بهبود بازشناسی چهره با یک تصویر از هر فرد به روش تولید تصاویر مجازی توسط شبکه‌های عصبی

ندا داداشی، فاطمه عبدالعلی و سیدعلی سیدصالحی
دانشکدهٔ مهندسی پزشکی، دانشگاه صنعتی امیرکبیر، تهران، ایران

چکیده:

در این مقاله با استفاده از تولید تصاویر مجازی به کمک شبکه‌های عصبی، مسئلهٔ بازشناسی چهره با یک تصویر از هر فرد مورد توجّه قرار گرفته است. برای جداسازی اطّلاعات شخص از حالت و تخمین مانیفولدهای زیرفضاهای مربوطه، از یک شبکهٔ عصبی تحلیل‌گر غیرخطی اطلاعات چهره استفاده شده است. به‌منظور افزایش تعداد نمونه‌های تعلیم در شبکهٔ طبقه‌بندی کننده، به کمک مانیفولدهای تخمین زده شده، تصاویر مجازی از چهره‌های نرمال موجود در پایگاه داده اصلی تولید شده است. با طراحی ساختارهای مختلفی از شبکه‌های عصبی مصنوعی به‌منظور استخراج مؤلّفه‌های زیرفضاهای اطّلاعات فرد و حالت، کیفیت چهره‌های مجازی و درنتیجه درصد صحّت بازشناسی در شبکهٔ طبقه‌بندی کننده بهبود می‌یابد. برای تخمین بهتر مانیفولدهای اطّلاعات شخص و بهبود قدرت تعمیم، یک روش تعلیم برمبنای خوشه‌بندی بدون سرپرستی ارائه شده است. با به کارگیری این روش و تعلیم شبکهٔ طبقه‌بندی کننده به کمک تصاویر مجازی حاصل، درصد صحّت بازشناسی ٪۸۳/۶۳ روی دادگان آزمون حاصل شده که نسبت به مدل مرجع و روش PCA به‌ترتیب دارای بهبود ٪۱۲/۷۳ و ٪۲۶/۳۶ است.

واژگان کلیدی: بازشناسی چهره، یک تصویر از هر فرد، تصاویر مجازی، یادگیری مانیفولد، شبکهٔ عصبی، تنوعات حالت

۱- مقدمه

هدف از بازشناسی چهره، شناسایی[1] و یا تأیید هویت[2] اشخاص از تصاویر و یا تصاویر ویدیویی آنها با استفاده از پایگاه داده از چهره‌های این افراد می‌باشد. فن‌آوری بازشناسی چهره به‌دلیل داشتن مزایایی ازجمله دقت بالا و تهاجم پایین، در مواردی مانند امنیت اطلاعات، اجرا و نظارت بر قانون، کارت های هوشمند، کنترل دسترسی و غیره مورد استفاده قرار می‌گیرد (Zhao, et al., 2003؛ Chellappa, et al., 1995؛ Daugman, 1997). علی‌رغم تحقیقات انجام شدهٔ بسیار طی چند دههٔ گذشته، به‌دلیل پیچیدگی ذاتی مسئله، بازشناسی چهره هنوز با مشکل روبه‌روست. این مشکلات ناشی از تنوعات چهرهٔ انسان به‌دلیل تغییرات شرایط نورپردازی، پوشیدگی و حالت است (Zhao, et al., 2003).

در اکثر روش‌های به کار رفته، فرض بر این است که تصاویر متعدد از هر فرد در حالت‌های مختلف برای تعلیم موجود است (Zhao, et al., 2003). درحالی که در بسیاری از کاربردهای واقعی، جمع‌آوری تصاویر متعدد از هر فرد با محدودیت‌هایی مواجه است و علاوه‌بر آن هزینه‌های ذخیره و زمان پردازش داده‌ها نیز افزایش می‌یابد (Tan, et al., 2006).

روش‌هایی که برای تشخیص فرد در حالت‌های مختلف تنها از یک تصویر وی برای تعلیم سیستم استفاده می‌کنند، موسوم به "روش‌های یک نمونه از هر شخص[3] " می‌باشند. در این روش‌ها محدودیتی برای تصاویر آزمون

[1] Identify
[2] Verify
[3] One sample per person

نبوده و این تصاویر می‌توانند در حالات و شرایط نورپردازی متفاوت باشند. باید توجّه کرد که در شرایط یک تصویر از هر فرد، کارآیی بسیاری از روش‌ها کاهش یافته و حتی برخی از روش‌ها با شکست مواجه می‌شوند. بنابراین، در سال‌های اخیر، این زمینه علی‌رغم مشکلات بیشتر، مورد توجّه قرار گرفته است (*Tan, et al., 2006*).

تاکنون برای حل این مشکل از روش‌های مختلفی استفاده شده است. در برخی روش‌های مورد استفاده از ویژگی‌های موضعی برای تشخیص افراد استفاده شده است. به‌طور مثال، در (*Gao, et al., 2005*) از ویژگی‌های نقاط گوشه‌ای جهت‌دار[1] و در (*Manjunath, et al., 1992*؛ *Wiskott, et al., 1997*؛ *Lades, et al., 1993*؛ *Kepenekci, et al., 2002*) نظریهٔ انطباق گراف مورد استفاده قرار گرفته است. همچنین، روش زیرفضای احتمالاتی محلی (*Martinez, 2002*)، شبکه‌های عصبی خودسازمانده (*Lawrence, et al., 1997*؛ *Tan, et al., 2005*)، مدل مخفی مارکوف (*Le, et al., 2004*)، روش توسعه یافتهٔ آنالیز متمایزگر خطّی[2] (*Chen, et al., 2004*؛ *Huang, et al., 2003*)، ویژگی‌های ترکیبی موضعی (*Lam and Yan, 1998*)، بازشناسی چهره با الگوهای باینری موضعی (*Ojala, et al., 2002*) و بازشناسی برمبنای ویژگی‌های فرکتالی (*Komleh, et al., 2001*) نیز در این دسته قرار دارند.

درمقابل، گروهی دیگر از روش‌ها از کل چهره به‌عنوان ورودی استفاده می‌کنند. در (*Wu and Zhou, 2002*) از رویکرد آنالیز مؤلّفه‌های اساسی به‌منظور غنی‌کردن اطّلاعات چهره استفاده شده است. در (*Yang, et al., 2004*) نیز روش آنالیز مؤلّفه‌های اساسی دوبعدی به کار گرفته شده و در تخمین ماتریس کووریانس به جای بردارهای یک‌بُعدی از ماتریس دوبعدی تصاویر استفاده شده است. در روش پیشنهادی (*Jung, et al., 2004*) با استفاده از مدل نوفه با سه پارامتر، تصاویر جدید چهره (مشابه تصاویر مخدوش شده) به‌منظور بازشناسی تولید شده‌اند. در (*Wang, et al., 2005*) به‌دلیل وجود تنها یک تصویر در هر کلاس، پراکنش بین کلاسی قابل تخمین نبوده و بنابراین فرض‌شده که کلاس‌های مختلف دارای تنوعات درون کلاسی مشابه هستند. درنتیجه در رویکرد پیشنهادی، دانش پیشین پراکنش بین کلاسی سایر افراد نیز برای تخمین پراکنش داخل کلاسی مورد استفاده قرار می‌گیرد.

یکی از رویکردهای اصلی برای حل مسأله یک تصویر از هر شخص با استفاده از ویژگی‌های عمومی[3]، افزایش تعداد داده‌های تعلیم می‌باشد. بدین منظور از داده‌های مجازی استفاده می‌شود که این داده‌ها می‌توانند نمایش های[4] جدید از تنها تصویر موجود و یا تصاویر مجازی جدید باشند. در (*Frade, et al., 2005*) با استفاده از فیلترهای خطی و غیرخطی نمایش‌های جدید ایجاد شده‌اند. در (*Martinez, 2003 and 2004*) از آشفتگی[5] تصویر برای ایجاد نمایش‌های جدید استفاده شده است. در (*Chen, et al., 2004*) روش آنالیز مؤلفه‌های اساسی دوبعدی (*Wu and Zhou, 2002*) به کمک تصاویر n بُعدی توسعه و بهبود یافته است. در (*Beymer and Poggio, 1996*؛ *Vetter, 1998*؛ *Niyogi, et al., 1998*) به‌منظور حل مسئله بازشناسی چهره با حالت متغیر، تصاویر مجازی با استفادهٔ دانش پیشین راجع به تصاویر چهره تولید شده است.

در این مقاله با استفاده از قابلیت‌های شبکه‌های عصبی در پردازش غیرخطی سیگنال، ساختارهایی برای آنالیز مؤلفه‌های اساسی غیرخطی پیشنهاد شده است. با استفاده از ساختارهای پیشنهادی و یک پایگاه دادهٔ عمومی شامل تصاویر افراد در حالات و شرایط نورپردازی مختلف، اطلاعات افراد از حالت جدا شده و این اطلاعات برای تخمین مانیفولدهای زیرفضاهای مربوط به اطلاعات فرد و حالت مورد استفاده قرار گرفته است. با ترکیب غیرخطی مؤلفه‌های حالات مختلف و مؤلفه‌های حالت نرمال شخص، می‌توان تخمینی از حالات مختلف فرد به‌دست آورد. در شبکهٔ طبقه‌بندی کننده، تصاویر نرمال و تصاویر مجازی سنتز شده به‌عنوان دادهٔ تعلیم و تصاویر واقعی اشخاص در حالات مختلف به‌عنوان دادهٔ آزمون به کار رفته‌اند. ساختار این مقاله به شرح زیر است:

در بخش دوم، جنبه‌های نظری جداسازی اطلاعات شخص از حالت به‌صورت مستقل از یکدیگر مورد بررسی قرار می‌گیرد. در بخش سوم به معرفی ساختار شبکه‌های عصبی پیشنهادی پرداخته و به دنبال آن در بخش چهارم، نحوهٔ ارزیابی نتایج ارائه می‌گردند. درنهایت بخش پنجم شامل جمع‌بندی و نتیجه‌گیری است.

[1] Directional Corner Points
[2] Linear discriminant analysis
[3] global
[4] Representation
[5] Perturbation

۲- جداسازی اطلاعات فرد از حالت

تصویر چهره حاوی دو دسته اطلاعات مستقل " اطلاعات شخص" و "اطلاعات حالت" می‌باشد که اطلاعات شخص هویت شخص و اطلاعات حالت چگونگی حالت و وضعیت شخص را منتقل می‌کنند. با ترکیب غیرخطی این دو دسته اطلاعات تصویر شخص در حالت‌های گوناگون و یا حالت یکسان افراد مختلف شکل می‌گیرد. مغز انسان به‌خوبی قادر به تفکیک این دو دسته اطلاعات بوده و با دیدن چهرهٔ یک شخص خاص در حالت‌های مختلف قادر به شناسایی فرد و با دیدن چهره اشخاص ناشناس در حالت‌های گوناگون قادر به تشخیص حالت افراد می‌باشد.

در بازشناسی چهره با استفاده از یک تصویر از هر فرد، از هر شخص فقط یک تصویر نرمال برای تعلیم موجود بوده و تصاویر آزمون در حالت‌های مختلف می‌باشند. بنابراین به مدل بازشناس فقط یک تصویر نرمال تعلیم داده می‌شود. مشکل اصلی در اینجا عدم توانایی یک تصویر برای ایجاد تعمیم لازم برای بازشناسی می‌باشد. به عبارت دیگر یک تصویر از هر شخص قادر به شکل‌دهی مناسب مانیفولد حالات مختلف فرد نمی‌باشد. رویکرد پیشنهادی تخمین مجازی حالت‌های دیگر فرد و استفاده از حالت‌های مجازی تولید شده برای تعلیم مناسب مانیفولدهای حالات مختلف این فرد به مدل یادگیرنده می‌باشد. چنین مدلی پس از یادگیری، بهتر از مدل یادگیرنده فقط با یک تصویر قادر به بازشناسی حالت مختلف فرد خواهد بود.

اگر بردار n بُردار n بُعدی تصویر فرد در یک حالت باشد، این تصویر در فضای m بُعدی ورودی با یک نقطه نشان داده می‌شود. با تغییر حالت این فرد، نقطه در فضای ورودی حرکت کرده و زیرمانیفولد غیرخطی ایجاد شده از حرکت این نقطه در فضای ورودی ***مانیفولد تغییرات حالت*** این شخص نامیده می‌شود. با تغییر حالات و تغییر افراد در فضای ورودی، یک مانیفولد غیرخطی پوشش داده می‌شود که در برگیرنده زیرمانیفولدهای مربوط به حالات مختلف هر فرد در فضای ورودی خواهد بود. برای تشخیص افراد از چهره داده شده، باید تعلق هر فرد را به هر کدام از این زیر مانیفولدها تعیین شود.

حال اگر $\bar{x}$ حالت نرمال شخص و قرارگرفته بر روی زیرمانیفولد مربوطه باشد، می‌توان رابطه (۱) را به‌صورت زیر نوشت

$$\bar{X}_j = \bar{x} + v_j \qquad (۱)$$

که در آن $\bar{X}_j$ نشان‌دهنده حالت ***j***ام و v_j تغییراتی است که در فضای ورودی بر حالت نرمال برای به‌دست آمدن حالت جدید اثر می‌کند.

برای تخمین خوب وکارآمد مانیفولد تغییر حالت باید v_j به‌خوبی تخمین زده شود. همچنین باید جداسازی v_j و $\bar{x}$ در صورت در دسترس بودن $\bar{X}_j$ به‌خوبی انجام پذیرد. این اثر ترکیب غیرخطی سیگنال ها از منابع مختلف می باشد که برای جداسازی آنها باید مؤلفه‌های مستقل غیرخطی ایجاد کنندهٔ حالات مختلف در فرد را داشته و از ترکیب آن مؤلّفه‌ها با حالت نرمال فرد، حالت های مختلف او را با شباهت‌های بالا با حالت واقعی، ایجاد کرد.

شبکه‌های عصبی مصنوعی ابزارهای بالقوه برای طبقه‌بندی‌کننده و سیستم‌های تصمیم‌گیری می‌باشند. اطلاعات بر روی نورون‌ها و وزن‌ها گسترده شده است. پس از تعلیم شبکه‌های عصبی، می‌توان از آنها در کاربردهای مختلف بینایی ماشین ازجمله پیش‌پردازش (تعیین لبه، ترمیم و پالایش تصاویر)، استخراج ویژگی، حافظهٔ انجمنی (ذخیره و بازیابی اطلاعات) و بازشناسی الگو *(سیدصالحی ۱۳۸۳)* استفاده کرد.

پردازش گسترده و موازی اطّلاعات روی نورون ها و به همراه استفاده از توابع غیرخطّی نرم و مرزهای تصمیم فازی امکان پردازش توأم تأثیرات غیرخطی متقابل را در شبکه‌های عصبی فراهم می‌آورد. به این ترتیب با تعریف کردن مناسب منابع مختلف سیگنال و چگونگی تعامل آن‌ها نسبت به یکدیگر از طریق تعلیم دادگان به شبکه‌های عصبی مصنوعی مختلف، می توان شرایط چنین پردازش غیرخطی سیگنال‌ها را فراهم کرد.

بنابراین برای رسیدن به این هدف می توان از یک شبکهٔ عصبی خودانجمنی به‌منظور آنالیز مؤلّفه های اساسی غیرخطی که مؤلّفه‌های حالات را از افراد جدا می‌کند، استفاده کرد. برای به‌دست آوردن این مؤلفه ها، شبکه با تعداد زیادی داده که شامل افراد مختلف با حالات مختلف هستند، تعلیم داده می شود تا مانیفولدهای غیرخطی تغییر حالات را به‌خوبی تخمین زده شود.

بیان مناسب مؤلفه های مستقل در فضای مؤلفه‌ها در بهبود عملکرد سیستم، نقش به‌سزایی دارد. زیرمانیفولد مربوط به افراد به‌صورت جداگانه و زیرمانیفولد مربوط به حالت ها نیز جداگانه و به نحو مطلوب باید بیان شده و سیستم تعلیم داده شود. برای بیان افراد می‌توان از بیان تک‌بیت به این صورت که برای هر فرد یک نورون در خروجی در نظر گرفته شده و درصورت تشخیص فرد، عدد ۱ به آن نورون اختصاص یافته و یا به‌صورت باینری که برای هر فرد یک کد باینری در نظر گرفته شود، استفاده کرد. تعلیم مناسب این دو زیرفضا به شبکه، نیازمند تعداد کافی از دادگان مناسب می‌باشد تا نقاط کافی از مانیفولدهای تغییر حالات و تغییر افراد را ایجاد کرده و تخمین مانیفولد

را دقیق‌تر نماید. به هر حال، به‌دلیل کمبود داده در مسئله یک تصویر از هر فرد، از یک پایگاه دادهٔ عمومی که شامل تصاویر تعلیم متفاوت با تصاویر آزمون است، استفاده می‌شود. شبکهٔ عصبی به‌منظور استخراج اطّلاعات متمایزکننده تعلیم داده می‌شود. از آنجایی که تصاویر چهرهٔ انسان‌ها تنوعات درون فردی[1] مشابه دارد، اطّلاعات متمایزکنندهٔ فرد هدف با توجّه به اطّلاعات سایر افراد قابل یادگیری است *(Wang, et al., 2005)*.

۳- ساختار شبکهٔ عصبی

در این بخش به معرفی دادگان و ساختار شبکه‌های عصبی مورد استفاده در این تحقیق می‌پردازیم.

۳-۱- معرفی دادگان تعلیم و آزمون

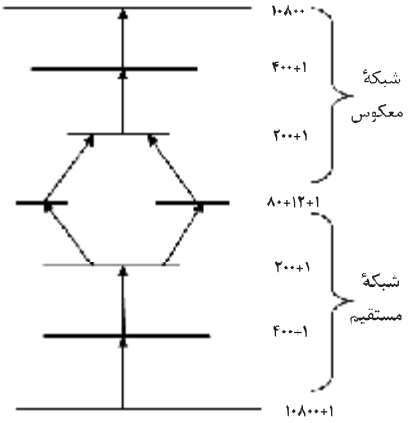
در این تحقیق بخشی از پایگاه داده AUT که در دانشکده برق دانشگاه صنعتی امیرکبیر تهیه شده‌است، به‌عنوان پایگاه دادهٔ عمومی مورد استفاده قرار گرفته‌است. این دادگان شامل ۹۶۰ تصویر از ۸۰ نفر می‌باشد. به ازای هر شخص دوازده تصویر موجود می‌باشد که تغییرات این تصاویر شامل جهت چهره (نگاه روبه‌رو، نگاه به راست و چپ، نگاه به بالا و پایین)، ظاهر چهره با عینک، حالات احساسی چهره (لبخند،اخم، تعجب، چشمان بسته) و تفاوت در میزان روشنایی(نور از راست و نور از چپ) در نظر گرفته شده است. تصاویر به‌صورت سیاه و سفید با ۲۵۶ سطح خاکستری و دارای وضوح ۱۲۰×۹۰ می باشند. پایگاه دادهٔ اصلی که آزمایش‌ها بر روی آن انجام می‌شود، شامل ده نفر و دوازده تصویر از هر فرد می‌باشد که از این پس این پایگاه داده ZD نامیده می‌شود. این حالت‌ها مشابه حالت‌های پایگاه داده عمومی می‌باشد. این پایگاه داده در دانشکدهٔ مهندسی پزشکی دانشگاه صنعتی امیرکبیر تهیه شده و تمام تصاویر در پس‌زمینهٔ یکسان تهیه شده و تنوع سنی نیز تا حد ممکن در آن رعایت شده است.

۳-۲- شبکهٔ عصبی مستقیم - معکوس

در (شکل ۱) نخستین ساختار شبکهٔ عصبی پیشنهادی برای جداسازی اطلاعات شخص از حالت، نشان داده شده است. ساختار پیشنهادی یک شبکهٔ عصبی، خودانجمنی است. با دادن تصاویر در حالات مختلف افراد (۸۰ نفر پایگاه داده AUT در ۱۲ حالت از هر شخص) در ورودی، اطّلاعات حالت و شخص در لایهٔ گلوگاه از هم جدا می‌شوند.

در این شبکه لایهٔ پنهان اوّل و دوم وظیفهٔ نگاشت اطّلاعات از فضای ورودی به فضای ویژگی را داشته و دو قسمت طبقهٔ خروجی جهت طبقه‌بندی افراد و حالات می باشد. این شبکه با الگوریتم پس‌انتشار خطا و به‌صورت با سرپرستی تعلیم دیده است. به این صورت که برای هر شخص و هر حالت کُدی در خروجی به‌عنوان خروجی مطلوب تعریف شده است. این کدها به‌صورت تک‌بیتی می‌باشد و در هر حالت و یا شخص یکی از بیت های مربوطه به شخص و حالت یک می شود. به‌عبارت دیگر برای هر تصویر دو نورون، یک نورون در قسمت حالت خروجی و یک نورون در قسمت شخص خروجی یک می‌شود.

نقش قسمت دوم این شبکه به‌طور دقیق معکوس شبکهٔ مستقیم می‌باشد. لایه‌های پنهان در این شبکه وظیفهٔ نگاشت معکوس[2] اطّلاعات از فضای ویژگی به فضای تصویر را دارند. لایهٔ ورودی این شبکه مشابه لایهٔ خروجی شبکهٔ مستقیم بوده و متشکل از کدهایی می‌باشد که نمایندهٔ فرد و حالت آن می باشند. لایهٔ خروجی تصویر مرتبط با کد ورودی را می‌سازد. این شبکه نیز به‌صورت با سرپرستی تعلیم دیده است.



(شکل ۱): شبکهٔ مستقیم- معکوس.

پس از تعلیم شبکه، مانیفولد غیرخطی زیرفضای تغییر حالت تخمین زده شده است. از این مانیفولد برای تولید حالات مجازی اشخاص موجود در پایگاه داده ZD[3] استفاده شده است. بدین منظور چهره‌های نرمال موجود در این پایگاه دادهٔ در ورودی شبکه قرار داده می‌شود. کد به‌دست آمده در قسمت افراد نشان‌گر میزان شباهت این فرد به هر کدام از ۸۰ نفر پایگاه داده AUT می‌باشد.

۳-۲-۱- تولید تصاویر مجازی و تعلیم شبکهٔ طبقه‌بندی کننده

پس از تعلیم شبکه، از آن می‌توان برای تولید تصاویر مجازی نمونه‌های پایگاه ZD استفاده کرد. بدین منظور

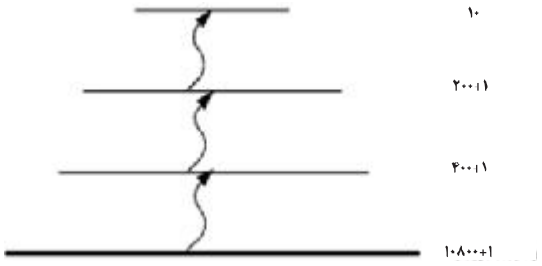
[1] intra-subject

[2] Decoding

[3] Zamani Database

چهرهٔ نرمال فرد در ورودی شبکه قرار داده می‌شود. با ثابت نگه‌داشتن کد ایجاد شده در قسمت اطّلاعات اشخاص و تغییر کد حالت در قسمت اطّلاعات حالت، تصاویر جدید در خروجی تولید می‌شوند.

برای آزمودن تصاویر مجازی تولید شده از یک شبکهٔ طبقه‌بندی کننده استفاده شده که ساختار آن در (شکل ۲) نشان داده شده است. این شبکه دارای دو لایهٔ پنهان بوده، لایهٔ ورودی شبکه دارای تعداد نورون‌های برابر پیکسل‌های تصویر (۱۰۸۰۰ نورون)، لایهٔ پنهان اوّل دارای ۴۰۰ نورون و لایهٔ پنهان دوم دارای ۲۰۰ نورون با تابع غیرخطی سیگموید می‌باشد.



(شکل ۲): شبکهٔ طبقه‌بندی کننده با بیان خروجی تک بیت.

در لایهٔ خروجی از بیان تک‌بیتی استفاده شده است. به این صورت که برای هر طبقه (شخص) یک نورون در خروجی در نظر گرفته شده است که در هنگام تعلیم آن نورون برابر یک و بقیهٔ نورون‌ها برابر صفر در نظر گرفته می‌شود. به‌عنوان مثال برای شخص اوّل کد 0000000001 و برای شخص دوم کد 0000000010 در نظر گرفته شده‌است.

جهت تعلیم این شبکه، چهره‌های نرمال واقعی و دوازده چهرهٔ مجازی تولید شده برای هر شخص به‌عنوان داده‌های تعلیم و یازده چهرهٔ واقعی هر شخص به‌عنوان داده‌های آزمون در نظر گرفته شده است. برای اطمینان از اینکه شبکه بیش از اندازه تعلیم ندیده باشد، جایی‌که درصد صحّت بر روی داده‌های آزمون مقدار بیشینه را داشته و پس از آن روند کاهشی داشته باشد، تعلیم متوقف می‌شود.

در (شکل‌های ۳ و ۴)، تصاویر مجازی تولیدی متناظر با دو ورودی مختلف نمایش داده شده است. همان‌طور که مشاهده می‌شود، با وجود اینکه حالات تصاویر مجازی به‌خوبی ایجاد شده است، ولی میزان شباهت تصاویر تولیدی به شخص اصلی بسیار پایین می‌باشد. به نظر می‌رسد که بیان افراد برحسب افراد پایگاه داده عمومی به‌خوبی انجام نشده و بنابراین استخراج ویژگی رضایت‌بخش نیست.

(شکل ۳): تصاویر مجازی تولیدی برای شخص اوّل از پایگاه داده ZD به‌دست آمده توسط شبکهٔ مستقیم- معکوس.

(شکل ۴): تصاویر مجازی تولیدی برای شخص دوم از پایگاه داده ZD به‌دست آمده توسط شبکهٔ مستقیم- معکوس.

۳-۳- تولید کُد شخص به‌صورت بدون سرپرستی

در (شکل ۵) ساختار اصلاحی پیشنهادی برای جداسازی بهتر اطّلاعات حالت از چهره نشان داده شده است. این ساختار یک شبکهٔ خودانجمنی است که نورون‌های لایهٔ میانی آن به دو بخش بدون سرپرستی و با سرپرستی تقسیم بندی شده است. به نظر می‌رسد که کد توصیف کنندهٔ شخص، بیان مناسبی برای اطّلاعات شخص نبوده و شبکه در این قالب قادر به شکل‌دهی درست مانیفولدهای زیرفضای شخص نمی‌باشد. بنابراین در این ساختار استخراج کد شخص را به عهدهٔ شبکه گذاشته و انتظار می‌رود کد بهینه را برای توصیف شخص در قسمت بدون سرپرستی استخراج کند. در بخش باسرپرستی، برای هر حالت کد خاصی تخصیص می‌یابد و خطا به لایه‌های زیرین پس‌انتشار می‌یابد و در قسمت بدون سرپرستی لایهٔ گلوگاه، شبکه به‌صورت بدون سرپرستی کد مناسب اشخاص را ایجاد می‌کند.

لایهٔ گلوگاه شامل نودودو نورون است که دوازده نورون برای اطّلاعات حالت و هشتاد نورون برای اشخاص تخصیص یافته است. در مرحلهٔ تعلیم، چهره‌های موجود در پایگاه داده AUT به ورودی شبکه داده می‌شود. خروجی مطلوب همان تصویر ورودی می‌باشد. در دو لایهٔ اوّل (بخش کدینگ)، اطّلاعات شخص و حالت جدا شده و در قسمت دی‌کدینگ اطّلاعات تجزیه شده جهت بازسازی چهره، بازترکیب می‌شوند. درواقع بخش کدینگ را می‌توان

به‌صورت یک تابع غیرخطی پیوسته r درنظر گرفت که تصاویر وروی x را از فضای پیکسل به فضای کد نگاشت می‌کند و بُردار y بدین ترتیب تولید می‌شود:

$$y = r(x) \qquad r : R^n \rightarrow R^m \qquad (۲)$$

دراین رابطه، n بُعد فضای پیکسل و m بعد فضای کد است. بخش دی‌کدینگ را نیز می‌توان به‌صورت تابع غیرخطی پیوسته g درنظر گرفت که فضای کدها را به فضای پیکسل نگاشت می‌کند تا تصاویر بازسازی‌شده $\hat{x}$ تولید شوند.

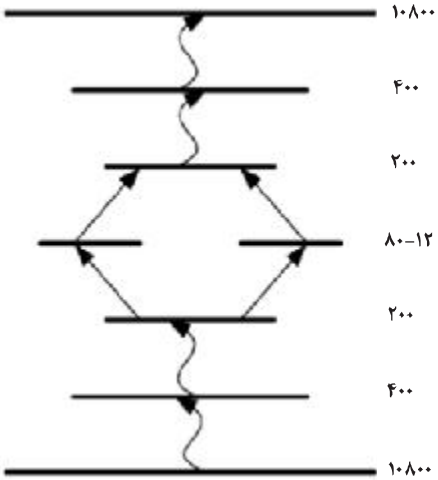
$$\hat{x} = g(y) \qquad g : R^m \rightarrow R^n \qquad (۳)$$

از ترکیب روابط (۲) و (۳) ، رابطه زیر به‌دست می‌آید:

$$\hat{x} = g(r(x)) \qquad (۴)$$

شبکه باید با تنظیم وزن‌ها، خطای خروجی را کمینه کند که درنهایت برای هر فرد کدهای مناسب را در قسمت بدون سرپرستی ایجاد و تصویر را به نحو صحیحی در خروجی بسازد. در این شبکه تابع هزینه به‌صورت $C = \sum_i (x_i - \hat{x}_i)^2$ می‌باشد.

از آنجایی‌که برای حالت‌های یکسان در افراد مختلف کد حالت یکسان است، انتظار می‌رود قسمت مشترک در این تصاویر که مربوط به حالت چهره می‌باشد، در قسمت با سرپرستی جدا شده و باقی اطّلاعات در قسمت بدون سرپرستی، کد شخص را ایجاد کند.



(شکل ۵): ساختار شبکهٔ عصبی پیشنهادی جهت تولید خودکار کد شخص.

در تعلیم شبکه تفاضل خطای خروجی و خروجی مطلوب برای اصلاح وزن‌ها پس‌انتشار می‌شود. در لایهٔ گلوگاه در قسمت با سرپرستی این خطا با تفاضل خطای کد مطلوب حالت و کد ایجاد شده جمع شده و پس‌انتشار ادامه می یابد. در (شکل ۶) تصاویر مجازی تولیدی توسط شبکه پیشنهادی نشان داده شده است.

همان‌طور که مشاهده می‌شود، تصاویر مجازی به‌صورت قابل قبول ایجاد نشده‌اند. در تحلیل عدم بازسازی مناسب حالت تصاویر می‌توان گفت که چون در قسمت بدون سرپرستی قیدی جهت عدم عبور اطّلاعات حالت وجود نداشته، بنابراین اطّلاعات حالت نیز در شکل گیری کدهای این قسمت دخیل بوده‌اند. این امر موجب شده‌است که حالت تصاویر بازسازی شده حالت میانگینی از تمامی حالت‌ها باشد. برای حلّ این مشکل باید اصلاحاتی انجام شود تا از عبور اطّلاعات حالت از این قسمت جلوگیری کند و جداسازی اطّلاعات شخص از حالت دقیق‌تر انجام شود.

۳-۴- شبکهٔ عصبی با خوشه‌بندی بدون سرپرستی لایهٔ اطّلاعات شخص

همان‌گونه که درقبل نیز اشاره شد، اشکال عمده در ساختارهای قبلی عدم وجود قید در ایجاد کد شخص توسط شبکه و تولید کدهای متفاوت جهت نمایش یک شخص در حالات مختلف است. در (شکل ۷)، طرح خوشه‌بندی بدون سرپرستی اطّلاعات افراد نمایش داده شده است. در این قسمت به‌دلیل عدم وجود قید کافی، اطّلاعات شخص به‌خوبی جدا نشده و قسمتی از اطلاعات حالت نیز در کد شخص تأثیرگذار می‌باشد؛ بنابراین برای بیان بهتر کد شخص، ساختار جدیدی ارائه شده است. در این ساختار تلاش می‌شود که شبکه به‌صورت بدون سرپرستی، حالت‌های مختلف هر فرد را در یک خوشه قرار داده و برای حالت‌های مختلف شخص کد یکسانی را تخصیص دهد.

(شکل ۶): تصاویر مجازی تولید شده توسط روش ایجاد کد شخص به‌صورت بدون سرپرستی.

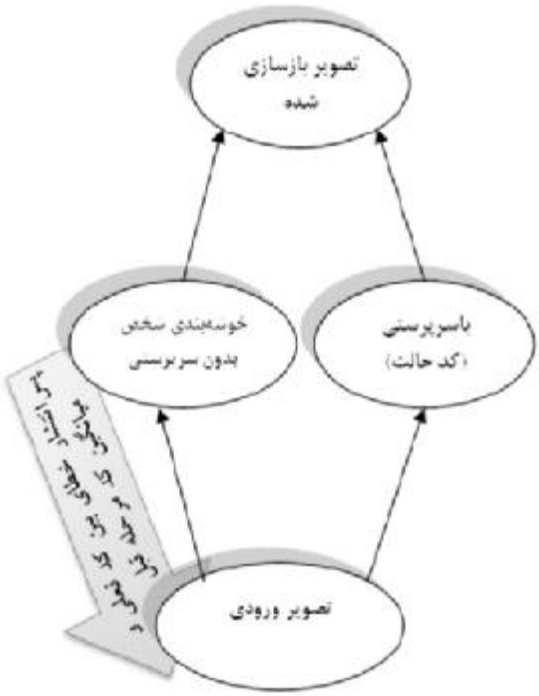
در این شبکه نیز مانند ساختارهای قبل، ورودی نورون j ام از حاصل‌ضرب ماتریس ورودی $\overline{o}$ در وزن متناظر w_{ij} به‌دست می‌آید. ($net_j = \sum_i w_{ij} o_i$) همچنین از تابع فعال‌سازی غیرخطی سیگموید استفاده شده است و θ نیز بیانگر مقدار بایاس است. منظور از y_i خروجی نورون‌های لایهٔ پنهان اول و دوم شبکهٔ مستقیم می‌باشد.

$$y_j = \frac{1}{1 + \exp(net_j - \theta)} \qquad (۵)$$

این شبکه نیز به‌صورت خودانجمنی تعلیم می‌بیند و نورون‌های لایهٔ گلوگاه برای جداسازی مؤلّفه‌های فرد از حالت، به دو بخش تقسیم شده است. قسمت حالت همانند طرح قبل با تخصیص کد بوده و به‌صورت با سرپرستی تعلیم می‌بیند. تفاوت این طرح با طرح قبل در قسمت ایجاد کد اشخاص می‌باشد. در ساختار پیشنهادی، در هر بار تعلیم تصاویر یک شخص، متوسط مقادیر خروجی نورون‌های بخش خوشه‌بندی شخص در لایهٔ گلوگاه توسط رابطهٔ (۶) به‌دست آمده و به‌عنوان کد شخص در مرحلهٔ بعد تعلیم استفاده می‌شود.

$$M_{i+1} = (\gamma \cdot M_i) + (1-\gamma) \cdot P \qquad (۶)$$

در این رابطه، M نشان‌گر کد به هنگام شده برای شخص، P خروجی نورون‌ها در بخش خوشه‌بندی و γ ضریب میزان تأثیر کد جدید (در اینجا ۰/۹۵ درنظر گرفته شده) می‌باشد. اندیس i بیان‌گر تصاویر هر فرد در ۱۲ حالت متفاوت است که بدین ترتیب شبکه این تصاویر را در یک خوشه قرار می‌دهد. بنابراین علاوه‌بر پس‌انتشار خطای خروجی، در هر مرحله خطای بین کد تولید شده و کد مرحله قبل تعلیم نیز بر روی وزن‌های مسیر پس‌انتشار می‌شود. بدین ترتیب شبکه با استفاده از دادگان موجود برای هر شخص، کد خاصی را در لایهٔ میانی تولید کرده و بدون نیاز به کد از پیش تعیین شده، لایهٔ شخص را به خوشه‌های مجزایی دسته‌بندی کند.



(شکل ۷): طرح خوشه‌بندی بدون سرپرستی اطلاعات افراد.

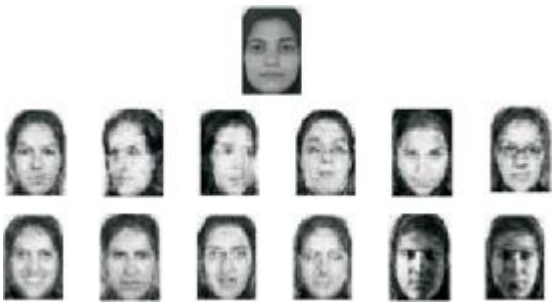
همانند بخش‌های قبل، از وزن های این روش نیز در تولید تصاویر مجازی استفاده شده است. تصاویر مجازی تولیدی در (شکل ۸) نشان داده شده است. همان‌طور که مشاهده می‌شود، در تصاویر مجازی حاصل حالت‌ها بهتر ایجاد شده‌اند، اما هنوز برخی حالات به‌وضوح، شکل نگرفته‌اند. از این تصاویر جهت تعلیم شبکهٔ طبقه بندی کننده استفاده شده است.

(شکل ۸): تصویر مجازی تولید شده توسط شبکهٔ خوشه‌بندی افراد.

۳-۵- شبکهٔ عصبی با خوشه‌بندی بدون سرپرستی در هر دو لایهٔ اطّلاعات شخص و حالت

در این ساختار، از مزایای روش خوشه‌بندی بدون سرپرستی، در لایهٔ اطّلاعات حالت نیز بهره برده شده است. شبکه با استفاده از دادگان موجود برای هر حالت نیز کد خاصی را در لایهٔ میانی تولید کرده و بدون نیاز به کد از پیش تعیین شده، لایهٔ اطّلاعات حالت را به خوشه‌های مجزایی دسته‌بندی می‌کند. همچنین به‌منظور استخراج مؤلفه‌های مشترک بهینه، بیت‌های حالت به هشت و بیت‌های افراد به شانزده کاهش یافته است. با افزایش کیفیت تصاویر مجازی تولیدی و افزایش درصد صحّت بازشناسی نشان داده شده است که با تعداد نورون کمتر و خوشه‌بندی در هر دو بخش اطّلاعات فرد و حالت، مؤلّفه‌های مشترک بهتری حاصل می‌شود. درحقیقت این اثر مشابه هرس‌کردن شبکه می‌باشد، زیرا تعداد زیاد نورون‌های لایهٔ پنهان مضر بوده و موجب ایجاد ابرصفحهٔ زائد می‌شود. در اینجا نیز ساختار شبکه دارای افزونگی است و با تعداد نورون کمتر می‌توان به مؤلّفه‌های مشترک بهتری دست پیدا کرد. پس از استخراج مؤلّفه‌های مشترک، قدرت تعمیم شبکه و کیفیت تصاویر مجازی حاصل افزایش یافته، درنتیجه درصد صحت بازشناسی نیز افزایش می‌یابد. درحقیقت، نحوهٔ شکل‌گیری مؤلّفه‌های غیرخطی مربوط به

دو زیرفضا به‌طور مستقل از یکدیگر بر قدرت تعمیم شبکه بسیار تأثیرگذار است. تصاویر مجازی تولید شده به کمک این روش، در (شکل ۹) نمایش داده شده است. همان‌گونه که مشاهده می‌شود، اگرچه در شباهت تصاویر مجازی حاصل به شخص مورد نظر تغییری حاصل نشده، اما کلیهٔ حالات به‌وضوح، شکل گرفته‌اند.



(شکل ۹): تصویر مجازی تولید شده توسط روش خوشه‌بندی در هر دو سمت افراد و حالات.

۴- نتایج شبیه‌سازی

در این بخش در ابتدا به‌عنوان مدل مرجع، یک شبکهٔ عصبی جلوسوی چندلایه به‌طور مستقیم با تصاویر یک حالت نرمال از هر فرد تعلیم دیده و نتایج عملکرد بازشناسی آن روی تصاویر سایر حالات این افراد مورد ارزیابی قرار می‌گیرند. همچنین از آنجا که از روش PCA در مقالات به‌عنوان یک مبنای محک و ارزیابی نامبرده می‌شود، در این مقاله نیز این روش با عنوان روش محک انتخاب شده و مورد مقایسه قرار می‌گیرد که جزئیات پیاده‌سازی این روش در بخش ۴-۲ آمده است. در بخش ۴-۳ به بیان نتایج حاصل از تعلیم شبکه طبقه‌بندی کننده توسط تصاویر مجازی حاصل از شبکه‌های تحلیل‌گر پیشنهادی می‌پردازیم و نتایج بازشناسی حاصل با دو روش قبلی مقایسه می‌شوند.

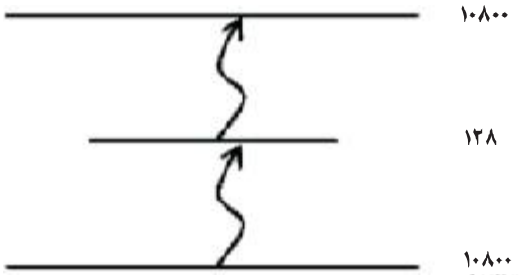
۴-۱- نتایج بازشناسی مدل مرجع

در گام اوّل و به‌عنوان مدل مرجع، شبکهٔ عصبی طبقه‌بندی کننده (شکل ۲) را تنها با تصاویر نرمال ده نفر موجود در پایگاه دادهٔ ZD تعلیم می‌دهیم و سپس آزمون مدل با تصاویر افراد در یازده حالت دیگر انجام می‌شود. بنابراین در این حالت تعداد تصاویر تعلیم و آزمون به‌ترتیب برابر با ۱۰ و ۱۱۰ می‌باشد. نتایج حاصل از این روش در سطر دوم جدول ۱ آمده است. همان‌گونه که مشاهده می‌شود، درصد صحت بازشناسی این مدل مرجع بر روی دادگان آزمون برابر با ۷۰/۹٪ است و بیان‌گر این است که این مدل بازشناسی با تعلیم یک تصویر از هر فرد توانایی ایجاد تعمیم کافی برای بازشناسی کامل سایر تنوعات حالت چهره را ندارد و به عبارت دیگر قادر به شکل دهی مناسب مانیفولد نواحی تصمیمی که شامل حالات مختلف فرد باشند، نیست.

۴-۲- روش محک

برای پیاده‌سازی PCA، یک شبکهٔ عصبی خودانجمنی خطی با روش پس‌انتشار خطا تعلیم داده می‌شود (شکل ۱۰). این شبکه دارای سه لایه است که لایهٔ ورودی و خروجی آن دارای تعداد نورون‌های یکسانی می‌باشند و همان‌طور که در ساختار شبکه نشان داده شده است، تعداد نورون‌های لایهٔ ورودی و خروجی برابر ۱۰۸۰۰ می باشد که از حاصل‌ضرب تعداد پیکسل‌ها در عرض و ارتفاع تصویر به‌دست می‌آید. لایهٔ میانی این شبکه دارای ۱۲۸ نورون خطی است. در این روش، تمامی ۱۲۰ تصویر موجود در پایگاه دادهٔ ZD جهت استخراج ویژگی به این شبکه تعلیم داده می‌شوند. پس از پایان تعلیم، شبکه تصاویر با بعد ۱۰۸۰۰ را به بعد ۱۲۸ در لایهٔ پنهان فشرده می‌کند. از مقادیر نورون‌های لایهٔ پنهان به‌عنوان ویژگی استخراج شده از تصاویر در شبکهٔ طبقه‌بندی کننده استفاده می‌شود.

ویژگی‌های استخراج شده از تصاویر با حالت نرمال هر شخص به‌عنوان دادهٔ تعلیم و ویژگی‌های استخراج شده از تصاویر با حالت‌های مختلف به‌عنوان دادهٔ آزمون شبکه طبقه‌بندی کننده مورد استفاده قرار می‌گیرند. بنابراین در این مرحله تعداد تصاویر تعلیمی ۱۰ مورد و تعداد تصاویر آزمون ۱۱۰ مورد می‌باشد. به‌کمک روش PCA درصد صحّت بازشناسی ٪۵۷/۲۷ بر روی دادگان آزمون و درصد صحّت ٪۱۰۰ بر روی دادگان تعلیم حاصل شده است.



(شکل ۱۰): ساختار شبکه خودانجمنی استفاده شده جهت پیاده‌سازی PCA.

۴-۳- نتایج تعلیم شبکهٔ طبقه‌بندی‌کننده با تصاویر مجازی حاصل از شبکه‌های تحلیل‌گر پیشنهادی

به‌منظور بزرگ کردن دادگان تعلیم، شبکهٔ عصبی تحلیل‌گر را با دادگان AUT[۱] تعلیم داده و سپس با دادن تصاویر نرمال پایگاه داده ZD به ورودی این شبکه و با تغییر کد حالت، تصاویر مجازی این افراد در حالات مختلف تولید می‌شوند. سپس از تصاویر مجازی حاصل نیز جهت تعلیم شبکهٔ طبقه‌بندی کننده استفاده می‌شود. نتایج بازشناسی ساختارهای مختلف شبکه تحلیل‌گر در (جدول ۱) خلاصه شده است. بدین ترتیب تعداد دادگان تعلیم و آزمون موجود برای هر شخص با احتساب تصاویر مجازی، به‌ترتیب برابر سیزده و یازده می‌شود (تصاویر تعلیم، تصاویر نرمال هر شخص به‌علاوه چهره‌های مجازی سنتز شده در دوازده حالت مختلف می‌باشند و تصاویر آزمون تصاویر واقعی هر شخص در یازده حالت دیگر افراد غیر از حالت نرمال است). همان‌طور که در (جدول ۱) نشان داده شده است، درصد صحّت بازشناسی برای دادگان تعلیم در تمامی روش‌ها برابر ٪۱۰۰ است. در آخرین ساختار پیشنهادی درصد صحت بازشناسی برای دادگان آزمون برابر با ٪۸۳/۶۳ می‌باشد که نسبت به مدل مرجع دارای رشد ٪۱۲/۷۳ است. همچنین درصد صحت بازشناسی تمامی ساختارهای پیشنهادی در مقایسه با روش محک یعنی PCA بسیار بالاتر است. عملکرد آخرین ساختار پیشنهادی (تولید تصاویر مجازی به روش خوشه‌بندی بدون سرپرستی اطّلاعات افراد و حالات) در مقایسه با روش PCA دارای بهبود ٪۲۶/۳۶ است.

۵- جمع‌بندی و نتیجه‌گیری

در این مقاله، ساختارهای مختلف شبکه عصبی جهت حل مسئله بازشناسی چهره با یک تصویر از هر فرد پیشنهاد شده است. با علم به این واقعیت که کارآیی روش‌های متداول با وجود محدودیت یک تصویر از هر فرد کاهش می‌یابد، به سنتز تصاویر مجازی پرداخته شده تا از تصاویر حاصل به‌عنوان دادگان تعلیم شبکه طبقه‌بندی کننده استفاده شود. از روش یادگیری عمومی جهت تخمین مانیفولدهای غیرخطی تغییر حالات و شخص استفاده شده است. به کمک این مانیفولدها تصاویر مجازی چهره‌های نرمال پایگاه داده ZD را تولید نمودیم. با بهبود ساختارهای پیشنهادی و واقعی‌تر شدن بیان مؤلفه‌ها، درصد صحت بازشناسی به ٪۸۳/۶۳ افزایش می‌یابد. با مقایسه این مقدار و درصد صحت بازشناسی روش محک، می‌توان به این نتیجه رسید که روش پیشنهادی یک روش امیدبخش برای کاربردهای واقعی است.

باید توجّه کرد که تصاویر مورد استفاده برای تعلیم شبکه تحلیل‌گر، دادگان AUT می‌باشد که شامل تصاویر هشتاد نفر در دوازده حالات مختلف است، درحالی‌که شبکه عصبی مغز انسان با تصاویر هزاران نفر و در حالات بسیار بیشتری تعلیم می‌بینند. بنابراین برخی از پژوهش‌گران به‌منظور تخمین بهتر مانیفولد تغییر حالت، ایده استفاده از تصاویر ویدیویی را پیشنهاد کرده‌اند که در این صورت به حالات بیشتری از هر فرد دسترسی خواهیم داشت. تعلیم مناسب فضاهای چهرهٔ اشخاص و حالات به کمک شبکه نیازمند تعداد کافی از دادگان مناسب می باشد تا نقاط کافی از مانیفولدهای تغییر حالات و تغییر افراد را ایجاد کرده و تخمین مانیفولد را دقیق‌تر نماید. با تخمین مناسب مانیفولدهای تغییر حالت در چهره و به‌دست آوردن محل تلاقی آن با مانیفولدهای تغییر شخص، حالات مجازی شخص به‌صورت مناسب به‌دست می‌آید.

تشکّر و قدردانی

بدین وسیله از آقای دکتر کریم فائز و متخصصان آزمایشگاه پردازش تصویر و شناسایی الگوی دانشکدهٔ برق دانشگاه صنعتی امیرکبیر که دادگان AUT را در اختیار ما قرار دادند، تشکّر و قدردانی می‌گردد. همچنین از خانم مهندس زمانی که دادگان ZD را در دانشکدهٔ مهندسی پزشکی جمع‌آوری کردند، سپاس‌گزاری می‌شود.

(جدول ۱): نتایج درصد صحّت بازشناسی کلیه روش‌های به کار رفته.

روش به کار رفته	تعداد تصاویر تعلیم	تعداد تصاویر آزمون	درصد صحت دادگان تعلیم	درصد صحت دادگان آزمون	تعداد تصاویر آزمون صحیح تشخیص داده شده
روش PCA (روش محک)	۱۰	۱۱۰	۱۰۰٪	۵۷/۲۷٪	۶۳
تعلیم شبکه طبقه‌بندی کننده فقط با تصاویر نرمال افراد	۱۰	۱۱۰	۱۰۰٪	۷۰/۹٪	۷۸
تولید تصاویر مجازی با استفاده از شبکه عصبی مستقیم_معکوس	۱۳۰	۱۱۰	۱۰۰٪	۶۲/۷۲٪	۶۹
تولید تصاویر مجازی به روش تولید کد شخص بصورت بدون سرپرستی	۱۳۰	۱۱۰	۱۰۰٪	۶۰/۹۱٪	۶۷
تولید تصاویر مجازی به روش خوشه‌بندی بدون سرپرستی اطلاعات افراد	۱۳۰	۱۱۰	۱۰۰٪	۸۲/۷۳٪	۹۱
تولید تصاویر مجازی به روش خوشه‌بندی بدون سرپرستی اطّلاعات افراد و حالات	۱۳۰	۱۱۰	۱۰۰٪	۸۳/۶۳٪	۹۲

۶- مراجع

سیدصالحی، سیدعلی. افزایش کارآیی بازشناخت الگوی شبکه‌های عصبی جلوسو از طریق توسعه روش‌هایی برای دوسویه کردن عملکرد آنها، ۱۳۸۳، دانشگاه صنعتی امیرکبیر، دانشکدهٔ مهندسی پزشکی. گزارش طرح پژوهشی.

Beymer, D., Poggio, T., 1996. Face recognition from one example view. Science, 272 (5250).

Vetter, T., 1998. Synthesis of novel views from a single face image. Int. J. Comput. Vision, 28 (2), pp. 102-116.

Chellappa, R., Wilson, C.L., Sirohey, S., 1995. Human and machine recognition of faces: a survey. Proc. IEEE, 83 (5), pp. 705-740.

Chen, S.C., Liu, J., Zhou, Z., 2004. Making FLDA applicable to face recognition with one sample per person. Pattern Recognition, 37 (7), pp. 1553-1555.

Chen, S.C., Zhang, D.Q., Zhou, Z., 2004. Enhanced (PC)2A for face recognition with one training image per person. Pattern Recognition Lett., 25 (10), pp. 1173-1181.

Daugman, J., 1997. Face and gesture recognition: overview. IEEE Trans. Pattern Anal. Mach. Intell., 19 (7), pp. 675-676.

Frade, F., et al., 2005. Repr esen tati onal oriented component analysis (ROCA) for face recognition

with one sample image per training class. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, June, pp. 266-273.

Gao, Y., Qi, Y., 2005. Robust visual similarity retrieval in single model face databases. Pattern Recognition, 38 (7), pp. 1009-1020.

Huang, J., et al., 2003. Component-based LDA method for face recognition with one training sample. AMFG, pp. 120-126.

Jung, H.C., Hwang, B.W, Lee, S.W., 2004. Authenticating corrupted face image based on noise model. Proceedings of the Sixth IEEE International

Wang, J., Plataniotis, K.N., Venetsanopoulos, A.N., 2005. Selecting discriminant eigenfaces for face recognition. Pattern Recognition Lett., 26 (10), pp. 1470-1482.

Wiskott, L., et al., 1997. Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach Intell., 19 (7), pp. 775-779.

Wu, J., Zhou, Z., 2002. Face recognition with one training image per person. Pattern Recognition Lett., 23 (14), pp.1711-1719.

Tan, X., et al., 2005. Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft kNN ensemble. IEEE Trans. Neural Networks, 16 (4), pp. 875-886.

Yang, J., et al., 2004. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Anal. Mach. Intell., 26 (1), pp. 131-137.

Zhao, W., et al., 2003. Face recognition: a literature survey. ACM Comput. Surv., December Issue, 2003. pp. 399-458.

ندا داداشی مدرک کارشناسی خود را در مهندسی پزشکی- بیوالکتریک از دانشگاه صنعتی امیرکبیر در سال ۱۳۸۴ و مدرک کارشناسی ارشد را در همان رشته از دانشگاه صنعتی امیرکبیر در سال ۱۳۸۷ دریافت کرده است. وی اکنون مشغول به تحصیل در دوره دکتری مهندسی پزشکی- بیوالکتریک در دانشگاه صنعتی امیرکبیرمی‌باشد. زمینه‌های پژوهشی مورد علاقهٔ وی بازشناسی الگو، زیست‌سنجی و شبکه‌های عصبی مصنوعی می‌باشد.

نشانی رایانامک ایشان عبارت است از:

ndadashi@gmail.com

فاطمه عبدالعلی در سال ۱۳۸۷ در رشتهٔ مهندسی پزشکی، گرایش بیوالکتریک در مقطع کارشناسی از دانشگاه صنعتی امیرکبیر فارغ‌التحصیل شده است. سپس دورهٔ کارشناسی ارشد را در همان رشته تا سال ۱۳۸۹ ادامه داده است. زمینه‌های پژوهشی مورد علاقهٔ وی پردازش سیگنال با بهره‌گیری از هوش مصنوعی، مدل‌سازی عملکرد مغز و شبکه‌های عصبی مصنوعی می‌باشد.

نشانی رایانامک ایشان عبارت است از:

abdolali.fatemeh@gmail.com

Conference on Automatic Face and Gesture Recognition, pp. 272- 277.
Kepenekci, B., Tek, F.B., Bozdagi, Akar., 2002. Occluded face recognition based on Gabor wavelets. ICIP 2002, September 2002, Rochester, NY, MP-P3.10.

Komleh, H.E., Chandran, V., Sridharan, S., 2001. Robustness to expression variations in fractal-based face recognition. Proceedings of ISSPA-01, vol. 1, Kuala Lumpur, Malaysia, 13-16 August, pp. 359-362.

Lades, M., et al., 1993. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput., 42 (3), pp. 300-311.

Lam, K., Yan, H., 1998. An analytic-to-holistic approach for face recognition based on a single frontal view. IEEE Trans. Pattern Anal. Mach. Intell., 20 (7), pp. 673-686.

Lawrence, S., et al., 1997. Face recognition: a convolutional neural-network approach. IEEE Trans. Neural Networks, 8 (1), pp. 98-113.

Le, H., Li, H., 2004. Recognizing frontal face images using hidden Markov models with one training image per person. Proceedings of the 17th International Conference on Pattern Recognition (ICPR04), vol. 1, pp. 318-321.

Manjunath, B.S., Chellappa, R., Malsburg, C.V.D., 1992. A feature based approach to face recognition. Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 373-378.

Martinez, A.M., 2002. Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class. IEEE Trans. Pattern Anal. Mach. Intell., 25 (6), pp. 748-763.

Martinez, A.M., 2003. Recognizing expression variant faces from a single sample image per class. Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR), pp. 353-358.

Niyogi, P., Girosi, F., Poggio, T., 1998. Incorporating prior information in machine learning by creating virtual examples. Proc. IEEE, 86 (11), pp. 2196-2209.

Ojala, T., Pietikinen, M., Menp, T., 2002. Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell., 24, pp. 971-987.
Tan, X., Chen, S., Zhou, Z., Zhang, F., 2006. Face recognition from a single image per person: A survey. Pattern Recognition, 39, 2006, pp. 1725 - 1745.

سیدعلی سیدصالحی مدرک کارشناسی خود را در مهندسی برق از دانشگاه صنعتی شریف در سال ۱۳۶۱، کارشناسی ارشد را در مهندسی برق از دانشگاه صنعتی امیرکبیر در سال ۱۳۶۷ و دکتری خود را در مهندسی برق- بیوالکتریک از دانشگاه تربیت مدرس در سال ۱۳۷۴ دریافت نموده است. وی درحال حاضر دانشیار دانشکدهٔ مهندسی پزشکی دانشگاه صنعتی امیرکبیر می‌باشد. زمینه‌های پژوهشی مورد علاقهٔ ایشان پردازش و بازشناسی گفتار، شبکه‌های عصبی مصنوعی و زیستی، مدل‌سازی عملکرد مغز و پردازش خطی و غیرخطی سیگنال می‌باشد.

نشانی رایانامک ایشان عبارت است از:

ssalehi@aut.ac.i