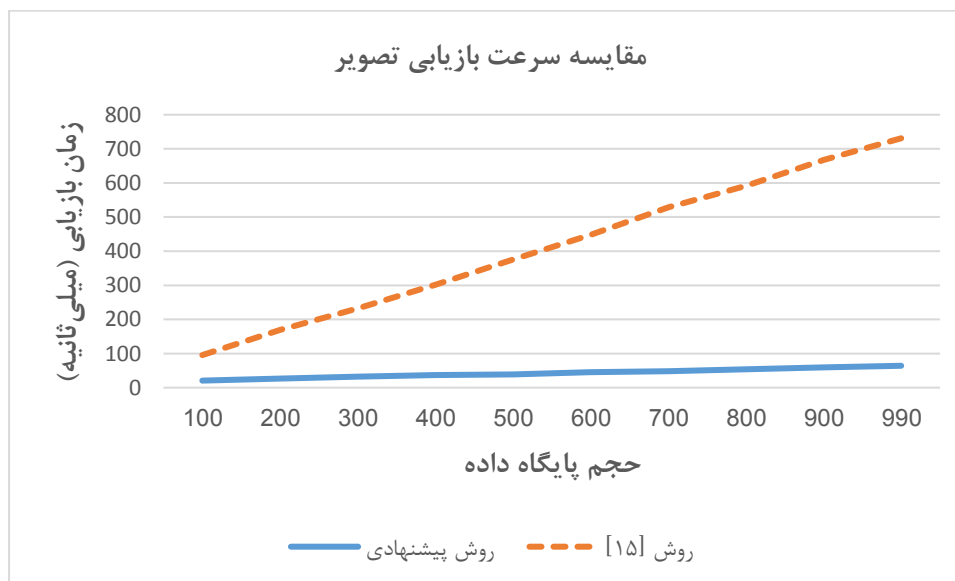
(الف) تصویر ورودی از مجموعه آزمون

شکل (۴-۱۳): شناسایی نادرست تصویر چهره در روش ارائه شده

جدول (۴-۵): بررسی زمانی روش پیشنهادی

مدت زمان	بخش الگوریتم
۱,۵ ثانیه	آشکارسازی و جداسازی چهره
۱۱,۱ ثانیه	ذخیره کل تصاویر پایگاه داده در یک ماتریس
۱۶,۶ ثانیه	استخراج ویژگی از کل پایگاه داده
۵۱,۴ میلی ثانیه	محاسبه نزدیک‌ترین شباهت به ازای یک تصویر از مجموعه آزمون با کل تصاویر پایگاه داده
۶۴,۱ میلی ثانیه	مدت زمان بازیابی یک تصویر از آزمون

همچنین الگوریتم ارائه شده [۱۵] در سیستمی که در این پایان‌نامه استفاده شده، اجرا کرده و از جنبه سرعت با روش پیشنهادی مورد مقایسه قرار گرفته است. شکل (۴-۱۴) زمان بازیابی یک تصویر آزمون را با حجم‌های متفاوت پایگاه داده این دو روش نشان می‌دهد. همان‌طور که مشاهده می‌کنید، روش پیشنهادی توانسته با سرعت قابل توجهی نسبت به روش [۱۵] تصویر آزمون را بازیابی کند. همچنین اگر حجم پایگاه داده افزایش یابد، نمودار شکل (۴-۱۲) نشان می‌دهد که سرعت روش پیشنهادی قابل توجه خواهد بود.



شکل (۴-۱۴): مقایسه زمان بازیابی یک تصویر آزمون در حجم‌های مختلف پایگاه داده روش پیشنهادی با روش [۱۵]



الگوریتم پیشنهادی مرجع [۱۶] با استفاده از پایگاه داده sqlite و زبان برنامه نویسی C++ و C# در محیط Visual Studio 2015 در سیستمی با مشخصات CPU:core i7 و RAM: 8GB پیاده سازی و اجرا کرده است. این روش با استفاده از پایگاه داده FERET و ۹۹۰ تصویر نمونه‌ی آموزشی به دقت ۹۵/۹۵ و سرعت بازیابی با زمانی حدود ۱۵۸ میلی ثانیه برای هر تصویر از مجموعه آزمون دست یافته است.

با توجه به جدول زیر، عملکرد روش پیشنهادی نسبت به روش [۱۵] هم از جنبه دقت و هم سرعت بهتر بوده است.

جدول (۴-۶): مقایسه روش پیشنهادی با روش‌های دیگر

روش	پایگاه داده	مجموعه آزمون	تعداد نفرات مورد آزمایش	تعداد نمونه‌های آموزش	طول بردار ویژگی	دقت(درصد)
[۱۳]	FERET	fb	۱۱۹۶	حاوی تصاویر تکراری	۲۳۰۱	۹۷
[۱۵]	FERET	fb	۹۹۰	۱	گزارش نشده	۹۰/۴
[۱۶]	FERET	fb	۹۹۰	۱	۲۶۰۱	۹۵/۹۵
[۱۷]	FERET	fb	۲۰۰	۱	گزارش نشده	۹۰/۵
[۱۹]	FERET	fb	۲۰۰	۱	گزارش نشده	۹۳/۹
روش پیشنهادی	FERET	fb	۲۰۰	۱	۱۱۲	۹۷/۵
روش پیشنهادی	FERET	fb	۹۹۰	۱	۱۱۲	۹۲/۷

روش ارائه شده در مقایسه با روش [۱۶] اگرچه اختلاف کمی با نرخ بازیابی آن دارد اما این روش به تصاویر با وضوح بالا نیاز دارد. همچنین این نکته قابل ذکر است که روش پیشنهادی در این پایان‌نامه، به وضوح تصاویر حساس نیست.

روش [۱۷] و روش [۱۹] با استفاده از ۲۰۰ تصویر که به ازای هر فرد، تنها یک تصویر چهره در پایگاه داده موجود بوده، مورد ارزیابی قرار گرفته‌اند. لذا برای اینکه شرایط مقایسه یکسان باشد، در جدول (۴-۶)، نرخ بازیابی روش ارائه شده در این پایان‌نامه، با ۲۰۰ تصویر از پایگاه داده FERET، گزارش شده است. روش پیشنهادی نسبت به دو روش [۱۷ و ۱۹] دارای عملکرد بهتری است. همچنین روش ارائه شده در برخورد با چالش یک نمونه آموزشی در یک پایگاه داده حجیم، توانسته به نرخ بازیابی خوبی دست یابد.

#### ۴-۶- جمع‌بندی

در این فصل به ارزیابی روش پیشنهادی و بررسی نتایج پرداخته شده است. در ابتدا ناحیه چهره با الگوریتم Viola-Jones به دست آورده شد. سپس با استفاده از روش تجزیه نامنفی ماتریس (NMF)، ویژگی‌های چهره استخراج شدند. برای تعیین مقدار مناسب مرتبه‌ی این روش، الگوریتم را چندین بار و با مقدار  $I$  های مختلف مورد آزمایش قرار گرفت. در نهایت مقدار  $I$  مناسب و همچنین اندازه تصویر ورودی به دست آورده شد. با اعمال پیش‌پردازش‌های مناسب توانستیم نرخ بازیابی را نزدیک به ۳ درصد افزایش دهیم. با توجه به آزمایش‌های انجام شده و استفاده از چند معیار دیگر، ضریب همبستگی به عنوان معیار شباهت در این تحقیق استفاده شد. این روش با طول بردار ویژگی ۱۱۲ و زمان بازیابی ۶۴٫۱ میلی ثانیه توانسته تصویر آزمون را با دقت نزدیک به ۹۳ درصد بازیابی کند. این نکته قابل ذکر است که این روش به تصاویر با وضوح پایین، حساس نیست، و همین امر نشان‌دهنده توانایی این روش در کاربردهای امنیتی و نظارتی است. علاوه بر این، روش پیشنهادی از ظرفیت بالایی برخوردار است.

چرا که با افزایش حجم پایگاه داده، نرخ بازیابی با شیب اندکی کاهش می‌یابد. اما ایرادی که به این روش می‌توان گرفت، حساس بودن آن به زاویه چهره است که در شناسایی تصویر چهره تأثیر دارد.

روش پیشنهادی از جنبه طول بردار ویژگی، نرخ و زمان بازیابی با روش‌های دیگر مقایسه شد. نتایج مقایسه نشان می‌دهد که این روش، با این‌که کمترین طول بردار ویژگی را داشته اما توانسته به نرخ بازیابی نسبتاً خوبی دست یابد. علاوه بر این، روش پیشنهادی از نظر زمان بازیابی به ازای یک تصویر آزمون، عملکرد بهتری نسبت به سایرین داشته است.



فصل پنجم

نتیجہ گیری

## ۵-۱- خلاصه و جمع‌بندی

هدف این پایان‌نامه، ارائه روشی برای استخراج ویژگی از تصاویر چهره در پایگاه داده حجیم است، که علاوه بر دقت مناسب، بتواند با سرعت خوبی تصاویر چهره را نیز بازیابی کند.

در این تحقیق، با کمک روش تجزیه ماتریس نامنفی، ویژگی‌های اساسی در ساختار چهره استخراج می‌شوند. روش پیشنهادی، بر اساس مجموعه تصاویر چهره موجود در پایگاه داده مورد بررسی، دو ماتریس پایه (دیکشنری<sup>۱</sup>) و ماتریس وزن را نتیجه می‌دهد. ماتریس پایه دارای چندین نسخه از عناصر چهره مانند بینی، دهان و دیگر قسمت‌های صورت است، که این نسخه‌ها در مکان‌ها و یا شکل‌های مختلف قرار دارند. با ترکیب این عناصر و ماتریس وزن از یک چهره، می‌توان تغییرات مختلف از آن چهره را تولید کرد. با توجه به همین نکته، به منظور تشکیل دقیق‌تر عناصر ساختاری، دیکشنری جداگانه‌ای برای نیمه بالا و پایین چهره تشکیل شد. همچنین از سه پیش پردازش؛ برابرسازی هیستوگرام، یکسان‌سازی هیستوگرام وفقی با محدودیت کنتراست و تنظیم شدت نور تصویر به منظور بهبود کیفیت و روشنایی تصاویر، استفاده شد و در بهبود نتیجه نیز تأثیر داشت. در نهایت، برای مقایسه بین بردارهای ویژگی به دست آمده با بهره‌گیری از معیار ضریب همبستگی، هویت چهره شناسایی شد.

در این پایان‌نامه، با استفاده از مجموعه پایگاه داده FERET که حاوی ۹۹۰ تصویر است و به ازای هر شخص یک تصویر موجود می‌باشد، الگوریتم ارائه شده، مورد ارزیابی قرار گرفت. نرخ بازیابی این روش حدوداً نزدیک به ۹۳ درصد است و سرعت بازیابی آن به ازای یک تصویر چهره از مجموعه آزمون، تقریباً برابر با ۶۴٫۱ میلی ثانیه می‌باشد. از مزیت‌های روش پیشنهادی می‌توان به این نکته اشاره کرد که این روش، نیازی به وضوح بالای تصاویر ندارد. لذا در کاربردهایی که وضوح تصاویر پایین است؛ مانند دوربین‌های نظارتی و امنیتی، این روش می‌تواند به خوبی عمل کند. نکته دیگر این‌که طول بردار

---

<sup>1</sup> Dictionary

ویژگی‌های استخراج شده نسبت به سایر روش‌های موجود، بسیار کم است و نیازی به استفاده از روش‌های کاهش ابعاد ندارد. بنابراین همین امر عامل مؤثری در سرعت بازیابی تصاویر است. همچنین با توجه به نمودار (۴-۹) در فصل چهارم، با افزایش حجم پایگاه داده، نرخ بازیابی با شیب اندکی کاهش می‌یابد. این نتیجه حاکی از آن است که روش پیشنهادی می‌تواند در حجم بالا، از کارایی مناسبی نیز برخوردار باشد.

## ۵-۲- پیشنهاد ادامه کار

- از آن جایی که مرحله اول سیستم شناسایی چهره، آشکارسازی چهره است؛ لذا قسمتی از محاسبات سیستم را این مرحله به عهده دارد. اگر بتوان کاری کرد که این مرحله با حفظ دقت بالا، سریع‌تر انجام شود؛ می‌توان روش پیشنهادی را از نظر نرخ بازیابی و بالا بردن سرعت بهبود داد.
- الگوهای محلی زیادی در چهره مانند نواحی لب، بینی و دیگر اجزای صورت وجود دارد؛ که از این رو می‌توان با تمرکز روی این ویژگی‌ها و با استفاده از روش‌های استخراج ویژگی محلی مانند LBP روش را بهبود داد.
- روش پیشنهادی به چرخش سر یا زاویه چهره حساس است. لذا مقاوم سازی روش به بهبود نتیجه کمک می‌کند.
- انتخاب مقادیر اولیه مناسب برای ماتریس پایه و وزن در روش NMF، می‌تواند در نرخ بازیابی تأثیرگذار باشد. مقداردهی اولیه به عنوان یک مسأله باز مطرح بوده و نیاز به مطالعه عمیق‌تری دارد.
- اعمال پیش‌پردازش‌های مناسب‌تر مانند همومورفیک، می‌تواند در بهبود نرخ بازیابی مؤثر باشد.
- در این تحقیق، تصاویر به صورت عرضی نصف شدند. اما باید تقسیم‌های مختلفی بررسی شوند تا بهترین آن لحاظ گردد. به عنوان مثال از زیر بینی، تصویر به دو قسمت تقسیم شود یا تک

تک اجزای چهره جدا شده و به طور جداگانه روی هر کدام، روش پیشنهادی اعمال شود؛ می‌تواند در بهبود نرخ بازیابی تأثیرگذار باشد.

- هنگامی که آشکارسازی چهره انجام شد، بعضی از اجزای چهره، مانند نوک بینی، لب و ... شناسایی، و برای هر کدام یک فریم گذاشته شود. سپس هنگام مقایسه با تصاویر دیگر، این فریم‌ها روی فریم‌های تصاویر دیگر تطبیق داده شود. همین امر باعث بهبود نتیجه خواهد شد.



- [1] Jain, A.K., Ross, A. and Prabhakar, S., 2004. An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1), pp.4-20.
- [2] Zhao, W., Chellappa, R., Phillips, J., Rosenfeld, A., 2003. Face recognition: a literature survey, *ACM Computing Surveys (CSUR)*, 35(4), pp. 399–458.
- [3] Pentland, A., 2000. Looking at people: sensing for ubiquitous and wearable computing. *IEEE Transaction on Pattern Analysis & Machine Intelligence*, (1), pp. 107-119.
- [4] Hassaballah, M. and Aly, S., 2015. Face recognition: challenges, achievements and future directions. *IET Computer Vision*, 9(4), pp.614-626.
- [5] 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, pp. I-I.
- [6] Rouhi, R., Amiri, M. and Irannejad, B., 2012. A review on feature extraction techniques in face recognition. *Signal & Image Processing*, 3(6), pp.1.
- [7] Belhumeur, P. N., Hespanha, J., and Kriegman, D. J., 1997. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *PAMI*, 19(7), pp.711–720.
- [8] Turk, M., and Pentland, A., 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), pp.71–86.
- [9] Liu, C., and Wechsler, H., 2002. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *TIP*, 11(4), pp.467–476.
- [10] Kittler, J., Hilton, A., Hamouz, M., and Illingworth, J., 2004. 3d assisted face recognition: a survey of 3d imaging, modelling and recognition approaches. In *ECCV*, pp.469–481.
- [11] Lu, J., Liong, V.E., Zhou, X. and Zhou, J., 2015. Learning compact binary face descriptor for face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(10), pp.2041-2056.
- [12] Martínez, A.M. and Kak, A.C., 2001. Pca versus lda. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2), pp.228-233.
- [13] Ahonen, T., A. Hadid and M. Pietikainen., 2006. Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence* 28(12), pp.2037-2041.

[14] Purnomo, F., Suhartono, D., Shodiq, M., Susanto, A., Raharja, S. and Kurniawan, R.W., 2015. November. Face recognition using Gabor wavelet and non-negative matrix factorization. In *SAI Intelligent Systems Conference (IntelliSys)*, IEEE. , pp.788-792.

[۱۵] بخشی، م، م، حسن پور، ح، فاتح، م، "استخراج ویژگی‌های مکان-فرکانسی جهت بازیابی تصویر چهره از پایگاه داده حجیم تصاویر"، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۸، شماره ۲، صفحه ۵۰۹-۵۱۷، تابستان ۱۳۹۷.

[۱۶] نوریان خ، خسروی ح، (۱۳۹۶)، "تعیین هویت بلادرنگ چهره با استفاده از ویژگی‌های مکان فرکانس چهره در مجموعه‌های بزرگ"، *سومین کنفرانس پردازش سیگنال و سیستم‌های هوشمند ایران*، دانشگاه صنعتی شاهرود.

[17] Gao, Q.X., Zhang, L. and Zhang, D., 2008. Face recognition using FLDA with single training image per person. *Applied Mathematics and Computation*, 205(2), pp.726-734.

[18] Golub, G.H. and Van Loan, C.F., Matrix Computations (Johns Hopkins U. Press, Baltimore, Md., 1983). *S. Tominaga and BA Wandell*.

[19] 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*.

[20] Wu, X., He, R. and Sun, Z., 2015, June. A lightened cnn for deep face representation. In *2015 IEEE Conference on IEEE Computer Vision and Pattern Recognition (CVPR)* (Vol. 4, p. 5).

[21] Khosravi, M.H., Hassanpour, H. and Ahmadifard, A., 2018. A content recognize ability measure for image quality assessment considering the high frequency attenuating distortions. *Multimedia Tools and Applications*, 77(6), pp.7357-7382.

[22] Amiri, S.A. and Hassanpour, H., 2012. A preprocessing approach for image analysis using gamma correction. *Department of Computer Engineering Shahrood University of Technology Iran*.

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

[24] Chang, D. Ch., Wu, W. R., 1998., Image Contrast Enhancement Based on a Histogram Transformation of Local Standard Deviation, *IEEE Transaction on Modeling Imaging*, Vol.17, No.4, pp. 518-531.

[25] Yoon, H., Han, Y., Hahn, H., 2009., Image Contrast Enhancement based Sub-histogram Equalization Technique without Over-equalization Noise, *World Academy of Science, Engineering and Technology*, Vol. 3, pp. 132-138.

[26] Grag, R., Mittal, B., Grag, Sh., 2011., Histogram Equalization Techniques For Image Enhancement, *International Journal of Electronics & Communication Technology (IJECT)*, Vol 2, No. 1.

[27] Gonzalez, R. C., Woods, R. E., Eddins S. L., 2006., Digital Image Processing Using MATLAB, *Prentice Hall*.

[۲۸] راد، ر.، (۱۳۹۶)، رساله دکتری: "برچسب‌زنی خودکار تصاویر بر مبنای تجزیه ماتریس نامنفی به صورت چند منظری". دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف.

[29] Wild, S., Curry, j. & Dougherty, A., 2004. Improving nonnegative matrix factorizations through structured initialization. *IEEE Trans. Pattern Recognition*, vol. 37, no. 11, pp. 2217-2232.

[30] Rezaei, M. and Derhami, V., 2019. Improving LNMF Performance of Facial Expression Recognition via Significant Parts Extraction using Shapley Value. *Journal of AI and Data Mining*, 7(1), pp.17-25.

[31] Li, S., Hou, X., Zhang, H., and Cheng, Q., 2001. Learning spatially localized, parts-ased representation. In *IEEE Computer Vision and Pattern Recognition*, Martigny, Switzerland, pp. 207–212.

[32] Nikolaus, R., 2007. Learning the Parts of Objects using Non-negative Matrix Factorization. *MMer Team*.

[33] Lee, Daniel D., H. Sebastian Seung., 1999. , Learning the parts of objects by non-negative matrix factorization, *Nature* 401(6755), p.788.

[34] Wang, Y., et al. 2005. Non-negative matrix factorization framework for face recognition, *International Journal of Pattern Recognition and Artificial Intelligence* 19.04. pp.495-511.

[35] Gillis, N., 2014. The why and how of nonnegative matrix factorization. *Regularization, Optimization, Kernels, and Support Vector Machines*, 12(257).

[36] Kim, J. and Park, H., 2011. Fast nonnegative matrix factorization: An active-set-like method and comparisons. *SIAM Journal on Scientific Computing*, 33(6), pp.3261-3281.

[37] Lee, D. D., and Seung, H. S., 2000. Algorithms for non-negative matrix factorization. *Advances in Neural Information Processing Systems 13*, pp.556–562.

[38] Cichocki, A. and Phan, A.H., 2009. Fast local algorithms for large scale nonnegative matrix and tensor factorizations. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 92(3), pp.708-721.

[39] Berry, M.W., Browne, M., Langville, A.N., Pauca, V.P. and Plemmons, R.J., 2007. Algorithms and applications for approximate nonnegative matrix factorization. *Computational statistics & data analysis*, 52(1), pp.155-173.

[40] Lin, C.-J., 2007. *Projected gradient methods for nonnegative matrix factorization*, *Neural Comput.*, 19, pp. 2756–2779.

- [41] Asuero A., Sayago A., and Gonzalez A., 2006. The correlation coefficient: An overview *Critical reviews in analytical chemistry*. 36(1), pp. 41-59.
- [42] Taylor R., 1990. Interpretation of the correlation coefficient: a basic review, *Journal of diagnostic medical sonography*. 6(1), pp. 35-39.
- [43] Philips, P. J., Moon, H., Rizvi, S. A., and Rauss, P. J., 2000., The ferret evaluation methodology for face-recognition algorithms. *PAMI*, 22(10), pp.1090–1104.

## واژه‌نامه

K nearest neighbor	K نزدیک‌ترین همسایه
Step size	اندازه گام
Occlusion	انسداد
Stationary Point	ایستا
Cascade	آبشاری
AdaBoost	آدابوست
Contrast Limited Adaptive Histogram	برابرسازی هیستوگرام وفقی با محدودیت کنتراست
Short Time Fourier Transform	تبدیل فوریه زمان کوتاه
Non-negative Matrix Factorization	تجزیه ماتریس نامنفی
Singular Value Decomposition	تجزیه مقدار منفرد
Fisher Linear Discriminant Analysis	تجزیه و تحلیل خطی فیشر
Principal Component Analysis	تجزیه و تحلیل مؤلفه‌های اصلی
Independent Component Analysis	تحلیل مؤلفه مستقل
Illumination variation	تغییرات روشنایی
Pose variation	تغییرات زاویه
Cumulative density function	توزیع تجمعی
Fisherface	چهره فیشر
Eigenface	چهره ویژه
Sparse	خلوت، تنک
handcrafted	دست‌ساز
Dictionary	دیکشنری
Wash out	رنگ پریدگی
Expanding Sample Method	روش گسترش نمونه
Zeng	زنگ
Biometric	زیست‌سنجی
Convolutional Neural Network	شبکه عصبی کانولوشن
Image Intensity	شدت روشنایی تصویر
Correlation Coefficient	ضریب همبستگی
Classifiers	طبقه‌بند کننده‌ها

factor	عامل
Multiplicative Update rules	قواعد به‌روزرسانی ضربی
Convolve	پیچش
Gradient Descent	کاهش گرادیان
Chi square	کای دو
Alternating Least Square	کمترین مربعات متناوب
Vector Quantization	کوانتیزاسیون برداری
Gao	گاوو
Round	گرد
Local Binary Pattern	الگوی دودویی محلی
Alternating Non-negative Least Square	مربعات نامنفی متناوب
Nearest Neighbor	نزدیک‌ترین همسایه
Down Sampling	نمونه‌برداری کم
Resolution	وضوح
Low Resolution	وضوح پایین
Holistic Feature	ویژگی سراسری
Local Feature	ویژگی محلی
Haar feature	ویژگی هار
Kernel	هسته
Histogram Orientation Gradient	هیستوگرام گرادیان جهت دار
Transfer Learning	یادگیری انتقال
Histogram equalization	یکسان‌سازی هیستوگرام

## جدول اختصارات

KNN	K Nearest Neighbor	K نزدیک ترین همسایه
Clahe	Contrast Limited Adaptive Histogram	برابرسازی هیستوگرام وفقی با محدودیت کنتراست
STFT	Short Time Fourier Transform	تبدیل فوریه زمان کوتاه
NMF	Non-negative Matrix Factorization	تجزیه ماتریس نامنفی
SVD	Singular Value Decomposition	تجزیه مقدار تکین
FLDA	Fisher Linear Discriminant Analysis	تجزیه و تحلیل خطی فیشر
PCA	Principal Component Analysis	تجزیه و تحلیل مؤلفه‌های اصلی
ICA	Independent Component Analysis	تحلیل مؤلفه مستقل
CNN	Convolutional Neural Network	شبکه عصبی کانولوشن
ANLS	Alternating Non-negative Least Square	کمترین مربعات نامنفی متناوب
VQ	Vector Quantization	کوانتیزاسیون برداری
LBP	Local Binary Pattern	الگوی دودویی محلی
HOG	Histogram Orientation Gradient	هیستوگرام گرادیان جهت‌دار

## **Abstract**

Over the past two decades, face recognition (FR) has become a popular area of research in computer vision and one of the most successful application of image analysis and understanding. Face recognition is used in a variety of areas, including security, image processing, video surveillance and criminal identification.

Face recognition challenges such as, occlusion, pose variation, illumination, facial expressions, large database and a single sample per person. Most of the methods perform well, when their database is small, or multiple samples per person are existed.

In this thesis, the main focus is on three challenges of database capability, single sample per person and speed a test image retrieval. In this research a method is proposed to extract features from facial images. By using Non-negative Matrix Factorization (NMF), the basic features of the face structure is extracted. The matrix of images is decomposed to basis matrix ( $W$ ) and weight matrix ( $H$ ). The basis images contain several versions of mouths, noses and other facial parts, where the various versions are in different locations or forms. The variability of a whole face is generated by combining the element of matrix  $W$  and matrix  $H$ . Hence, to recognize a facial image in the database, searching is on the feature vectors of them.

In this research, in order to more precisely form structural elements, a separate a basis matrix is formed for the upper and lower half of the set of facial images. Also three pre-processing methods is used to improve to quality of the images, e.g. histogram equalization, Image intensity and contrast limited adaptive histogram.

The FERET database contains 990 images and only one image per person. It is used to evaluate the proposed method. Experimental results show that the retrieval speed of a test image is faster than the other methods, however the recognition rate is closed to 93%.

**Keywords:** feature extraction, large database, single sample per person, face recognition, retrieval speed, non-negative matrix factorization





Shahrood University of Technology

Faculty of Computer Engineering and Information Technology  
M.Sc. Thesis in Artificial Intelligence Engineering

**Feature Extraction in a large database for face recognition**

By:  
Fatemeh Nikan

Supervisor:  
Prof. H. Hassanpour

Jan 2019