

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

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

چکیده:

در این مقاله با استفاده از تولید تصاویر مجازی به کمک شبکه‌های عصبی، مسئله بازشناسی چهره با یک تصویر از هر فرد مورد توجه قرار گرفته است. برای جداسازی اطلاعات شخص از حالت و تخمين مانیفولدهای زیرفصاها مربوطه، از یک شبکه عصبی تحلیل گر غیرخطی اطلاعات چهره استفاده شده است. به منظور افزایش تعداد نمونه‌های تعلیم در شبکه طبقه‌بندی کننده، به کمک مانیفولدهای تخمين زده شده، تصاویر مجازی از چهره‌های نرمال موجود در پایگاه داده اصلی تولید شده است. با طراحی ساختارهای مختلفی از شبکه‌های عصبی مصنوعی به منظور استخراج مؤلفه‌های زیرفصاها اطلاعات فرد و حالت، کیفیت چهره‌های مجازی و درنتیجه درصد صحبت بازشناسی در شبکه طبقه‌بندی کننده بهبود می‌یابد. برای تخمين بهتر مانیفولدهای اطلاعات شخص و بهبود قدرت تعمیم، یک روش تعلیم برمبنای خوشبندی بدون سرپرستی ارائه شده است. با به کارگیری این روش و تعلیم شبکه طبقه‌بندی کننده به کمک تصاویر مجازی حاصل، درصد صحبت بازشناسی $83/63\%$ روی دادگان آزمون حاصل شده که نسبت به مدل مرجع و روش PCA به ترتیب دارای بهبود $12/73\%$ و $26/36\%$ است.

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

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)*. در حالی که در بسیاری از کاربردهای واقعی، جمع آوری تصاویر متعدد از هر فرد با محدودیت‌هایی مواجه است و علاوه بر 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"^۳ می‌باشند. در این روش‌ها محدودیتی برای تصاویر آزمون

۱- مقدمه

هدف از بازشناسی چهره، شناسایی^۱ و یا تأیید هویت^۲ اشخاص از تصاویر و یا تصاویر ویدیویی آنها با استفاده از پایگاه داده از چهره‌های این افراد می‌باشد. فن اوری بازشناسی چهره بهدلیل داشتن مزایایی از جمله دقت بالا و تهاجم پایین، در مواردی مانند امنیت اطلاعات، اجرا و نظارت بر قانون، کارت‌های هوشمند، کنترل دسترسی و غیره مورد استفاده قرار می‌گیرد *(Zhao, et al., 2003; Daugman, 1997; Chellappa, et al., 1995)*. علی‌رغم تحقیقات انجام شده بسیار طی چند دهه گذشته، بهدلیل پیچیدگی ذاتی مسئله، بازشناسی چهره هنوز با مشکل روبروست. این مشکلات ناشی از تنوعات چهره انسان

¹ Identify

² Verify

³ One sample per person

یکی از رویکردهای اصلی برای حل مسئله یک تصویر از هر شخص با استفاده از ویژگی‌های عمومی^۳، افزایش تعداد داده‌های تعلیم می‌باشد. بدین منظور از داده‌های مجازی استفاده می‌شود که این داده‌ها می‌توانند نمایش‌های^۴ 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 آشنونگ^۵ تصویر برای ایجاد نمایش‌های جدید استفاده شده است. در (*Chen, et al., 2004*) روش آنالیز مؤلفه‌های اساسی دو بعدی (*Wu and Zhou, 2002*) به کمک تصاویر ۱۱ بعدی توسعه و بهبود یافته است. در (*Reymer and Poggio, 1996*؛ *Niyogi, et al., 1998*؛ *Vetter, 1998*؛ *Manjunath, et al., 1992* و در (*Kepenekci, Wiskott, et al., 1997*؛ *Lades, et al., 1993* و در (*Tan, et al., 2005*؛ *Lawrence, et al., 1997* و در (*Huang, et al., 2004*؛ *Chen, et al., 2004* و در (*Lam and Yan, 2003* و در (*Ojala, et al., 2002* و در (*Komleh, et al., 2001*) نیز در این دسته قرار دارند).

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 داده 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 به دست آورد. در شبکه طبقه‌بندی کننده، تصاویر نرمال و تصاویر مجازی سنتز شده به عنوان داده تعلیم و تصاویر واقعی اشخاص در حالات مختلف به عنوان داده آزمون به کار رفته‌اند. ساختار این مقاله به شرح زیر است:

در بخش دوم، جنبه‌های نظری جداسازی اطلاعات شخص از حالت به صورت مستقل از یکدیگر مورد بررسی قرار می‌گیرد. در بخش سوم به معروفی ساختار شبکه‌های عصبی پیشنهادی پرداخته و به دنبال آن در بخش چهارم، نحوه ارزیابی نتایج رائمه می‌گردد. در نهایت بخش پنجم شامل جمع‌بندی و نتیجه‌گیری است.

³ global

⁴ Representation

⁵ Perturbation

نبوده و این تصاویر می‌توانند در حالات و شرایط نورپردازی متفاوت باشند. باید توجه 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 توجه 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^۶ و در (*Manjunath, et al., 1992* و در (*Kepenekci, Wiskott, et al., 1997*؛ *Lades, et al., 1993* و در (*Tan, et al., 2002*) نظریه اطباق گراف مورد استفاده قرار گرفته است. همچنین، روش زیرفضای احتمالاتی محلی (Martinez, 2002)، شبکه‌های عصبی خودسازمانده (Huang, et al., 2004) و در (*Le, et al., 2004*)، روش توسعه یافته آنالیز مخفی مارکوف (Chen, et al., 2004) متایزگر خطی^۷ (Lam and Yan, 1998)، بازناسی چهره با الگوهای بایزی موضعی (Ojala, et al., 2002) و بارشناستی برمنبای ویژگی‌های فرکتالی (Komleh, et al., 2001) نیز در این دسته قرار دارند.

در مقابل، گروهی دیگر از روش‌ها از کل چهره به عنوان ورودی استفاده می‌کنند. در (*Wu and Zhou, 2002*) از رویکرد آنالیز مؤلفه‌های اساسی به منظور غنی کردن اطلاعات چهره استفاده شده است. در (*Yang, et al., 2004*) نیز روش آنالیز مؤلفه‌های اساسی دو بعدی به کار گرفته شده و در تخمین ماتریس کوواریانس به جای بردارهای یک بعدی از ماتریس دو بعدی تصاویر استفاده شده است. در روش پیشنهادی (*Jung, et al., 2004*) با استفاده از مدل نویه با سه پارامتر، تصاویر جدید چهره (مشابه تصاویر مخدوش شده) به منظور بازناسی تولید شده‌اند. در (*Wang, et al., 2005*) به دلیل وجود تنها یک تصویر در هر کلاس، پراکنش بین کلاسی قابل تخمین نبوده و بنابراین فرض شده که کلاس‌های مختلف دارای تنواعات 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.

⁶ Directional Corner Points

⁷ Linear discriminant analysis



برای تخمین خوب و کارآمد مانیفولد تغییر حالت باید v_j به خوبی تخمین زده شود. همچنین باید جداسازی \bar{x}_j و \bar{x}_i در صورت در دسترس بودن \bar{x}_i به خوبی انجام پذیرد. این اثر ترکیب غیرخطی سیگنال ها از منابع مختلف می باشد که برای جداسازی آنها باید مؤلفه های مستقل غیرخطی ایجاد کننده حالات مختلف در فرد را داشته و از ترکیب آن مؤلفه ها با حالت نرمال فرد، حالت های مختلف او را با شاهد های بالا با حالت واقعی، ایجاد کرد.

شبکه های عصبی مصنوعی ابزارهای بالقوه برای طبقه بندی کننده و سیستم های تصمیم گیری می باشند. اطلاعات بر روی نورون ها و وزن ها گسترش داده است. پس از تعلیم شبکه های عصبی، می توان از آنها در کاربردهای مختلف بینایی ماشین از جمله پیش پردازش (تعیین لبه، ترمیم و پالایش تصاویر)، استخراج ویژگی، حافظه انجمی (ذخیره و بازبایی اطلاعات) و بازناسی الگو (سید صالحی ۱۳۹۳) استفاده کرد.

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

بنابراین برای رسیدن به این هدف می توان از یک شبکه عصبی خودانجمنی به منظور آنالیز مؤلفه های اساسی غیرخطی که مؤلفه های حالت را از افراد جدا می کند، استفاده کرد. برای به دست آوردن این مؤلفه ها، شبکه با تعداد زیادی داده که شامل افراد مختلف با حالت مختلف هستند، تعلیم داده می شود تا مانیفولد های غیرخطی تغییر حالت را به خوبی تخمین زده شود.

بیان مناسب مؤلفه های مستقل در فضای مؤلفه ها در بهبود عملکرد سیستم، نقش بهسازی دارد. زیرمانیفولد مربوط به افراد به صورت جداگانه و زیرمانیفولد مربوط به حالت ها نیز جداگانه و به نحو مطلوب باید بیان شده و سیستم تعلیم داده شود. برای بیان افراد می توان از بیان تکییت به این صورت که برای هر فرد یک نورون در خروجی در نظر گرفته شده و در صورت تشخیص فرد، عدد ۱ به آن نورون اختصاص یافته و یا به صورت با اینتری که برای هر فرد یک کد با اینتری در نظر گرفته شود، استفاده کرد. تعلیم مناسب این دو زیرفضا به شبکه، نیازمند تعداد کافی از دادگان مناسب می باشد تا نقاط کافی از مانیفولد های تغییر حالت و تغییر افراد را ایجاد کرده و تخمین مانیفولد

۲- جداسازی اطلاعات فرد از حالت

تصویر چهره حاوی دو دسته اطلاعات مستقل "اطلاعات شخص" و "اطلاعات حالت" می باشد که اطلاعات شخص هویت شخص و اطلاعات حالت چگونگی حالت و وضعیت شخص را منتقل می کنند. با ترکیب غیرخطی این دو دسته اطلاعات تصویر شخص در حالت های گوناگون و یا حالت یکسان افراد مختلف شکل می گیرد. مغز انسان به خوبی قادر به تفکیک این دو دسته اطلاعات بوده و با دیدن چهره یک شخص خاص در حالت های مختلف قادر به شناسایی فرد و با دیدن چهره اشخاص ناشناس در حالت های گوناگون قادر به تشخیص حالت افراد می باشد.

در بازناسی چهره با استفاده از یک تصویر از هر فرد، از هر شخص فقط یک تصویر نرمال برای تعلیم موجود بوده و تصاویر آزمون در حالت های مختلف می باشند. بنابراین به مدل بازناس اطلاعات فقط یک تصویر نرمال تعلیم داده می شود. مشکل اصلی در اینجا عدم توانایی یک تصویر برای ایجاد تعمیم لازم برای بازناسی می باشد. به عبارت دیگر یک تصویر از هر شخص قادر به شکل دهنده مناسب مانیفولد حالت مختلف فرد نمی باشد. رویکرد پیشنهادی تخمین مجازی حالت های دیگر فرد و استفاده از حالت های مجازی تولید شده برای تعلیم مناسب مانیفولد های حالت مختلف این فرد به مدل یادگیرنده می باشد. چنین مدلی پس از یادگیری، بهتر از مدل یادگیرنده فقط با یک تصویر قادر به بازناسی حالت مختلف فرد خواهد بود.

اگر بردار \bar{x}_j بعدی تصویر فرد در یک حالت باشد، این تصویر در فضای m بعدی ورودی با یک نقطه نشان داده می شود. 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 می شود. با تغییر حالات و تغییر افراد در فضای ورودی، یک مانیفولد غیرخطی پوشش داده می شود که در برگیرنده زیرمانیفولد های 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 زیرمانیفولد ها تعیین شود.

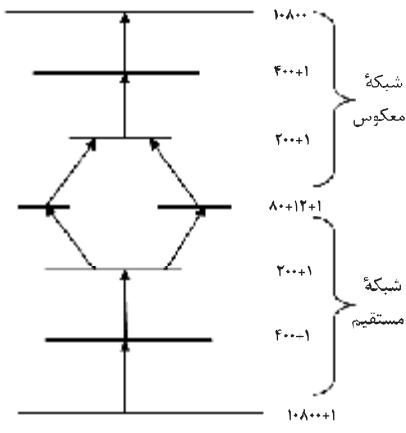
حال اگر \bar{x}_j حالت نرمال شخص و قرار گرفته بر روی زیرمانیفولد مربوطه باشد، می توان رابطه (۱) را به صورت زیر نوشت

$$(1) \quad \bar{X}_j = \bar{x} + v_j$$

که در آن \bar{x}_j نشان دهنده حالت j ام و v_j تغییراتی است که در فضای ورودی بر حالت نرمال برای به دست آمدن حالت جدید اثر می کند.

قسمت طبقه خروجی جهت طبقه‌بندی افراد و حالات می‌باشد. این شبکه با الگوریتم پس انتشار خطوط و به صورت با سرپرستی تعلیم دیده است. به این صورت که برای هر شخص و هر حالت کدی در خروجی به عنوان خروجی مطلوب تعریف شده است. این کدها به صورت تک‌بیتی می‌باشد و در هر حالت و یا شخص یکی از بیت‌های مربوطه به شخص و حالت یک می‌شود. به عبارت 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 به طور دقیق معکوس شبکه مستقیم می‌باشد. لایه‌های پنهان در این شبکه وظيفة نگاشت معکوس^۲ اطلاعات از فضای ویژگی به فضای تصویر را دارند. لایه ورودی این شبکه مشابه لایه خروجی شبکه مستقیم بوده و متشکل از کدهایی می‌باشد که نماینده فرد و حالت آن می‌باشدند. لایه خروجی تصویر مرتبط با کد ورودی را می‌سازد. این شبکه نیز به صورت با سرپرستی تعلیم دیده است.



(شکل ۱) شبکه 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 استفاده شده است. بدین منظور چهره‌های نرمال موجود در این پایگاه داده در ورودی شبکه قرار داده می‌شود. کد به دست آمده در قسمت افراد نشان‌گر میزان شباهت این فرد به هر کدام از ۸۰ نفر پایگاه داده AUT می‌باشد.

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

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

را دقیق‌تر نماید. به هر حال، به دلیل کمبود داده در مسئله یک تصویر از هر فرد، از یک پایگاه داده عمومی که شامل تصاویر تعلیم متفاوت با تصاویر آزمون است، استفاده می‌شود. شبکه عصبی به منظور استخراج اطلاعات متمایز‌کننده تعلیم داده می‌شود. از آنجایی که تصاویر چهره انسان‌ها تنوعات درون فردی^۳ مشابه دارد، اطلاعات متمایز‌کننده فرد هدف با توجه به اطلاعات سایر افراد قابل یادگیری است (Wang, et al., 2005).

۳- ساختار شبکه 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 شده است، به عنوان پایگاه داده عمومی مورد استفاده قرار گرفته است. 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 رو به رو، نگاه به راست و چپ، نگاه به بالا و پایین)، ظاهر چهره با عینک، حالات احساسی چهره (لبخند، اخمش، تعجب، چشمان بسته) و تفاوت در میزان روشنایی (نور از راست و نور از چپ) در نظر گرفته شده است. تصاویر به صورت سیاه و سفید با ۲۵۶ سطح خاکستری و دارای وضوح ۹۰×۱۲۰ می‌باشند. پایگاه داده 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 دانشکده مهندسی پزشکی دانشگاه صنعتی امیرکبیر تهیه شده و تمام تصاویر در پس زمینه یکسان تهیه شده و تنوع سنی نیز تا حد ممکن در آن رعایت شده است.

۳-۲- شبکه عصبی مستقیم - معکوس

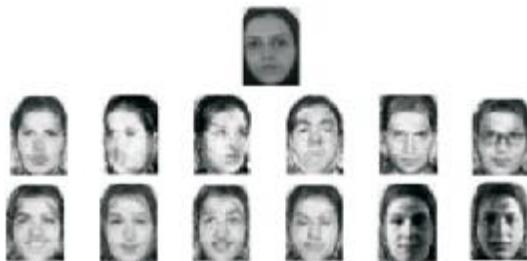
در (شکل ۱) نخستین ساختار شبکه عصبی پیشنهادی برای جداسازی اطلاعات شخص از حالت، نشان داده شده است. ساختار پیشنهادی یک شبکه عصبی، خودانجمنی است. با دادن تصاویر در حالات مختلف افراد (۸۰ نفر پایگاه داده AUT در ۱۲ حالت از هر شخص) در ورودی، اطلاعات حالت و شخص در لایه گلوگاه از هم جدا می‌شوند.

در این شبکه لایه پنهان اول و دوم وظيفة نگاشت اطلاعات از فضای ورودی به فضای ویژگی را داشته و دو

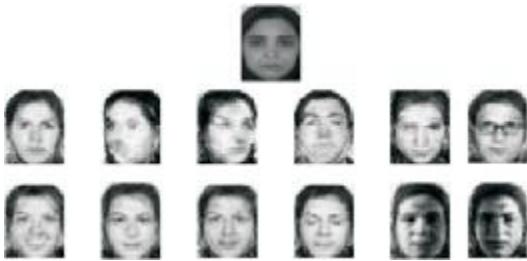
² Decoding
³ Zamani Database

^۱ intra-subject





(شکل ۳): تصاویر مجازی تولیدی برای شخص اول از پایگاه داده ZD به دست آمده توسط شبکه مستقیم- معکوس.



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

۳-۳-۳- تولید گد شخص به صورت بدون سرپرستی

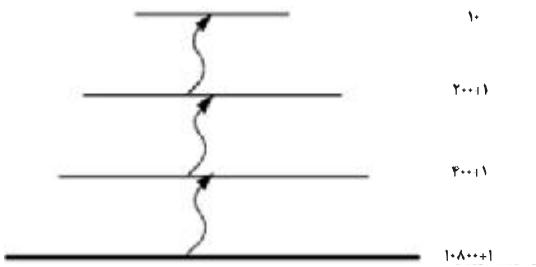
سرپرستی

در (شکل ۵) ساختار اصلاحی پیشنهادی برای جداسازی بهتر اطلاعات حالت از چهره نشان داده شده است. این ساختار یک شبکه خودآجمنی است که نورون های لایه میانی آن به دو بخش بدون سرپرستی و با سرپرستی تقسیم بندی شده است. به نظر می رسد که کد توصیف کننده شخص، بیان مناسبی برای اطلاعات شخص نبوده و شبکه در این قالب قادر به شکل دهنده درست مانیفولدهای زیرفضایی شخص نمی باشد. بنابراین در این ساختار استخراج کد شخص را به عهده شبکه گذاشته و انتظار می رود که بهینه را برای توصیف شخص در قسمت بدون سرپرستی استخراج کند. در بخش پسرپرستی، برای هر حالت کد خاصی تخصیص می یابد و خطاب به لایه های زیرین پس انتشار می یابد و در قسمت بدون سرپرستی لایه گلوگاه، شبکه به صورت بدون سرپرستی کد مناسب اشخاص را ایجاد می کند.

لایه گلوگاه شامل نودودو نورون است که دوازده نورون برای اطلاعات حالت و هشتاد نورون برای اشخاص تخصیص یافته است. در مرحله تعلیم، چهره های موجود در پایگاه داده AUT به ورودی شبکه داده می شود. خروجی مطلوب همان تصویر ورودی می باشد. در دو لایه اول (بخش کدینگ)، اطلاعات شخص و حالت جدا شده و در قسمت دی کدینگ اطلاعات تجزیه شده جهت بازسازی چهره بازترکیب می شوند. درواقع بخش کدینگ را می توان

چهره نرمال فرد در ورودی شبکه قرار داده می شود. با ثابت نگهداشتن کد ایجاد شده در قسمت اطلاعات اشخاص و تغییر کد حالت در قسمت اطلاعات حالت، تصاویر جدید در خروجی تولید می شوند.

برای آزمودن تصاویر مجازی تولید شده از یک شبکه طبقه بندی کننده استفاده شده که ساختار آن در (شکل ۲) نشان داده شده است. این شبکه دارای دو لایه پنهان بوده، لایه ورودی شبکه دارای تعداد نورون های برابر پیکسل های تصویر (۱۰۸۰۰ نورون)، لایه پنهان اول دارای ۴۰۰ نورون و لایه پنهان دوم دارای ۲۰۰ نورون با تابع غیرخطی سیگموید می باشد.



(شکل ۲): شبکه طبقه بندی کننده با بیان خروجی تک بیت.

در لایه خروجی از بیان تک بیتی 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 برای یک و بقیه نورون ها برای صفر در نظر گرفته می شود. به عنوان مثال برای شخص اول کد ۰۰۰۰۰۰۰۰۰۱ در نظر گرفته شده است. برای شخص دوم کد ۰۰۰۰۰۰۰۰۱۰ در نظر گرفته شده است.

جهت تعلیم این شبکه، چهره های نرمال واقعی و دوازده چهره مجازی تولید شده برای هر شخص به عنوان داده های تعلیم و یازده چهره واقعی هر شخص به عنوان داده های آزمون در نظر گرفته شده است. برای اطمینان از اینکه شبکه بیش از اندازه تعلیم ندیده باشد، جایی که در صد صحت بر روی داده های آزمون مقدار بیشینه را داشته و پس از آن روند کاهشی داشته باشد، تعلیم متوقف می شود.

در (شکل های ۳ و ۴)، تصاویر مجازی تولیدی متناظر با دو ورودی مختلف نمایش داده شده است. همان طور که مشاهده می شود، با وجود اینکه حالات تصاویر مجازی به خوبی ایجاد شده است، ولی میزان شباهت تصاویر تولیدی به شخص اصلی بسیار پایین می باشد. به نظر می رسد که بیان افراد بر حسب افراد پایگاه داده عمومی به خوبی انجام نشده و بنابراین استخراج ویژگی رضایت بخش نیست.

همان‌طور که مشاهده می‌شود، تصاویر مجازی به صورت قابل قبول ایجاد نشده‌اند. در تحلیل عدم بازسازی مناسب حالت تصاویر می‌توان گفت که چون در قسمت بدون سرپرستی قیدی جهت عدم عبور اطلاعات حالت وجود نداشت، بنابراین اطلاعات 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 حل این مشکل باید اصلاحاتی انجام شود تا از عبور اطلاعات حالت از این قسمت جلوگیری کند و جداسازی اطلاعات شخص از حالت دقیق‌تر انجام شود.

۴-۳- شبکه عصبی با خوشبندی بدون

سرپرستی لایه اطلاعات شخص

همان‌گونه که در قیل 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 خوشبندی بدون سرپرستی اطلاعات افراد نمایش داده شده است. در این قسمت بدليل عدم وجود قید کافی، اطلاعات 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 به صورت بدون سرپرستی، حالت‌های مختلف هر فرد را در یک خوش قرار داده و برای حالت‌های مختلف شخص کد یکسانی را تخصیص دهد.



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

در این شبکه نیز مانند ساختارهای قبل، ورودی نورون \bar{z} از حاصل ضرب ماتریس ورودی \bar{o} در وزن منتظر w_{ij} به دست 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 \bar{y} خروجی نورون‌های لایه پنهان اول و دوم شبکه مستقیم می‌باشد.

$$y_j = \frac{1}{1 + \exp(net_j - \theta)} \quad (5)$$

به صورت یکتابع غیرخطی پیوسته \bar{t} درنظر گرفت که تصاویر ورودی x را از فضای پیکسل به فضای کد نگاشت می‌کند و بُردار y بدین ترتیب تولید می‌شود:

$$y = r(x) \quad r : R^n \rightarrow R^m \quad (2)$$

در این رابطه، n بعد فضای پیکسل و m بعد فضای کد است. بخش دیگرینگ را نیز می‌توان به صورت تابع غیرخطی پیوسته g درنظر گرفت که فضای کدها را به فضای پیکسل نگاشت می‌کند تا تصاویر بازسازی شده \hat{x} تولید شوند.

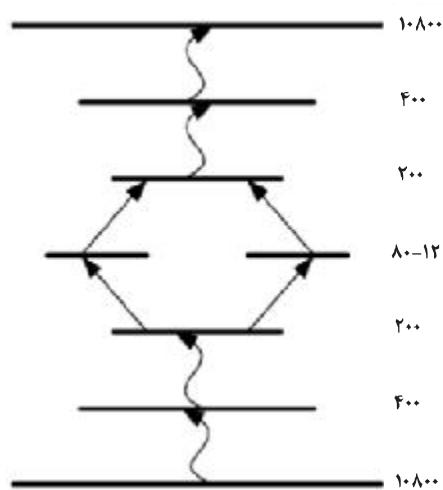
$$\hat{x} = g(y) \quad g : R^m \rightarrow R^n \quad (3)$$

$$\hat{x} = g(r(x)) \quad (4)$$

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 به صورت

$$C = \sum_i (x_i - \hat{x}_i)^2$$

از آنجایی که 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 اطلاعات در قسمت بدون سرپرستی، کد شخص را ایجاد کند.



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

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

همانند بخش‌های قبل، از وزن‌های این روش نیز در تولید تصاویر مجازی استفاده شده است. تصاویر مجازی تولیدی در (شکل ۸) نشان داده شده است. همان طور که مشاهده می‌شود، در تصاویر مجازی حاصل حالت‌ها بهتر ایجاد شده‌اند، اما هنوز برخی حالات به‌وضوح، شکل نگرفته‌اند. از این تصاویر جهت تعلیم شبکه طبقه‌بندی کننده استفاده شده است.



(شکل ۸): تصویر مجازی تولید شده توسط شبکه خوش‌بندی افراد.

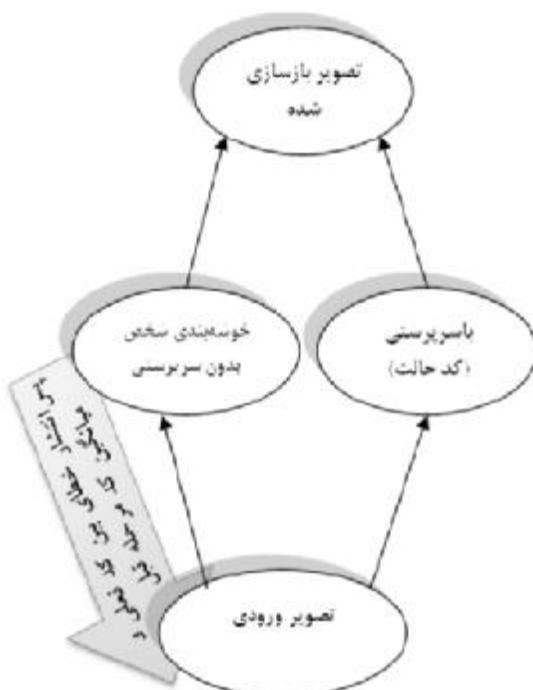
۵-۳- شبکه عصبی با خوش‌بندی بدون سرپرستی در هر دو لایه اطلاعات شخص و حالت

در این ساختار، از مزایای روش خوش‌بندی بدون سرپرستی، در لایه اطلاعات حالت نیز بهره برده است. شبکه با استفاده از دادگان موجود برای هر حالت نیز کد از خاصی را در لایه میانی تولید کرده و بدون نیاز به کد از پیش تعیین شده، لایه اطلاعات حالت را به خوش‌بندی مجازی دسته‌بندی می‌کند. همچنین به‌منظور استخراج مؤلفه‌های مشترک بهینه، بیت‌هایی حالت به هشت و بیت‌هایی افراد به شانزده کاهاش یافته است. با افزایش کیفیت تصاویر مجازی تولیدی و افزایش درصد صحت بازشناسی نشان داده شده است که با تعداد نورون کمتر و خوش‌بندی در هر دو بخش اطلاعات فرد و حالت، مؤلفه‌های مشترک بهتری حاصل می‌شود. در حقیقت این اثر مشابه هرس کردن شبکه می‌باشد، زیرا تعداد زیاد نورون‌های لایه پنهان مضر بوده و موجب ایجاد ابرصفحه زائد می‌شود. در اینجا نیز ساختار شبکه دارای افزونگی است و با تعداد نورون کمتر می‌توان به مؤلفه‌های مشترک بهتری دست پیدا کرد. پس از استخراج مؤلفه‌های مشترک، قدرت تعمیم شبکه و کیفیت تصاویر مجازی حاصل افزایش یافته، درنتیجه درصد صحت بازشناسی نیز افزایش می‌باشد. در حقیقت، نحوه شکل‌گیری مؤلفه‌های غیرخطی مربوط به

این شبکه نیز به صورت خودانجمنی تعلم می‌بیند و نورون‌های لایه گلوگاه برای جداسازی مؤلفه‌های فرد از حالت، به دو بخش تقسیم شده است. قسمت حالت همانند طرح قبل با تخصیص کد بوده و به صورت با سرپرستی تعلم می‌بیند. تفاوت این طرح با طرح قبل در قسمت ایجاد کد اشخاص می‌باشد. در ساختار پیشنهادی، در هر بار تعلم تصاویر یک شخص، متوسط مقادیر خروجی نورون‌های بخش خوش‌بندی شخص در لایه گلوگاه توسط رابطه (۶) به دست آمده و به عنوان کد شخص در مرحله بعد تعلم استفاده می‌شود.

$$M_{i+1} = (\gamma \cdot M_i) + (1 - \gamma) \cdot P \quad (6)$$

در این رابطه، M نشان‌گر کد به هنگام شده برای شخص، P خروجی نورون‌ها در بخش خوش‌بندی و γ ضریب میران تأثیر کد جدید (در اینجا 0.95) در نظر گرفته شده می‌باشد. اندیس i بیان‌گر تصاویر هر فرد در ۱۲ حالت متفاوت است که بدین ترتیب شبکه این تصاویر را در یک خوش‌بندی قرار می‌دهد. بنابراین علاوه بر پس‌انتشار خطای خروجی، در هر مرحله خطای بین کد تولید شده و کد مرحله قبل تعلم نیز بر روی وزن‌های مسیر پس‌انتشار می‌شود. بدین ترتیب شبکه با استفاده از دادگان موجود برای هر شخص، کد خاصی را در لایه میانی تولید کرده و بدون نیاز به کد از پیش تعیین شده، لایه شخص را به خوش‌بندی مجازی دسته‌بندی کند.

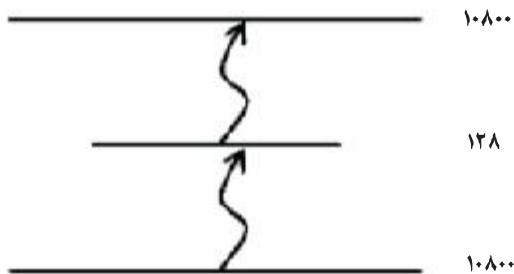


(شکل ۷): طرح خوش‌بندی بدون سرپرستی اطلاعات افراد.

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 تعییم 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.

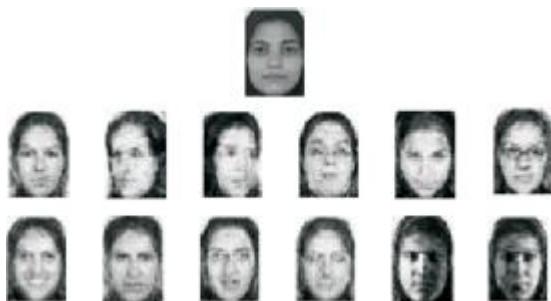
۲-۴- روش محک

برای پیاده‌سازی PCA، یک شبکه عصبی خودانجمنی خطی با روش پس‌انتشار خط‌خط تعلیم داده می‌شود (شکل ۱۰). این شبکه دارای سه لایه است که لایه ورودی و خروجی آن دارای تعداد نورون‌های یکسانی می‌باشند و همان‌طور که در ساختار شبکه نشان داده است، تعداد نورون‌های لایه ورودی و خروجی برابر ۱۰۸۰۰ می‌باشد که از حاصل ضرب تعداد پیکسل‌ها در عرض و ارتفاع تصویر به دست می‌آید. لایه میانی این شبکه دارای ۱۲۸ نورون خطی است. در این روش، تمامی ۱۲۰ تصویر موجود در پایگاه داده ZD جهت استخراج ویژگی به این شبکه تعلیم داده می‌شوند. پس از پیان تعلیم، شبکه تصاویر با بعد ۱۰۸۰۰ را به بعد ۱۲۸ در لایه پنهان فشرده می‌کند. از مقادیر نورون‌های لایه پنهان به عنوان ویژگی استخراج شده از تصاویر در شبکه طبقه‌بندی کننده استفاده می‌شود. ویژگی‌های استخراج شده از تصاویر با حالت نرمال هر شخص به عنوان داده تعلیم و ویژگی‌های استخراج شده از تصاویر با حالت‌های مختلف به عنوان داده آزمون شبکه طبقه‌بندی کننده مورد استفاده قرار می‌گیرند. بنابراین در این مرحله تعداد تصاویر تعلیمی ۱۰۰ مورد و تعداد تصاویر آزمون ۱۱۰ مورد می‌باشد. به کمک روش PCA درصد صحت بازشناسی ۵۷/۲۷٪ بر روی دادگان آزمون و درصد صحت ۱۰۰٪ بر روی دادگان تعلیم حاصل شده است.



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

دو زیرفضا به طور مستقل از یکدیگر بر قدرت تعییم 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 کلیه حالات بهوضوح، شکل گرفته‌اند.



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

۴- نتایج شبیه‌سازی

در این بخش در ابتدا به عنوان مدل مرجع، یک شبکه عصبی جلوسوی چندلایه به طور 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 به عنوان 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 ۲-۴ آمده است. در بخش ۳-۴ 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 اول و به عنوان مدل مرجع، شبکه عصبی طبقه‌بندی کننده (شکل ۲) را تنها با تصاویر نرمال ده نفر موجود در پایگاه داده 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 به ترتیب برابر با ۱۰ و ۱۱۰ می‌باشد. نتایج حاصل از این روش در سطر دوم جدول ۱ آمده است. همان‌گونه که مشاهده می‌شود، درصد



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 توجه کرد که تصاویر مورد استفاده برای تعلیم شبکه تحلیل‌گر، دادگان AUT می‌باشد که شامل تصاویر هشتاد نفر در دوازده حالات مختلف است، درحالی که شبکه عصبی مغز انسان با تصاویر هزاران نفر و در حالات بسیار بیشتری تعلیم می‌بیند. بنابراین برخی از پژوهش‌گران به منظور 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 چهره اشخاص و حالات به کمک شبکه نیازمند تعداد کافی از دادگان مناسب می‌باشد تا نقاط کافی از مانیفولد را دقیق تر نماید. با تخمین مناسب تخمین مانیفولد را دقیق تر نماید. با تخمین مناسب مانیفولد را تغییر حالت در چهره و به دست آوردن محل تلاقی آن با مانیفولد را تغییر شخص، حالات مجازی شخص به صورت مناسب به دست می‌آید.

تشکر و قدردانی

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

۴-۳-۴- نتایج تعلیم شبکه طبقه‌بندی کننده با تصاویر مجازی حاصل از شبکه‌های تحلیل‌گر پیشنهادی

بهمنظور بزرگ کردن دادگان تعلیم، شبکه عصبی تحلیل‌گر را با دادگان AUT 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 شبکه 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، به ترتیب برابر سیزده و یازده می‌شود (تصاویر تعلیم، تصاویر نرمال هر شخص به علاوه 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 صحت 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 خوشبندی بدون سربستی اطلاعات 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 به عنوان دادگان تعلیم شبکه طبقه‌بندی کننده استفاده شود. از روش یادگیری عمومی جهت تخمین مانیفولد راهی غیرخطی تغییر حالات و شخص استفاده شده است. به کمک این مانیفولد راه تصاویر مجازی چهره‌های

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

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

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. Representation 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

۶- مراجع

سیدصالحی، سیدعلی. افزایش کارآیی بازشناخت الگوی شبکه‌های عصبی جلوسو از طریق توسعه روش‌هایی برای دوسریه کردن عملکرد آنها، ۱۳۸۳، دانشگاه صنعتی امیرکبیر، دانشکده 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

فصلنامه
پژوهیزی



- 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
امیرکبیر می‌باشد. زمینه‌های پژوهشی مورد علاقه وی بازشناسی
الگو، زیست‌سنجه و شبکه‌های عصبی مصنوعی می‌باشد.
نشانی رایانمک ایشان عبارت است از:

ndadashi@gmail.com



فاطمه عبدالعلی در سال ۱۳۸۲ در رشته
مهندسی پزشکی، گرایش بیوالکتریک در
مقطع کارشناسی از دانشگاه صنعتی
امیرکبیر فارغ‌التحصیل شده است. سپس
دوره کارشناسی ارشد را در همان رشته تا
سال ۱۳۸۹ ادامه داده است. زمینه‌های
پژوهشی مورد علاقه وی پردازش سیگنال با بهره‌گیری از هوش
مصنوعی، مدل‌سازی عملکرد مغز و شبکه‌های عصبی مصنوعی
می‌باشد.

نشانی رایانمک ایشان عبارت است از:

abdolali.fatemch@gmail.com

سال ۱۳۹۰ شماره ۱ پیاپی ۱۵

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., Maenp, 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 در حال حاضر دانشیار دانشکده
مهندسی پزشکی دانشگاه صنعتی امیرکبیر می‌باشد. زمینه‌های
پژوهشی مورد علاقه 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.ir



سال ۱۳۹۰ شماره ۱ پیاپی ۱۵

