

# آموزش بازشناسی چهره با استفاده از ترکیب شبکه‌های عصبی

## مبتنی بر اختلاط خبره‌ها

رضا ابراهیم پور<sup>۱</sup>، نینا طاهری مخصوص<sup>۲</sup>، علیرضا حاجیانی<sup>۳</sup>

### چکیده

در این مقاله، به آموزش بازشناسی چهره مبتنی بر اختلاط خبره‌ها پرداخته شده است و ایده اساسی آن اضافه کردن واحد گشتاور به ساختار اختلاط خبره‌ها با هدف افزودن کارایی بازشناسی چهره می‌باشد. این واحد به هر طبقه بند ساختار اختلاط خبره‌ها نسبت داده می‌شود. برای استخراج ویژگی از تجزیه مؤلفه‌های اساسی و برای بازشناسی از روش بهبود یافته اختلاط خبره‌ها با اضافه کردن واحد گشتاور بهره گرفته شده است. ضمناً در ساختار اختلاط خبره‌های به کار گرفته شده به جای به کارگیری شبکه پرسپترون خطی در خبره‌های ساختار اختلاط خبره‌های مرسوم از شبکه پرسپترون چندلایه استفاده شده است. در آزمایش‌ها سه ساختار متفاوت بر روی مجموعه تصاویر Yale و ORL در مقایسه با روش ارایه شده، مورد بررسی قرار گرفته شده است. نتایج آزمایش‌ها حاکی از این است که ساختار ارایه شده با سرعت بیشتری به کارایی بالاتری رسیده است.

**کلمات کلیدی** بازشناسی چهره، اختلاط خبره‌ها، واحد گشتاور، تجزیه مؤلفه‌های اساسی

### ۱- مقدمه

بر ظاهر<sup>۱</sup> که به صورت آماری چهره را مورد پردازش قرار می‌دهند و دیگر روش‌های مبتنی بر مدل<sup>۲</sup> که به صورت هندسی عمل می‌نمایند [۱].

بارزترین تفاوت میان روش‌های بازشناسی چهره در چگونگی استخراج و نمایش خصوصیات چهره می‌باشد. در روش‌های مبتنی بر ظاهر، این خصوصیات شامل یک بردار ویژگی تشکیل شده از مقادیر خاکستری پیکسل‌های چهره است که به صورت مستقیم یا غیرمستقیم در بازشناسی چهره استفاده می‌شود. از معروف‌ترین روش‌های مبتنی بر ظاهر به آنالیز مؤلفه‌های اساسی<sup>۳</sup> می‌توان اشاره کرد که به PCA مشهور است [۲ و ۳]. کارایی بالای این روش در کاهش ابعاد، سبب شده است از آن به صورت مستقیم یا غیرمستقیم در روش‌های دیگر استفاده شود [۴]. در روش PCA فرض بر این است که داده‌ها دارای توزیع گوسین هستند. برای داده‌هایی با توزیع گوسین، تمام اطلاعات مربوط به وابستگی داده‌ها در ماتریس کواریانس داده‌ها موجود است (ماتریس کواریانس بیان‌کننده ارتباطات آماری درجه اول است). ولی در روش دیگر، یعنی تجزیه

بازشناسی چهره در سالیان اخیر بسیار مورد توجه قرار گرفته است تا بدانجا که بازشناسی چهره را به یکی از اصلی‌ترین زمینه‌های تحقیقاتی در بینایی ماشین، بازشناسی الگو و یادگیری ماشین تبدیل کرده است. در بازشناسی چهره، سیستم با توجه به چهره‌های آموزش دیده شده، یک چهره را که شباهت بیشتری به چهره مورد نظر داشته باشد، انتخاب کرده و به عنوان جواب نهایی در نظر می‌گیرد. روش‌های موجود برای بازشناسی چهره را می‌توان به دو دسته کلی تقسیم کرد: یکی روش‌های مبتنی

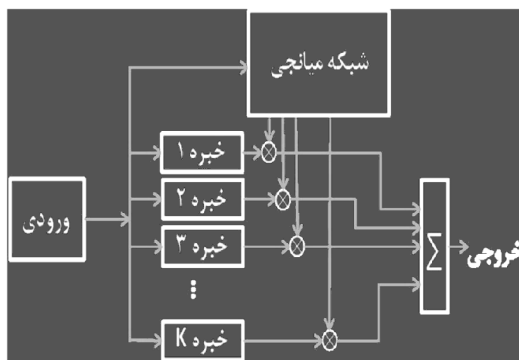
مقاله در تاریخ ۸۷/۱۰/۳ دریافت و در تاریخ ۸۷/۱۲/۵ به تصویب نهایی رسید.

<sup>۱</sup> استادیار، دانشکده برق و کامپیوتر، دانشگاه تربیت دبیر شهید رجایی - پست الکترونیکی rebrahimpour@srttu.edu

<sup>۲</sup> کارشناسی ارشد، دانشکده برق و کامپیوتر، دانشگاه تربیت دبیر شهید رجایی

<sup>۳</sup> دانشجوی کارشناسی، رشته برق الکترونیک، دانشکده برق و کامپیوتر، دانشگاه تربیت دبیر شهید رجایی

یادگیری با سرپرست صورت می گیرد. خبره ها، هر یک به صورت با سرپرست آموزش داده می شوند. هر کدام از خبره ها به تنهایی می توانند پاسخ مطلوبی را ارائه نمایند. همچنین خبره ها به وسیله روش یادگیری خود سازماندهی، آموزش می بینند که چگونه با یکدیگر ترکیب شوند تا بهترین جواب را تولید نمایند یا به عبارت دیگر تلاش می کنند تا یک مشارکت خوبی در فضای ورودی پیدا کنند تا هر طبقه بند هم به خوبی عمل نماید و هم کارایی کل گروه بهینه گردد.



شکل ۱ ساختار مدل اختلاط خبره ها

در شکل ۱ ساختار اختلاط خبره های مرسوم، با  $K$  خبره نشان داده شده است [۱۵]. در این ساختار واحدی به نام شبکه میانجی<sup>۱۵</sup> وجود دارد که به عنوان یک عامل تنظیم کننده تأثیر هر خبره، در جواب نهایی عمل می نماید. در اینجا فرض بر این است که هر خبره در ناحیه مربوط به خود به خوبی عمل می کند. هدف نهایی، ترکیب این خبره ها با توجه به کارایی آنها می باشد. در این ساختار دو مدل از یادگیری وجود دارد: یکی یادگیری با سرپرست که درون هر خبره صورت می گیرد و دیگر یادگیری بدون سرپرست که در بخش شبکه میانجی انجام می شود.

آماری درجه اول است). ولی در روش دیگر، یعنی تجزیه مؤلفه های مستقل<sup>۴</sup> که به ICA معروف است [۵ و ۶]، نیاز به چنین فرضی نیست و سعی در کمینه کردن ارتباطات میان داده های در مرتبه دوم و بالاتر دارد [۷]. برخلاف دو روش گذشته که به صورت بدون سرپرست<sup>۵</sup> عمل می کردند، در روش تجزیه تفکیک کننده خطی<sup>۶</sup> [۸] که به LDA مشهور است، به صورت با سرپرست<sup>۷</sup> و با استفاده از اطلاعات مرتبط با کلاس چهره ها، طبقه بندها آموزش می بینند. هدف نهایی در این روش پیدا کردن یک مجموعه از بردارهای پایه است تا که هنگامی که چهره ها بر روی آنها تصویر شوند، فاصله میان نمونه های یک کلاس حداقل و فاصله میان نمونه های کلاس های مختلف بیشینه گردد. در مقابل روش های مبتنی بر مدل، سعی در یافتن ویژگی های خاصی از چهره است که بتواند چهره های افراد مختلف را از یکدیگر متمایز کند (مانند فاصله میان دو چشم، فاصله چشم و بینی و...). این روش ها خود به دو دسته کلی تقسیم می شوند: یکی مدل های مبتنی بر گراف های خوشه ای کشسان<sup>۸</sup> و دیگر مدل های مبتنی بر ظاهر پویا<sup>۹</sup> [۹-۱۲].

از سوی دیگر ساختارهای مبتنی بر ماشین های مشاور<sup>۱۰</sup> و الگوریتم های یادگیری تجمعی<sup>۱۱</sup> که در مبحث یادگیری ماشین مطرح است، مکانیزم هایی جهت ترکیب نمودن روش های ذکر شده ارائه می دهند [۱۳-۱۵]. الگوریتم های یادگیری تجمعی را می توان از لحاظ ساختاری به دو دسته کلی تقسیم نمود: یکی ساختارهای ایستا که در این دسته از روش ها، پاسخ طبقه بندهای پایه (خبره ها) به وسیله مکانیزمی ترکیب می شود که شامل سیگنال ورودی نیست و دیگر ساختارهای پویا که در این دسته، سیگنال ورودی به صورت مستقیم در فرایند ترکیب جواب خروجی طبقه بندها تأثیر دارد. مدل اختلاط خبره ها<sup>۱۲</sup> از جمله این ساختارها به حساب می آید.

## ۲- اختلاط خبره ها

در ابتدا یک شبکه پیمانته ای را در نظر بگیرید که فرآیند آموزش از طریق ترکیب یک روش خود سازماندهی<sup>۱۴</sup> که ذاتاً یک روش بدون سرپرست است، به همراه یک روش

### ۳- ساختار ارایه شده اختلاط خبره ها با افزودن واحد گشتاور

در ساختار ارایه شده در این مقاله با تقسیم هوشمندانه فضای ورودی بین طبقه بندهای پایه به ایجاد گوناگونی در آنها می پردازد. در این ساختار همانند [۱۷-۱۹] از شبکه های عصبی پرسپترون چند لایه<sup>۱۶</sup> به عنوان طبقه بند پایه و شبکه میانجی به جای شبکه های عصبی پرسپترون تک لایه<sup>۱۷</sup> استفاده شده است. همچنین اضافه کردن واحد گشتاور به روابط اصلاح وزن ها جهت کاربرد در بازشناسی چهره، از جمله بهبودهایی است که در ساختار اختلاط خبره ها داده شده است. این بهبود جهت افزایش سرعت همگرایی و گسترش آن برای حالت تخصیص وزن شبکه میانجی در مورد هر خبره به تعداد کلاس های الگو است.

در این ساختار، یادگیری طبقه بندها از نوع یادگیری رقابتی است یعنی هنگام یادگیری، طبقه بندی که خطای کمتری بر روی نمونه یادگیری دارد تشویق می شود. در واقع شبکه میانجی دو وظیفه دارد: اول این که فضای ورودی را به صورت هوشمندانه بین طبقه بندها تقسیم کند، دیگری این که با توجه به توانمندی طبقه بند برای طبقه بندی صحیح الگوی ورودی، وزنی به آن تخصیص دهد. همزمان با یادگیری طبقه بندها، شبکه میانجی یاد می گیرد چگونه وزن مربوط به نظر هر طبقه بند را به صورت تابعی از الگوی ورودی محاسبه کند.

در ادامه این بخش ابتدا با تفصیل بیشتری در مورد واحد گشتاور و تأثیر این واحد در روند اصلاح وزن شبکه بحث می شود و سپس تغییراتی که متوجه روابط و فرمول های ساختار اختلاط خبره های مرسوم (بدون واحد گشتاور و با شبکه های عصبی پرسپترون یک لایه به عنوان طبقه بند و شبکه میانجی) می شود مورد بررسی قرار می گیرد.

#### ۳-۱ واحد گشتاور

با اضافه نمودن واحد گشتاور در معادله تصحیح ضرایب، تحرک بیشتری را می توان در میزان تغییر ضرایب وزنی شبکه ایجاد کرد. این واحد در ابتدا باعث تغییرات زیادی در

ضرایب وزنی می شود، لیکن با کاهش مقدار تغییرات، میزان تغییرات به علت وجود واحد گشتاور نیز کاهش می یابد. در نتیجه احتمال درگیر شدن شبکه در کمینه های محلی در ابتدا کمتر است زیرا واحد گشتاور، تغییرات را فراتر از افزایش های محلی بر تابع انرژی در جهت روند کلی کاهش انرژی باعث می شود. واحد گشتاور در تسریع همگرایی در شیب های کم عمق بسیار مؤثر است زیرا باعث می شود شبکه در جهت رسیدن به جواب در مسیر سرازیری سرعت گیرد. بستر انرژی ممکن است دارای دره های تنگ و طولانی باشیب کم باشد که نهایتاً به نقطه کمینه ختم می شوند. همگرایی در طول این دره ها کند است و الگوریتم معمولاً در حالی که به سوی جواب حرکت می کند در عرض دره ها نوسان می یابد در این حالت تسریع در حرکت بدون افزایش احتمال پریدن و گذشتن از نقطه کمینه دشوار است افزایش واحد گشتاور در این موارد بسیار مؤثر است [۱۶ و ۲۲]. واحد گشتاور و نحوه به کارگیری آن در روابط اصلاح وزن ساختار اختلاط خبره های به کار گرفته شده در بخش ۳-۲ آمده است.

#### ۳-۲ ساختار اختلاط خبره های بهبود یافته

همان طور که گفته شد در ساختار ارایه شده از شبکه های عصبی پرسپترون، چند لایه در خبره ها و شبکه میانجی استفاده شده است. حال فرض کنید اگر خروجی شبکه آم برای هر بردار ورودی به صورت  $O_i$  و خروجی سیستم مرکب به صورت  $O_T$  نشان داده شود داریم:

$$O_T = \sum_{i=1}^L g_i \cdot O_i \quad (1)$$

که در آن  $L$  تعداد طبقه بندهای پایه و  $g_i$  وزن تخصیص داده شده به نظر شبکه آم (توسط شبکه میانجی) هستند.  $g_i$  تخمینی از احتمال پسین آن است که شبکه آم بتواند الگو را به درستی طبقه بندی کند. شبکه میانجی با توجه به این احتمال، وزن  $g_i$  را به آن تخصیص می دهد. تعداد نرون های لایه خروجی شبکه میانجی، با تعداد طبقه بندهای پایه برابر است.  $g_i$  به صورت زیر تابعی از الگوی ورودی و وزن های یادگیری است.

گشتاور و  $h_i$  مرتبط با مجموع مربعات خطای طبقه بند نام بوده و با رابطه ۷ تعریف می شود.

$$h_i = \frac{g_i \exp(-\frac{1}{2}(y - O_i)^T (y - O_i))}{\sum_j g_j (-\frac{1}{2}(y - O_j)^T (y - O_j))} \quad (7)$$

#### ۴- آزمایش ها

برای ارزیابی این ساختار و دیگر ساختارهایی که برای بازشناسی چهره به چالش کشیده شده اند، نیاز به مجموعه داده های استاندارد است تا بتوان آنها را با روش های موجود مقایسه کرد. به این منظور در این مقاله از بانک چهره Yale و ORL استفاده شده است.



شکل ۲ افراد بالا تصاویر یک نفر از بانک چهره یال و افراد پایین یک نفر از افراد بانک چهره ۲۰ /

بانک چهره ORL تعداد ۴۰۰ تصویر چهره از ۴۰ شخص متفاوت را در خود دارد. این تصاویر در زمان های گوناگون و نورپردازی های متفاوت گرفته شده است. این نمونه ها شامل حالات چهره متفاوت (خندان، چشمان بسته) و همراه تصاویر با عینک و بدون عینک برای هر فرد می باشد و تمامی این تصاویر از یک موقعیت مکانی چهره گرفته

$$g_i = \frac{\exp(O_{g_i})}{\sum_{j=1}^L \exp(O_{g_j})} \quad (2)$$

با تعریف فوق برای  $g_i$ ، مجموع وزن های تخصیص داده شده به طبقه بندهای پایه، برابر ۱ است. بزرگتر بودن مقدار  $g_i$  برای ۱ شبکه به معنی توانمندتر بودن آن برای طبقه بندی صحیح الگوی ورودی است. در حالت مرزی، مقدار یک برای  $g_i$  به معنی این است که فقط با استفاده از خروجی طبقه بند نام می توان کلاس الگوی ورودی را به درستی تعیین کرد و مقدار ۰ (صفر) برای  $g_i$  به معنی آن است که نظر طبقه بند نام در مورد کلاس الگوی ورودی نادرست است. در بقیه حالت ها، کلاس الگوی ورودی با توجه به ترکیب خطی نظرات کلیه طبقه بندهای موجود در سیستم پیدا می شود (رابطه ۱).

اصلاح وزن طبقه بندهای پایه و شبکه میانجی با واحد گشتاور به صورت روابط ۳، ۴، ۵ و ۶ است.

$$\Delta w(n)_y = \alpha \Delta w(n-1)_y + \eta_e h(n)_i (y - O(n)_i)(O(n)_i(1 - O(n)_i))Oh(n)_i^T \quad (3)$$

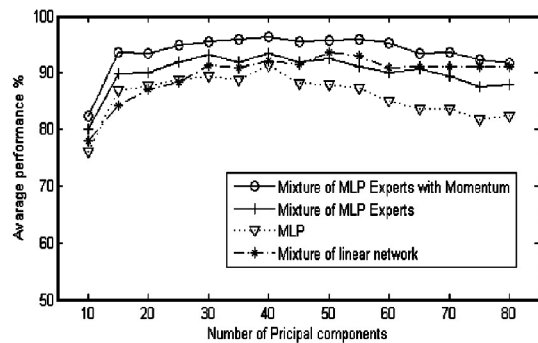
$$\Delta w(n)_h = \alpha \Delta w(n-1)_h + \eta_e h(n)_i w(n)_y^T (y - O(n)_i)(O(n)_i(1 - O(n)_i))Oh(n)_i(1 - Oh(n)_i) \quad (4)$$

$$\Delta w(n)_{yg} = \alpha \Delta w(n-1)_{yg} + \eta_g (h(n) - g(n)) (O(n)_g(1 - O(n)_g))Oh(n)_g^T \quad (5)$$

$$\Delta w(n)_{hg} = \alpha \Delta w(n-1)_{hg} + \eta_g w(n)_{yg}^T (h(n) - g(n)) (O(n)_g(1 - O(n)_g))Oh(n)_g(1 - Oh(n)_g)x_i \quad (6)$$

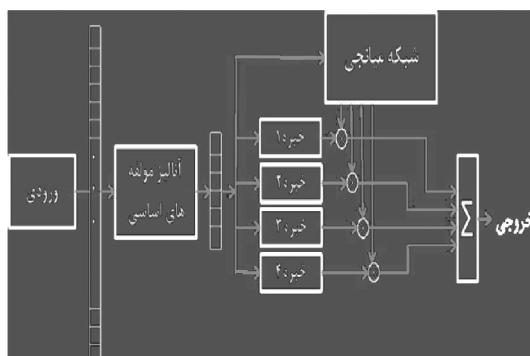
که در آن،  $y$  بردار خروجی مطلوب الگو،  $\eta_e$  و  $\eta_g$  به ترتیب نرخ یادگیری برای شبکه میانجی و خبره ها است.  $\alpha$  واحد

محاسباتی اهمیت دارد تعداد کمتری از مؤلفه ها را انتخاب نمود. شکل ۳ تفاوت تصاویر چهره را در انتخاب تعداد ۱۰ مؤلفه اساسی و ۸۰ مؤلفه اساسی برای پنج تصویر از یکی از افراد بانک چهره ORL را نمایش می دهد. می توان مشاهده کرد که هر چه تعداد مؤلفه های اساسی بیشتر شود جزئیات بیشتری از تصاویر چهره را در تصاویر بازسازی شده می توان دید.



شکل ۴ کارایی سیستم نسبت به تعداد مؤلفه های اساسی

در شکل ۵ می توان تأثیر اعمال و نحوه بازنمایی تصاویر چهره به خبره ها و همچنین به شبکه میانجی را پس از بردن به فضای PCA و کاهش بُعد تصاویر مشاهده کرد. همان طور که مشخص است، تمام تصاویر چهره به همه خبره ها و شبکه میانجی داده می شود. از تصاویر پنج ORL پنج نمونه از هر فرد برای آموزش و پنج نمونه دیگر برای آزمایش و از تصاویر Yale پنج نمونه از هر فرد برای آموزش و شش نمونه برای آزمایش ساختار استفاده می شود.



شکل ۵ ساختار ارایه شده مبتنی بر اختلاط خبره ها و نحوه بازنمایی تصاویر چهره به خبره ها و شبکه میانجی

شده است. همچنین بانک چهره Yale شامل ۱۶۵ تصویر از ۱۵ شخص متفاوت می باشد. برای هر شخص ۱۱ تصویر گوناگون قرار داده شده است که شامل حالات نور مرکزی، عینک، خوشحالی، نور چپ، بدون عینک، حالت عادی، نور راست، ناراحتی، خواب آلودگی، هیجان زدگی و چشم بسته می باشد. همان طور که مشخص است، این بانک اطلاعاتی با وجود کوچک بودن شامل حالات گوناگونی برای هر شخص است که باعث ارزیابی مناسب روش ها می گردد. شکل ۲ تصاویر نمونه از بانک چهره ORL و Yale را نمایش می دهد.

#### ۴-۱ نحوه بازنمایی تصاویر چهره به ساختار ارایه شده

برای راحتی کار و کم کردن حجم تصاویر در بازنمایی تصاویر چهره به هر خبره ابتدا تصاویر بانک چهره ORL و Yale (آموزش و آزمایش) به سایز  $48 \times 48$  تبدیل شده اند. پس از تغییر سایز به  $48 \times 48$  تصاویر چهره به فضای PCA برده شده اند. برای تعیین تعداد مؤلفه های اساسی انتخابی مؤثرتر در بازنمایی به ساختار، ساختار با تعداد ۱۰، ۲۰، ۳۰، ...، ۸۰ مؤلفه امتحان شده است.



شکل ۳ بازسازی تصاویر با بزرگترین و کوچکترین تعداد مؤلفه های اساسی استفاده شده در آزمایش ها

با توجه به نتایج آزمایش های انجام شده شکل ۴، تعداد ۴۰ مؤلفه اساسی با توجه به الگوریتم یادگیری و مجموعه آموزشی و پارامترهای یادگیری، بهترین نرخ بازشناسی را داشته و این تعداد مؤلفه اساسی به خبره ها بازنمایی شده است. همچنین در شکل ۴ می توان مشاهده کرد که کارایی ساختار با تعداد مؤلفه های کمتر از ۴۰ نیز با اختلاف ناچیزی دارای کارایی های مشابه است، در سیستمی که بار

#### ۴-۲ ارزیابی ساختار ارایه شده

برای ارزیابی این ساختار دو دسته آزمایش انجام شده است. در دسته اول آزمایش ها، برای حالتی صورت گرفته است که ساختار شامل تعدادی شبکه عصبی پرسپترون چندلایه (به عنوان خبره) می باشد برای آن که تصویری از تأثیر گوناگونی ایجاد شده در این ساختار و اصلاح وزن های ساختار با واحد گشتاور، بر میزان بهبود نرخ بازشناسی چهره سیستم ترکیب شبکه های عصبی پرسپترون چندلایه مبتنی بر اختلاط خبره ها با واحد گشتاور داشته باشیم، آزمایش ها برای شبکه عصبی پرسپترون چندلایه، ترکیب شبکه های عصبی پرسپترون خطی و شبکه های عصبی پرسپترون چندلایه مبتنی بر اختلاط خبره ها بدون واحد گشتاور نیز تکرار شده است. در این ساختارها فقط وزن های تصادفی اولیه، منشأ ایجاد گوناگونی در خبره ها هستند و از قاعده میانگین برای ترکیب نتایج خروجی طبقه بندها استفاده شده است. در تمام روش ها، برای تأثیر تصادفی بودن وزن های اولیه شبکه ها ناچیز باشد، هر آزمایش ۱۰ بار تکرار شده است و نتایج ارایه شده، میانگین این ۱۰ تکرار است. لازم به ذکر است در این ۱۰ تکرار، الگوهای آموزشی و آزمایشی یکسان است.

دسته دوم آزمایش ها، با اهداف:

- انتخاب تعداد بهینه مؤلفه های اساسی در مرحله استخراج ویژگی برای بازنمایی به خبره ها (بخش ۴-۱)
- انتخاب تعداد مناسبی از خبره ها در ساختار ارایه شده
- انتخاب پارامترهای ساختاری بهینه برای هر خبره
- انتخاب پارامترهای آموزش بهینه مانند تعداد تکرارهای آموزش<sup>۱۸</sup>

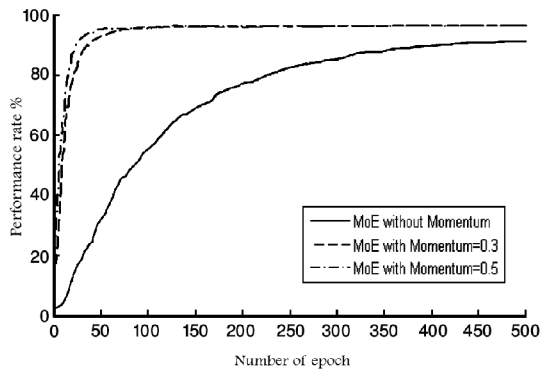
انجام شده که البته نتایج تمامی آزمایش های این تحقیق در این مقاله آورده نشده است. در ادامه خلاصه نتایج هر دو دسته از آزمایش ها به صورت نمودار و جداول آورده شده است.

جدول ۱ جزئیات ساختارهای به چالش کشیده شده و نرخ بازشناسی چهره به دست آمده مجموعه داده های ORL و Yale

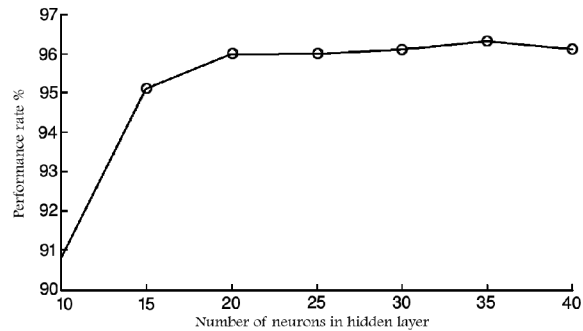
ساختار شبکه	شبکه پرسپترون چندلایه	اختلاط خبره ها	اختلاط خبره ها با شبکه چندلایه	ساختار ارایه شده
ORL توپولوژی	۴۰:۸۰:۴۰	۴ خبره	خبره: ۴۰:۲۰:۴۰ میانجی: ۴۰:۱۰:۴	خبره: ۴۰:۲۰:۴۰ میانجی: ۴۰:۱۰:۴
	۴۰:۶۰:۱۵	۴ خبره	خبره: ۴۰:۲۰:۱۵ میانجی: ۴۰:۱۰:۴	خبره: ۴۰:۲۰:۱۵ میانجی: ۴۰:۱۰:۴
نرخ یادگیری	۰/۱	در خبره ها: ۰/۱ در میانجی: ۰/۵	در خبره ها: ۰/۱ در میانجی: ۰/۵	در خبره ها: ۰/۱ در میانجی: ۰/۵
	Yale	%۸۶	%۹۰/۵	%۹۷
ORL	%۹۲/۵	%۹۵/۵	%۹۶	%۹۸/۵

پس از انتخاب ۴۰ مؤلفه اساسی از تصاویر چهره برای استخراج بهترین ساختار در شبکه ها و همچنین انتخاب پارامترهای یادگیری در همه شبکه ها آزمایش های زیادی را انجام شده است، جدول ۱ بهترین ساختار و بهترین پارامتر را با بالاترین نرخ بازشناسی چهره در هر ساختار به چالش کشیده شده، نشان می دهد. از این پس در آزمایش ها پارامترهای یادگیری  $\eta_g = 0.1$ ،  $\eta_e = 0.5$  و واحد گشتاور  $\alpha_e = \alpha_g = 0.5$  می باشد.

مسئله ۴۰ کلاسه و ۱۵ کلاسه را حل کنند، به همین علت تعداد نودهای (هر نود برای شناسایی یک نفر) لایه خروجی طبقه بندها برای دو بانک چهره به کار رفته، متفاوت است (لایه خروجی خبره ها ۱۵ نود برای بانک چهره Yale و لایه خروجی خبره ها ۴۰ نود برای بانک چهره ORL می باشد).



شکل ۷ نمودار نرخ بازشناسی چهره بر حسب تعداد تکرارهای آموزش ساختار بر روی بانک چهره RL



شکل ۶ نمودار نرخ بازشناسی بر حسب نودهای لایه میانی

شبکه های عصبی پرسپترون چند لایه ای که در خبره ها و شبکه میانجی ساختار به کار گرفته شده اند، دارای یک لایه پنهان می باشند. لایه ورودی این شبکه ها دارای ۴۰ نود و لایه میانی خبره ها دارای ۲۰ نود است (شکل ۶ نمودار) و لایه میانی در شبکه میانجی دارای ۱۰ نود می باشد. چون بانک چهره ORL دارای تصاویر ۴۰ نفر و Yale دارای تصاویر ۱۵ نفر می باشند طبقه بندها باید به ترتیب یک

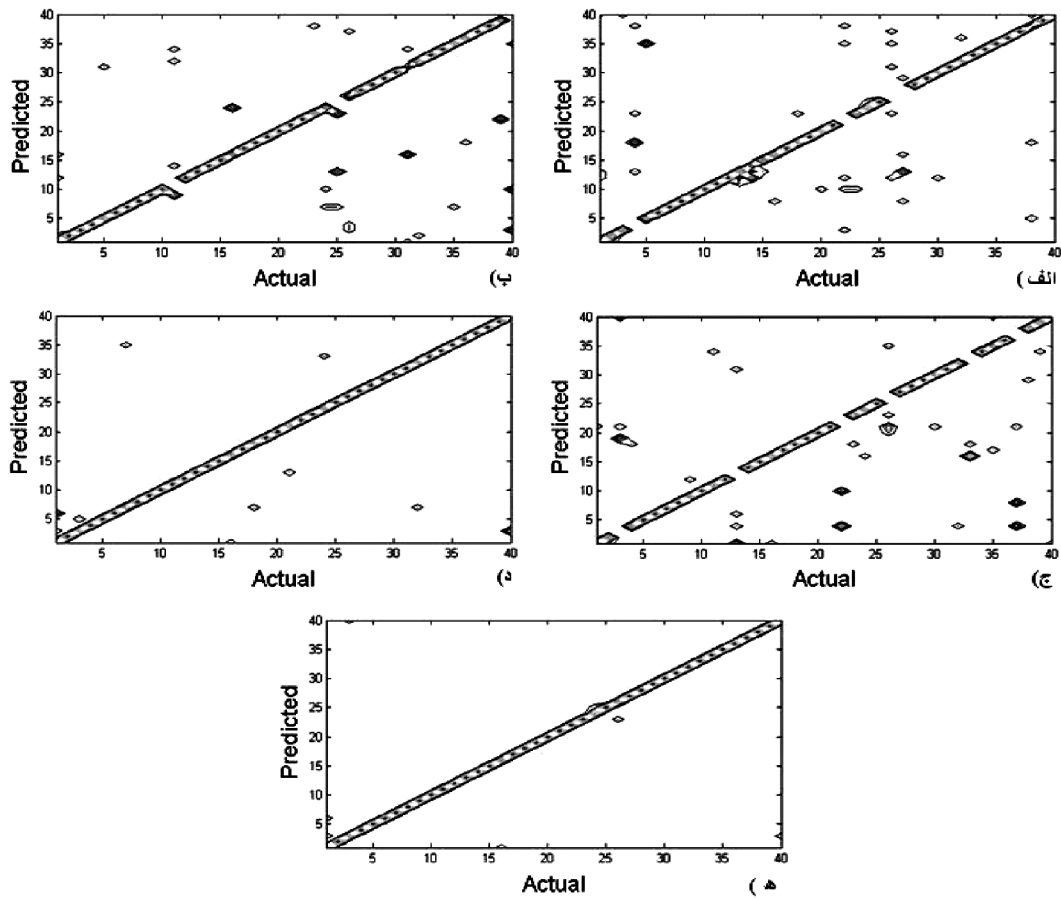
جدول ۲ مختصری بر روش های بازشناسی چهره و نرخ بازشناسی چهره به دست آمده بر مجموعه داده های Yale و ORL

مجموعه داده	مرجع و روش	نرخ بازشناسی %	توضیحات ساختار
Yale	[۲۴] ساختن ماشین های کمیته برای بازشناسی چهره (FRCM)، شبکه های عصبی، گراف کشسان (EGM)، ماشین های بردارهای پشتیبان (SVM)،	بالاترین کارایی را FRCM با نرخ بازشناسی ۸۶٫۱٪ نسبت به دیگر روش های به کار گرفته شده دارد	در این روش معتبرسازی به کار گرفته شده است.
	[۲۵] ترکیب مبتنی بر فیلد تصادفی مارکف (MRF)	۹۶٫۱۱٪	۵ تصویر برای آموزش و ۶ تصویر برای آزمایش به کار گرفته شده است.
	[۲۶] مدل مارکف پنهان	۸۷٪	
ORL	[۲۶] شیه 2DHMM	۹۵٪	زمان کلاس بندی و زمان آموزش داده نشده.
	[۲۷] SVM با درخت دودویی	۹۱٫۲۱٪ برای SVM و ۸۴٫۸۶٪ برای NCC	در این کار مقایسه ای بین SVM و روش eigenface استاندارد با استفاده از NCC (نزدیک ترین مرکز کلاس بندی) انجام گرفته است.
	[۲۸] بهینه شده Pairwise (O-PWC) SVM coupling	۹۵٫۱۳٪ برای PWC و ۹۶٫۷۹٪ برای O-PWC	۲۰۰ نمونه تصادفی برای آموزش و ۲۰۰ نمونه دیگر برای آزمایش به کار گرفته شده است.
	[۲۹] چندین SVM + شبکه عصبی میانجی	۹۷٫۹٪	میانگین زمان پردازش ۰٫۲۲ ثانیه گزارش شده است.
	[۳۰] PDBNN	۹۶٪	زمان بازشناسی چهره ۱ ثانیه و زمان آموزش شبکه ۲۰ دقیقه

پیشنهاد شده در این قسمت دارای نرخ بازشناسی بهتری بر روی بانک چهره Yale و ORL از دیگر ساختارهای به چالش کشیده شده است، با این تذکر که با سرعت بیشتر و تعداد تکرار کمتر به این کارایی می رسد.

واحد گشتاور در روابط اصلاح وزن ساختار اختلاط خیره ها با هدف افزایش سرعت همگرایی اضافه شد. برای درستی این ادعا، آزمایش هایی با واحدهای گشتاور متفاوت و همچنین بدون وارد کردن این واحد به ساختار انجام شده است. نتایج این آزمایش ها در شکل ۷ دیده می شود که نشان می دهد با  $\alpha_e = \alpha_g = 0.5$  می توان سریع تر (با تعداد تکرار کمتر) به نرخ بازشناسی چهره بالاتری دست یافت. در تمامی این آزمایش ها دیده می شود که کارایی روش





شکل ۸ الف تا د- ماتریس سردرگمی خیره ها و ه- ماتریس سردرگمی ساختار اختلاط خیره های ارایه شده بر روی بانک چهره ORL.

ماتریس های سردرگمی هر یک از خیره های ساختار ارایه شده و همچنین ماتریس سردرگمی کلی ساختار بر روی بانک چهره Yale و ORL است. توجه شود که نقاطی که بر روی قطر اصلی ماتریس سردرگمی کلی ساختار نیستند نشان دهنده تصاویری هستند که درست بازشناسی نشده اند.

برای نشان دادن چگونگی پخش خطاها در همه کلاس ها که روشی برای نمایش گوناگونی خیره ها است، از ماتریس سردرگمی<sup>۱۹</sup> بر روی مجموعه های آزمایش استفاده شده است. ماتریس سردرگمی اطلاعاتی در مورد کلاس بندی واقعی و کلاس بندی پیش بینی شده توسط طبقه بندها به دست می دهد [۲۳]. شکل ۸ و ۹ نمایش گرافیکی از

همان طور که مشخص است تعداد این نقاط بسیار کم است و این نشان دهنده همان نرخ بازشناسی بالایی است که در آزمایش های قبلی نیز ساختار به آن رسیده بود.

### ۵- جمع بندی و نتیجه گیری

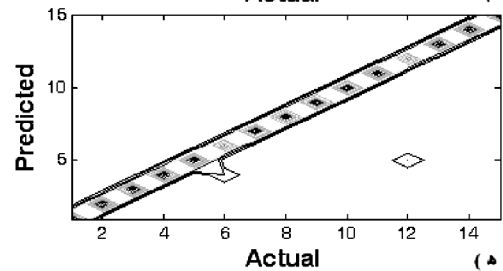
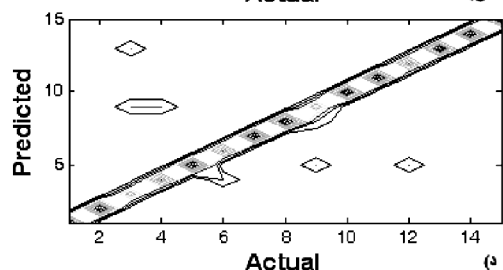
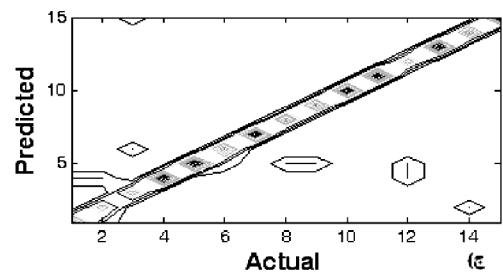
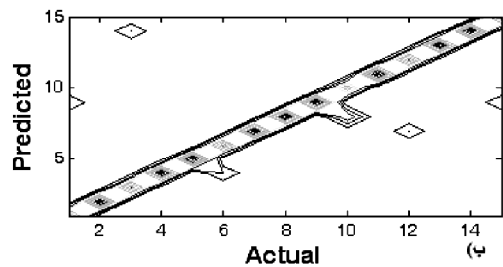
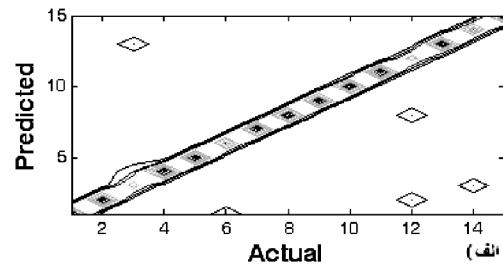
در این مقاله جهت بازشناسی چهره از ترکیب شبکه عصبی مبتنی بر اختلاط خبره ها استفاده شده است. در ساختار ارائه شده، ایده و چگونگی به کارگیری واحد گشتاور در روش پیشنهاد شده و اختلاط خبره ها با شبکه های پرسپترون چند لایه آورده شده است. وظیفه این واحد تسریع در همگرایی ساختار و تضمین بیشتر برای فرار از مینیمم های محلی می باشد. نتایج آزمایش های انجام شده و جدول ۲ نشان دهنده سرعت همگرایی بیشتر و بهبود نتایج بازشناسی چهره برای روش پیشنهادی در مقایسه با روش های مشابه موجود می باشد.

### تقدیر و تشکر

این پروژه با حمایت های مادی مرکز تحقیقات مخابرات ایران صورت گرفته است.

### ۶- پی نوشت

- <sup>1</sup> Appearance Based Methods
- <sup>2</sup> Model Based Methods
- <sup>3</sup> Principal Component Analysis (PCA)
- <sup>4</sup> Independent Component Analysis (ICA)
- <sup>5</sup> Unsupervised
- <sup>6</sup> Linear Discriminate Analysis (LDA)
- <sup>7</sup> Supervised
- <sup>8</sup> Elastic Bunch Graph (EBG)
- <sup>9</sup> Active Appearance Model (AAM)
- <sup>10</sup> Committee Machines
- <sup>11</sup> Ensemble Learning Algorithm
- <sup>12</sup> Mixture of Experts (ME)
- <sup>13</sup> Momentum term
- <sup>14</sup> Self Organizing
- <sup>15</sup> Gating Network
- <sup>16</sup> Multi-Layer Perceptron Network
- <sup>17</sup> Linear network
- <sup>18</sup> Training epochs
- <sup>19</sup> Confusion matrix



شکل ۹ الف تا د- ماتریس سردرگمی خبره ها و ه- ماتریس سردرگمی ساختار اختلاط خبره های ارائه شده بر روی بانک چهره Yale

## مراجع

- [1] Lu J., Plataniotis K. and Venetsanopoulos A., "Face Recognition Using LDA-Based Algorithms", IEEE Transaction on Neural Networks, Vol. 14, No. 1, 2003, pp. 195-200.
- [2] Turk M. and Pentland A., "Face Recognition Using Eigenfaces", the IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, USA, 1991, pp 586-591.
- [3] Sirovich L. and Kirby M., "Low-dimensional procedure for the characterization of human faces", Journal of Optical Society of America, Vol. 4, 1987, pp. 519-524.
- [4] Romdahani S., *Face recognition using principal component analysis*, Survey, 1999.
- [5] Bartlett M., Movellan J. and Sejnowski T., "Face Recognition by Independent Component Analysis", IEEE Transaction on Neural Networks, Vol. 13, No. 6, 2002, pp. 1450-1464.
- [6] Liu C. and Wechsler H., "Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition s", the Second International Conference on Audio and Video-based Biometric Person Authentication, Washington D.C., USA, 2000, pp. 211-216.
- [7] Baek K., Draper B.A., Beveridge J.R. and She K., "PCA vs ICA: A comparison on the FERET data set", Joint Conference on Information Sciences, Durham, NC, 2002.
- [8] Lu J., Plataniotis K., Venetsanopoulos A. and Li S., "Ensemble-Based Discriminant Learning With Boosting for Face Recognition", IEEE transaction on neural networks, Vol. 17, NO. 1, 2006.
- [9] Wiskott L., Fellous J., Krueger N. and Malsburg C., "Face Recognition by Elastic Bunch Graph Matching", Chapter 11 in Intelligent Biometric Techniques in Fingerprint and Face Recognition, CRC Press, 1999, pp. 355-396.
- [10] Wiskott L., Fellous J., Krueger N. and Malsburg C., "Face Recognition by Elastic Bunch Graph Matching", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, 1999, pp. 776-779.
- [11] Cootes T. and Taylor C., "Statistical Models of Appearance for Computer Vision", Technical Report, University of Manchester, 1999.
- [12] Cootes T., Walker K. and Taylor C., "View-Based Active Appearance Models", IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France, 2000, pp. 227-232.
- [13] Freund Y. and Schapire R., "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", Journal of Computer and System Sciences, Vol. 55, No. 1, 1997, pp. 119-139.
- [14] Meir R. and Raetsch G., "An Introduction to Boosting and Leveraging", Advanced Lectures on Machine Learning, 2003, pp. 118-183.
- [15] Haykin S., *Neural Networks: A Comprehensive Foundation*, book, Prentice Hall PTR Upper Saddle River, USA, 1998.
- [16] Jacobs R., Jordan M., Nowlan S., and Hinton G., "Adaptive Mixtures of Local Experts", Neural Comput3, 1991, pp. 79-87.
- [17] Ebrahimpour R., Kabir E. and Yousefi M., "Teacher-Directed Learning in View-independent Face Recognition with Mixture of Experts Using Overlapping Eigenspaces", Computer Vision and Image Understanding, Vol. 111, 2008, pp. 195-206.
- [18] Ebrahimpour R., Kabir E. and Yousefi M., "Teacher-Directed Learning in View-independent Face Recognition with Mixture of Experts Using Single-view Eigenspaces", Journal of Franklin Institute, Vol. 345, 2008, pp. 87-101.
- [19] Ebrahimpour R., Kabir E., Esteky H. and Yousefi M., "View-independent Face Recognition with Mixture of Experts, Neurocomputing", Vol. 71, 2008, pp. 1103-1107.
- [20] Ebrahimpour R., Kabir E., Esteky H. and Yousefi M., "A Mixture of Multilayer Perceptron Experts Network For Modeling Face/Nonface Recognition in Cortical Face Processing Regions", Intelligent Automation and Soft Computing, Vol. 14, No. 2, 2008, pp. 145-156.
- [21] Ebrahimpour R., Kabir E. and Yousefi M., "Face Detection Using Mixture of MLP Experts", Neural Processing Letters, Vol. 26, No. 1, 2007, pp. 69-82.

- [22] Beale R. and Jackson T., "*Neural computing: an introduction*", Institute of Physics Publishing, **1998**.
- [23] Kohavi R. and Provost F., "*Glossary of terms*", Machine Learning, 30, **1998**, pp.271-274.
- [24] Tang H., Lyu M. and King I., "*Face recognition committee machine*", In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, April 6-10, **2003**, pp. 837-840.
- [25] Huang R., Pavlovic V. and Metaxas D. N., "*A hybrid face recognition method using Markov random fields*", **2004**, pp. 157-160.
- [26] Samaria F. and Harter A.C., "*Parameterisation of a stochastic model for human face identification*", Second IEEE Workshop Applications of Computer Vision, **1994**.
- [27] Jonsson K., Kittler J., Li Y. P. and J. Matas, "*Support vector machines for face authentication*", in Pridmore T. and Elliman D. , editors, British Machine Vision Conference, **1999**, pp. 543-553.
- [28] Guo G.D., Zhang H.J. and Li S.Z., "*Pairwise face recognition*", In Proceedings of 8th IEEE International Conference on Computer Vision, Vancouver, Canada. July 9-12, **2001**.
- [29] Kim K.I., Jung K. and Kim J., "*Face recognition using support vector machines with local correlation kernels*" , International Journal of Pattern Recognition and Artificial Intelligence, Vol. 16 No. 1, **2002**, pp. 97-111.
- [30] Lin S.H., Kung S.Y. and Lin L.J., "*Face recognition/detection by probabilistic decision-based neural network*", IEEE Trans. Neural Networks, vol. 8, **1997**, pp. 114-132.