



ارائه یک روش استخراج ویژگی از تصاویر چهره مبتنی بر اعمال تبدیل روی ویژگی‌های به دست آمده از شبکه‌های عصبی کانولوشن

مرتضی علی‌احمدی* و روح‌الله دیانت

گروه مهندسی IT، دانشکده فنی و مهندسی، دانشگاه قم، قم، ایران

چکیده

در این مقاله، یک روش استخراج ویژگی از داده ارائه شده است. ایده پیشنهادی، کلی بوده و قابل به کارگیری در استخراج ویژگی از هر نوع داده است. در این روش، بردار ویژگی آزمون به ویژگی‌های موجود در همه دسته‌ها اضافه و سپس تبدیل مناسبی روی مجموعه ویژگی‌های هر دسته (با احتساب بردار آزمون اضافه شده)، اعمال می‌شود. نحوه اعمال تبدیل و مجموعه اقدامات بعد از آن، به نحوی صورت می‌گیرد که موجب می‌شود ویژگی‌های موجود در دسته‌ای که داده آزمون درواقع متعلق به آن است، دچار آسیب چندانی نشود و در مقابل، ویژگی‌های دسته‌هایی که داده آزمون متعلق به آنها نیست، دچار تخریب شوند. به طور شهودی می‌توان گفت، این امر، منجر به افزایش نرخ پذیرش بدروستی (TP) در الگوریتم‌های دسته‌بندی یا شناسایی الگو می‌شود. به عنوان یک نمونه، ایده پیشنهادی، در مسئله شناسایی چهره با استفاده از شبکه‌های عصبی کانولوشن (CNN)، به عنوان یک پس‌پردازش و ویژگی‌های حاصل، به عنوان ویژگی‌های نهایی، برای عملیات شناسایی چهره به کار گرفته شد. نتایج پیاده‌سازی، نشان دهنده بهبود حدود ۳/۴٪ روی پایگاه داده LFW است.

وازگان کلیدی: استخراج ویژگی، شبکه‌های عصبی کانولوشن، تبدیل موجک، تبدیل فوریه

Introducing a method for extracting features from facial images based on applying transformations to features obtained from convolutional neural networks

Morteza ali ahmadi* & Rouhollah dianat

Department of IT Engineering, Faculty of Technical and Engineering,
The University of Qom Qom, Iran

Abstract

In pattern recognition, features are denoting some measurable characteristics of an observed phenomenon and feature extraction is the procedure of measuring these characteristics. A set of features can be expressed by a feature vector which is used as the input data of a system. An efficient feature extraction method can improve the performance of a machine learning system such as face recognition in the image field.

Most of the feature extraction methods in facial images are categorized as geometric feature extractor methods, linear transformation-based methods and neural network-based methods. Geometric features include some characteristics of the face such as the distance between the eyes, the height of the nose and the width of the mouth. In the second category, a linear transformation is applied to the original data and displaces them to a new space called feature space. In the third category, the last layer in the network, which is used for categorization, is removed, and the penultimate layer output is used as the extracted features. Convolutional Neural Networks (CNNs) are one the most popular neural networks and are used in recognizing and verifying the face images, and also, extracting features.

* Corresponding author

*نویسنده عهده دار مکاتبات

• تاریخ ارسال مقاله: ۱۳۹۷/۰۱/۲۱ • تاریخ پذیرش: ۱۳۹۸/۰۶/۱۱ • تاریخ انتشار: ۱۵/۰۹/۱۳۹۹ • نوع مطالعه: پژوهشی

سال ۱۳۹۹ شماره ۳ پیاپی ۴۵



The aim of this paper is to present a new feature extraction method. The idea behind the method can be applied to any feature extraction problem. In the proposed method, the test feature vector is accompanied with the training feature vectors in each class. Afterward, a proper transform is applied on feature vectors of each class (including the added test feature vector) and a specific part of the transformed data is considered. Selection of the transform type and the other processing, such as considering the specific part of the transformed data, is in such a way that the feature vectors in the actual class are encountered with less disturbing than the other ones. To meet this goal, two transformations, Fourier and Wavelet, have been used in the proposed method. In this regard, it is more appropriate to use transformations that concentrate the energy at low frequencies. The proposed idea, intuitively, can lead to improve the true positive (TP) rate. As a realization, we use the idea in CNN-based face recognition problems as a post-processing step and final features are used in identification. The experimental results show up to 3.4% improvement over LFW dataset.

Keywords: Feature extraction - Convolutional neural networks - Wavelet transform - Fourier transform.

بردارهای Fisher بوده است. این روش‌ها حتی بدون انجام

تراسازی و گام تعیین لندهای دارند، کارایی مناسبی دارند. در مقاله فعلی، رویکردنی جهت استخراج ویژگی، ارائه شده است. ایده بیان شده، عمومی بوده که به طور بالقوه قابلیت به کارگیری را در هر مسئله دسته‌بندی و شناسایی الگو دارد. ایده کلی پیشنهادی آن است که داده آزمون به همه دسته‌های موجود در گالری (افراد)، اضافه شود. در ادامه، تبدیلی روی داده‌های هر دسته که اکنون شامل داده آزمون نیز است، اعمال می‌کنیم. نحوه اعمال تبدیل و فرایندهای انجام شده بعد از آن درنهایت منجر به آن می‌شود که اگر داده آزمون به دسته نامناسب یعنی به دسته‌ای که داده متعلق به آن نیست، اضافه شود، اثر مخرب ایفا کرده و بدین ترتیب، احتمال خطأ (یعنی تشخیص نادرست دسته) را کمتر می‌کند.

در مقابل، افزودن داده آزمون به دسته مناسب، یعنی به دسته‌ای که داده درواقع به متعلق به آن است، اثر تحریبی کمتری دارد. بدین ترتیب، میزان خطأ در شناسایی کاهش می‌یابد. شکل (۱) ایده رویکرد پیشنهادی را به شیوه ترسیمی نشان می‌دهد. برای سادگی بیان، ویژگی‌های دو بعدی، در نظر گرفته شده‌اند تا قابل ترسیم باشند. شکل (۱-الف)، وضعیت اولیه را نشان می‌دهد. اضافه کردن داده آزمون به همه دسته‌ها و اعمال تبدیل، منجر می‌شود که دسته که داده آزمون به طور واقعی متعلق به آن است، به نمونه آزمون T نزدیک‌تر شود، زیرا با میزان تخریب کمتری مواجه شده است. در مقابل، بقیه دسته‌ها، از نمونه آزمون دورتر می‌شوند. این ویژگی درنهایت، منجر به افزایش دقت دسته‌بندی و تولید خطای کمتر، خواهد شد؛ لذا انتظار داریم استفاده از این روش، منجر به بهبود پارامتر TPIR^۴ شود. درخصوص معیار FPIR^۵، تأثیر الگوریتم قابل پیش‌بینی

⁴ True positive identification rate (TPIR)

⁵ False positive identification rate (FPIR)

۱- مقدمه

شناسایی چهره^۱، یکی از مهم‌ترین مباحث مطرح در حوزه تحلیل تصاویر چهره است. یک الگوریتم شناسایی چهره، یک تصویر چهره را به عنوان ورودی دریافت و تعلق یا عدم تعلق تصویر به افراد موجود را در یک گالری، تعیین می‌کند. مهم‌ترین قسمت یک الگوریتم شناسایی چهره، مرحله استخراج ویژگی است. ایده در نظر گرفته شده می‌بایست به صورتی باشد که منجر به تولید ویژگی‌های باشد که نسبت به موانعی چون تغییر وضعیت یا حالت چهره، سن، شرایط نوری مختلف و ...، مقاوم شود. الگوریتم‌های فراوانی برای استخراج ویژگی از تصویر چهره ارائه شده‌اند که در بخش ۲، مهم‌ترین آنها مرور خواهند شد. یک دسته از رویکردهایی که در سال‌های اخیر مورد توجه بوده و دقت‌های بالایی در استفاده از آنها در حوزه‌های مختلف شناسایی الگو، گزارش شده، روش‌های مبتنی بر شبکه‌های عصبی عمیق است. در حوزه شناسایی چهره، روش‌های مبتنی بر شبکه عصبی، کانولوشن (CNN)^۲، بسیار کارا بوده و با دقت‌های بالایی، عمل شناسایی را انجام می‌دهند. به عنوان نمونه می‌توان نتایج مسابقات ILSVRC^۳ را یاد کرد که در آن، دقت الگوریتم‌های دسته‌بندی مختلف روی مجموعه داده هزار دسته‌ای، آزمایش می‌شوند. در مسابقه ILSVRC CNN 2012 (۲۰۱۲)، شبکه برگزارشده در سال ۲۰۱۲، پیشنهادشده توسط کریزفسکی و دیگران، به خطای دسته‌بندی Top-5 برابر ۱۵/۳٪ دست یافت که از الگوریتم رتبه دوم این مسابقه ۱۰/۹٪ بهتر بود [۱]. گفتنی است که الگوریتم رتبه دوم، سنتی بوده و مبتنی بر الگوریتم SIFT و

¹ Face recognition

² Convolutional neural network (CNN)

^۳ ImageNet large scale visual recognition challenge (ILSVRC)



آن که در بخش پیاده سازی، الگوریتم پیشنهادی روی رویکرد مبتنی بر CNN اعمال شده، تعداد روش های بیشتری در این حوزه، مورد بررسی قرار گرفته است. در یک نگاه کلی، اغلب روش های معروف استخراج ویژگی را از چهره می توان در یکی از سه دسته زیر قرار داد:

- روش های مبتنی بر ویژگی های هندسی.^۱
- روش های مبتنی بر تبدیلات خطی.^۲
- روش های مبتنی بر شبکه های عصبی

۱-۲- روش های مبتنی بر استخراج ویژگی های هندسی

در اغلب روش های اولیه استخراج ویژگی از تصاویر چهره، ویژگی های به کار گرفته شده، از نوع ویژگی های هندسی بوده اند. منظور از ویژگی های هندسی، مواردی نظیر فاصله چشم ها از یکدیگر، ارتفاع بینی، عرض دهان... است. در جدول (۱) برای نمونه، سه روش از این دسته، مرور شده است.

(جدول-۱): چند روش هندسی استخراج ویژگی تصاویر چهره
(Table-1): Some geometrical face feature extraction methods

روش	توضیحات
برونلی و پوگیو [۲]	ابتدا تصویر چهره را با روش یادشده در [۳] تراز ^۳ کرده و سپس، ویژگی های هندسی مانند طول و عرض بینی، موضع دهان و شکل چانه به دست آوردن.
قیمیر و همکاران [۴]	ابتدا نقاط اصلی چهره با EBGM ^۴ (در [۵]) به دست آمده اند. از آنجایی که حالات چهره، با متراکم شدن چهره همراه است، برای ریاضی مجموعه متراکمی از نقاط چهره، روش KLT ^۵ ([۶]) را به کار برند. نتیجه این ریاضی، منجر به ایجاد ویژگی های هندسی (نقطه، خط و مثلث) شد. متایمترین ویژگی ها، با به کار گیری AdaBoost ([۷]) بد همراه دسته بندی ELM ^۶ ([۸]) انتخاب شدند. در انتهای، ویژگی های هندسی به نام FER، از خطوط تقویت شده و ملت ها به دست آمدند.
کوآردا و دیگران [۹]	مجموعه ای از ویژگی های هندسی با استفاده از SVM ^۷ . الگوریتم های ژنتیک و روشی به نام mRmR ^۸ پیشنهاد شد. SVM به کار گرفته شده، مبتنی برتابع هسته خطی،تابع پایه شعاعی ^۹ و پرسپترون چند لایه ای می باشد.

¹ Geometric features

² Linear transformations

³ Align

⁴ Elastic bunch graph matching (EBGM)

⁵ Kanade-Lucas-Tomaci (KLT)

⁶ Extreme learning machine (ELM)

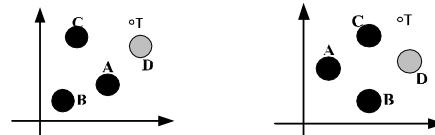
⁷ Minimum redundancy maximum relevance (mRmR)

⁸ Radial basis function

نیست؛ چون، تصویر آزمون، متعلق به هیچ یک از افراد نیست و اعمال الگوریتم، ویژگی های موجود در همه دسته ها را چار تخریب می کند. البته در بخش پیاده سازی نشان دادیم که در این وضعیت هم، در مجموع، مقدار FPIR مناسبی به دست آمده است (تعریف و توضیح این دو پارامتر در بخش ۴-۲ بیان خواهد شد).

در این نوشتار، کاربرد ایده بیان شده در حوزه شناسایی چهره بیان شده و همچنین، فرض شده است که داده های ورودی به الگوریتم، ویژگی های به دست آمده به سیله یک CNN باشد.

ساختر ارائه شده در ادامه به شرح زیر است: با توجه به آن که الگوریتم پیشنهادی، روی ویژگی های مستخرج از تصاویر چهره از CNN^{۱۰} موردنظر از این بحث است، در بخش ۲، مرور کلی بر روش های استخراج ویژگی از تصاویر چهره و به خصوص، رویکردهای مبتنی بر CNN خواهیم داشت. در بخش بعد، روش پیشنهادی، ابتدا به صورت کلی بیان و سپس، برای استخراج ویژگی در حوزه تصاویر چهره، سفارشی شده است. بخش ۴ به بررسی و تحلیل نتایج حاصل از پیاده سازی رویکرد پیشنهادی در حوزه شناسایی چهره می پردازد. بخش پنجم نیز به نتیجه گیری اختصاص دارد.



(الف) (ب)

(شکل-۱): بردار ویژگی مربوط به دسته فرضی A، B، C و D و همچنین بردار ویژگی مربوط به داده آزمون T قبل (الف) و بعد از اعمال روش پیشنهادی (ب) می تواند به صورت نشان شده در شکل باشند. این امر، منجر به افزایش دقیقت شناسایی می شود. در واقع متعلق به دسته D است؛ لذا، بردار ویژگی مربوط به دسته با رنگ دیگری نشان داده شده است.

(Figure-1): Feature vectors of 4 hypothetical class A, B, C and D accompanying test feature vector T

I) before and II) after applying the proposed method can be as showed in figure. This leads to increased accuracy of identification. Since, T actually belongs to the class D, so the feature vector of the class D is shown in a different color.

۲- آشنایی با روش های استخراج ویژگی از تصاویر چهره

در این قسمت، روش های مختلف استخراج ویژگی از تصاویر چهره به صورت کلی مورد بررسی قرار می گیرند. با توجه به

یکی از کاربردهای مهم شبکه‌های عصبی، استخراج ویژگی از تصاویر چهره، آمده است.

یکی از کاربردهای مهم شبکه‌های عصبی، استخراج ویژگی از داده ورودی است. بدین منظور، لایه آخر موجود در شبکه را که برای دسته‌بندی به کار می‌رود، حذف کرده و خروجی لایه ماقبل آخر را به عنوان ویژگی‌های استخراج شده به کار می‌برند. درواقع، در اینجا، شبکه عصبی نقش یک استخراج‌کننده ویژگی را دارد. در حوزه‌های شناسایی الگو مانند شناسایی چهره، به طور معمول شبکه عصبی به کار می‌رود.

(جدول-۳): روش‌های استخراج ویژگی مبتنی بر شبکه‌های عصبی عمیق

(Table-3): Deep neural network-based feature extraction methods

توضیحات	نام روش
در این پژوهش، شبکه‌های عصبی کانولوشن (CNN) ^۹ چندگانه‌ای پیشنهاد شده است که هر یک از شبکه‌ها وظیفه استخراج ویژگی‌هایی با نام DeepID از قسم خاصی از چهره را داشته و دسته‌بندی، براساس کنار هم قراردادن تمام DeepID‌ها صورت گرفته است. همین نویسنده‌گان در سال‌های بعد شبکه‌های DeepID2 [16]، DeepID3 [17] و DeepID2 [18] را ارائه کرده‌اند که در آنها بهبودهایی چون استفاده از شبکه‌های عمیق‌تر، داده‌های آموزشی بیشتر و درنظر گرفتن پس‌پردازش‌هایی چون کاستن از اختلاف بین ویژگی‌های تصاویر یک فرد و افراد دیگر اختلاف بین ویژگی‌های افراد مختلف، داده شده است.	DeepID [15]
در DeepFace، نخست برای تراز کردن چهره از مدل سازی سه بعدی مبتنی بر تبدیل افین ^{۱۰} استفاده کرده و ثانیا برای استخراج ویژگی و دسته‌بندی، CNN ^{۱۱} لایه‌ای به کار برند که برخی لایه‌های آن، به جای لایه‌های کانولوشن استاندارد، لایه‌هایی با اتصال محلی ^{۱۲} بدون اشتراک وزن ^{۱۳} هستند.	DeepFace [19]
در این مقاله CNN بسیار عمیق (۳۷ لایه) جهت استخراج ویژگی‌های مقاوم از تصاویر چهره، ارائه شده است. سه معماری A، B و D براساس [21] در نظر گرفته می‌شود. معماری A شامل ۱۱ بلوک بوده که هر یک شامل یک عملگر خطی (لایه کانولوشن و یا لایه به طور کامل متصل ^{۱۴}) و یک یا چند عملگر غیرخطی (لایه ReLU و لایه Max pooling) است. معماری B و D مشابه با A بوده با این تفاوت که به ترتیب دارای دو و پنج لایه کانولوشن اضافی هستند. آخرین لایه به طور کامل متصل در بلوک یازدهم، برای آموزش یک دسته‌بند به کار می‌رود.	VGGFace [20]

⁹ Convolutional neural network (CNN).

¹⁰ Affine transformation

¹¹ Locally connected layers

¹² Weight sharing

¹³ Fully connected

۲-۲- روش‌های استخراج ویژگی مبتنی بر تبدیلات خطی

مبناهای کلی این دسته، اعمال یک تبدیل خطی روی داده‌های اولیه بوده که در ادامه به فضای جدید که فضای ویژگی^۱ نام دارد، منتقل می‌شود. اگر داده اولیه، در ماتریس مشاهده^۲ $X = AX$ در قرار داشته باشد، به ماتریس مشاهده جدید A ماتریس ویژگی می‌شود. تفاوت بین رویکردها، ناشی از تفاوت در نحوه محاسبه A است. ماتریس A شامل بردارهای پایه فضای جدید (فضای ویژگی) است. در جدول (۲) بعضی از مهم‌ترین روش‌های استخراج ویژگی مبتنی بر تبدیلات خطی در حوزه تصاویر چهره، آورده شده‌اند.

(جدول-۲): چند روش استخراج ویژگی مبتنی بر تبدیلات خطی

(Table-2): Some linear transform-based feature extraction methods

توضیحات	نام روش
ماتریس A طوری به دست آورده می‌شود که پراکندگی بین داده‌ها مانع از تغییر مولفه اصلی (PCA) ^۳ استفاده شده است (در PCA، ماتریس تبدیل بر مبنای مانع از تغییر مولفه اصلی (PCA) است). با این تفاوت که به جای پراکندگی به صورت نسبت پراکندگی بین دسته‌ای ^۴ به پراکندگی درون دسته‌ای ^۵ تعریف شده است.	EigenFace [10]
در اینجا نیز، معیار محاسبه A ، مانع از تغییر مولفه اصلی (PCA) است. با این تفاوت که به جای PCA، FLDA ^۶ جایگزین شده است. در FLDA، پراکندگی به صورت نسبت پراکندگی بین دسته‌ای ^۷ به پراکندگی درون دسته‌ای ^۸ تعریف شده است.	FisherFace [12]
اگر داده اولیه در X و تصویر آزمون، در بردار y قرار داشته باشد، به دنبال تبدیل مانند $y = Ax$ هستیم، برداری است که نقش هر عنصر X (هر تصویر آموزشی) را در ساخت y مشخص می‌کند. اگر تصویر آزمون، متعلق به فرد درون گالری باشد، ضرایب مربوط به آن فرد، سیار بزرگتر از ضرایب مربوط به تصاویر سایر افراد گالری می‌باشد.	نمایش تنک ^۹ [13] و [14]

۲-۳- روش‌های مبتنی بر شبکه‌های عصبی

بیشتر رویکردهای امروزی استخراج ویژگی از تصاویر چهره، از شبکه‌های عصبی و به خصوص، شبکه‌های عصبی عمیق، استفاده می‌کنند. در جدول (۳) چند شبکه عصبی عمیق

¹ Feature space

² Observation matrix

³ Transformation matrix

⁴ Principal component analysis (PCA)

⁵ Fisher linear discriminant analysis (FLDA)

⁶ Between-class scatter

⁷ Within-class scatter

⁸ Sparse representation



سایر دسته‌ها می‌شود (جزئیات بیشتر در بخش توضیح الگوریتم پیشنهادی (۲-۲-۳)). بدین ترتیب یک مجموعه آموزشی جدید به وجود می‌آید.

ب- با توجه به تغییر مجموعه آموزشی، دسته‌بند را دوباره با داده آموزشی جدید، آموزش بده.

ت- اکنون داده آزمون را به دسته‌بند جدید بده و عمل دسته‌بندی را روی آن انجام بده.

نکته: در صورتی که برای دسته‌بندی، از معیارهای فاصله یا انطباق استفاده شود، مراحل ب و ت به صورت زیر انجام می‌شود:

ب- فاصله داده آزمون را با داده نماینده هر دسته به دست بیاور.

ت- داده آزمون را به دسته‌ای منسوب کن که کمترین فاصله را با آن دسته داشته باشد.

در بخش‌های آینده، نشان خواهیم داد که الگوریتم بیان شده توانسته این ایده را محقق سازد.

۲-۳- بیان الگوریتم ارائه شده و توضیحاتی در خصوص بعضی مراحل آن

در این قسمت، متن الگوریتم پیشنهادی، بیان می‌شود و سپس، توضیحاتی در خصوص بعضی مراحل آن داده می‌شود.

۲-۳-۱- بیان متن الگوریتم

وروودی: الگوریتم دارای دو دسته ورودی به شرح زیر است.

- مجموعه داده‌های آموزشی که هر یک، متعلق به یکی از C دسته هستند. فرض می‌کنیم دسته نام ، دارای n_i نمونه باشد. هر نمونه، d بعدی در نظر گرفته می‌شود. بدین ترتیب، بهارای هر دسته نام ، n_i نمونه d بعدی خواهیم داشت.

- داده آزمون که طبیعتاً آن هم d بعدی است و آن را با Γ نمایش می‌دهیم.

الف- داده آزمون را به هر یک از دسته‌ها اضافه کن. اکنون دسته نام ، n_i+1 نمونه d بعدی خواهد داشت. به بیان دیگر می‌توان گفت بهارای هر دسته نام ، یک ماتریس از داده‌ها به ابعاد $(n_i+1) \times d$ خواهیم داشت.

ب- اکنون بهارای هر دسته نام کارهای زیر را انجام بده

۳- توضیح روش پیشنهادی

در این قسمت، در ابتدا، ایده رویکرد پیشنهادی توضیح داده می‌شود. در ادامه، الگوریتم کلی کار بیان و توضیحاتی در خصوص بعضی از مراحل الگوریتم، بیان می‌شود. در انتهای این بخش، الگوریتم کلی بیان شده در کاربرد شناسایی چهره، سفارشی می‌شود.

۱-۳- توضیح ایده روش پیشنهادی

در رویکردهای متقابل دسته‌بندی، دو مجموعه مجزای "داده‌های آموزشی"^۱ و "داده‌های آزمون"^۲ در نظر گرفته می‌شود. با استفاده از داده‌های آموزشی، یک دسته‌بند ساخته و سپس، از دسته‌بند برای داده‌های آزمون استفاده می‌شود. در چنین ساختاری، طبیعتاً نباید هیچ یک از داده‌های آزمون در مجموعه داده‌های آموزشی وجود داشته باشد.

در این مقاله، به مسئله بیان شده در بالا، از جنبه دیگری نگریسته شده است. فرض کنید یک دسته‌بند با استفاده از تعدادی داده آموزشی، ساخته شده باشد و همچنین، یک داده به عنوان داده آزمون در اختیار داشته باشیم.

ایده‌ای که مبنای روش پیشنهادی است آن است که به طور شهودی می‌توان گفت اگر داده آزمون به دسته واقعی، یعنی دسته‌ای که داده آزمون متعلق به آن است، اضافه شود، اتفاق خاصی نمی‌افتد؛ اما اگر این داده به دسته‌های دیگر تخصیص یابد، پارامترهای آماری آن دسته‌ها دچار اسیب و خدشه می‌شود. نکته دیگر آن است که در چنین ساختاری، چون اثر داده آزمون به همه دسته‌ها اضافه می‌شود، مشکل بیان شده در بالا، در خصوص عدم اشتراک داده‌های آزمون و آموزشی پیش نمی‌آید.

بهمنظور عملی ترکردن این ایده، مراحل کلی زیر پیشنهاد می‌شود:

- الف- داده آزمون را به همه دسته‌ها اضافه کن و سپس، تبدیلی روی آن‌ها انجام بده و بخشی از آن را که در کاربرد فعلی مفید است، انتخاب کن. منظور از مفیدبودن، آن است که اگر این بخش یادشده به داده‌های دسته‌ها اضافه شود، منجر به عدم تخریب اطلاعات دسته‌ای می‌شود که تصویر آزمون به آن متعلق است و در مقابل، موجب تخریب اطلاعات

¹ Training data

² Test data

نتایج پیاده‌سازی‌های انجام شده، رضایت‌بخش بوده است. به نظر می‌رسد در اینجا، گرفتن تبدیل معکوس، هر چند راه کار ایده‌آل و اصلی نیست، اما تا اندازه‌ای فضای این داده‌ها را به فضای مکانی، نزدیک می‌کند. برای مثال، در تبدیل فوریه چون محاسبه معکوس بر مبنای توان‌های نمایی با علامت مثبت، انجام می‌پذیرد، بعضی از توان‌های نمایی با علامت منفی (که در تبدیل فوریه به کار برده می‌شوند) را خنثی می‌کند.

ب-۵- اکنون، داده‌ای را که در قسمت الف به دسته اضافه شده است، حذف کن و داده به دست آمده در قسمت

ب-۴- جایگزین آن کن. اکنون، دسته همچنان دارای $(n_i + 1)$ نمونه d عنصری خواهد بود.

ب-۶- میانگین نمونه‌های دسته را محاسبه کن و آن را به عنوان نماینده این دسته در نظر بگیر. این میانگین را با r_i نمایش بده.

مراحل الگوریتم پیشنهادی در شکل (۲) به صورت شبه کد نیز نشان داده شده است.

(شکل-۲): شبکه کد مربوط به الگوریتم پیشنهادی
(Figure-2): Pseudo-code of the proposed algorithm

Algorithm 1:

Input:

tr : Training data (tr_{ij} : j th training data for class i (d -dimensional vector))
 td : Test data,
 C : Number of classes
 n_i : Number of samples in Class i
 $Trans$: A suitable transform (Fourier, Wavelet)
 $InvTrans$: Inverse transform

- a. for $i=1, 2, \dots, C$:
 - a.1: $O_{\{n_i\} \times d} \leftarrow [tr_{i1}; tr_{i2}; \dots; tr_{in_i}]$
 - a.2: $O'_{\{n_i+1\} \times d} \leftarrow [tr_{i1}; tr_{i2}; \dots; tr_{in_i}; td]$
 - a.3: $T_i(:, j) \leftarrow Trans\{O'_i(:, j)\}$ (for $j=1, 2, \dots, d$)
 - a.4: $S_i(:, j) \leftarrow Sort(T_i(:, j), 'descend')$
 - a.5: $D_i \leftarrow [S_i(1, end) S_i(2, end) \dots S_i(d, end)]$
 - a.6: $I_i = InvTrans\{D_i\}$
 - a.7: $O''_{\{n_i\} \times d} \leftarrow DeleteRow(O'_{\{n_i+1\} \times d}, td)$
 - a.8: $O''_{\{n_i\} \times d} \leftarrow AppendRow(O''_{\{n_i\} \times d}, I_i)$
- end

Output: $O''_{\{n_i\} \times d}$

در بخش نتایج پیاده‌سازی، کارایی این الگوریتم در کاربرد شناسایی چهره نشان داده است. در آن بخش، با انجام

ب-۱- روی هر بعد (هر ستون) ماتریس داده، به صورت مجزا، یک تبدیل مناسب، مانند فوریه، DCT^۱ یا موجک اعمال کن. منظور از مناسب بودن، استفاده از تبدیلاتی است که تمرکز انرژی در فرکانس‌های پایین داشته باشند و هرچه یک تبدیل، تمرکز انرژی بیشتری در مؤلفه‌های پایین داشته باشد، در مجموع می‌تواند نتیجه بهتری را بدهد. در هر حال، ورودی تبدیل، ماتریس داده با ابعاد $n_i + 1 \times d$ و خروجی تبدیل، یک ماتریس داده با ابعاد $n_i + 1 \times d$ است؛ در این رابطه، n_i ، تعداد نمونه‌های خروجی تبدیل است.

ب-۲- در هر ستون، مقادیر ماتریس داده جدید (تبدیل یافته)، را به ترتیب نزولی مرتب کن.

ب-۳- مقدار آخرین مؤلفه را در هر ستون در نظر بگیر. به این ترتیب، یک بردار d عنصری از این مقادیر تشکیل می‌شود.

ب-۴- واضح است که بردار d بعده به دست آمده در قسمت ب-۳ در حوزه‌ای متفاوت با حوزه داده‌های آموزشی (حوزه مکان) قرار دارد. از طرف دیگر، نحوه به دست آمدن داده‌های این بردار، به شکل انتخاب مقدار (کمینه) از تبدیلات اعمال شده روی ستون‌ها است، لذا نمی‌توان گفت که این داده‌ها، حتی در فضای تبدیل یافته (برای مثال فوریه) قرار دارند. به هر حال، واضح است که می‌بایست در این مرحله، تبدیلی اعمال شود که فضای بردار به دست آمده را به فضای مکانی^۲ نزدیک گرداند. با توجه به آن که در فرایند محاسبه این بردار، سیگماهایی مشکل از ضرایب نمایی مختلف و همچنین عمل گر غیر خطی، کمینه‌سازی مورد استفاده قرار گرفته است، در عمل، به دست آوردن چنین تبدیلی بسیار دشوار و حتی غیر ممکن به نظر می‌رسد. در پیاده‌سازی‌های انجام شده، در این مرحله، یک تبدیل معکوس مورد استفاده قرار گرفته است. به عنوان مثال، اگر تبدیل انجام گرفته در معکوس فوریه روی بردار d بعده، انجام گرفته است.

^۱ Discrete cosine transform (DCT)

^۲ به عنوان مثال اگر ماتریس داده با ابعاد 1000×5 باشد، با اعمال ۱۰۰۰ تبدیل فوریه سه نقطه‌ای که هر کدام روی یک ستون ماتریس اعمال می‌شود، به ماتریس داده خروجی 3×1000 می‌رسیم.

^۳ Spatial



زیرا در آنها با اعداد غیر صفر مواجه هستیم. اگر از فرض ایده‌آل تمام صفر بودن بردار هم، کوتاه بیاییم، با توجه به اینکه داده‌های دسته با هم نزدیکی و شباهت زیادی دارند، اضافه کردن یک داده غیر متعلق به دسته، اثر تخریبی خود را بر جای می‌گذارد.

۳-۳- استفاده از الگوریتم ارائه شده در

شناسایی چهره

در الگوریتم یادشده در قسمت قبل، هیچ‌گونه قید و شرطی در خصوص نحوه استخراج ویژگی و یا فرایند دسته‌بندی، گذاشته نشده است. لذا، ایده، کلی بوده و روی هر روش دسته‌بندی یا شناسایی الگو، قابل اعمال است. در این قسمت، نسخه سفارشی‌شده این الگوریتم را در کاربرد شناسایی چهره، بیان می‌کنیم.

همان‌گونه که در بخش ۲ بیان شد، رویکردهای مبتنی بر CNN، در حوزه‌های مربوط به شناسایی الگو، دقت‌های بسیار مطلوب دارند. همچنین، در همان بخش بیان شد که در حوزه شناسایی چهره، به طور معمول از این شبکه‌ها به عنوان استخراج‌کننده ویژگی استفاده، سپس، ویژگی‌های به دست آمده، برای شناسایی به یک دسته‌بند داده می‌شود. درنتیجه، ورودی دسته‌بند، ویژگی‌های به دست آمده از بلاک CNN است.

الگوریتم یادشده در بخش ۳-۲ با توجه به نکات زیر می‌تواند در نقش یک الگوریتم شناسایی چهره ظاهر شود:

- در مسأله شناسایی چهره، تعدادی تصویر از هر فرد، در اختیار داریم. این مجموعه تصاویر، در اصطلاح گالری نامیده می‌شود.

- در یک الگوریتم شناسایی چهره، افراد موجود در گالری، نقش دسته‌ها را ایفا می‌کنند. اگر تعداد افراد n نفر باشند، شناسایی چهره، یک مسأله دسته‌بندی n دسته‌ای است.

- فرض کنید از فرد مشخص در گالری، k تصویر در اختیار باشد. هر یک از این تصاویر به طور مجزا به بلاک CNN داده شده و بردار ویژگی مربوط به آن، به دست می‌آید. بدین ترتیب اگر از فرد n ، n تصویر در اختیار باشد، به ازای این فرد، n بردار ویژگی به دست می‌آید. بردارهای ویژگی متعلق به تصاویر افراد داخل گالری، مجموعه داده آموزشی تشکیل می‌دهند.

آزمایش‌های متعدد نشان داده‌ایم که به کارگیری ویژگی‌های به دست آمده از رویکرد پیشنهادی، منجر به دستیابی به کارایی بهتر شده است.

۲-۲- توضیحاتی در خصوص مراحل الگوریتم

با توجه به وضوح الگوریتم، توضیح خاصی در خصوص مراحل اجرای آن، لازم به نظر نمی‌رسد؛ تنها به چند نکته، اشاره می‌شود:

- در مرحله ب-۱- تبدیلی نظیر فوریه، DCT یا موجک روی داده، اعمال شده است. نوع تبدیل و اندازه خروجی آن، می‌تواند موارد متعددی باشد. در بخش پیاده‌سازی، تبدیلات فوریه و موجک^۱ مورد ارزیابی قرار گرفته است.

- ممکن است، این سؤال مطرح شود که می‌توان هیچ تبدیلی اعمال نکرد و تنها خود داده آزمون را به هر دسته اضافه کرد و سپس عمل محاسبه نماینده و فاصله را انجام داد؛ اما آزمایش‌های انجام‌شده نشان داد که این کار، نتایج مناسبی به دنبال نخواهد داشت و لذا، از پرداختن به این موضوع، صرف نظر شد.

- همان‌گونه که در بخش توضیح ایده روش پیشنهادی (۳-۱) بیان شد، یکی از مراحل مهم ایده پیشنهادی، اعمال تبدیل و درنظر گرفتن بخش مفید تبدیل است. در این قسمت، منظور از مفید بودن را توضیح می‌دهیم. فرض کنید داده‌های مربوط به یک دسته، به طور کامل مانند یکدیگر و بر هم منطبق باشند. همچنین، فرض کنید داده آزمون نیز، در واقع، مربوط به همین دسته و باز، مساوی داده‌های دسته باشد. در این صورت، اگر در هر ستون، تبدیل فوریه گرفته شود فقط مؤلفه نخست تبدیل، غیر صفر بوده و سایر مؤلفه‌ها برابر صفر هستند. درنتیجه، بردار \mathbf{d} عنصری تشکیل شده در مرحله ب-۳ الگوریتم، بردار تمام صفر است؛ اما اگر، داده آزمون، جزو آن دسته نباشد، مؤلفه آخر (کوچک‌ترین مؤلفه)، بردار غیر صفر می‌شود و در واقع، اثر مخرب دارد. گفتنی است که در حالت تعلق داده آزمون به دسته صحیح (تمام صفر بودن بردار)، میانگین تحت الشعاع قرار می‌گیرد که به علت اضافه شدن یک نمونه به دسته و در نتیجه، تقسیم بر $n+1$ به جای تقسیم بر n است؛ اما تأثیر این امر در دسته واقعی به مراتب، کمتر از دسته‌های دیگر است؛

^۱ Wavelet

خروجی: k (شماره دسته‌ای (فردی) که منجر به ایجاد فاصله مینیمم شده است). در ادامه این بخش، نخست، توصیف شرایط آزمایش (شامل مجموعه داده، سخت افزار و نرم افزار به کار رفته) انجام می‌شود. در ادامه، معیارهای ارزیابی کارایی یک الگوریتم شناسایی چهره بیان شده و در انتهای، با بیان نتایج آزمایش‌های متعدد، نشان می‌دهیم به کاربردن رویکرد پیشنهادی به عنوان استخراج ویژگی، منجر به بهبود مناسبی در کارایی شناسایی چهره خواهد شد.

۴-۱- توصیف مجموعه داده، سخت افزار و نرم افزار

در جدول (۴)، ۴ مجموعه داده مورد استفاده، توضیح داده شده‌اند. همچنین، در شکل‌های (۳ و ۴) تصاویر مربوط به بعضی از آنها، نشان داده شده است. در جدول (۵) نیز، مشخصات سخت افزار و همچنین ابزارهای نرم افزاری مورد استفاده در پیاده‌سازی‌ها، آورده شده است.

(جدول-۴): مجموعه داده‌های مورد استفاده
(Table-4): Datasets used in the paper

توضیحات	نام
شامل تصاویر چهل نفر و از هر نفر ده تصویر است. تصاویر در وضعیت‌های نوری و حالت چهره مختلف گرفته شده‌اند. ابعاد تصاویر، 92×112 می‌باشد [22].	AT&T
شامل ۳۷۵۵ تصویر از ۲۷۶ نفر با ابعاد 480×640 است که در وضعیت‌های مختلف چهره و شرایط نوری مختلف، گرفته شده‌اند [23].	MUCT
تصاویر از وب گرفته شده‌اند و لذا یک مجموعه داده کنترل شده و جالشی است. ابعاد تصاویر، 250×250 است. برای ۱۶۸۰ نفر دو و یا بیشتر تصویر وجود دارد. تعداد کل این تصاویر ۹۱۶۴ است. همچنین ۴۰۶۹ نفر، فقط دارای یک تصویر هستند. برای آنکه محاسبه TPIR ساده‌تر انجام شود، فقط تصاویر مربوط به ۱۶۸۰ نفر در پیاده‌سازی مورد استفاده شده است [24].	LFW
شامل پانصد تصویر از صد نفر (هر نفر پنج تصویر) است. تصاویر، مربوط به افراد مجموعه LFW است. در مواردی که تعداد تصاویر فرد در LFW کمتر از پنج بود، با جستجو در وب، تصاویر مربوط به آن فرد، تا رسیدن به پنج تصویر، تکمیل شده است.	100LFW

• یک الگوریتم شناسایی چهره می‌باشد در مواجهه با یک تصویر مجهول، تعیین کند که این تصویر متعلق به چه فردی در گالری است. در رویکردهای CNN، تصویر مجهول نیز به بلاک CNN داده شده و بردار ویژگی مربوط به آن محاسبه می‌شود. این بردار ویژگی، همان داده آزمون الگوریتم ارائه شده‌ما، است. با توجه به ملاحظات بالا، اکنون می‌توان از این الگوریتم، به عنوان یک پردازش‌گر ویژگی‌های به دست آمده به وسیله CNN، جهت شناسایی چهره استفاده کرد.

۴- نتایج پیاده‌سازی

در این بخش، کارایی رویکرد پیشنهادی برای استخراج ویژگی را در حوزه شناسایی چهره^۱ مورد بررسی و تجزیه و تحلیل قرار می‌دهیم. با انجام آزمایش‌ها متعدد نشان خواهیم داد، ویژگی‌های به دست آمده توسط الگوریتم یادشده، منجر به دست یابی به نتایج بهتر در حوزه شناسایی چهره شده‌اند. یک الگوریتم شناسایی چهره، دارای دو ورودی است:

- ✓ یک گالری، شامل تصاویر تعدادی از افراد مشخص.
 - ✓ یک تصویر آزمون، با نام فرضی Γ .
- هدف الگوریتم آن است که مشخص کند نخست آیا این تصویر متعلق به فردی داخل گالری است یا خیر. در صورت تعلق، الگوریتم باید فرد را نیز مشخص کند. مرافق کلی یک الگوریتم شناسایی چهره به صورت زیر است:

- ✓ بهایی هر تصویر موجود در گالری، با استفاده از یک روش استخراج ویژگی، یک بردار ویژگی به دست بیاور.
- ✓ میانگین بردار ویژگی‌های مربوط به تصاویر هر فرد، را محاسبه کن. فرض کنید بهایی فرد شماره m_i ، میانگین، برابر m_i باشد.
- ✓ فاصله داده آزمون با هر کدام از m_i ‌ها را به دست بیاور.

$$D_i = dist(\Gamma, m_i) \quad (1)$$

- ✓ اگر D_i بیشتر از یک مقدار آستانه τ باشد، تصویر آزمون را متعلق به هیچ فردی ندان. در غیر این صورت، تصویر را به دسته‌ای (فردی) که کمترین D_i را موجب شده منسوب کن.

$$k = \arg_i \min D_i \quad (2)$$

¹ Face recognition

در این رابطه، T ، تعداد کل تصاویر غیر متعلق به گالری است که برای آزمایش به الگوریتم داده شده و F ، تعدادی از این تصاویر است که الگوریتم آنها را بهاشتباه، متعلق به گالری دانسته است. پارامتر τ ، مقدار آستانه قرارداده شده روی معیار فاصله است که طبیعتاً وابسته به آستانه انتخابی τ است؛ همچنین، پارامتر $TPIR$ نیز از رابطه زیر به دست می آید:

$$TPIR(R, \tau) = \frac{P(R, \tau)}{S} \quad (4)$$

در این رابطه S ، تعداد تصاویر متعلق به گالری است که برای آزمایش به الگوریتم داده شده است. P نیز تعدادی از این تصاویر است که در آنها، الگوریتم به درستی فرد مورد نظر را مشخص کرده است.

R ، پارامتر رتبه^۳ است. به طور معمول یک الگوریتم شناسایی چهره به جای یک نتیجه، R نامزد بهترین دقت تعلق، خروجی می دهد. هر چه R بیشتر باشد، مقدار $TPIR$ بیشتر خواهد بود. مشابه $FPIR$ ، τ مقدار آستانه است. طبیعتاً، زمانی که رتبه برابر یک باشد، فقط یک فرد به عنوان نتیجه شناسایی، اعلام می شود.

یک الگوریتم شناسایی چهره، زمانی الگوریتم مناسبی است که اگر تصویر آزمون متعلق به هیچ فرد داخل گالری نباشد آن را اعلام کند و اگر تصویر، متعلق به فردی در گالری باشد، به درستی آن فرد را تشخیص دهد. پارامتر $FPIR$ ، دقت الگوریتم در تشخیص تعلق یا عدم تعلق تصاویر به گالری را می سنجد و $TPIR$ بیان گر میزان کارایی الگوریتم در تشخیص فرد صحیح است.

پارامتر $1-TPIR$ در اصطلاح "نرخ تشخیص رد بهاشتباه" (FNIR)^۴ نامیده می شود. $FNIR(R, \tau)$ بیان گر آن است که از بین تصاویر آزمون که در واقع متعلق به فردی در داخل گالری هستند، الگوریتم با چه نسبتی فرد مورد نظر را به درستی تشخیص نداده است.

واضح است که هر دو پارامتر $TPIR$ و $FPIR$ در بازه $[0, 1]$ هستند. البته، برای این که یک الگوریتم شناسایی چهره عملکرد مناسبی داشته باشد، می بایستی $FPIR$ مقدار کم (نزدیک به صفر) و $TPIR$ ، مقدار بزرگ (نزدیک به یک) داشته باشد.

³ Rank

⁴ False negative identification rate (FNIR)

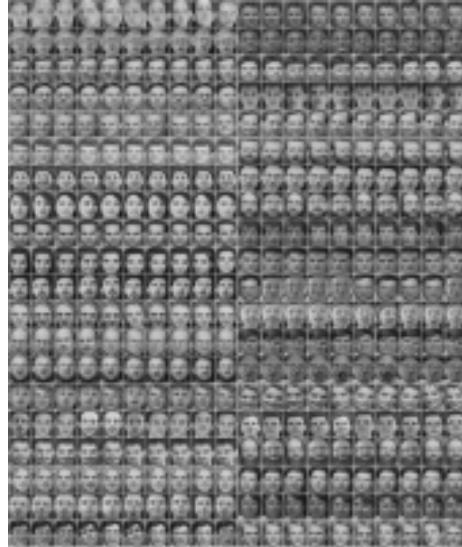
(جدول-۵): سخت افزار و نرم افزار مورد استفاده

(Table-5): Hardware and software used in the paper

نرم افزار	سخت افزار
MATLAB 2017-MatConvNet ([25])	CPU: Corei7 (7700k)- RAM: 16GB-Hard: (SSD-512GB)-Cache: 8MB-GPU: 1070 (8 GB)

(شکل-۳): مجموعه تصاویر پایگاه داده AT&T

(Figure-3): Collection of AT&T database images



(شکل-۴): بعضی تصاویر مجموعه داده LFW (جمع آوری شده توسط نویسنده)

(Figure-4): Some of the LFW database images (collected by authors)



۴-۲- معیارهای ارزیابی کارایی

برای سنجش کارایی یک الگوریتم شناسایی چهره، دو معیار "نرخ تشخیص پذیرش بهاشتباه" ($FPIR$) و "نرخ تشخیص پذیرش به درستی" ($TPIR$) در نظر گرفته می شوند [27]. معیار $FPIR$ به صورت زیر تعریف می شود:

$$FPIR(\tau) = \frac{F(\tau)}{T} \quad (3)$$

¹ False positive identification rate

² True positive identification rate

فرض می‌کند، تصاویر ورودی رنگی و با اندازه Face $224 \times 224 \times 3$ باشند و لذا، روی هر چهار داده، یک تغییر اندازه^۱، انجام شده است؛ به علاوه، داده AT&T بهصورت تصاویر خاکستری هستند. با سه تکرار مؤلفه‌های خاکستری، یک تصویر، که هر پیکسل دارای سه مقدار است، ساخته و این داده جدید، به شبکه اعمال شد.

در ادامه، نتایج یک سری از آزمایش‌ها و پیاده‌سازی‌ها بیان شده و مورد تحلیل و بررسی قرار می‌گیرند. با توجه به مطالب یادشده، برای انجام پیاده‌سازی، حالات زیر در نظر گرفته شده است:

- بهمنظور مقایسه بردار ویژگی آزمون و بردار ویژگی‌های مربوط به هر فرد در داده آموزشی، سه معیار فاصله کسینوسی، همبستگی و فاصله اقلیدسی درنظر گرفته است.
- سه حالت برای تبدیل به کار گرفته شده است: بدون استفاده از تبدیل (ویژگی‌های به دست آمده از CNN)، با استفاده از تبدیل فوریه و تبدیل موجک.
- سه معیار برای ارزیابی درنظر گرفته شده است: ERR ، $FPIR$ و $TPIR$.
- چهار مجموعه داده برای ارزیابی به کار برد شده‌اند: $100LFW$, $MUCT$, $AT\&T$ و LFW .

بدین ترتیب تعداد حالات بسیار زیادی برای پیاده‌سازی وجود دارند که با توجه محدودیت‌های موجود، تنها بخشی از حالات بیان شده، تحت قالب آزمایش‌های زیر، مورد ارزیابی قرار گرفته‌اند:

با توجه به آن‌که، هدف، بررسی تأثیر پس‌برداش ارائه شده است، مقایسه روی دو حالت استفاده از پس‌برداش و عدم استفاده از آن، انجام گرفته و نتایج حاصله در جداول، آورده شده است.

در جدول (۶) تأثیر ویژگی‌های به دست آمده از اعمال تبدیلات مختلف، در افزایش کارایی الگوریتم شناسایی چهره مبتنی بر شبکه VGGFace نشان داده شده است. به کارگیری تبدیلات فوریه و موجک، منجر به حفظ کارایی و یا افزایش آن در بعضی موارد شده است. جدول (۷) مانند جدول (۶) است، با این تفاوت که به جای فاصله کسینوسی، فاصله همبستگی درنظر گرفته است. نتایج مناسب روش پیشنهادی روی انواع مختلف معیارهای سنجش شباهت، نشان‌دهنده آن است که بهبود کارایی، ناشی از ایده ارائه شده بوده و ارتباطی با نوع معیار سنجش شباهت مورد استفاده ندارد.

⁴ Resizing

در پیاده‌سازی‌های انجام شده، میانگین وزن دار $TPIR$ و $FPIR$ محاسبه شده است. این میانگین وزن دار بر حسب تعداد تصاویر متعلق به افراد داخل گالری و تصاویر غیر مرتبط درنظر گرفته شده و تحت عنوان نمره نهایی (FS^۲) استفاده شده است. این پارامتر، برآیند کارایی کلی الگوریتم را نشان می‌دهد.

بهمنظور محاسبه یک مقدار مناسب برای آستانه τ از معیار نرخ خطای معادل (EER^۳) استفاده شده است. در محاسبه EER، ابتدا مقداری از τ را پیدا می‌کنیم که به ازای آن $(FNIR(R, \tau) = FPIR(\tau))$ یک مقدار ثابت در نظر گرفته می‌شود. اگر این مقدار آستانه را $\hat{\tau}$ بنامیم $EER = FNIR(R, \hat{\tau}) = FPIR(\hat{\tau})$. عدد $\hat{\tau}$ می‌تواند به عنوان یک مقدار مناسب برای آستانه، محسوب شود.

۴-۳- نتایج پیاده‌سازی، بررسی و تحلیل

در این بخش به بیان، بررسی و تحلیل نتایج الگوریتم ارائه شده در بخش ۳-۳ می‌پردازیم. نکات زیر در خصوص این پیاده‌سازی قابل ذکر است:

- CNN مورد استفاده برای استخراج ویژگی‌های اولیه در این قسمت، همان شبکه VGGFace ارائه شده به وسیله پرخی است. این شبکه با ۲۶۲۲ نفر و از هر نفر هزار تصویر، آموزش داده شده است (در جمع ۲۶۲۲۰۰ تصویر). جزئیات مربوط به لایه‌ها و نحوه ساخت شبکه در [20] ارائه شده است.
- در پیاده‌سازی انجام شده، خروجی لایه ۳۶ به عنوان بردار ویژگی، درنظر گرفته است. با توجه به اینکه، تعداد عناصر این بردار ویژگی، ۴۰۹۶ عدد است، درنهایت، بردار ویژگی که در این پژوهش برای هر تصویر مورد استفاده قرار می‌گیرد، ۴۰۹۶ مقداری خواهد بود.
- ارزیابی کارایی با استفاده از هر دو معیار $TPIR$ و $FPIR$ انجام گرفته است.
- هر چهار مجموعه داده $AT\&T$, $MUCT$, LFW و $VGG-100LFW$ مورد ارزیابی قرار گرفته‌اند. شبکه

¹ Final score (FS)² Equal error rate (ERR)

³ در عمل، برای داده‌های واقعی، ممکن است هیچ‌گاه این دو معیار، دقیقاً با هم مساوی نشوند؛ لذا روال به این صورت است که ابتدا به ازای مقداری مختلف... τ ... $FPIR$ و $FNIR$ را محاسبه می‌کنند. سپس، $\hat{\tau}$ را برابر با مقداری می‌گیرند که کمترین تفاضل را بین $FPIR$ و $FNIR$ موجب شده باشد.



¹ Linear Discriminant Analysis (LDA)

² Kernel Fisher Analysis (KFA)

³ Butterfly algorithm

به منظور نشان دادن کارایی الگوریتم در مواجهه با روش‌های غیر مبتنی بر شبکه‌های عصبی، در جدول (۱۱) نتایج حاصل از روش‌های استخراج ویژگی^۱ LDA و KFA^۲ نیز آورده شده است (برای آشنایی با LDA و KFA به ترتیب به [۲۸] و [۲۹] مراجعه شود). مجموعه داده AT&T در این آزمایش درنظر گرفته شده است. نتایج جدول نیز حاکی از کارایی مناسب الگوریتم ارائه شده است. گفتنی است با توجه به دقت بسیار کم روش‌های سنتی روی داده‌های کنترل نشده، این آزمایش روی سایر مجموعه داده‌ها، انجام نشده است.

در انتهای این بخش، درخصوص سربار محاسباتی ایده ارائه شده، توضیحاتی ارائه می‌شود. قسمت محاسبه تبدیلات در الگوریتم، از نظر زمان اجرا وضعيت مناسبي دارد. به عنوان مثال، محاسبه تبدیل فوریه یا معکوس آن، با الگوریتم مشهور پروانه^۳ است که با مرتبه زمانی $O(dN_i \log N_i)$ انجام است. در رابطه یادشده، N_i تعداد نقاط خروجی تبدیل و d ، تعداد ویژگی‌ها هستند. بقیه بخش‌های الگوریتم نیز شامل مواردی مانند مرتب‌سازی و پیمایش نقاط بوده و مرتبه‌های زمانی مناسبی دارند. از نظر سربار حافظه‌ای نیز با توجه به آن که داده آزمون به هر دسته اضافه شده است، سربار بدنسبه کوچکی تحمیل می‌شود که این سربار مناسب با تعداد دسته‌ها است. با توجه به موارد بیان شده، ایده پیشنهادی، دارای سربار معقول و مناسبی است.

۵- نتیجه‌گیری

در این مقاله، روشی برای استخراج ویژگی‌های تصاویر چهره به منظور استفاده از آنها در شناسایی چهره ارائه شد. ایده پیشنهادی، کلی بوده و قابل به کارگیری در هر روش استخراج ویژگی و هر الگوریتم دسته‌بندی یا شناسایی الگو است.

در الگوریتم ارائه شده، داده آزمون به همه دسته‌های موجود در گالری، اضافه شده و تبدیلی روی داده‌های هر دسته که اکنون شامل داده آزمون نیز است، اعمال می‌کنیم. فرآیند تبدیل و سایر پردازش‌های انجام‌شده بعد از آن، باعث می‌شود که اگر داده آزمون به دسته نامناسب یعنی به دسته‌ای که داده متعلق به آن نیست، اضافه شود، اثر مخرب ایفا کرده و بدین‌ترتیب، احتمال خطأ (یعنی تشخیص

ایده بیان شده منجر به بهبود وضعیت TPIR و یا به بیان معادل FNIR شده است. در مقابل، FPIR رویکرد ارائه شده در بعضی موارد، وضعیت نامناسب‌تری نسبت به حالت بدون تبدیل دارد. بدین‌ترتیب به کارگیری الگوریتم پیشنهادی زمانی مناسب است که باید تصاویر آزمون، متعلق به گالری هستند و الگوریتم فقط می‌خواهد فردی را که تصاویر متعلق به او است، در گالری بیابد. در شکل‌های (۵ و ۶)، نمودار FPIR و FNIR به ازای ۲۰ مقدار آستانه مختلف از صفر تا یک با گام ۰/۰۵ و با فرض استفاده از فاصله‌های همبستگی و اقلیدسی، رسم شده است. داده 100LFW ۱۰۰ مورد استفاده قرار گرفته و با توجه به نمودارهای رسم شده، پارامتر EER (محل تقاطع نمودارهای FPIR و FNIR)، به دست آمده و در جدول‌های (۸ و ۹) آورده شده است. جدول (۱۰) مشابه جدول‌های (۶ و ۷) است؛ با این تفاوت که فاصله اقلیدسی به جای فاصله‌های کسینوسی و همبستگی جایگزین شده است. با توجه به آن که از نظر تئوری، فاصله اقلیدسی می‌تواند هر مقداری در بازه [۰, +∞) باشد، امکان درنظر گرفتن آستانه، فراهم نیست. در این حالت، تصویر چهره آزمون، به فردی در داخل گالری منسوب می‌شود که کمترین فاصله را موجب شده باشد. به علت عدم درنظر گرفتن مقدار آستانه، پارامتر FPIR قابل محاسبه نبوده و لذا در جدول، صرفاً معیار TPIR آورده شده است. نتایج حاصل از اعمال فاصله اقلیدسی نیز نشان‌دهنده کارایی بهتر رویکرد ارائه شده است.

نکته قابل توجه، آن است که نتایج پیاده‌سازی الگوریتم پیشنهادی، هم روی دادگان کنترل شده (آزمایشگاهی) مانند AT&T و MUCT و هم روی مجموعه داده‌های واقعی (کنترل شده و جمع‌آوری شده از وب) یعنی LFW و 100LFW نتایج مناسبی داشته است و این امر خود، مؤیدی بر کاربودن ایده ارائه شده است؛ بهخصوص آن که نتایج الگوریتم، روی مجموعه داده چالشی‌ای چون LFW، بهبود چشم‌گیری داشته است؛ در ضمن گفتنی است که نتیجه رویکرد پیشنهادی روی داده‌های آزمایشگاهی که بدون اعمال روش نیز دقت بالایی دارند، به طور کامل مناسب بوده و کارایی را پایین نیاورده است. در داده AT&T دقت شناسایی بدون اعمال تبدیل، ۱۰۰٪ بوده که بعد از اعمال تبدیل، این دقت هیچ کاهشی نداشته است. برای داده MUCT حتی افزایش بسیار جزئی در کارایی داشته‌ایم (در جدول (۱۰)).



نادرست دسته) را کمتر می‌کند؛ همچنین، افزودن داده آزمون به دسته مناسب، یعنی به دسته‌ای که داده در واقع متعلق به آن است، اثر تخریبی کمتری دارد. در اعمال الگوریتم پیشنهادی، دو نوع تبدیل معروف (فوریه و موجک) مورد استفاده قرار گرفته‌اند. به‌منظور نشان دادن توانایی روش بیان شده، کارایی آن در بهبود ویژگی‌های استخراج شده از

شبکه VGGFace که یکی از قوی‌ترین CNN‌های ارائه شده در حوزه پردازش تصاویر چهره است، نشان داده شده است؛ به علاوه، کارایی مناسب رویکرد ارائه شده، روی داده‌های چالشی (مانند LFW) و روی داده‌های آزمایشگاهی (مانند AT&T) اثبات شده، به طوری که بروی LFW، بهبودی در حدود ۳/۴٪ داشته است.

(جدول-۶): تأثیر ویژگی‌های به دست آمده از اعمال تبدیلات مختلف در کارایی الگوریتم شناسایی چهره مبتنی بر شبکه VGGFace .

(Table -6): Effect of the features obtained from applying different transformations on the efficiency of the face recognition algorithm based on the VGGFace network (R=1 and the cosine distance is considered by $\tau = 0.3$. In the wavelet transform, the DB11 has been used.)

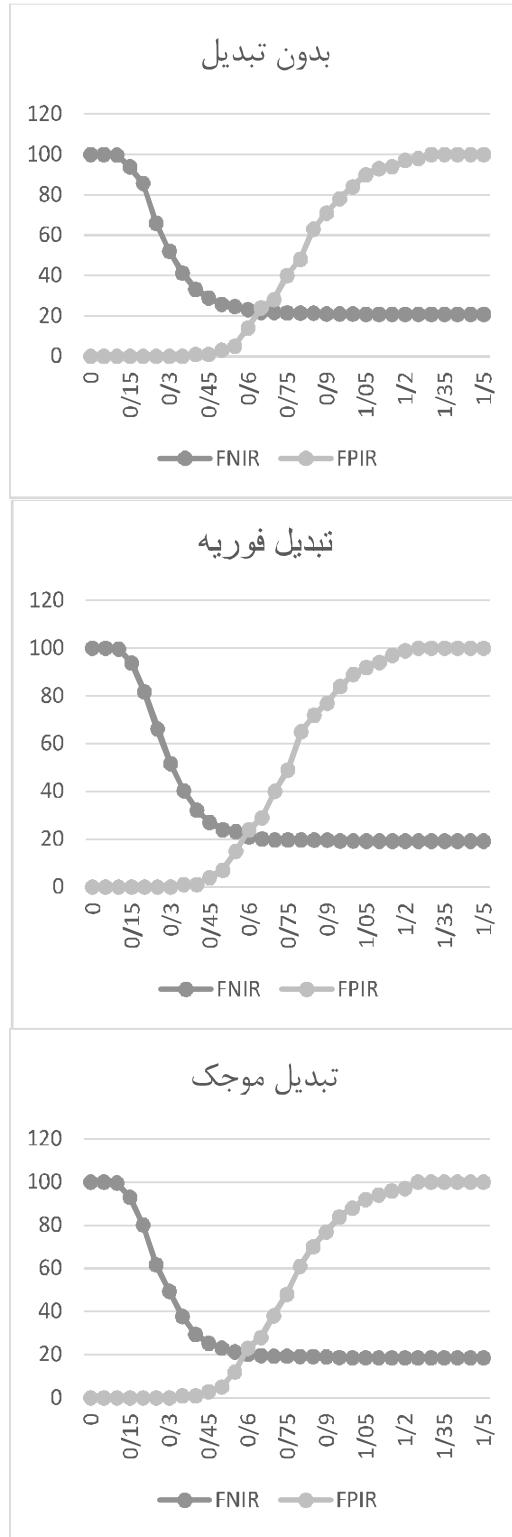
Final score	FPIR	TPIR	بدون تبدیل
100	0	100	AT&T
99.74	2	99.97	MUCT
86.52	8	86.22	LFW
90.5	0	81	100LFW
<hr/>			
Final score	FPIR	TPIR	تبدیل فوریه
100	0	100	AT&T
99.74	2	99.97	MUCT
86.57	8	86.27	LFW
90.5	0	81	100LFW
<hr/>			
Final score	FPIR	TPIR	تبدیل موجک
100	0	100	AT&T
99.74	2	99.97	MUCT
90.33	18	89.655	LFW
91.1	0	82.2	100LFW

(جدول-۷): تأثیر ویژگی‌های به دست آمده از اعمال تبدیلات مختلف در کارایی الگوریتم شناسایی چهره مبتنی بر شبکه VGGFace .

(Table -7): Effect of the features obtained from applying different transformations on the efficiency of the face recognition algorithm based on the VGGFace network (R=1 and the correlation distance is considered by $\tau = 0.5$. In the wavelet transform, the DB11 has been used.)

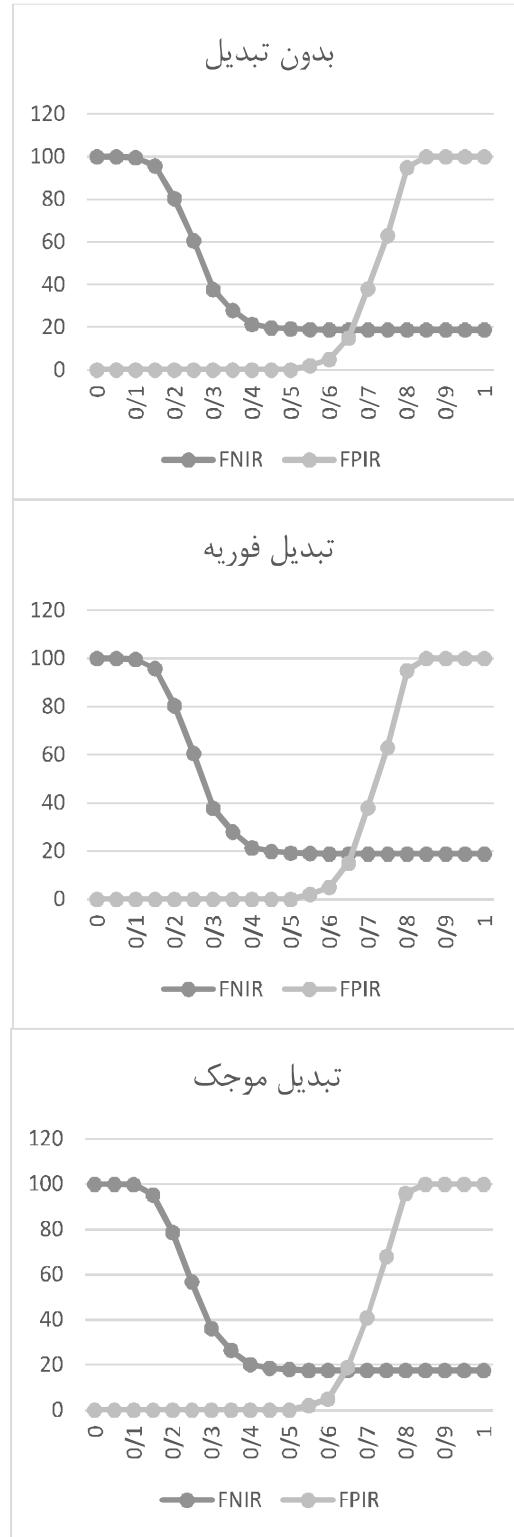
Final score	FPIR	TPIR	بدون تبدیل
100	0	100	AT&T
99.62	3	99.97	MUCT
87.28	17	87.52	LFW
90.4	0	80.8	100LFW
<hr/>			
Final score	FPIR	TPIR	تبدیل فوریه
100	0	100	AT&T
99.62	3	99.97	MUCT
87.28	17	87.52	LFW
90.4	0	80.8	100LFW
<hr/>			
Final score	FPIR	TPIR	تبدیل موجک
100	0	100	AT&T
99.51	4	99.97	MUCT
89.06	29	90.05	LFW
91	0	82	100LFW

ارائه یک روش استخراج ویژگی از تصاویر چهار مبنی بر اعمال تبدیل روی ویژگی‌های به دست آمده از شبکه‌های عصبی کانولوشن



(شکل-۶): FPIR و FNIR بهزای بیست مقدار آستانه مختلف از صفر تا ۱/۵ با گام ۰/۰۵ با فرض استفاده از فاصله اقلیدسی از صفر تا ۱/۵ با گام ۰/۰۵ با فرض استفاده از فاصله همبستگی R=1 و داده ۱۰۰LFW استفاده شده است).

(Figure-6): FPIR and FNIR for 20 different threshold values from zero to 1.5 with step 0.05 assuming the use of Euclidian distance ($R=1$ and 100LFW is used.)



(شکل-۵): FPIR و FNIR بهزای بیست مقدار آستانه مختلف از صفر تا یک با گام ۰/۰۵ با فرض استفاده از فاصله همبستگی از صفر تا یک با گام ۰/۰۵ با فرض استفاده از فاصله همبستگی R=1 و داده ۱۰۰LFW استفاده شده است).

(Figure-5): FPIR and FNIR for 20 different threshold values from zero to one with step 0.05 assuming the use of correlation distance ($R=1$ and 100LFW is used.)

(Table-11): Results of the proposed algorithm on features obtained from non-neural networks-based methods with different transformations. AT&T dataset has been considered. (R=1, Euclidian distance and DB11 wavelet transform have been used)

روش	بدون تبدیل	تبدیل فوریه	تبدیل موجک
LDA	93.5	93.5	93.75
KFA	96.25	97	97

6- References

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097-1105, 2012.
- [2] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE transactions on pattern analysis and machine intelligence*, vol. 15, no. 10, pp. 1042-1052, 1993.
- [3] R. J. Baron, "Mechanism of human facial recognition," *International Journal of Man Machine Studies*, vol. 15, pp. 137-178, 1981.
- [4] D .Ghimire, J. Lee, Z.-N. Li, and S. Jeong, "Recognition of facial expressions based on salient geometric features and support vector machines," *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 7921-7946, 2017
- [5] L. Wiskott, N. Krüger, N. Kuijper, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 775-779, 1997.
- [6] J. yves Bouguet, "Pyramidal implementation of the lucas kanade feature tracker," *Intel Corporation, Microprocessor Research Labs*, 2000.
- [7] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," in *European conference on computational learning theory*, 1995: Springer, pp. 23-37.
- [8] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513-529, 2012.
- [9] W. Ouarda, H. Trichili, A. M. Alimi, and B. Solaiman, "Face recognition based on geometric features using Support Vector Machines," in *Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of*, 2014: IEEE, pp. 89-95.
- [10] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Computer Vision and Pattern Recognition, 1991. Proceedings*

(جدول-۸): مقایسه کارایی الگوریتم ارائه شده با اعمال تبدیلات مختلف بر حسب پارامتر EER. روی داده R=1 (100LFW) ، فاصله همبستگی و تبدیل موجک DB11 مورد استفاده قرار گرفته اند.

(Table-8): Comparison of the efficiency of the proposed algorithm with different transformations according to the EER parameter over 100LFW data (R=1, correlation distance and DB11 wavelet transform have been used).

100LFW	بدون تبدیل	تبدیل فوریه	تبدیل موجک
EER	16.9	16.9	18.3

(جدول-۹): مقایسه کارایی الگوریتم ارائه شده با اعمال تبدیلات مختلف بر حسب پارامتر EER. روی داده R=1(100LFW) ، فاصله اقلیدسی و تبدیل موجک DB11 مورد استفاده قرار گرفته اند.

(Table-9): Comparison of the efficiency of the proposed algorithm with different transformations according to the EER parameter over 100LFW data (R=1, Euclidian distance and DB11 wavelet transform have been used)

100LFW	بدون تبدیل	تبدیل فوریه	تبدیل موجک
EER	22.9	22.5	21.6

(جدول-۱۰): تأثیر ویژگی های به دست آمده از اعمال تبدیلات مختلف در کارایی الگوریتم شناسایی چهره مبتنی بر شبکه، فاصله اقلیدسی و تبدیل موجک DB11 مورد استفاده قرار گرفته اند.

(Table-10): Effect of the features obtained from applying different transformations on the efficiency of the face recognition algorithm based on the VGGFace network (R=1, Euclidian distance and DB11 wavelet transform have been used)

تبدیل فوریه	TPIR
AT&T	100
MUCT	99.95
LFW	91.02
100LFW	80.8

تبدیل موجک	TPIR
AT&T	100
MUCT	99.95
LFW	91.63
100LFW	81.4

بدون تبدیل	TPIR
AT&T	100
MUCT	99.92
LFW	87.19
100LFW	79.2

(جدول-۱۱): نتایج الگوریتم ارائه شده با فرض استفاده از ویژگی های به دست آمده از روش های غیر مبتنی بر شبکه های AT&T وی باشد (R=1، فاصله اقلیدسی و تبدیل موجک DB11 مورد استفاده قرار گرفته اند)



- [22] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proceedings of 1994 IEEE workshop on applications of computer vision*, 1994: IEEE, pp. 138-142.
- [23] S. Milborrow, J. Morkel, and F. Nicolls, "The MUCT landmarked face database," *Pattern Recognition Association of South Africa*, vol. 201, no. 0, 2010.
- [24] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," in *Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*, 2008.
- [25] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015: ACM, pp. 689-692.
- [26] GTDLBench. *The Database of Faces (AT&T)* [Online]. Available: https://git-dsl.github.io/GTDLBench/datasets/att_face_dataset/.
- [27] P. Grother and M. Ngan, "Performance of face identification algorithms," *NIST Inter-agency Internal Report 8009*, 2014.
- [28] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, no. 2, pp. 179-188, 1936.
- [29] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K.-R. Mullers, "Fisher discriminant analysis with kernels," in *Neural networks for signal processing IX: Proceedings of the 1999 IEEE signal processing society workshop (cat. no. 98th8468)*, 1999: Ieee, pp. 41-48.
- مرتضی علی احمدی تحصیلات**
کارشناسی خود را در رشته علوم ریاضی در
دانشگاه قم، در سال ۱۳۹۱ و همچنین
تحصیلات کارشناسی ارشد خود را در
رشته مهندسی فناوری اطلاعات، در همان
دانشگاه، در سال ۱۳۹۳ به پایان رسانده و همچنین در حال
حاضر دانشجوی دکترای رشته مهندسی فناوری اطلاعات
است. زمینه پژوهشی و مورد علاقه وی، پردازش صدا و
تصویر و داده‌کاوی است.
نشانی رایانمۀ ایشان عبارت است از:
ma.ahmadi@qom.ac.ir



CVPR'91., IEEE Computer Society Conference on, 1991: IEEE, pp. 586-591.

- [11] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of educational psychology*, vol. 24, no. 6, p. 417, 1933.
- [12] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711-720, 1997.
- [13] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210-227, 2009.
- [14] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 372-386, 2012.
- [15] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1891-1898.
- [16] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in *Advances in neural information processing systems*, 2014, pp. 1988-1996.
- [17] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2892-2900.
- [18] Y. Sun, D. Liang, X. Wang, and X. Tang, "Deepid3: Face recognition with very deep neural networks," *arXiv preprint arXiv:1502.00873*, 2015.
- [19] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701-1708.
- [20] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in *BMVC*, 2015, vol. 1, no. 3, p. 6.
- [21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

روح الله دیانت تحصیلات کارشناسی ارشد
و دکتراخود را به ترتیب در سال‌های
۱۳۸۲ و ۱۳۸۹ در دانشگاه صنعتی شریف
به اتمام رسانده و هم‌اکنون استادیار
دانشگاه قم (گروه مهندسی کامپیوتر و فناوری اطلاعات)
است. زمینه اصلی فعالیت‌وی، پردازش چندرسانه‌ای (صدا،
تصویر و ویدئو) است.
نشانی ریاضیات ایشان عبارت است از:
rdianat@qom.ac.ir

