

## شناسایی چهره مقاوم نسبت به انسداد و تخریب بر مبنای نمایش مشارکتی و کورآنتروپی

شادی کلالی<sup>۱</sup>، یحیی فرقانی<sup>۲</sup> و مجید وفایی جهان<sup>۳</sup>

### چکیده

مدل تُنک مبتنی بر معیار شباهت کورآنتروپی، نوعی مدل طبقه‌بندی یا شناسایی چهره مبتنی بر روش نمایش تُنک است که نسبت به نویز و انسداد در داده‌های آزمون، مقاوم است. در این مدل، ترکیبی خطی از تصاویر آموزشی، به نحوی تعیین می‌شود که بر مبنای معیار کورآنتروپی، بیشترین شباهت را با داده آزمون داشته باشد و تُرم ۱ بردار ضرایب این ترکیب خطی، حداقل باشد. تُرم ۱، مشتق‌ناپذیر است و لذا، نمی‌توان برای حل این مدل، از روش‌های کارآمد مبتنی بر گرادینان استفاده کرد. برای ساده‌سازی این مدل و حل سریعتر آن با روش‌های مبتنی بر گرادینان، ضرایب ترکیب خطی، نامنفی در نظر گرفته شده است. قید نامنفی بودن ضرایب ترکیب خطی، قید محدودکننده‌ای است که در صحت طبقه‌بندی، تاثیر منفی می‌گذارد. در این مقاله، برای رفع این مشکل، بجای تُرم ۱ از تُرم ۲ بردار ضرایب ترکیب خطی، استفاده می‌شود و دو روش سریع برای حل مدل جدید ارائه می‌گردد. به تعبیر دیگر، مدل پیشنهادی، مدل شناسایی چهره مبتنی بر نمایش مشارکتی است که از مفهوم کورآنتروپی برای مقاوم شدن مدل در برابر نویز و انسداد استفاده کرده است. آزمایش‌های انجام شده نشان می‌دهد که مدل پیشنهادی، نرخ صحت طبقه‌بندی و زمان اجرای بهتری نسبت به مدل نمایش تُنک مبتنی بر کورآنتروپی با ضرایب نامنفی دارد.

**کلیدواژه‌ها:** طبقه‌بندی مبتنی بر نمایش تُنک، نمایش مشارکتی، کورآنتروپی، طبقه‌بندی مقاوم، انسداد، تخریب.

### ۱ مقدمه

اساساً اختلاف و تنوع زیاد در چهره افراد، به گونه‌ای است که نمی‌توان به راحتی چهره‌ها را در دسته‌ها و گروه‌های مشخصی طبقه‌بندی کرد. علاوه بر آن، این فرایند تحت تاثیر تغییرات زیادی است. تغییر سن باعث تغییر چهره می‌گردد. در ضمن ممکن است، تغییر در چهره به دلیل شرایط تصویربرداری باشد، این شرایط می‌تواند شامل تغییرات در شدت نور و نیز چگونگی قرار گرفتن فرد یا زاویه تصویربرداری از چهره، تغییر ژست، تاری تصویر و رزولوشن پایین باشد، تغییرات دیگری هم ممکن است از طریق تخریب، نویز و انسداد بر روی چهره<sup>۱</sup> با پوشش‌هایی نظیر کلاه، شال گردن، ماسک، عینک آفتابی، موی صورت، روسری و آرایش صورت به وجود آید، که در هر صورت، باعث مشکلات اساسی در شناسایی تصویر چهره می‌گردد. البته تمرکز این مقاله، بر روی انسداد و تخریب تصاویر چهره است.

طبقه‌بندی یا شناسایی خودکار چهره، به سبب کاربردهای فراوان آن، به یکی از موضوعات تحقیقاتی مهم در زمینه بینایی ماشین تبدیل شده است و روش‌های بسیاری برای آن ارائه گردیده است. برای شناسایی چهره ورودی، از اطلاعات موجود در بانک تصاویر استفاده می‌شود. این بانک شامل تصویر چهره افراد مختلف می‌باشد.

شناسایی خودکار چهره، فرایندی پیچیده و دشوار است.

این مقاله در بهمن‌ماه سال ۹۷ دریافت، در اردیبهشت‌ماه ۹۸ بازنگری و در مردادماه همان سال پذیرفته شد.

<sup>۱</sup> کارشناس ارشد مهندسی کامپیوتر از دانشگاه آزاد اسلامی واحد مشهد

رایانامه: [kalali.shadi@gmail.com](mailto:kalali.shadi@gmail.com)

<sup>۲</sup> گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد مشهد

رایانامه: [kalali.shadi@gmail.com](mailto:kalali.shadi@gmail.com)

رایانامه: [vafaeijahanmajid@gmail.com](mailto:vafaeijahanmajid@gmail.com)

<sup>۱</sup>Occlusion

بالا<sup>۸</sup> برده می‌شود و براساس داده‌های آموزشی استاندارد و روش KDA<sup>۹</sup> [۱۰]، فضایی تعیین می‌شود که داده‌های هم‌کلاس به هم نزدیکتر باشند و داده‌های هر کلاس از داده‌های کلاس دیگر دورتر باشد. سپس جهات اصلی تغییرات انسداد، براساس اختلاف تصاویر آموزشی استاندارد و تصاویر آموزشی دارای انسداد در فضای ویژگی با بعد بالا، به کمک روش KPCA<sup>۱۱</sup> [۱۱] تعیین می‌شود و به فضای KDA منتقل می‌شوند. در پایان، تصویر ورودی که ویژگی‌هایش با روش LBP استخراج شده است، پس از انتقال به فضای KDA، بصورت ترکیب خطی<sup>۱۲</sup> از داده‌های آموزشی در فضای KDA (دیکشنری اصلی) و جهات اصلی تغییرات انسداد که با KPCA تعیین شده است و به فضای LDA منتقل شده است (دیکشنری توسعه یافته) بیان می‌شود.

چهار انتقاد به مدل KED وارد است. انتقاد اول: در مدل KED، به مولفه‌های اصلی تغییرات انسداد، وزن مشابهی اختصاص داده شده است. در حالی که مولفه‌های اصلی تغییرات انسداد، وزن متفاوتی دارند که متناسب است با مقدار ویژه هر مولفه اصلی. انتقاد دوم: برخلاف باور مدل KED، بازسازی یک تصویر دارای انسداد، فقط با ترکیب تصاویر بدون انسداد و مولفه‌های اصلی یا جهت‌های اصلی تغییرات انسداد میسر نیست بلکه بازسازی یک تصویر دارای انسداد، با ترکیب تصاویر بدون انسداد، میانگین تغییرات انسداد و مولفه‌های اصلی یا جهت‌های اصلی تغییرات انسداد میسر می‌شود. انتقاد سوم: اهمیت و توانایی دیکشنری اصلی و دیکشنری توسعه یافته برای بازسازی تصویر آزمون، ضرورتاً یکسان نیست. در حالیکه در مدل KED، اهمیت این دو دیکشنری برای بازسازی تصاویر آزمون، یکسان در نظر گرفته شده است. انتقاد چهارم: زمان اجرای طبقه‌بندی مبتنی بر نمایش  $\mathbb{T}^k$  که در مدل KED استفاده شده است بیشتر از نمایش مشارکتی است. ما اخیراً در مدل KED توسعه یافته [۱۲]، در پاسخ به این چهار انتقاد وارد به مدل KED، مدل جدیدی ارائه دادیم.

در تمام مدل‌های تاکنون اشاره شده، برای نمایش تصویر ورودی به صورت ترکیب خطی از تصاویر آموزشی، بردار ضرایب ترکیب خطی، به نحوی تعیین می‌شود که  $\mathbb{T}^k$  دوم اختلاف تصویر ورودی و ترکیب خطی تصاویر آموزشی، حداقل شود. متأسفانه در صورت وجود انسداد یا تخریب و نویز در تصویر ورودی، این اختلاف به شدت تحت تاثیر پیکسل‌های مربوط به انسداد یا تخریب و نویز قرار می‌گیرد و تعیین صحیح بردار ضرایب ترکیب خطی را دچار مشکل می‌کند. برای رفع این مشکل، بجای محاسبه اندازه اختلاف تصویر ورودی و ترکیب خطی تصاویر آموزشی با استفاده از معیار  $\mathbb{T}^k$ ، از معیار کورآنتروپی استفاده شده است [۱۳، ۱۴]. در مدل  $\mathbb{T}^k$  مبتنی بر کورآنتروپی یا CESRC<sup>۱۱</sup>، با

یکی از مدل‌های طبقه‌بندی که توفیق زیادی در شناسایی چهره کسب کرده است مدل طبقه‌بندی مبتنی بر نمایش  $\mathbb{T}^k$  یا SRC [۱] است. ایده اصلی این مدل این است که تصویر هر چهره ورودی، به صورت ترکیب خطی  $\mathbb{T}^k$  از تصاویر چهره آموزشی قابل بیان است و بزرگترین ضرایب این ترکیب خطی متعلق به تصاویری است که در دسته یکسان با تصویر چهره ورودی قرار دارد. براین اساس، می‌توان به دسته‌ای که تصویر ورودی به آن تعلق دارد دست یافت. منظور از ترکیب خطی  $\mathbb{T}^k$ ، ترکیبی با بردار ضرایبی با حداقل  $\mathbb{T}^k$  می‌باشد.

در مدل طبقه‌بندی نمایش مشارکتی یا CRC<sup>۲</sup> [۲]، بجای حداقل کردن  $\mathbb{T}^k$  ضرایب داده‌های آموزشی، از  $\mathbb{T}^k$  استفاده می‌شود. اگر چه این روش نمی‌تواند ترکیب  $\mathbb{T}^k$  از تصاویر آموزشی بدست آورد لیکن، بصورت تحلیلی و بصورت تجربی نشان داده شده است که صحت مدل CRC، حتی بیشتر از صحت مدل SRC است و زمان اجرای آن نیز، بهتر از مدل SRC است [۳]. در مدل SACRC<sup>۴</sup> [۴] ترکیب  $\mathbb{T}^k$  و  $\mathbb{T}^k$  بردار ضرایب ترکیب خطی حداقل شده است که منجر به بهبود بیشتر صحت طبقه‌بندی شده است.

در نوع جدیدی از مدل SRC، یعنی مدل K نزدیکترین همسایه خطی محلی یا LLK<sup>۵</sup> [۵]، هنگام تعیین ترکیب خطی<sup>۱۲</sup> از تصویر ورودی، به تصاویر آموزشی که فاصله گوسی کمتری به تصویر ورودی دارند، اهمیت بیشتری داده می‌شود. صحت طبقه‌بندی این مدل، بهتر از مدل پایه SRC گزارش شده است.

در روش طبقه‌بندی نمایش  $\mathbb{T}^k$  مقاوم [۶]، تصویر ورودی به صورت مجموع بردار مجهول نویز و ترکیب خطی  $\mathbb{T}^k$  از تصاویر آموزشی بیان می‌شود. این ترکیب خطی به نحوی تعیین می‌شود که اندازه بردار نویز و ضرایب ترکیب خطی، حداقل باشد.

در روش طبقه‌بندی نمایش  $\mathbb{T}^k$  توسعه یافته یا ESRC<sup>۶</sup> [۷]، ابتدا تغییرات درون کلاسی هر چهره براساس داده‌های آموزشی استاندارد (بدون انسداد) و غیراستاندارد (با انسداد) تعیین می‌شود و سپس تصویر ورودی به صورت ترکیب خطی  $\mathbb{T}^k$  از تصاویر آموزشی استاندارد بعلاوه تغییرات درون کلاسی بیان می‌شود. دقت مدل ESRC بهتر از دقت مدل پایه SRC است و زمان اجرای آن بدتر از زمان اجرای مدل پایه SRC است [۷].

مدل طبقه‌بندی مبتنی بر یادگیری دیکشنری کرنلی توسعه یافته یا KED<sup>۸</sup> [۸]، نوع جدیدی از مدل ESRC است. در این مدل، ابتدا با استفاده از روش LBP<sup>۹</sup> [۹]، از کلیه تصاویر آموزشی استخراج ویژگی می‌شود. سپس این داده‌ها به فضای ویژگی با بعد

<sup>1</sup> Sparse Representation for Classification

<sup>2</sup> Collaborative Representation for Classification

<sup>3</sup> Sparse Augmented Collaborative Representation for Classification

<sup>4</sup> Locally Linear KNN

<sup>5</sup> Extended SRC

<sup>6</sup> Kernel Extended Dictionary Learning

<sup>7</sup> Local Binary Pattern

<sup>8</sup> High-dimensional feature space

<sup>9</sup> Kernel Discriminant Analysis

<sup>10</sup> Kernel Principle Component Analysis

<sup>11</sup> Correntropy SRC

## ۱-۲ مدل SRC

مجموعه داده‌های آموزشی  $X = [x_1, x_2, \dots, x_n]$  را در نظر بگیرید که  $x_i \in R^m$  و  $m$ ، بعد یا تعداد ویژگی‌های هر داده است. فرض بر این است که برچسب داده  $x_i$  معلوم و برابر با  $c_i \in [1, C]$  است. برای تعیین برچسب داده آزمون  $y \in R^m$  براساس مجموعه آموزشی  $X$  در مدل SRC [۱]، نخست، داده آزمون با استفاده از مدل زیر بصورت ترکیب خطی تکندهای آموزشی درمی‌آید:

$$\min_{\alpha} \|y - X\alpha\|^2 + \lambda \|\alpha\|_1, \quad (1)$$

که  $\lambda \geq 0$  ابرپارامتر از پیش تنظیم شده مدل است. هر چه  $\lambda$  بزرگتر باشد اهمیت جمله دوم مدل نسبت به جمله اول مدل، بیشتر می‌شود. جمله اول مدل، داده آزمون  $y$  را بصورت ترکیب خطی داده‌های آموزشی درمی‌آورد که  $\alpha_i$  ضریب  $i$ -امین داده آموزشی در این ترکیب خطی است. از آنجا که بی‌نهایت ترکیب خطی مختلف از داده آزمون  $y$  وجود خواهد داشت، جمله دوم مدل، از بین این ترکیب‌های خطی، ترکیب تُنکی را انتخاب می‌کند. در واقع مقدار  $\lambda$ ، میزان تُنک بودن این ترکیب خطی را کنترل می‌کند. انتظار می‌رود که در صورت تنظیم صحیح ابرپارامتر  $\lambda$ ، داده آزمون  $y$ ، بصورت ترکیب خطی تعداد اندکی از داده‌های آموزشی هم‌کلاس با خودش درآید. مدل (۱) را می‌توان به روش‌های مختلفی چون FISTA [۲۰] حل کرد. پس از به دست آوردن بردار ضرایب بهینه، برچسب نمونه آزمون به روش زیر تعیین می‌شود:

فرض کنید  $\delta_c(\alpha)$  برداری باشد که درایه  $i$ -ام آن برابر با  $\alpha_i$  است اگر  $c_i$ ، برابر با  $c$  باشد و در غیر این صورت، مقدار درایه  $i$ -ام آن، برابر با صفر است. برچسب داده آزمون  $y$  برابر با برچسب کلاسی می‌شود که مقدار باقیمانده زیر را حداقل کند:

$$r_c(y) = \|y - X\delta_c(\alpha)\|^2. \quad (2)$$

## ۲-۲ مدل CRC

مدل CRC [۲] بصورت زیر می‌باشد:

$$\min_{\alpha} F = \|y - X\alpha\|^2 + \lambda \|\alpha\|^2. \quad (3)$$

تابع هدف مدل CRC برخلاف تابع هدف مدل SRC، مشتق‌پذیر است. لذا، برای حل مدل (۳)، کفایت معادله زیر حل شود:

$$\frac{\partial F}{\partial \alpha} = 0. \quad (4)$$

داریم:

$$0 = \frac{\partial F}{\partial \alpha} = -2X^T(y - X\alpha) + 2\lambda\alpha,$$

$$\rightarrow \alpha = Zy, \quad (5)$$

که در آن،

$$Z = (X^T X + \lambda I)^{-1} X^T. \quad (6)$$

هر یک از پیکسل‌های تصویر به‌طور متفاوت رفتار می‌شود و بدین طریق، تأکید بیشتری بر روی پیکسل‌هایی دارد که مرتبط با نمونه-های آموزش همان کلاس تصویر ورودی هستند. به این معنا که اگر در تصویر ورودی، خطای تخریب و انسداد وجود داشته باشد، پیکسل‌های تخریب و مسدود شده، سهم کمتری در اندازه اختلاف تصویر ورودی و ترکیب خطی تصاویر آموزشی خواهند داشت و لذا، سهم چنین پیکسل‌هایی در ارزیابی برچسب تصویر ورودی نیز کمتر می‌شود. در این مدل، ترکیب خطی تصاویر آموزشی، به نحوی تعیین می‌شود که نُرم ۱ بردار ضرایب ترکیب خطی، حداقل باشد. البته در مرجع [۱۳]، برای کاهش زمان اجراء، قید نامنفی بودن ضرایب خطی نیز در نظر گرفته شده است تا نُرم ۱ بردار ضرایب ترکیب خطی به مجموع ضرایب ترکیب خطی تبدیل شود و مدل ساده‌تری بدست آید. البته قید نامنفی بودن ضرایب ترکیب خطی، قید محدودکننده‌ای است که در صحت طبقه‌بندی مدل، تأثیر منفی می‌گذارد. در این مقاله، برای رفع این مشکل، ترکیب خطی تصاویر آموزشی، به نحوی تعیین می‌شود که نُرم ۲ بردار ضرایب ترکیب خطی، حداقل باشد. به تعبیر دیگر، مدل پیشنهادی، مدلی مبتنی بر مدل CRC است که از مفهوم کورآنتروپی برای مقاوم شدن مدل در برابر نویز و انسداد استفاده کرده است. دو الگوریتم سریع برای حل مدل پیشنهادی ارائه می‌شود که زمان اجرای بهتری نسبت به الگوریتم حل مدل نمایش تُنک مبتنی بر کورآنتروپی [۱۳] دارد و صحت طبقه‌بندی آن نیز بهتر از مدل نمایش تُنک مبتنی بر کورآنتروپی با ضرایب نامنفی است.

از بین مدل‌های مقاوم یادشده، از مدل‌های مقاوم ESRC، KED و KED توسعه‌یافته زمانی می‌توان استفاده کرد که در مجموعه داده‌های آموزشی، داده‌های بدون انسداد و با انسداد موجود باشد درحالی‌که، در مدل مقاوم CESRC و مدل مقاوم پیشنهادی، نیازی به وجود داده‌های دارای انسداد در مرحله آموزش نیست.

درضمن، لازم به ذکر است که معیار کورآنتروپی برای مقابله با نویز در مسائل مختلف دیگری همچون خوشه‌بندی [۱۵]، استخراج ویژگی [۱۶]، یادگیری رتبه-پایین [۱۷، ۱۸] و تجزیه نامنفی ماتریس [۱۹] نیز بکارگرفته شده است.

در ادامه، در بخش ۲، مدل SRC، CRC و CESRC با تفصیل بیشتری شرح داده می‌شود. سپس، در بخش ۳، مدل پیشنهادی ارائه می‌شود. در بخش ۴، نتایج آزمایش‌ها گزارش می‌شود و در بخش ۵، نتیجه‌گیری می‌شود.

## ۲ چند مدل طبقه‌بندی مبتنی بر نمایش

در این بخش، نخست مدل‌های طبقه‌بندی SRC، CRC و سپس مدل طبقه‌بندی مقاوم نسبت به انسداد CESRC معرفی می‌شود. مدل طبقه‌بندی CESRC، مدلی مبتنی بر مدل SRC است و مدل پیشنهادی این مقاله، مبتنی بر مدل SRC است.

و برای مقدار ثابت  $z$ ، مقدار بیشینه در نقطه  $p = -g(z)$  حاصل خواهد شد.

با جایگزینی رابطه (۹) در مدل (۸)، مدل زیر حاصل می‌شود:

$$J_{CESR} = \max_{a,p} \sum_{j=1}^m \left( p_j (y_j - \sum_{i=1}^n X_{ij} a_i)^2 - \varphi(p_j) \right) - \lambda \sum_{i=1}^n a_i$$

subject to  $a_i \geq 0, i = 1, 2, \dots, n.$

(۱۰)

که  $p = [p_1, \dots, p_m]^T$  متغیر کمکی است. بنا به قضیه ۱، در صورت ثابت فرض شدن مقدار  $a$ ، مقدار بهینه  $p_j$  با استفاده از رابطه زیر بدست می‌آید:

$$p_j^{t+1} = -g(y_j - \sum_{i=1}^n X_{ij} a_i^t). \quad (11)$$

که  $t$ ، به تکرار  $t$ -ام اشاره می‌کند. بر اساس، مدل (۱۰)، در صورت ثابت فرض شدن مقدار  $p$ ، مقدار بهینه  $a$  با استفاده از رابطه زیر بدست می‌آید:

$$a^{t+1} = \arg \max_a (y - Xa)^T \text{diag}(p^{t+1})(y - Xa) - \lambda \sum_i a_i, \text{ subject to } \forall i: a_i \geq 0, \quad (12)$$

که  $\text{diag}(\cdot)$  عملگری است که بردار را تبدیل به ماتریس قطری می‌کند. بنابراین، می‌توان از روش تناوبی برای حل مدل (۱۰) استفاده کرد یعنی، بصورت متناوب مقدار بهینه  $p$  و  $a$  با استفاده از روابط (۱۱) و (۱۲) تعیین می‌شود تا همگرایی حاصل شود.

مسئله بهینه‌سازی (۱۲) یک مسئله حداقل مربعات خطی وزن‌دار با قید غیرمنفی است و  $-p^{t+1}$  به عنوان وزن در این مسئله ظاهر شده است. با تبدیل متغیر  $\hat{X} = \text{diag}(\sqrt{-p^{t+1}})X$  و  $\hat{y} = \text{diag}(\sqrt{-p^{t+1}})y$  مسئله بهینه‌سازی محدب (۱۳) بازنویسی کرد.

$$\min_a \left( \frac{\lambda}{2} - \hat{X}^T \hat{y} \right)^T a + \frac{1}{2} a^T \hat{X}^T \hat{X} a, \quad \text{s.t. } \forall i: a_i \geq 0. \quad (13)$$

از آنجایی که  $\hat{X}^T \hat{X}$  نیمه معین مثبت است، مسئله (۱۳) محدب است و می‌توان جواب بهینه آنرا با روش‌های حل مسائل بهینه‌سازی محدب بدست آورد. نهایتاً، پس از به دست آوردن بردار ضرایب پراکنده مبتنی بر کورآنتروپی با استفاده از روش تکراری یادشده، برچسب نمونه آزمون به روش زیر تعیین می‌شود:

فرض کنید  $\delta_c(\alpha)$  برداری باشد که درایه  $i$ -ام آن برابر با  $\alpha_i$  است اگر  $c_i$  برابر با  $c$  باشد و در غیر این صورت، مقدار درایه  $i$ -ام آن، برابر با صفر است. برچسب داده آزمون  $\gamma$  برابر با برچسب کلاسی می‌شود که مقدار کورآنتروپی باقیمانده، یعنی مقدار زیر را بیشینه کند:

$$r_c(y) = g(y - X\delta_c(x)). \quad (14)$$

در صورتی که قید نامنفی بودن ضرایب در نظر گرفته نمی‌شود، مدل (۱۳) بصورت زیر درمی‌آید [۱۴]:

$$\min_a \frac{\lambda}{2} \|a\|_1 + \hat{y}^T \hat{X} a + \frac{1}{2} a^T \hat{X}^T \hat{X} a, \quad (15)$$

در این صورت، با توجه به مشتق‌ناپذیر بودن  $\|\cdot\|_1$ ، برای حل مدل (۱۵)، باید از الگوریتمی مثل الگوریتم FISTA استفاده کرد.

از آنجا که ماتریس  $Z$  مستقل از داده آزمون است محاسبه رابطه (۶)، فقط یکبار و در مرحله آموزش انجام می‌شود و برای محاسبه  $\alpha$ ، کفایت حاصلضرب ماتریس  $Z$  در بردار آزمون  $\gamma$  محاسبه شود. در حالی که در مدل SRC، برای محاسبه  $\alpha$  به‌ازای هر داده آزمون  $\gamma$ ، مدل SRC باید مجدداً با استفاده از الگوریتمی مثل الگوریتم FISTA حل شود. لذا، روش CRC سریعتر از روش SRC است.

## ۲-۳ مدل طبقه‌بندی مقاوم نسبت به نویز و انسداد

### CESRC

مدل طبقه‌بندی نمایش تک مبتنی بر کورآنتروپی یا CESRC [۱۳، ۱۴] به صورت زیر است:

$$J_{CESR} = \max_a \sum_{j=1}^m g(y_j - \sum_{i=1}^n a_i X_{ij}) - \lambda \|a\|_1, \quad (V)$$

که در آن  $g(z) = \exp(-\frac{z^2}{2\sigma^2})$  تابع کرنل گوسی است و  $\sigma > 0$  پارامتر عرض کرنل است. مقدار پارامتر عرض کرنل، در کارایی الگوریتم تاثیر بسزایی دارد و اغلب، به صورت تجربی تعیین می‌شود. اگر  $\sigma$  کوچک باشد سرعت کاهش  $g(y_j - \sum_{i=1}^n a_i X_{ij})$  به‌ازای افزایش  $(y_j - \sum_{i=1}^n a_i X_{ij})^2$  بیشتر می‌شود. در صورتی که مقدار  $(y_j - \sum_{i=1}^n a_i X_{ij})^2$  بزرگ باشد، پیکسل  $j$ -ام تاثیر کمی در تابع هدف مدل (۷) خواهد داشت و لذا، تاثیر کمی در تعیین بردار ضرایب ترکیب خطی و تعیین برچسب تصویر آزمون خواهد داشت. بزرگ بودن مقدار  $(y_j - \sum_{i=1}^n a_i X_{ij})^2$  به این معنی است که فاصله پیکسل  $j$ -ام تصویر آزمون از ترکیب خطی پیکسل‌های متناظرش در تصاویر آموزشی زیاد است. هرچه این فاصله بیشتر باشد احتمال نویزی بودن پیکسل  $j$ -ام تصویر آزمون بیشتر می‌شود. در واقع، می‌توان چنین تعبیر کرد که  $g(y_j - \sum_{i=1}^n a_i X_{ij})$  برابر با احتمال نویزی نبودن پیکسل  $j$ -ام است. بر اساس این تعبیر، تابع  $g(\cdot)$ ، تابع توزیع احتمال نویزی نبودن پیکسل است و پارامتر  $\sigma$ ، مقدار انحراف معیار آن است.

حل مسئله کمینه‌سازی نرم یک با کارایی مناسب، یکی از زمینه‌های تحقیقاتی می‌باشد [۲۰]. اگر در مسئله بهینه‌سازی (۷)، عبارت نرم یک موجود نبود، می‌توانستیم آن را توسط روش‌های مبتنی بر گرادیان حل کنیم. در مرجع [۱۳]، با اعمال قید غیرمنفی بودن بردار ضرایب  $a$ ، مسئله (۷)، هموار شده و بصورت زیر درآمده است:

$$J_{CESR} = \max_a \sum_{j=1}^m g(y_j - \sum_{i=1}^n X_{ij} a_i) - \lambda \sum_{i=1}^n a_i, \quad (8)$$

s.t.  $a_i \geq 0, i = 1, 2, \dots, n.$  سپس، برای حل مدل (۸)، از تکنیک بهینه‌سازی نیمه محدب و حداکثر امید استفاده شده است که در ادامه، شرح داده می‌شود.

قضیه ۱ [۲۱]: یک تابع مزدوج محدب  $\varphi$  از  $g(x)$  وجود دارد، به طوری که:

$$g(z) = \max_p \left( p \frac{\|z\|^2}{\sigma^2} - \varphi(p) \right), \quad (9)$$

### ۳ مدل پیشنهادی

در این بخش، نخست انگیزه ارائه مدل پیشنهادی مطرح می‌شود. سپس، مدل پیشنهادی، که مدل طبقه‌بندی مقاوم نسبت به نویز و انسداد و مبتنی بر مدل CRC است ارائه می‌شود و الگوریتمی برای حل آن ارائه می‌شود. نهایتاً در مورد اهمیت مدل پیشنهادی، بحث می‌شود.

### ۳-۱ انگیزه ارائه مدل پیشنهادی

همانطور که ملاحظه شد در مدل CESRC [۱۳, ۱۴]، ترکیبی خطی از تصاویر آموزشی به نحوی تعیین می‌شود که براساس معیار کورآنتروپی، به داده آزمون شبیه باشد و نرم بردار ضرایب ترکیب خطی، حداقل باشد. البته در مرجع [۱۳]، برای کاهش زمان اجراء، قید نامنفی بودن ضرایب خطی نیز در نظر گرفته شده است تا نرم بردار ضرایب ترکیب خطی به مجموع ضرایب ترکیب خطی تبدیل شود و مدل ساده‌تری بدست آید. قید نامنفی بودن ضرایب ترکیب خطی، قید محدودکننده‌ای است که در صحت طبقه‌بندی مدل، تاثیر منفی می‌گذارد. در این مقاله، برای رفع این مشکل، ترکیب خطی تصاویر آموزشی، به نحوی تعیین می‌شود که نرم بردار ضرایب ترکیب خطی، حداقل باشد. در این صورت، با توجه به مشتق‌پذیر بودن نرم ۲، نیازی به استفاده از قید نامنفی بودن ضرایب، برای ساده‌سازی مدل پیشنهادی نیست. در واقع، در این مقاله، پس از استفاده از نرم ۲ بجای نرم ۱ در مدل CESRC، دو الگوریتم سریع و سریع‌تر برای حل مدل پیشنهادی ارائه می‌شود. نوآوری اصلی این مقاله، ارائه همین الگوریتم سریع‌تر است.

### ۳-۲ معرفی مدل پیشنهادی و روش حل آن

مدل پیشنهادی که CECRC نامیده می‌شود بصورت زیر می‌باشد:

$$J_{CECR-RLS} = \max_a \sum_{j=1}^m g(y_j - \sum_{i=1}^n X_{ij} a_i) - \lambda \|a\|_2^2, \quad (۱۶)$$

بنا به قضیه (۱)، مدل (۱۶) را می‌توان بصورت زیر نوشت:

$$\hat{J}_{CECR-RLS} = \max_{a,p} \sum_{j=1}^m \left( p_j (y_j - \sum_{i=1}^n X_{ij} a_i)^2 - \varphi(p_j) \right) - \lambda \|a\|_2^2 \quad (۱۷)$$

که متغیر  $p = [p_1, \dots, p_m]^T$  یک بردار متغیر کمکی است. بنا به قضیه ۱، در صورت ثابت فرض شدن مقدار  $a$ ، مقدار بهینه  $p_j$  با استفاده از رابطه زیر بدست می‌آید:

$$p_j^{t+1} = -g(y_j - \sum_{i=1}^n X_{ij} a_i^t), \quad (۱۸)$$

که  $t$  بیانگر  $t$ -امین تکرار است. در صورت ثابت فرض شدن مقدار  $p$  در مدل (۱۷)، مقدار بهینه  $x$  با استفاده از رابطه زیر بدست می‌آید:

$$a^{t+1} = \arg \max_a (y - Xa)^T \text{diag}(p)(y - Xa) - \lambda a^T a, \quad (۱۹)$$

که  $\text{diag}(\cdot)$  عملگری است که بردار را تبدیل به ماتریس قطری می‌کند. بنابراین، می‌توان از روش تناوبی برای حل مدل (۱۷)

استفاده کرد یعنی، بصورت متناوب مقدار بهینه  $p$  و  $a$  با استفاده از روابط (۱۸) و (۱۹) تعیین می‌شود تا همگرایی حاصل شود. با معلوم فرض شدن مقدار بردار ضرایب  $a$ ، مساله (۱۸)، یک مقداردهی و انتساب ساده برای تعیین مقدار بردار وزن  $p$  است. مساله بهینه‌سازی (۱۹) نیز یک مساله بهینه‌سازی کمترین مربعات خطای وزن دار (با بردار وزن  $-p$ ) است. ابتدا مساله بهینه‌سازی (۱۹) را به صورت مساله بهینه‌سازی محدب (۲۰) بازنویسی می‌کنیم:

$$\max_a a^T X^T \text{diag}(p) X a + y^T \text{diag}(p) y - 2a^T X^T \text{diag}(p) y - \lambda a^T a \quad (۲۰)$$

با استفاده از تغییر متغیر  $\hat{X} = \text{diag}(\sqrt{-p^{t+1}}) X$  و  $\hat{y} = \text{diag}(\sqrt{-p^{t+1}}) y$  بصورت زیر درمی‌آید:

$$a^T \hat{X}^T \hat{X} a + \hat{y}^T \hat{y} - 2a^T \hat{X}^T \hat{y} - \lambda a^T a = a^T (\hat{X}^T \hat{X} + \lambda I) a + \hat{y}^T \hat{y} - 2a^T \hat{X}^T \hat{y}. \quad (۲۱)$$

بنابراین، مساله بهینه‌سازی (۲۰) را به صورت مساله (۲۲) بازنویسی می‌کنیم:

$$\min_a f = a^T (\hat{X}^T \hat{X} + \lambda I) a - 2a^T \hat{X}^T \hat{y} + \hat{y}^T \hat{y}. \quad (۲۲)$$

برای کمینه‌سازی مساله (۲۲) باید بصورت زیر از تابع هدف آن مشتق گرفته و برابر با صفر قرار دهیم:

$$0 = \frac{\partial f}{\partial a} = 2(\hat{X}^T \hat{X} + \lambda I) a - 2\hat{X}^T \hat{y} + 0 = 2((\hat{X}^T \hat{X} + \lambda I) a - \hat{X}^T \hat{y}). \quad (۲۳)$$

در نتیجه، مقدار بهینه  $a$ ، با حل دستگاه معادلات خطی (۲۳) و طبق رابطه زیر بدست می‌آید:

$$a = (\hat{X}^T \hat{X} + \lambda I)^{-1} \hat{X}^T \hat{y}. \quad (۲۴)$$

با توجه به اینکه  $\hat{X}^T \hat{X}$  یک ماتریس نیمه‌معین مثبت است، بنابراین،  $(\hat{X}^T \hat{X} + \lambda I)$  دارای معکوس است. لذا، روابط (۱۸) و (۱۹) که برای حل مدل CECRC پیشنهاد شد به صورت زیر بازنویسی می‌شود:

$$\begin{cases} p_j^{t+1} = -g(y_j - \sum_{i=1}^n X_{ij} a_i^t), & j = 1, 2, \dots, m; \\ a^{t+1} = (\hat{X}^T \hat{X} + \lambda I)^{-1} \hat{X}^T \hat{y}. \end{cases} \quad (۲۵)$$

البته اگر تعداد ویژگی‌ها یا تعداد پیکسل‌های تصاویر چهره ( $m$ )، خیلی کمتر از تعداد تصاویر ( $n$ ) باشد برای حل دستگاه معادلات (۲۳)، روش سریع‌تری پیشنهاد می‌شود. در چنین شرایطی، اگر دستگاه معادلات (۲۳) در ماتریس  $\hat{X}$  ضرب شود، دستگاه معادلات زیر بدست می‌آید:

$$(\hat{X} \hat{X}^T \hat{X} + \lambda \hat{X}) a - \hat{X} \hat{X}^T \hat{y} = 0. \quad (۲۶)$$

لذا،

$$(\hat{X} \hat{X}^T + \lambda I) \hat{X} a = \hat{X} \hat{X}^T \hat{y}. \quad (۲۷)$$

یا

$$(\hat{X} \hat{X}^T + \lambda I) \hat{a} = \hat{X} \hat{X}^T \hat{y}. \quad (۲۸)$$

که در آن،

$$\hat{a} = \hat{X}a. \quad (29)$$

با توجه به رابطه (۲۸) داریم:

$$\hat{a} = (\hat{X}\hat{X}^T + \lambda I)^{-1} \hat{X}\hat{X}^T \hat{y}. \quad (30)$$

پس از حل دستگاه (۲۸) و بدست آوردن  $\hat{a}$ ، مقدار  $a$  طبق رابطه (۲۹) به صورت زیر تعیین میشود:

$$a = \hat{X}^+ \hat{a}, \quad (31)$$

که  $\hat{X}^+$ ، شبه معکوس ماتریس  $\hat{X}$  می باشد که ماتریسی  $m \times n$  است. سرانجام، روابط (۱۸) و (۱۹) که برای حل مدل CECRC پیشنهاد شد به صورت زیر بازنویسی می شود:

$$\begin{cases} p_j^{t+1} = -g(y_j - \sum_{i=1}^n X_{ij} a_i^t), & j = 1, 2, \dots, m; \\ \hat{a}^{t+1} = (\hat{X}\hat{X}^T + \lambda I)^{-1} \hat{X}\hat{X}^T \hat{y}. \\ a^{t+1} = \hat{X}^+ \hat{a}^{t+1}. \end{cases} \quad (32)$$

دستگاه معادلات (۲۳)، یک دستگاه معادلات با  $n$  معادله و  $n$  مجهول است که برای حل آن، از روش حذف گوسی یا معکوس گیری از یک ماتریس  $n \times n$  مطابق با رابطه (۲۵) استفاده می شود. در حالی که، در صورت استفاده از رابطه (۳۲) بجای رابطه (۲۵) برای حل دستگاه معادلات (۲۳)، نیاز به محاسبه شبه معکوس یک ماتریس  $n \times m$  و حل دستگاه معادلات (۲۸) با  $m$  معادله و  $m$  مجهول است. اگر  $m$  خیلی کمتر از  $n$  باشد استفاده از رابطه (۳۲) برای حل مدل پیشنهادی، سریع تر از رابطه (۲۵) است چراکه حل دستگاه معادلات بسیار بزرگ  $n$  معادله و  $n$  مجهول یا محاسبه معکوس ماتریس بسیار بزرگ  $n \times n$  در رابطه (۲۵) بسیار زمان بر است.

الگوریتم ۲، فرآیند بهینه سازی مدل پیشنهادی CECRC و تعیین مقدار بهینه بردار ضرایب  $a$  را به طور خلاصه بیان می کند. پس از به دست آوردن بردار ضرایب  $a$ ، برچسب نمونه آزمون، دقیقاً مشابه با روش CESRC و به صورت زیر تعیین می شود:

فرض کنید  $\delta_c(\alpha)$  برداری باشد که درایه  $i$ -ام آن برابر با  $\alpha_i$  است اگر  $c_i$ ، برابر با  $c$  باشد و در غیر این صورت، مقدار درایه  $i$ -ام آن، برابر با صفر است. برچسب داده آزمون  $y$  برابر با برچسب کلاسی می شود که مقدار کورآنتروپی باقیمانده، یعنی مقدار زیر را حداکثر کند:

$$r_c(y) = g(y - X\delta_c(x)). \quad (33)$$

الگوریتم ۳، نحوه تعیین برچسب داده آزمون را بطور خلاصه بیان می کند.

الگوریتم ۲: الگوریتم حل مدل پیشنهادی CECRC.

ورودی:

- مجموعه نمونه های آموزشی در قالب ماتریس  $X$  و نمونه آزمون در قالب بردار  $y$
- پارامتر تنظیم  $\lambda$  و پارامتر هسته  $\sigma$
- پارامترهای شرط توقف ( $thr_a$  و  $iter_{max}$ )

خروجی:

- بردار ضرایب  $a$

۱- مقداردهی اولیه بردار ضرایب  $a$  به صورت تصادفی ( $a^1$ )

۲- حلقه تکرار با اندیس تکرار  $t = 0$

۱-۲- محاسبه بردار وزن  $p^{t+1}$  بر اساس رابطه (۲۵).

۲-۲- محاسبه  $\hat{X} = \text{diag}(\sqrt{-p^{t+1}})X$  و  $\hat{y} = \text{diag}(\sqrt{-p^{t+1}})y$

۲-۳- اگر تعداد ویژگی ها یا تعداد پیکسل های تصاویر چهره ( $m$ )، خیلی کمتر از تعداد تصاویر ( $n$ ) است بردار ضرایب  $a^{t+1}$  بر اساس رابطه (۳۲) محاسبه کن درغیراین صورت، از رابطه (۲۵) استفاده کن.

۲-۴-  $t = t + 1$

۲-۵- اگر  $\|a^{t+1} - a^t\|_2 \leq thr_a$  یا  $t == iter_{max}$ ، برو به پایان.

۳- پایان.

الگوریتم ۳: روش دسته بندی CECRC

ورودی:

- ماتریس داده های آموزش  $X \in \mathbb{R}^{m \times n}$  از کلاس  $C$
- داده آزمون  $y \in \mathbb{R}^m$
- مقدار پارامتر تنظیم تعمیم پذیری  $\lambda$  و پارامتر هسته  $\sigma$
- خروجی: برچسب داده آزمون  $y$

۱- نرمالیزه کرده داده های آموزشی.

۲- به دست آوردن بردار ضرایب بهینه  $a$  با استفاده از الگوریتم ۱.

۳- محاسبه کورآنتروپی باقی مانده، یعنی  $r_i(y)$ ، به ازای تمام کلاس ها  $C$ ،  $i = 1, 2, \dots, C$ ، مطابق با رابطه (۲۹).

۴- برچسب داده آزمون  $y$  برابر با برچسب کلاسی می شود که حداکثر کورآنتروپی باقی مانده را داشته باشد.

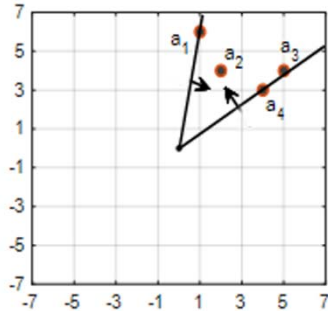
### ۳-۳ بحث در مورد اهمیت روش پیشنهادی و

#### محدودیت قید نامنفی بودن ضرایب

مدل CESRC مدلی مبتنی بر SRC است که از مفهوم کورآنتروپی برای مقاوم شدن مدل SRC نسبت به نویز و انسداد استفاده کرده است. در مدل CESRC، از نرم ۱ بردار ضرایب، استفاده شده است. از آنجا که نرم ۱، مشتق پذیر نیست امکان حل آن با روش های سریع مبتنی بر گرادیان میسر نبوده است و لذا، در مرجع [۱۳]، بردار ضرایب مدل CESRC نامنفی فرض شده است تا جمله نرم ۱ از تابع هدف مدل CESRC حذف شود و مدل ساده تری بدست آید.

در این مقاله، مدل CECRC مبتنی بر CRC ارائه شد که از مفهوم کورآنتروپی برای مقاوم شدن مدل CRC نسبت به نویز و

مدل CECRC قید مناسبی نیست. لذا، در مدل پیشنهادی این مقاله نیز، قید نامنفی بودن ضرایب لحاظ نشد.



شکل ۱- فضای پوش نامنفی ۴ داده در فضای دو بُعدی (هر داده، با یک نقطه توپر نشان داده شده است).

#### ۴ نتایج تجربی

در این بخش، با استفاده از مجموعه داده‌های استاندارد AR [۲۴]، روش پیشنهادی CECRC با روش‌های CESRC [۱۳]، SRC [۱]، CRC [۲] و SACRC [۴] مقایسه می‌شود. لازم به ذکر است که از بین مدل‌های شناسایی چهره مقاوم، از مدل‌های مقاوم ESRC، KED و KED توسعه یافته زمانی می‌توان استفاده کرد که در مجموعه داده‌های آموزشی، داده‌های بدون انسداد، با انسداد، بدون نویز و با نویز موجود باشد در حالی که، در مدل مقاوم CESRC و مدل مقاوم پیشنهادی، نیازی به وجود داده‌های دارای انسداد و نویز در مرحله آموزش نیست. از آنجا که در این مقاله، فرض شده است که داده‌های آموزشی، فاقد انسداد و نویز هستند، امکان مقایسه روش پیشنهادی با روش‌هایی مثل روش ESRC، KED و KED توسعه یافته (که قبلاً پیشنهاد داده‌ایم) نیست. در ادامه این بخش، نخست، مشخصات مجموعه داده AR ارائه می‌شود و نحوه پیکربندی مدل‌های شناسایی چهره و نحوه پیکربندی هر آزمایش ذکر می‌شود. سرانجام، نتایج حاصل از هر آزمایش ارائه می‌شود و مورد بررسی قرار می‌گیرد.

#### ۴-۱ معرفی مجموعه داده‌ها

##### ۴-۱-۱- مجموعه داده AR

مجموعه داده AR، شامل تصاویر ۱۱۶ فرد می‌باشد. از هر فرد، مجموعاً ۱۳ تصویر در جلسه اول و ۱۳ تصویر در جلسه دوم گرفته شده است. ابعاد هر تصویر، ۷۶۸×۵۷۶ می‌باشد. قبل از طبقه‌بندی یا شناسایی چهره، باید تصاویر چهره، به درستی تنظیم شده<sup>۱</sup> و برش زده شوند. در این مقاله، همانند بسیاری از مقالات مرتبط، از نشانه‌های اختصاصی چهره برای تنظیم تصاویر چهره استفاده شد. به بیان دقیق‌تر، ابتدا روش کشف چهره Viola-

انسداد استفاده کرد. در مدل CECRC، از نرم ۲ بردار ضرایب، استفاده شد. با توجه به مشتق‌پذیر بودن نرم ۲، نیازی به استفاده از قید نامنفی بودن ضرایب، برای ساده‌سازی مدل پیشنهادی نبود و دو الگوریتم مبتنی بر گرادینت سریع برای حل مدل پیشنهادی ارائه شد.

در ادامه، در مورد اثر نامطلوب نامنفی فرض شدن بردار ضرایب در مدل CESRC بحث می‌شود. قبلاً اثبات شده است که تصاویر یک فرد تحت نورپردازی‌های متفاوت و حالت‌های مختلف ژست صورت، بر روی یک زیرفضای با بعد کم که زیرفضای صورت نامیده می‌شود قرار می‌گیرند [۲۲، ۲۳]. چهار تصویر صورت از یک فرد (یک کلاس) را در نظر بگیرید. بدون از دست دادن کلیت بحث، فرض کنید که این چهار تصویر صورت، بر روی یک زیر فضای دو بُعدی قرار داشته باشند. به عبارت دیگر، زیر فضای صورت را یک زیر فضای دو بُعدی در نظر بگیرید. در شکل ۱، هر یک از این تصاویر صورت، با یک نقطه در زیر فضای یاد شده نشان داده شده است. فضای پوش نامنفی (یا ترکیب خطی نامنفی) این تصاویر یا این نقاط، یک قطاع از این فضای دو بُعدی است که توسط حاشیه‌ای‌ترین نقاط در راستای مبدا مختصات ایجاد می‌شود. این قطاع در شکل ۱، توسط فلش‌هایی مشخص شده است. همانطور که مشاهده می‌شود، فضای پوش این نقاط، قسمت محدودی از زیر فضای صورت است. به عبارت دیگر، داده‌های خارج از قطاع یاد شده که در زیر فضای صورت هستند را نمی‌توان بصورت ترکیب خطی داده‌های داخل قطاع نوشت. از طرفی، همانطور که قبلاً گفته شد تصاویر یک فرد تحت نورپردازی‌های متفاوت و حالت‌های مختلف ژست صورت، بر روی یک زیر فضای صورت قرار می‌گیرند [۲۲، ۲۳]. بنابراین، هر داده تست هم کلاس با چهار داده نشان داده شده در شکل ۱ یا هر تصویر صورت دیگر از همان فرد مورد بحث، در فضای دربرگیرنده این چهار داده است. در حالی که، اگر این داده تست، نقطه‌ای خارج از قطاع یاد شده ولی در فضای صورت این فرد باشد نمی‌تواند به درستی با ترکیب خطی از داده‌های هم کلاسش بازسازی شود و خطای بازسازی آن، یعنی مقدار جمله اول مدل CESRC با قید نامنفی ضرایب، زیاد خواهد شد. در این صورت، از آنجا که مدل CESRC، به دنبال حداقل کردن خطای بازسازی است، ممکن است از داده‌های کلاس دیگر، برای کاهش بازسازی داده تست استفاده شود. لذا، از آنجا که داده‌های مشارکت کننده در بازسازی داده تست، تعیین کننده برچسب داده تست هستند، احتمال تشخیص نادرست بیشتر می‌شود.

لازم به ذکر است که در صورتی که ضرایب داده‌ها نامقید باشد، داده‌ها با ضرایب منفی نیز در تشکیل فضای پوش نامقید شرکت میکنند. لذا، فضای پوش نامقید چهار داده یاد شده، کل فضا را دربر می‌گیرد. یعنی، داده تست در خارج از قطاع یاد شده را می‌توان به خوبی با استفاده از این چهار داده، بازسازی کرد. بنابراین، قید نامنفی بودن ضرایب، چه در مدل CESRC و چه در

<sup>۱</sup>Align

دارای تخریب پراکنده و پیوسته با درجات تخریب مختلف از مجموعه داده AR را نشان می‌دهد. هر آزمایش تشخیص چهره معمولاً در ۴ رزولوشن و مقیاس متفاوت، یعنی مقیاس اصلی یا مقیاس ۱، مقیاس ۰٫۷۵، مقیاس ۰٫۵، و مقیاس ۰٫۲۵، انجام می‌شود و نتایج هر آزمایش در هر مقیاس، به صورت جداگانه گزارش می‌شود. برای مقایسه روش‌های طبقه‌بندی مختلف، باید آن‌ها را در بهترین پیکربندی ممکن مورد آزمایش قرار داد. برای این منظور، بهترین مقدار پارامتر  $\lambda$  از مجموعه مقادیر  $\{0, 0.00001, 0.0001, 0.001, 0.01, 0.05, 0.1\}$  و بهترین مقدار پارامتر  $\sigma$ ، از مجموعه مقادیر  $\{0, 0.00001, 0.0001, 0.001, 0.01, 0.05, 0.1, 1, 10\}$  انتخاب شد. به عبارت دیگر، هر بار مقدار  $\lambda$  و  $\sigma$  برابر با یکی از مقادیر دو مجموعه مقدار یادشده قرار داده شد و صحت هر روش به ازای هر پارامتر بررسی شد و نهایتاً بهترین صحت گزارش شد. برای انجام آزمایشات، از سیستمی با پردازنده Intel(R) Core(TM) i7، حافظه 16GB و نرم‌افزار Matlab 2016a استفاده می‌شود.



شکل ۲- تصاویر یک نفر در مجموعه داده AR: تصاویر گرفته شده در جلسه اول (تصاویر ردیف بالا) و تصاویر گرفته شده در جلسه دوم (تصاویر ردیف پایین).



شکل ۳- تصاویر یک نفر در مجموعه داده CAS-PEAL.

Jones [۲۵] اجرا شد. سپس، بر اساس محل چهره و با استفاده از روش ارائه شده در مرجع [۲۶]، مکان ۹ نشانه اختصاصی چهره تعیین شد و بر اساس آن، تصویر چهره، برش زده شد و بخش‌های زاید تصویر حذف شد. ابعاد تصاویر برش زده،  $120 \times 165$  می‌باشد. مجموعه تصاویر گرفته شده از یک نفر در مجموعه داده AR، در شکل ۲ نمایش داده شده است. تصاویر گرفته شده در هر جلسه، شامل تصاویر گرفته شده تحت شرایط نورپردازی‌های متفاوت و حالت‌های متفاوت ژست صورت و دو نوع انسداد یعنی انسداد با عینک دودی و شال گردن می‌باشد. مجموعه تصاویر آموزشی،  $116 \times 7 = 812$  تصویر چهره بدون نویز و انسداد گرفته شده در جلسه اول مجموعه داده AR می‌باشند و مجموعه تصاویر آزمون، از مجموعه تصاویر گرفته شده در جلسه دوم مجموعه داده AR انتخاب شدند.

#### ۴-۱-۲- مجموعه داده CAS-PEAL

مجموعه داده CAS-PEAL، شامل تصاویر صورت  $1024$  نفر در ابعاد  $120 \times 100$  می‌باشد که از روبرو گرفته شده است. در این مقاله، از تصاویر صورت  $100$  نفر نخست استفاده شد. تصاویر گرفته شده از هر فرد، شامل تصاویر گرفته شده تحت شرایط نورپردازی‌های متفاوت، حالت‌های متفاوت ژست صورت، تصاویر با پس زمینه متفاوت و تصاویر همراه با انسداد با کلاه می‌باشد. مجموعه تصاویر گرفته شده از یک نفر در مجموعه داده CAS-PEAL، در شکل ۳ نمایش داده شده است. مجموعه داده‌های آموزشی، فقط تصاویر چهره طبیعی مجموعه داده CAS-PEAL در نظر گرفته شدند.

#### ۴-۲- نحوه پیکربندی مدل‌ها و آزمایش‌ها

سه دسته آزمایش بر اساس سه نوع مجموعه داده آزمون تدارک دیده شد: مجموعه داده آزمون معمولی، مجموعه داده آزمون دارای انسداد طبیعی، و مجموعه داده آزمون تخریب شده. مجموعه داده آزمون معمولی، فقط شامل تصاویر فاقد انسداد و تخریب است. مجموعه داده آزمون دارای انسداد طبیعی مجموعه داده AR، شامل تصاویر چهره انسداد یافته توسط عینک دودی و شال گردن می‌باشد. مجموعه داده آزمون دارای انسداد طبیعی مجموعه داده CAS-PEAL، شامل تصاویر چهره انسداد یافته توسط کلاه می‌باشد. مجموعه داده آزمون تخریب شده، همان مجموعه داده آزمون معمولی است که روی هر یک، به روش مصنوعی و با درجات تخریب مختلف، تخریب پراکنده یا پیوسته اعمال شده است. برای تخریب مصنوعی پراکنده، به صورت تصادفی مقادیر شدت روشنایی درصد مشخصی از پیکسل‌های تصویر چهره با مقادیر تصادفی از توزیع یکنواخت، تخریب می‌شود. برای تخریب مصنوعی پیوسته نیز درصد مشخصی از بخش تصادفی تصویر چهره، با تصویر یک میمون تخریب می‌شود. شکل ۴ یک تصویر معمولی، یک تصویر دارای انسداد عینک و شال گردن و تصاویری

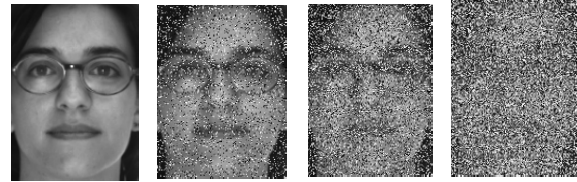


آزمون بدون انسداد و تخریب از مجموعه داده AR به تفکیک نوع تصاویر آزمون و همچنین بدون تفکیک تصاویر آزمون را نشان می‌دهد. تصاویر آزمون، شامل یک تصویر چهره طبیعی، سه تصویر چهره دارای ژست (لبخند، خشم و فریادزدن) و سه تصویر با شرایط نورپردازی متفاوت (نورپردازی از چپ، نورپردازی از راست، و نورپردازی از چپ و راست) می‌باشد. همانطور که مشاهده می‌شود، صحت روش پیشنهادی CECRC و روش CRC بیشتر یا برابر با صحت سایر روش‌ها شده است. قبلاً، در مرجع نیز نشان داده شده است که معمولاً روش CRC نسبت به روش SRC بهتر عمل می‌کند [۳]. دلیل عملکرد بهتر روش پیشنهادی CECRC نسبت به روش CESRC، نیز استفاده از قید محدودکننده نامنفی بودن ضرایب در مدل CESRC است که قبلاً در بخش ۳-۳ مورد بررسی قرار گرفت.

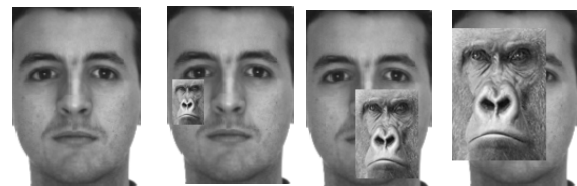
جدول ۲، صحت روش‌های مختلف طبقه‌بندی به‌ازای مجموعه داده آزمون بدون انسداد و تخریب از مجموعه داده CAS-PEAL به تفکیک نوع تصاویر آزمون و همچنین بدون تفکیک تصاویر آزمون را نشان می‌دهد. همانطور که مشاهده می‌شود، صحت روش پیشنهادی CECRC بیشتر یا برابر با صحت سایر روش‌ها شده است. همانطور که ملاحظه می‌شود، صحت طبقه‌بندی تصاویر چهره مجموعه داده CAS-PEAL تحت نورپردازی‌های مختلف، بسیار پایین‌تر از صحت طبقه‌بندی تصاویر چهره مجموعه داده AR تحت نورپردازی‌های مختلف شده است. دلیل این امر، استفاده از همه انواع تصاویر چهره بدون انسداد و نویز جلسه اول مجموعه داده AR به عنوان مجموعه آموزشی می‌باشد. در حالی که، فقط از تصاویر چهره طبیعی مجموعه داده CAS-PEAL به عنوان مجموعه داده آموزشی استفاده شد. در ضمن، صحت طبقه‌بندی مجموعه تصاویر با پس‌زمینه متفاوت توسط مدل پیشنهادی CECRC، خیلی بهتر از روش‌های غیرمقاوم SRC، CRC و SACRC شده است. دلیل این موضوع، این است که مدل پیشنهادی، با پیکسل‌های پس‌زمینه مانند پیکسل‌های انسداد برخورد می‌کند و همانطور که در ادامه هم ملاحظه خواهد شد مدل پیشنهادی CECRC بسیار بهتر از روش‌های طبقه‌بندی غیرمقاوم، تصاویر دارای انسداد را طبقه‌بندی می‌کند. در واقع، انسداد، هر چیزی است که با پیکسل‌های موجود در تصاویر آموزشی، تفاوت قابل ملاحظه داشته باشد. لذا، پیکسل‌های پس‌زمینه تصویر آزمون با پس‌زمینه متفاوت از تصاویر چهره طبیعی موجود در مجموعه داده آموزشی، همچون انسدادی بر روی زمینه تصاویر چهره طبیعی هستند.



(الف) (ب) (ج)



(د)



(ه)

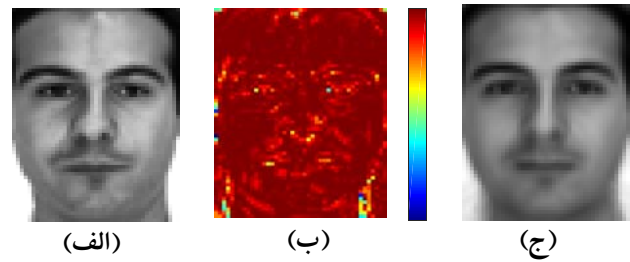
شکل ۴- الف) تصویر چهره بدون انسداد و تخریب. ب) تصویر دارای انسداد عینک دودی. ج) تصویر دارای انسداد شال‌گردن. د) تصویر دارای ۰، ۲۵، ۵۰ و ۷۵ درصد تخریب پراکنده (از چپ به راست). ه) تصاویر چهره دارای ۰، ۲۵، ۵۰ و ۷۵ درصد تخریب پیوسته (از چپ به راست) از مجموعه داده AR.

#### ۳-۴ آزمایش با داده‌های آزمون بدون انسداد و تخریب

شکل ۵، نتیجه اجرای روش پیشنهادی CECRC بر روی یکی از تصاویر آزمون بدون انسداد و بدون تخریب از کلاس ۱ را نشان می‌دهد. همانطور که گفته شد جمله  $g(y_j - \sum_{i=1}^n X_{ij}a_i)$  در تابع هدف مدل پیشنهادی، وزن پیکسل  $j$ -ام می‌باشد. وقتی  $\|y_j - \sum_{i=1}^n X_{ij}a_i\|$  بزرگ باشد به این معنی است که پیکسل  $j$ -ام در تصویر آزمون، احتمالاً پیکسلی نویزی یا تخریب شده است و برای همین، امکان بازسازی آن، با ترکیب خطی داده‌های آموزشی میسر نبوده است. در چنین شرایطی، مقدار  $g(y_j - \sum_{i=1}^n X_{ij}a_i)$  یا وزن پیکسل  $j$ -ام در تابع هدف مدل پیشنهادی، کم می‌شود و این پیکسل، مشارکت زیادی در تعیین مقدار بهینه  $a$  نخواهد داشت. وقتی تصویر آزمون، بدون انسداد و تخریب است، معمولاً به اغلب پیکسل‌ها، وزن تقریباً یکنواختی اختصاص می‌یابد چراکه اغلب پیکسل‌های تصویر آزمون بصورت ترکیب خطی داده‌های آموزشی درمی‌آیند. همان‌طور که در شکل ۵-ب) نیز مشاهده می‌شود، وزن غالب پیکسل‌ها تقریباً برابر هم شده است چراکه تصویر آزمون نشان داده شده در شکل ۵-الف)، فاقد نویز و تخریب بوده است. در شکل ۵-ج) تصویر حاصل از ترکیب خطی یادگرفته‌شده از داده‌های آموزشی نشان داده شده است که به خوبی مشابه تصویر آزمون نشان داده شده در شکل ۵-الف) شده است. سرانجام، جدول ۱، صحت روش‌های مختلف طبقه‌بندی از مجموعه داده

شکل ۶- (ج) تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی نشان داده شده است که اثری از تخریب و عکس میمون در آن وجود ندارد که بیانگر مقاوم بودن روش پیشنهادی نسبت به تخریب پیوسته است. جدول ۳، میانگین صحت روش‌های مختلف به‌ازای مجموعه داده آزمون از مجموعه داده AR نشان می‌دهد که بخشی تصادفی از هر داده آن، با تصویر یک میمون به اندازه ۲۵، ۵۰، ۷۵ یا ۹۰ درصد اندازه داده آزمون، بطور پیوسته تخریب شده است. همانطور که ملاحظه می‌شود روش پیشنهادی CECRC، همواره صحت بالاتری نسبت به سایر روش‌ها داشته است. در ضمن، به‌نظر می‌رسد زمانی که تخریب بیشتر می‌شود برتری صحت مدل پیشنهادی نسبت به سایر روش‌ها ملموس‌تر می‌گردد. لازم به ذکر است که با توجه به تصادفی بودن محل تخریب، هر آزمایش، ۱۰ بار با محل تصادفی متفاوتی که توسط تابع rand نرم افزار MATLAB تولید شد اجرا شده است و میانگین صحت بدست آمده، گزارش شده است. دلیل برتری روش پیشنهادی CECRC و روش CESRC نسبت به روش‌های SRC و CRC، استفاده روش‌های CECRC و CESRC از مفهوم کورآنتروپی برای مقاوم کردن مدل در برابر نویز و انسداد است. در این دو مدل مبتنی بر کورآنتروپی، با هر یک از پیکسل‌های تصویر به‌طور متفاوت رفتار می‌شود و بدین طریق، تاکید بیشتری بر روی پیکسل‌هایی دارد که مرتبط با نمونه‌های آموزش همان کلاس تصویر ورودی هستند. به این معنا که اگر در تصویر ورودی، خطای تخریب و انسداد وجود داشته باشد، پیکسل‌های تخریب و مسدود شده، سهم کمتری در اندازه اختلاف تصویر ورودی و ترکیب خطی تصاویر آموزشی خواهند داشت و لذا، سهم چنین پیکسل‌هایی در ارزیابی برچسب تصویر ورودی نیز کمتر می‌شود. دلیل برتری روش پیشنهادی CECRC نسبت به CESRC، استفاده روش CESRC از قید محدودکننده نامنفی بودن ضرایب برای حل آسان‌تر مدل است. قبلاً در بخش ۳-۳ مقاله، بطور مفصل به این موضوع پرداخته شد.

جدول ۴، میانگین صحت روش‌های مختلف به‌ازای مجموعه داده آزمون از مجموعه داده CAS-PEAL را نشان می‌دهد که بخشی تصادفی از هر داده آن، با تصویر یک میمون به اندازه ۲۵، ۵۰، ۷۵ یا ۹۰ درصد اندازه داده آزمون، بطور پیوسته تخریب شده است. همانطور که ملاحظه می‌شود باز هم روش پیشنهادی CECRC، همواره صحت بالاتری نسبت به روش‌های غیر مقاوم داشته است.



شکل ۵- نتیجه اجرای روش پیشنهادی بر روی یک داده آزمون از کلاس اول از مجموعه داده AR. (الف) تصویر بدون انسداد و تخریب داده آزمون. (ب) وزن یادگیری شده (توسط نوار رنگ، میزان وزن هر یک از پیکسل‌ها مشخص شده است). (ج) تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی.

جدول ۱- صحت روش‌های مختلف به‌ازای مجموعه داده آزمون بدون انسداد و تخریب مجموعه داده AR در ابعاد  $60 \times 43$ .

همه موارد	شرایط نورپردازی متفاوت	چهره‌های دارای ژست	چهره طبیعی	نوع چهره / روش
0.97	0.97	0.96	0.96	SRC
<b>0.99</b>	0.97	<b>1</b>	<b>1</b>	CRC
0.95	0.96	0.95	0.96	CESRC
0.97	0.97	0.96	<b>1</b>	SACRC
<b>0.99</b>	<b>0.99</b>	<b>1</b>	0.96	CECRC
				(proposed)

جدول ۲- صحت روش‌های مختلف به‌ازای مجموعه داده آزمون بدون انسداد و تخریب مجموعه داده CAS-PEAL در ابعاد  $120 \times 100$ .

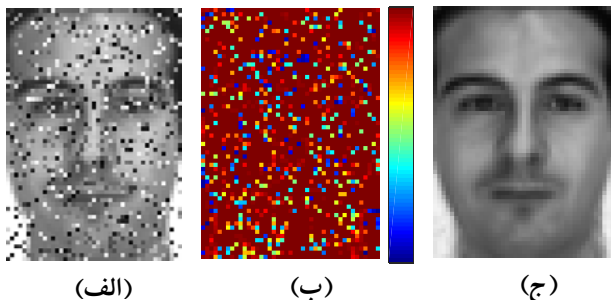
همه موارد	پس زمینه	شرایط نورپردازی متفاوت	چهره‌های دارای ژست	نوع چهره / روش
0.45	0.79	0.19	<b>0.91</b>	SRC
0.46	0.81	0.21	0.90	CRC
0.45	0.82	0.18	0.90	CESRC
0.47	0.81	0.21	<b>0.91</b>	SACRC
<b>0.48</b>	<b>0.84</b>	<b>0.22</b>	<b>0.91</b>	CECRC
				(proposed)

#### ۴-۴ آزمایش با داده‌های آزمون دارای تخریب پیوسته

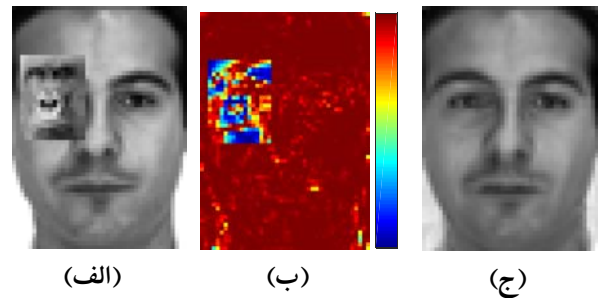
شکل ۶، نتیجه اجرای روش پیشنهادی بر روی یکی از تصاویر آزمون کلاس ۱ از مجموعه داده AR را نشان می‌دهد که بخشی تصادفی از آن، بطور پیوسته، با تصویر یک میمون به اندازه ۲۵ درصد تصویر آزمون تخریب شده است. همانطور که گفته شد  $g(y_j - \sum_{i=1}^n X_{ij}a_i)$  وزن پیکسل  $j$ -ام در تابع هدف مدل پیشنهادی است. اگر پیکسل  $j$ -ام داده آزمون نویزی باشد، انتظار داریم که وزن آن در تابع هدف مدل پیشنهادی، کم شود تا این پیکسل، مشارکت زیادی در تعیین مقدار بهینه  $a$  و برچسب داده آزمون نداشته باشد. همان‌طور که در شکل ۶- (ب) نیز مشاهده می‌کنید، وزن غالب پیکسل‌ها به جز عمده پیکسل‌های مربوط ناحیه تخریب شده، تقریباً برابر هم و مقدار بزرگی شده است. در

پیکسل‌های تخریب شده، تقریباً برابر هم و مقدار بزرگی شده است. در شکل ۷- (ج)، تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی نشان داده شده است که اثری از پیکسل‌ها نویزی در آن وجود ندارد که بیانگر مقاوم بودن روش پیشنهادی نسبت به تخریب پراکنده است. جدول ۵، میانگین صحت روش‌های مختلف به‌ازای یک مجموعه داده آزمون از مجموعه داده AR در مقیاس  $60 \times 43$  را نشان می‌دهد که ۲۵، ۵۰، ۷۵ یا ۹۰ درصد از پیکسل‌های هر داده آن، بطور تصادفی و پراکنده، تخریب شده‌اند. همانطور که ملاحظه می‌شود روش پیشنهادی CECRC، همواره صحت بالاتر و در یک مورد، صحتی برابر با سایر روش‌ها داشته‌است. دلیل این برتری، پیشتر، در بخش ۴-۳ نیز مطرح شد. در واقع، نتایج این بخش نیز موید همان استدلال‌هاست. در ضمن، به‌نظر می‌رسد زمانی که تخریب بیشتر می‌شود برتری صحت مدل پیشنهادی نسبت به سایر روش‌ها ملموس‌تر می‌گردد. لازم به ذکر است که با توجه به تصادفی بودن محل تخریب، هر آزمایش، ۱۰ بار با محل تصادفی متفاوتی که توسط تابع rand نرم افزار MATLAB تولید شد اجرا شده است و میانگین صحت بدست آمده گزارش شده است.

جدول ۶، میانگین صحت روش‌های مختلف به‌ازای یک مجموعه داده آزمون از مجموعه داده CAS-PEAL در مقیاس  $120 \times 100$  را نشان می‌دهد که ۲۵، ۵۰، ۷۵ یا ۹۰ درصد از پیکسل‌های هر داده آن، بطور تصادفی و پراکنده، تخریب شده‌اند. همانطور که ملاحظه می‌شود باز هم روش پیشنهادی CECRC، همواره صحت بالاتر یا برابر با سایر روش‌ها داشته‌است.



شکل ۷- نتیجه اجرای روش معرفی شده بر روی یک داده آزمون از کلاس اول از مجموعه داده AR با ۲۵ درصد تخریب پراکنده. (الف) تصویر داده آزمون با ۲۵ درصد تخریب پراکنده. (ب) وزن یادگیری شده (توسط نوار رنگ میزان وزن هر یک از پیکسل‌ها مشخص شده است). (ج) تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی.



شکل ۶- نتیجه اجرای روش پیشنهادی بر روی یک داده آزمون از کلاس اول از مجموعه داده AR با ۲۵ درصد تخریب پیوسته. (الف) تصویر داده آزمون دارای تخریب پیوسته ۲۵ درصد. (ب) وزن یادگیری شده (توسط نوار رنگ، میزان وزن هر یک از پیکسل‌ها مشخص شده است). (ج) تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی.

جدول ۳- میانگین صحت روش‌های مختلف به‌ازای مجموعه داده آزمون در مقیاس  $60 \times 43$  از مجموعه داده AR که بخشی تصادفی از هر داده آزمون، با تصویر یک میمون به اندازه ۲۵، ۵۰، ۷۵ یا ۹۰ درصد داده آزمون، بطور پیوسته تخریب شده است.

تخریب / روش	0.25	0.5	0.75	0.90
SRC	0.96	0.92	0.40	0.10
CRC	0.96	0.25	0.12	0.08
CESRC	0.96	0.92	0.52	0.18
SACRC	0.96	0.80	0.44	0.12
CECRC (proposed)	1	1	0.80	0.28

جدول ۴- میانگین صحت روش‌های مختلف به‌ازای مجموعه داده آزمون در مقیاس  $120 \times 100$  از مجموعه داده CAS-PEAL که بخشی تصادفی از هر داده آزمون، با تصویر یک میمون به اندازه ۲۵، ۵۰، ۷۵ یا ۹۰ درصد داده آزمون، بطور پیوسته تخریب شده است.

تخریب / روش	0.25	0.5	0.75	0.90
SRC	0.65	0.77	0.75	0.76
CRC	0.66	0.77	0.76	0.77
CESRC	0.98	0.94	0.86	0.78
SACRC	0.66	0.77	0.76	0.77
CECRC (proposed)	0.92	0.90	0.89	0.83

#### ۴-۵ آزمایش با داده‌های آزمون دارای تخریب پراکنده

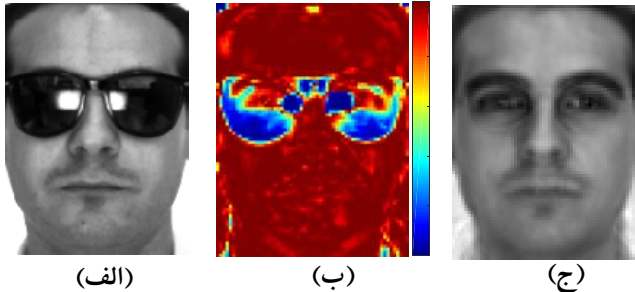
شکل ۷، نتیجه اجرای روش پیشنهادی بر روی یکی از تصاویر آزمون از کلاس ۱ از مجموعه داده AR را نشان می‌دهد که ۲۵ درصد از پیکسل‌هایش بطور تصادفی و پراکنده تخریب شده‌اند. همانطور که گفته شد  $g(y_j - \sum_{i=1}^n X_{ij} a_i)$  وزن پیکسل  $j$ -ام در تابع هدف مدل پیشنهادی است. اگر پیکسل  $j$ -ام داده آزمون نویزی باشد، انتظار داریم که وزن آن در تابع هدف مدل پیشنهادی، کم شود تا این پیکسل، مشارکت زیادی در تعیین مقدار بهینه  $a$  و برچسب داده آزمون نداشته باشد. همان‌طور که در شکل ۷- (ب) نیز مشاهده می‌شود، وزن غالب پیکسل‌ها به جز عمده پیکسل‌های مربوط

جدول ۵- میانگین صحت روش‌های مختلف به‌ازای یک مجموعه داده آزمون از مجموعه داده AR در مقیاس  $60 \times 43$  که ۲۵، ۵۰، ۷۵ یا ۹۰ درصد از هر داده آن، بطور تصادفی و پراکنده تخریب شده است.

تخریب / روش	0.25	0.5	0.75	0.90
SRC	0.92	0.72	0.24	0.08
CRC	0.92	0.20	0.12	0.08
CESRC	0.96	<b>0.96</b>	0.92	0.28
SACRC	0.96	0.60	0.28	0.12
CECRC (proposed)	1	<b>0.96</b>	<b>0.96</b>	<b>0.32</b>

جدول ۶- میانگین صحت روش‌های مختلف به‌ازای یک مجموعه داده آزمون از مجموعه داده CAS-PEAL در مقیاس  $100 \times 120$  که ۲۵، ۵۰، ۷۵ یا ۹۰ درصد از هر داده آن، بطور تصادفی و پراکنده تخریب شده است.

تخریب / روش	0.25	0.5	0.75	0.90
SRC	۱	1	۰,۴۳	۰,۰۳
CRC	۱	1	۰,۴۵	۰,۰۳
CESRC	۱	1	۰,۸۳۰۰	۰,۱۰۰۰
SACRC	۱	1	۰,۴۵	۰,۰۳
CECRC (proposed)	۱	1	۰,۹۶۰۰	۰,۱۷۰۰



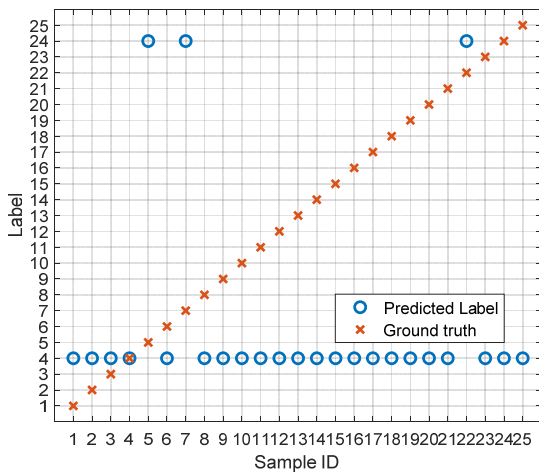
شکل ۸- نتیجه اجرای روش معرفی شده بر روی یک تصویر داده آزمون با انسداد عینک از کلاس اول از مجموعه داده AR. (الف) تصویر داده آزمون دارای انسداد عینک. (ب) وزن یادگیری شده (توسط نوار رنگ، میزان وزن هر یک از پیکسل‌ها مشخص شده است). (ج) تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی.

جدول ۷- صحت روش‌های مختلف به‌ازای مجموعه داده آزمون دارای انسداد عینک از مجموعه داده AR در چهار مقیاس مختلف.

مقیاس / روش	0.25	0.5	0.75	1
SRC	0.65	0.84	0.92	0.88
CRC	0.56	0.73	0.74	0.78
CESRC	0.80	<b>0.92</b>	0.92	0.94
SACRC	<b>0.88</b>	<b>0.92</b>	0.92	0.92
CECRC (proposed)	0.85	0.91	<b>0.96</b>	<b>0.96</b>

۴-۷ آزمایش با داده‌های آزمون دارای انسداد شال گردن از مجموعه داده AR

شکل ۹، نتیجه اجرای روش پیشنهادی بر روی یکی از تصاویر آزمون دارای انسداد شال گردن از کلاس ۴ از مجموعه داده AR را نشان می‌دهد. همانطور که گفته شد  $g(y_j - \sum_{i=1}^n X_{ij}a_i)$  وزن پیکسل ز-ام در تابع هدف مدل پیشنهادی است. اگر پیکسل ز-ام داده آزمون نویزی باشد، انتظار داریم که وزن آن در تابع هدف مدل پیشنهادی، کم شود تا این پیکسل، مشارکت زیادی در تعیین مقدار بهینه  $a$  و برچسب داده آزمون نداشته باشد. همان‌طور که در شکل ۸- (ب) نیز مشاهده می‌شود، وزن غالب پیکسل‌ها به جز پیکسل‌های مربوط به عینک، تقریباً برابر هم و مقدار بزرگی شده است. وزن پیکسل‌های عینک، کم شده است. این پیکسل‌ها، تاثیر چندانی در تعیین مقدار بهینه  $a$  و برچسب داده آزمون نخواهند داشت. لذا، روش پیشنهادی، به خوبی، نسبت به انسداد عینک، مقاومت نشان خواهد داد. شکل ۸- (ج) تصویر حاصل از ترکیب خطی یادگرفته شده از داده‌های آموزشی نشان داده شده است که به خوبی مشابه تصویر داده آزمون نشان داده شده در شکل ۸- (الف) شده است و اثری از عینک، در این تصویر وجود ندارد که خود، دال بر مقاوم بودن روش پیشنهادی نسبت به انسداد عینک می‌باشد. سرانجام، جدول ۷، صحت روش‌های مختلف به‌ازای مجموعه داده آزمون دارای انسداد عینک، در چهار مقیاس متفاوت را نشان می‌دهد. همان‌طور که مشاهده می‌شود، روش مقاوم CESRC و روش مقاوم پیشنهادی CECRC نسبت به روش‌های SRC و CRC، صحت بسیار بیشتری دارند. روش



شکل ۱۰- برچسب‌های تخمینی و واقعی برای مجموعه داده آزمون با انسداد شال‌گردن.

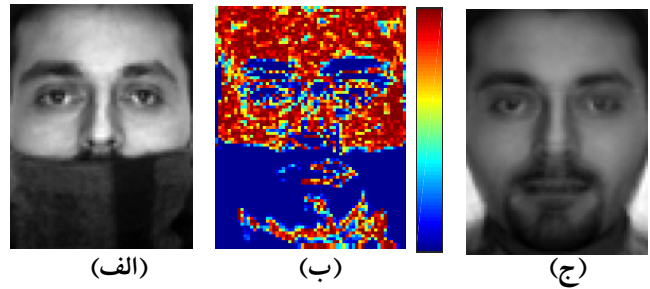
#### ۴-۸ آزمایش با داده‌های آزمون دارای انسداد کلاه از مجموعه داده CAS-PEAL

جدول ۹، صحت روش‌های مختلف به‌ازای مجموعه داده آزمون دارای انسداد کلاه از مجموعه داده AR را نشان می‌دهد. همانطور که مشاهده می‌شود، روش مقاوم CESRC و روش مقاوم پیشنهادی CECRC نسبت به روش‌های SRC و CRC، صحت بسیار بیشتری دارند. روش پیشنهادی CECRC به‌ازای سه مقیاس از چهار مقیاس، صحت بهتری نسبت به روش CESRC کسب کرده است. دلیل برتری روش پیشنهادی CECRC نسبت به بودن ضرایب برای حل آسان‌تر مدل است. قبلاً در بخش ۳-۳ مقاله، بطور مفصل به این موضوع پرداخته شد.

جدول ۷- صحت روش‌های مختلف به‌ازای مجموعه داده آزمون دارای انسداد کلاه از مجموعه داده CAS-PEAL.

روش	میانگین صحت
SRC	۰.۶۳۷۳
CRC	۰.۶۴۱۰۷
CESRC	۰.۶۳۹۴
SACRC	۰.۶۴۱۰۷
CECRC (proposed)	۰.۶۹۴۴۹

جدول ۶، میانگین صحت هر روش طبقه‌بندی را به‌ازای کل آزمایشات نشان می‌دهد. همانطور که ملاحظه می‌شود حتی با وجود آنکه روش پیشنهادی (همراه با روش CESRC)، برای تشخیص تصاویر دارای انسداد با شال‌گردن دچار مشکل می‌شوند، میانگین صحت روش پیشنهادی، بیشتر از روش‌های دیگر شده است. در واقع، این نتایج نشان می‌دهد که مفهوم کورآنتروپی، به خوبی توانسته است روش پیشنهادی CECRC و روش CESRC را نسبت به نویز و انسداد در تصاویر چهره مقاوم

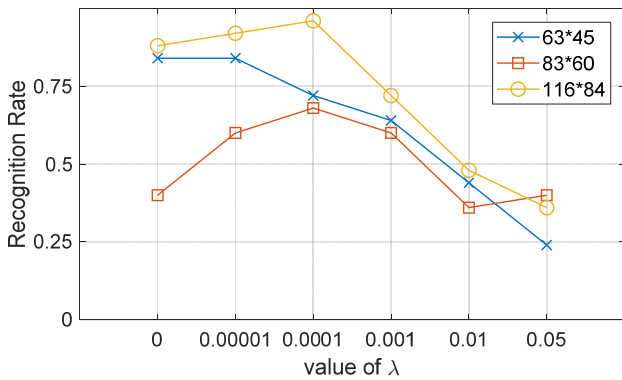


شکل ۹- نتیجه اجرای روش معرفی شده بر روی یک تصویر داده آزمون با انسداد شال‌گردن از کلاس چهارم از مجموعه داده AR. (الف) تصویر داده آزمون دارای انسداد شال‌گردن. (ب) وزن یادگیری‌شده (توسط نوار رنگ، میزان وزن هر یک از پیکسل‌ها مشخص شده است). (ج) تصویر حاصل از ترکیب خطی یادگرفته‌شده از داده‌های آموزشی.

جدول ۸، صحت روش‌های مختلف به‌ازای مجموعه داده آزمون دارای انسداد شال‌گردن از مجموعه داده AR، در چهار مقیاس متفاوت را نشان می‌دهد. همانطور که مشاهده می‌شود هیچ کدام از دو روش مقاوم CESRC و روش پیشنهادی CECRC نتوانسته‌اند به صحت طبقه‌بندی مطلوبی در مقایسه با سایر روش‌ها دست یابند. روش CRC بهترین نتیجه را کسب کرده است. برای درک بهتر دلیل کارایی کمتر روش پیشنهادی CECRC در حالت انسداد شال‌گردن، نمودار برچسب‌های تخمین زده‌شده و برچسب‌های واقعی برای ۲۵ داده آزمون دارای انسداد شال‌گردن از ۲۵ کلاس در شکل ۱۰ رسم شده است. با توجه به شکل ۱۰، برچسب اغلب داده‌های آزمون دارای انسداد شال‌گردن، به اشتباه، برابر با برچسب کلاس چهارم شده است. یعنی، در واقع تصویر چهره حاصل از بازسازی با استفاده از بردار ضرایب یادگرفته‌شده، میل به سمت تصاویر آموزشی کلاس چهارم دارد. نمونه‌ای از تصویر چهره داده کلاس چهارم در شکل ۹- (ج) نشان داده شد و همانطور که مشاهده می‌شود، تصاویر این کلاس دارای ریش هستند و در واقع قسمتی از شال‌گردن به دلیل شباهت به ریش، به سمت نمونه‌های این کلاس بایاس می‌شود. با حذف نمونه‌های کلاس چهارم، دقت روش معرفی‌شده به طور موثری تغییر کرده و افزایش می‌یابد و وجود تصویر چهره فرد با ریش منجر به کاهش کارایی این روش شده است. این مسئله بیانگر یک نقطه ضعف استفاده از عبارت کورآنتروپی در تابع هزینه می‌باشد.

جدول ۸- صحت روش‌های مختلف به‌ازای مجموعه داده آزمون دارای انسداد شال‌گردن در چهار مقیاس مختلف.

مقیاس / روش	0.25	0.5	0.75	1
SRC	0.40	0.44	0.48	0.48
CRC	0.44	<b>0.64</b>	<b>0.68</b>	<b>0.76</b>
CESRC	0.08	0.20	0.24	0.24
SACRC	<b>0.56</b>	<b>0.64</b>	0.60	0.64
CECRC (proposed)	0.32	0.44	0.48	0.48



شکل ۱۱- حساسیت صحت تشخیص روش پیشنهادی به پارامتر  $\lambda$  به ازای داده‌های آزمون دارای انسداد عینک مجموعه داده AR در سه مقیاس مختلف.

#### ۴-۱۰ حساسیت مدل پیشنهادی به پارامتر $\sigma$

پارامتر مهم دیگر روش معرفی شده، پارامتر سایز هسته گوسی یا  $\sigma$  است. همانطور که گفته شد اگر  $\|y_j - \sum_{i=1}^n X_{ij}a_i\|$  زیاد باشد یعنی امکان بازسازی پیکسل  $j$ -ام داده آزمون توسط داده‌های آموزشی چندان میسر نیست و احتمالاً این پیکسل، پیکسل نویزی است و باید وزن کمی به این پیکسل در تابع هدف مدل داده شود تا اثر چندانی در تعیین برجسته‌ترین داده آزمون نداشته باشد. اما مفاهیم  $\|y_j - \sum_{i=1}^n X_{ij}a_i\|$  "زیاد" و "وزن کم"، مفاهیمی فازی هستند که توسط تابع گوسی و پارامتر  $\sigma$  تعریف و تنظیم می‌شوند. اگر پارامتر  $\sigma$  خیلی کوچک در نظر گرفته شود، تقریباً همه پیکسل‌ها، نویزی تلقی می‌شوند و وزن کمی به آنها تعلق می‌گیرد و کارایی روش پیشنهادی، کاهش می‌یابد. اگر پارامتر  $\sigma$  خیلی بزرگ در نظر گرفته شود، تقریباً همه پیکسل‌ها، غیرنویزی تلقی می‌شوند و وزن زیادی به همه آنها تعلق می‌گیرد و بازهم کارایی روش پیشنهادی، کاهش می‌یابد. بنابراین، تنظیم دقیق پارامتر  $\sigma$ ، حیاتی است. شکل ۱۲ حساسیت صحت طبقه‌بندی روش پیشنهادی به پارامتر  $\sigma$  به ازای داده‌های آزمون دارای انسداد عینک مجموعه داده AR در سه مقیاس مختلف را نشان می‌دهد. لازم به ذکر است که در این آزمایش، مقدار سایر پارامترها بصورت زیر در نظر گرفته شد:

$$\lambda = 0.0001, iter_{max} = 250, thr_a = 0.0001.$$

همان‌طور که در شکل ۱۲ مشاهده می‌شود، با زیاد بزرگ شدن مقدار پارامتر  $\sigma$ ، نرخ صحت تشخیص کاهش می‌یابد و بالعکس با کاهش زیاد مقدار این پارامتر، نرخ صحت به شدت کاهش می‌یابد. در مجموع می‌توان گفت که در مقیاس‌های مختلف، غالباً زمانی که  $\sigma$  مقداری بین 0.05 تا 0.001 است بهترین نرخ صحت حاصل می‌شود.

کند. به همین دلیل، میانگین صحت این دو روش، بهتر از روش‌های SRC و CRC شده است که از مفهوم کورآنتروپی استفاده نکرده‌اند. دلیل برتری روش پیشنهادی CECRC نسبت به CESRC، استفاده روش CESRC از قید محدودکننده نامنفی بودن ضرایب برای حل آسانتر مدل است. قبلاً در بخش ۳-۳ مقاله، بطور مفصل به این موضوع پرداخته شد. به نظر می‌رسد که دلیل اینکه روش CESRC نتوانسته است به خوبی روش SACRC عمل کند نیز همان استفاده از قید محدودکننده نامنفی بودن ضرایب می‌باشد.

جدول ۶- میانگین صحت روش‌های مختلف به ازای کل آزمایشات این مقاله.

روش	میانگین صحت
SRC	0.6118
CRC	0.5324
CESRC	0.6465
SACRC	0.6665
CECRC (proposed)	<b>0.7476</b>

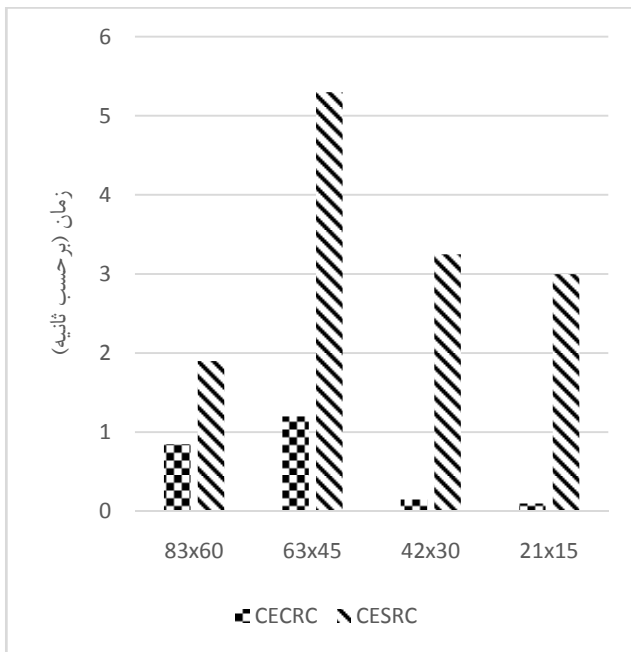
#### ۴-۹ حساسیت مدل پیشنهادی به پارامتر $\lambda$

پارامتر کنترل تعمیم‌پذیری  $\lambda$  در مدل پیشنهادی، تاثیر زیادی در کارایی روش دسته‌بندی دارد و میزان تعمیم‌پذیری و تعادل میان نرُم دو و مقاوم بودن را برقرار می‌کند. شکل ۱۱ حساسیت صحت طبقه‌بندی به پارامتر  $\lambda$  به ازای داده‌های آزمون دارای انسداد عینک مجموعه داده AR، در سه مقیاس مختلف را نشان می‌دهد. لازم به ذکر است که در این آزمایش، مقدار سایر پارامترها بصورت زیر در نظر گرفته شد:

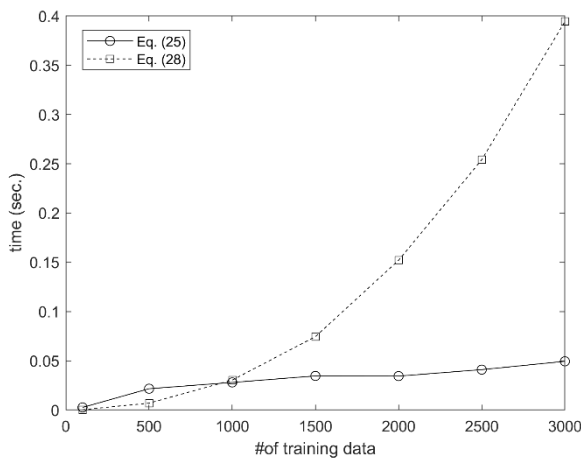
$$\sigma = 0.001, iter_{max} = 25, thr_a = 0.0001.$$

همان‌طور که در شکل ۱۱ مشاهده می‌شود، انتخاب مقدار مناسب پارامتر کنترل تعمیم‌پذیری  $\lambda$  می‌تواند به شدت در صحت طبقه‌بندی موثر باشد. با خیلی بزرگ شدن مقدار این پارامتر، نرخ صحت تشخیص، به شدت کاهش می‌یابد و بالعکس با کاهش زیاد مقدار این پارامتر نیز نرخ صحت کاهش می‌یابد. در مجموع می‌توان گفت که در مقیاس‌های مختلف، غالباً زمانی که  $\lambda$  مقداری بین 0.0001 تا 0.00001 است بهترین نرخ صحت حاصل می‌شود.

تعداد داده‌های آموزشی به بیش از ۱۰۰۰۰ داده افزایش می‌یابد اجرای رابطه (۳۲) سریع‌تر از رابطه (۲۵) می‌شود.



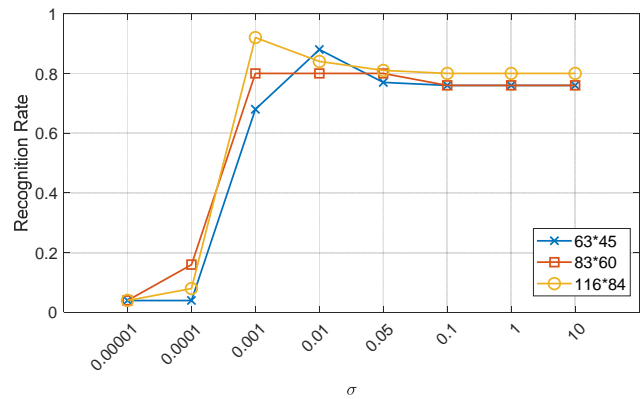
شکل ۱۳- مقایسه زمان اجرای روش CESRC و روش پیشنهادی CECRC بازای مجموعه داده AR.



شکل ۱۴- زمان اجرای دو رابطه پیشنهادی (۲۵) و (۲۸) برای حل مدل پیشنهادی CECRC برحسب تعداد داده‌های آموزشی.

## ۵ نتیجه‌گیری

در این مقاله، یک روش دسته‌بندی مقاوم برای دسته‌بندی تصاویر چهره در شرایط انسداد و تخریب به نام CECRC پیشنهاد شد. این روش، با تغییر در تابع هزینه مدل CESRC و استفاده از نرم دو بجای نرم یک بردار ضرایب و حذف قید غیر منفی بودن ضرایب معرفی گردید. به تعبیر دیگر، مدل پیشنهادی CECRC، مدلی مبتنی بر مدل CRC است که از مفهوم کورآنتروپی برای مقاوم شدن مدل نسبت به نویز و انسداد استفاده کرده است. نتایج حاصل از بررسی‌ها و آزمایشات انجام شده را می‌توان به صورت زیر خلاصه کرد:



شکل ۱۲- نمودار صحت تشخیص به ازای مقادیر مختلف پارامتر  $\sigma$  بر روی داده‌های انسداد عینک مجموعه داده AR.

## ۴-۱۱ زمان اجرا

در روش پیشنهادی CECRC، در هر تکرار، دستگاه معادلات خطی (۲۳) حل می‌شود. در این مقاله، برای حل این دستگاه معادلات، از روش سریع حذف گوسی به کمک عملگر تقسیم چپ برنامه Matlab (\) استفاده شد. در روش CESRC، در هر تکرار، یک مساله درجه دوم مقید حل می‌شود. برای حل این مساله، از دستور *quadprog* استفاده شد که از الگوریتم interior-point استفاده می‌کند. شکل ۱۳، زمان اجرای روش پیشنهادی CECRC را با روش CESRC به‌ازای مقیاس‌های متفاوت تصاویر مجموعه داده AR مقایسه می‌کند. همانطور که ملاحظه می‌شود زمان اجرای روش پیشنهادی، بهتر از روش CESRC است. لازم به ذکر است که زمان اجرای این دو روش مقاوم نسبت به تخریب و انسداد، بسیار بیشتر از روش‌های غیرمقاوم SRC و SACRC است.

در این مقاله، دو روش برای حل مدل پیشنهادی CECRC پیشنهاد شد. همانطور که گفته شد اگر تعداد ویژگی‌ها یا تعداد پیکسل‌های تصاویر چهره ( $m$ )، خیلی کمتر از تعداد تصاویر ( $n$ ) باشد برای حل دستگاه معادلات (۲۳) در مدل پیشنهادی، رابطه (۳۲) بجای رابطه (۲۵) پیشنهاد شد. دستگاه معادلات (۲۳)، یک دستگاه معادلات با  $n$  معادله و  $n$  مجهول است که برای حل آن، از روش حذف گوسی یا معکوس‌گیری از یک ماتریس  $n \times n$  مطابق با رابطه (۲۵) استفاده می‌شود. در حالی که، در صورت استفاده از رابطه (۳۲) برای حل مدل پیشنهادی، نیاز به محاسبه شبه معکوس یک ماتریس  $n \times m$  و حل دستگاه معادلات (۲۸) با  $m$  معادله و  $m$  مجهول است. اگر  $m$  خیلی کمتر از  $n$  باشد حل مدل پیشنهادی با استفاده از رابطه (۳۲) سریعتر از رابطه (۲۵) است چراکه حل دستگاه معادلات بسیار بزرگ  $n$  معادله و  $n$  مجهول یا محاسبه معکوس ماتریس بسیار بزرگ  $n \times n$  در رابطه (۲۵) بسیار زمان‌بر است. شکل ۱۴، زمان اجرای هر دو روش پیشنهادی را برحسب تعداد داده‌های آموزشی و به‌ازای تصاویر در مقیاس ۰/۲۵ نشان می‌دهد. همانطور که ملاحظه می‌شود زمانی که

- [7] W. Deng, J. Hu, and J. Guo, "Extended SRC: Undersampled Face Recognition via Intra-class Variant Dictionary," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 34, no. 9, pp. 1864–1870, 2012.
- [8] Ke-Kun Huang, Dao-Qing Dai, Chuan-Xian Ren, and Zhao-Rong Lai, "Learning Kernel Extended Dictionary for Face Recognition," *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, vol. 28, no. 5, pp. 1082–1094, 2017.
- [9] W. Liu, Y. Wang, and S. Li, "LBP Feature Extraction for Facial Expression Recognition," *Journal of Information & Computational Science*, vol. 8, no. 3, 2011.
- [10] S. Mika, "Fisher Discriminant Analysis with Kernels," in *IEEE Conference on Neural Networks for Signal Processing IX*, Madison, WI, USA, 1999, pp. 41–48.
- [11] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1–3, pp. 37–52, 1987.
- [۱۲] س. ی. ف. و. س. ج. م. ج. امان زاده، "بازبینی مدل شناسایی چهره مبتنی بر یادگیری دیکشنری کرنلی توسعه یافته"، نهمین کنفرانس فناوری اطلاعات و دانش (IKT 2017)، تهران، دانشگاه صنعتی امیرکبیر، ۱۳۹۶.
- [13] R. He, W.-S. Zheng, and B.-G. Hu, "Maximum correntropy criterion for robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1561–1576, 2011.
- [14] R. He, W.-S. Zheng, T. Tan, Z. J. I. t. o. p. a. Sun, and m. intelligence, "Half-quadratic-based iterative minimization for robust sparse representation," vol. 36, no. 2, pp. 261–275, 2014.
- [15] L. Wang and C. J. P. R. Pan, "Robust level set image segmentation via a local correntropy-based K-means clustering," vol. 47, no. 5, pp. ۲۰۱۴, ۱۹۲۵–۱۹۱۷.
- [16] C. Ding, J. Choi, D. Tao, L. S. J. I. t. o. p. a. Davis, and m. intelligence, "Multi-directional multi-level dual-cross patterns for robust face recognition," vol. 38, no. 3, pp. 518–531, 2016.
- [17] K. Guo, L. Liu, X. Xu, D. Xu, D. J. I. t. o. n. n. Tao, and l. systems, "Godec+: Fast and robust low-rank matrix decomposition based on maximum correntropy," vol. 29, no. 6, pp. 2323–2336, 2018.
- [18] T. Jin, R. Ji, Y. Gao, X. Sun, X. Zhao, and D. J. I. T. o. I. P. Tao, "Correntropy-Induced Robust Low-Rank Hypergraph," 2018.
- [19] L. Li, J. Yang, Y. Xu, Z. Qin, and H. Zhang, "Documents clustering based on max-correntropy nonnegative matrix factorization," in *Machine Learning and Cybernetics (ICMLC), 2014 International Conference on*, 2014, vol. 2, pp. 850–855: IEEE.
- [20] A. Y. Yang, S. S. Sastry, A. Ganesh, and Y. Ma, "Fast  $\ell_1$ -minimization algorithms and an application in robust face recognition: a review" EECS Department, University of California, Berkeley 2010.
- کارایی بیشتر روش معرفی شده در شرایط انسداد عینک و تخریب، در مقایسه با سایر روش‌های مقاوم و معمولی رقیب.
  - وجود قید غیرمنفی بودن بردار ضرایب در مدل CESRC، با محدود نمودن فضای پوش داده‌های هر کلاس، مقاوم بودن دسته‌بند را در شرایط تغییرات شدید نور و یا انسداد و تخریب، تحت تاثیر منفی قرار می‌دهد و حذف این قید می‌تواند موجب بهبود کارایی و مقاوم شدن تشخیص شود.
  - روش معرفی شده در شرایط بدون انسداد و تخریب نیز نتایج قابل توجهی را ارائه می‌دهد.
  - در شرایط تخریب و انسداد شدید، برتری روش پیشنهادی نسبت به سایر روش‌ها ملموس‌تر است (هر چند در این شرایط، به دلیل تخریب درصد زیادی از تصویر، دقت تشخیص به‌طور کلی پایین است).
  - زمان اجرای روش مقاوم معرفی شده در مقایسه با روش مقاوم CESRC بسیار بهتر است (با فرض استفاده از روش‌های مطرح شده برای حل مسئله بهینه‌سازی).
  - در روش پیشنهادی، از تابع گوسی برای تعیین وزن پیکسل‌ها استفاده شد. در آینده می‌توان به جای این تابع، از توابع دیگری مثل تابع  $l^1$  و Fair نیز برای وزن دهی استفاده کرد [۱۳] و اثر هر یک از این توابع وزن دهی را بررسی کرد.

## مراجع

- [1] 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.
- [2] R. Khaji, H. Li, R. A. Musleh, H. Li, and Q. Ali, "Collaborative Representation for Face Recognition based on Bilateral Filtering," *International Journal of Computer Science*, vol. 10, no. 2, 2013.
- [3] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?," in *2011 International conference on computer vision*, 2011, pp. 471–478: IEEE.
- [4] N. Akhtar, F. Shafait, and A. Mian, "Efficient classification with sparsity augmented collaborative representation," *Pattern Recogn.*, vol. 65, no. C, pp. 136–145, 2017.
- [5] Q. Liu and C. Liu, "A novel locally linear KNN model for visual recognition," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1329–1337.
- [6] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, Y. J. I. t. o. p. a. Ma, and m. intelligence, "Robust face recognition via sparse representation," vol. 31, no. 2, pp. 210–227, 2009.



- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [22] P. N. Belhumeur, J. P. Hespanha, D. J. J. I. T. o. P. A. Kriegman, and M. Intelligence, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," no. 7, pp. 711-72.۱۹۹۷, •
- [23] R. Basri, D. W. J. I. T. o. P. A. Jacobs, and M. Intelligence, "Lambertian reflectance and linear subspaces," no. 2, pp. 218-233, 2003.
- [24] A. M. Martinez and R. Benavente, *The AR face database*. 1998.
- [25] P. Viola and M. J. J. I. j. o. c. v. Jones, "Robust real-time face detection," vol. 57, no. 2, pp. 137-154, 2004.
- [26] M. Everingham, J. Sivic, A. J. I. Zisserman, and V. Computing, "Taking the bite out of automated naming of characters in TV video," vol. 27, no. 5, pp. 545-559, 2009.



**شادی کلالی** موفق به اخذ کارشناسی ارشد در رشته کامپیوتر از دانشگاه آزاد اسلامی واحد مشهد در سال ۱۳۹۷ شد. وی به عنوان معلم مشغول به فعالیت است. علاقمندی‌های وی، پردازش تصویر و بینایی ماشین میباشد.



**یحیی فرقانی** استادیار دانشگاه آزاد اسلامی واحد مشهد میباشد. وی در سال ۱۳۹۳ موفق به اخذ مدرک درجه دکتری در رشته کامپیوتر از دانشگاه فردوسی مشهد شد. وی پیشتر، مدرک کارشناسی و کارشناسی ارشد خود را از دانشگاه فردوسی مشهد و دانشگاه صنعتی اصفهان دریافت کرده بود. از جمله علاقمندی‌های وی، بهینه‌سازی، شناسایی الگو و پردازش تصویر و بینایی ماشین میباشد.



**مجید وفایی جهان** دانشیار دانشگاه آزاد اسلامی واحد مشهد می‌باشد. وی در سال ۱۳۸۸ موفق به اخذ مدرک درجه دکتری در رشته کامپیوتر از دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران شد. وی پیشتر، مدرک کارشناسی و کارشناسی ارشد خود را از دانشگاه فردوسی مشهد و دانشگاه صنعتی شریف دریافت کرده بود. از جمله علاقمندی‌های وی، مدلسازی، شبیه سازی کامپیوتری، تحلیل داده و فناوری بلاک‌چین می‌باشد.