

شناسایی تظاهرات احساسی درد در تصاویر چهره

محمدرضا نوروزیⁱ؛ منصور وفادوستⁱⁱ

چکیده

در این تحقیق با استفاده از تصاویر سطح خاکستری چهره، امکان طبقه بندی تظاهرات احساسی "درد" در مقابل سایر تظاهرات احساسی شامل حالت‌های "خنثی"، "خوشحالی"، "تعجب" و "عصبانیت" تحت بررسی قرار گرفته و سعی شده تا حالت درد در سه شدت مختلف درجه بندی گردد. ابزارهای استخراج ویژگی شامل مولفه‌های اساسی، گشتاورهای زرنیک و موجک گابور در کنار طبقه بندی کننده های شبکه عصبی پس انتشار خطا، ماشین بردار پشتیبان و آنالیز به وسیله تفکیک کننده های خطی مورد استفاده قرار گرفته است. در یک بانک اطلاعاتی شامل ۴۷۶ تصویر از ۳۰ نفر، دقت طبقه بندی حالات درد شدید، متوسط و کم به ترتیب برابر با ۷۷/۴۴ درصد، ۷۳/۷۱ درصد و ۷۱/۱۵ درصد به دست آمده است.

کلمات کلیدی

پردازش تصویر، تصویر چهره، تظاهرات احساسی چهره، شناسایی درد، استخراج ویژگی، طبقه بندی کننده.

Recognition of Pain Expression in Facial Images

Mohammadreza Noruzi; Mansour Vafadoost

ABSTRACT

A facial expression discrimination system has been developed which uses grayscale images and classifies the pains expression against the other facial expressions including: Neutral, Happiness, Anger and Surprise expressions. Principal Component Analysis (PCA), Zernike Moments and Gabor Wavelet Transform have been used for feature extraction and Support Vector Machine (SVM), Linear Discriminant Analysis (LDA) and Feed Forward Neural Network have been used for feature classification. A database including 476 face images from 30 persons has been used in this research. The result of classification is 77.44% for severe pain, 73.71% for moderate pain and 71.15% for low pain expression.

KEYWORDS

Image processing, Face image, Facial expression, Pain recognition, Feature extraction, Classification.

ⁱ دانشجوی دکترای رشته مهندسی پزشکی گرایش بیوالکترونیک، دانشگاه صنعتی امیرکبیر، دانشکده مهندسی پزشکی:

Email: noruzi@cic.aut.ac.ir

ⁱⁱ استادیار دانشگاه صنعتی امیرکبیر، دانشکده مهندسی پزشکی: Email: vmansur@aut.ac.ir

در تحقیق حاضر، روش‌های مختلفی برای استخراج ویژگی و طبقه بندی ویژگی‌ها به کار رفته است. در بخش استخراج ویژگی از تبدیل مولفه‌های اساسی (PCA: Principal Component Analysis)، گشتاورهای زرنیک (Zernike Moments) و تبدیل موجک گابور (Gabor Wavelet) و در بخش طبقه بندی، از شبکه عصبی جلو سو با آموزش به روش پس‌انتشار خطا (Back Propagation Neural Network)، آنالیز به روش تفکیک کننده‌های خطی (LDA: Linear Discriminant Analysis) و ماشین بردار پشتیبان (SVM: Support Vector Machine) استفاده شده است.

۲- تحقیقات مرتبط

در اینجا به تعدادی از روش‌های ارائه شده در زمینه شناسایی تظاهرات احساسی چهره اشاره می‌گردد:

گروه Edwards با در نظر گرفتن تصویر یکپارچه، مدلی از چهره را ارائه می‌کند. بر اساس جابجایی ۱۲۲ نقطه که به صورت دستی بر روی اجزای اصلی صورت چسبانده شده‌اند یک مدل استاتیکی از چهره تولید می‌گردد و با اعمال تبدیل مولفه‌های اساسی به همراه LDA، طبقه بندی نهایی انجام می‌گردد. بانک اطلاعاتی مورد استفاده شامل شش حالت احساسی از ۲۵ نفر بوده و نرخ شناسایی برابر با ۷۴ درصد اعلام گردیده است [۹].

گروه Hong از یک گراف ۱۶ نقطه‌ای استفاده می‌کند که در محل گره‌های گراف، ویولت گابور در ۸ جهت و ۵ فرکانس متفاوت محاسبه می‌گردد. در این روش ابتدا یک گراف مشخص بر روی تصویر ورودی منطبق می‌گردد. سپس تصاویر موجود در بانک اطلاعاتی با تصویر ورودی مقایسه شده و آن تصویری را که گراف مربوطه اش بیشترین انطباق را با گراف مربوط به تصویر ورودی داشته باشد انتخاب می‌گردد. نرخ شناسایی احساسات شش‌گانه برای افراد موجود در پایگاه اطلاعاتی برابر با ۸۹ درصد و برای افراد بیگانه برابر با ۷۳ درصد اعلام گردیده است [۱۸]. در روشی که بوسیله

Hara و Kobayashi ارائه شده، هندسه قرار گیری اجزای چهره نسبت به یکدیگر، معیار سنجش می‌باشد. بر اساس فاصله بین دو چشم، مکان ۳۰ نقطه اساسی روی چهره تعیین شده و در ادامه به کمک یک شبکه عصبی با قانون پس انتشار خطا، شش حالت احساسی طبقه بندی می‌شود. بانک اطلاعات شامل ۹۰ تصویر برای ۱۵ نفر می‌باشد. در این روش نرخ شناسایی برابر با ۸۵ درصد اعلام گردیده است [۲۲].

آشکارسازی تظاهرات احساسی چهره، از موضوعات مهم در زمینه ارتباط انسان با ماشین می‌باشد. در کنارحالت خنثی، روانشناسان عملکرد ظاهری چهره را به ۶ دسته اصلی تقسیم کرده‌اند که شامل حالت‌های: خوشحالی، ناراحتی، عصبانیت، ترس، تعجب و چندان می‌شود. آنها همچنین فعالیت‌های موضعی صورت را به ۴۴ دسته متمایز تقسیم کرده‌اند که واحدهای عملکردی (Action Units) نامیده می‌شوند [۱۰].

در شکل کلی دو گونه تصویر در طبقه بندی تظاهرات احساسی مورد استفاده قرار می‌گیرد که دسته اول شامل تصاویر ثابت و دسته دوم شامل تصاویر متحرک می‌باشد. بر اساس این دسته بندی، روش‌های مختلفی ارائه شده است که در دسته اول، استخراج ویژگی‌های ساختاری چهره و در دسته دوم استخراج ویژگی‌های حرکتی چهره مد نظر قرار گرفته است. البته حالت‌های ترکیبی نیز در بعضی از تحقیقات مشاهده شده است. از دیدگاه دیگر نیز می‌توان روش‌های اعمال شده را تقسیم بندی نمود. بعضی از روشها به تمامیت چهره پرداخته و کل تصویر را تحت پردازش قرار می‌دهند در حالی که بعضی دیگر، نواحی محدودی از چهره را بصورت جداگانه مورد آنالیز قرار داده و سپس بر اساس این نتایج محلی، عمل تصمیم گیری را انجام می‌دهند [۱۱].

آشکارسازی و تمایز تظاهرات احساسی در تصاویر ثابت معمولاً به وسیله مقایسه تصاویر ورودی با تصاویر مرجع انجام می‌گردد. تقریباً تحقق این مقصود بوسیله مقایسه مستقیم شدت روشنایی پیکسل‌های تصاویر ناممکن می‌باشد. در عمل، یکسری ویژگی از تصاویر مورد نظر استخراج شده و عمل مقایسه بر اساس این ویژگی‌ها انجام می‌گردد. استخراج بردار ویژگی معمولاً بوسیله تبدیل‌های ریاضی انجام شده و مقایسه این بردارها، توسط طبقه بندی کننده صورت می‌گیرد.

مطابق دسته بندی ارائه شده برای تظاهرات احساسی چهره، حالت درد جزء شش حالت اصلی طبقه بندی شده توسط روانشناسان نمی‌باشد، زیرا حالت "درد" و حالت‌های "بد آمدن"، "ترس" و "عصبانیت"، در تعدادی از واحدهای عملکردی مشترک می‌باشند [۷]. طبقه بندی تظاهرات درد در چهره از دیدگاه پزشکی و مراقبتی می‌تواند بسیار مهم و سودمند باشد. از این دیدگاه می‌توان حالت "درد" را در یک کلاس جداگانه فرض نموده و سایر احساسات را در مقابل آن در کلاس "غیر درد" طبقه بندی نمود. در این تحقیق سعی شده تا ضمن پیاده سازی این دیدگاه متفاوت، کلاس "درد" در سه شدت مختلف طبقه بندی

در روش ارائه شده توسط Kimura و Yachida، به وسیله اعمال مشتق گیر، لبه‌های تصویر استخراج می‌گردد و سپس بر روی این لبه‌ها، یک فیلتر نرم کننده اعمال می‌گردد تا جهت حرکت‌ها به همراه سرعت آنها به عنوان بردار ویژگی استخراج شود [۲۱]. همچنین در الگوریتم ارائه شده توسط Otsuka و Ohya بر اساس فلوی نوری مبتنی بر گرادیان، حرکت چشم راست و دهان بررسی می‌گردد. تصاویر مورد استفاده به وسیله یک دوربین دریافت می‌گردند که بر روی یک کلاه خود نصب شده و همیشه در مقابل فرد می‌باشد. با اعمال تبدیل فوریه دو بعدی بر روی بردار سرعت افقی و عمودی، یک بردار ویژگی ۱۵ بعدی استخراج می‌گردد [۲۸].

در روش ارائه شده به وسیله سید عربی، از تصاویر ویدئویی استفاده می‌گردد. در این روش نقاط مهم در تصویر اولیه ورودی، توسط کاربر علامت گذاری می‌گردند. سپس، بر اساس شباهت بین هر ناحیه با نواحی همسایه در تصاویر بعدی، نحوه جابجایی این نقاط در تصاویر متوالی استخراج می‌گردد. از این ویژگی برای جداسازی واحدهای عملکردی چهره استفاده شده است [۳۰].

در روش ارائه شده توسط گروه Lyons، ۲۴ نقطه به صورت دستی روی چهره مشخص گشته و در آنها تبدیل گابور گرفته می‌شود. سپس دامنه ضرایب گابور نمونه برداری شده و در یک بردار قرار می‌گیرند. پس از اعمال تبدیل PCA، با استفاده از LDA، عمل طبقه بندی انجام می‌گیرد. در این روش نرخ شناسایی تظاهرات احساسی برابر با ۹۲ درصد اعلام شده است [۲۳].

تعداد تحقیقاتی که به طور مشخص در زمینه حالت درد انجام شده باشند بسیار انگشت شمار بوده و همگی مربوط به سالهای اخیر می‌باشند. در اولین گزارش رسمی در این رابطه توسط گروه Hancock، از تبدیل گابور برای استخراج ویژگی و از شبکه عصبی برای طبقه بندی استفاده شده است. بانک اطلاعاتی مورد استفاده در این روش، شامل حالت درد به همراه سایر تظاهرات احساسی اصلی می‌باشد که از ۲۳ نفر تهیه شده است. دقت طبقه بندی حالت درد نسبت به سایر حالات برابر با ۸۰ درصد عنوان شده است [۱۶]. در تحقیق دیگر با استفاده از تبدیل مولفه‌های اساسی و طبقه بندی کننده ماشین بردار پشتیبان، حالت درد در نوزادان تازه متولد شده با طول عمر حد اکثر ۳۶ ساعت، استخراج شده است. برای این منظور یک بانک اطلاعاتی شامل ۲۰۴ تصویر از ۲۶ نوزاد تهیه شده است. درصد شناسایی درد برابر با ۸۸ درصد ثبت گردیده است [۳]. در آخرین تحقیق ارائه شده در زمینه درد، از تصاویر ویدئویی

برای طبقه بندی تظاهرات احساسی درد استفاده شده است. در این روش از تصاویر آموزشی که حاوی حالات درد می‌باشند، مولفه‌های اساسی استخراج شده و فضای ویژگی ساخته می‌شوند. سپس بر اساس رنگ پوست صورت، محل چهره در تصاویر ورودی جایابی شده و سپس به فضای ویژگی‌ها نگاشته می‌گردد. بر اساس فاصله بردارهای ویژگی با مقادیری از پیش تعیین شده، عمل طبقه بندی انجام می‌گردد. برای حالت درد نسبت به حالت غیر درد، دقت جایابی برابر با ۸۹ درصد عنوان شده است و در صورتی که از نواحی چشم‌ها و دهان نیز بردارهای ویژه جداگانه‌ای استخراج گردد، این دقت به ۹۲ درصد می‌رسد [۲۶]، [۲۵].

۳- ابزارهای ریاضی مورد استفاده

یکی از ابزارهای مورد استفاده، شبکه عصبی جلو سو با آموزش به روش پس‌انتشار خطا است. باتوجه به آنکه اطلاعات مربوط به شبکه‌های عصبی پس انتشار خطا در هر کتاب درسی مرتبط با این موضوع یافت می‌شود، از ارائه توضیحات مربوط به آن خود داری می‌گردد. مقدمات مربوط به شبکه‌های عصبی در مرجع [۱۲] و مباحث تکمیلی در مرجع [۱۷] قابل دسترسی می‌باشد. در همین راستا، اطلاعات مربوط به آنالیز به روش تفکیک کننده‌های خطی (LDA) در مراجع [۵] و [۸] قابل دسترسی است.

۳-۱- تبدیل مولفه‌های اساسی

در تبدیل مولفه‌های اساسی (PCA) محورهای یکه و متعامدی در فضای ویژگی یافت می‌شوند که بهترین باز نمایی را از سیگنال ورودی ارائه می‌دهند. این محورهای جدید را محورهای ویژه و بردار مشخصه آنها، بردار ویژه نامیده می‌شوند. تبدیل مولفه‌های اساسی از طریق محاسبه مقادیر ویژه ماتریس کوواریانس نمونه‌های ورودی محاسبه می‌شود. نکته قابل توجه آن است که در شرایط عملی، ممکن است رتبه ماتریس کوواریانس کامل نبوده و در نتیجه بعضی از مقادیر ویژه صفر باشند. یکی از روشهایی که مقادیر ویژه غیر صفر ماتریس کوواریانس را محاسبه می‌کنند، روش تجزیه به مقادیر تکین (SVD: Singular Value Decomposition) است [۵]. از این روش در این مقاله نیز استفاده شده است.

۳-۲- گشتاورهای زرنیک

در این روش از چند جمله‌ای‌های زرنیک استفاده کرده و یک فضای متعامد و کامل برای استخراج ویژگی تولید می‌گردد. این چند جمله‌ای‌ها از نوع اعداد مختلط بوده و متعامد می‌باشند. برای یک تابع دو بعدی $f(x, y)$ که در داخل دایره واحد

نگاشت شده است می توان گشتاور mn ام زرنیک را طبق رابطه ۱ بدست آورد [۱]، [۲۰].

$$A_{mn} = \frac{m+1}{\pi} \iint_{x,y} f(x,y) [V_{mn}(x,y)]^* \quad (۱)$$

که در این رابطه:

$$x^2 + y^2 < 1$$

$$m = 0, 1, 2, \dots, \infty \quad \& \quad |n| \leq m$$

همچنین در رابطه ۱، تفاضل m از قدر مطلق n برابر عدد زوج بوده و $V_{mn}(x,y)$ بیانگر چندجمله‌ای‌های زرنیک می باشد. در فضای قطبی، مقدار $V_{mn}(x,y)$ به صورت رابطه ۲ در داخل دایره واحد قابل بیان است.

$$V_{mn}(r,\theta) = R_{mn}(r) \exp(in\theta) \quad (۲)$$

با فرض $i = \sqrt{-1}$ مقدار $R_{mn}(r)$ با استفاده از روابط ۳ و ۴ قابل محاسبه است.

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s F(m,n,s,r) \quad (۳)$$

در این حالت تابع F طی رابطه ۴ تعریف می شود.

$$F(m,n,s,r) = \frac{(m-s)!}{s! \left(\frac{m+|n|}{2} - s\right)! \left(\frac{m-|n|}{2} - s\right)!} r^{m-2s} \quad (۴)$$

۳-۳- تبدیل گابور

تبدیل گابور یک تبدیل زمان-فرکانس است که در حالت دو بعدی مطابق رابطه ۵ نشان داده می شود.

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\{-0.5(x'^2/\sigma_x^2 + y'^2/\sigma_y^2)\} \exp\{i2\pi\omega_0 x'\} \quad (۵)$$

که در این رابطه:

$$x' = x \cos \theta + y \sin \theta$$

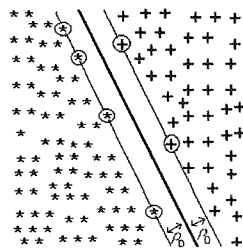
$$y' = -x \sin \theta + y \cos \theta$$

$g(x,y)$ نشان دهنده فیلتر دو بعدی به مرکز $(0,0)$ ، $i = \sqrt{-1}$ یک مقدار ثابت، ω_0 فرکانس مدولاسیون، θ زاویه مدولاسیون و σ_x, σ_y پهناهای موثر فیلتر می باشند. با اعمال تبدیل گابور به یک تصویر، محتویات فرکانسی تصویر که در پنجره زیر فیلتر قرار گرفته در راستای زاویه θ استخراج می شوند [۱۹]، [۲۴]. تبدیل گابور محتویات فرکانسی مرزها، بافت و همچنین جهت توزیع آنها را استخراج می کند. تبدیل گابور نه تنها می تواند مرزها، بلکه زاویه گسترش آنها را نیز استخراج نماید. وجود این خاصیت در تبدیل گابور آنرا به یک ابزار قدرتمند جهت آنالیز بافت تصاویر تبدیل کرده است [۴]، [۱۶]، [۲۹].

۳-۴- ماشین بردار پشتیبان

اساس کار طبقه بندی کننده به روش ماشین بردار پشتیبان (SVM) جداسازی خطی کلاسها در فضایی مناسب است. روش SVM برای مسائلی شامل دو کلاس ورودی، قابل استفاده است. این روش به دنبال یافتن ابر صفحه‌ای است که بتواند با بیشترین فاصله از دو کلاس، عمل طبقه بندی را انجام دهد.

اینک داده‌های دو کلاس که به صورت خطی و با خطایی کم در فضا قابل جدا شدن می باشند در نظر گرفته می شوند. هدف، یافتن بهترین ابر صفحه جداکننده دو کلاس است، به طوریکه بیشترین فاصله را از هر یک از کلاسها داشته باشد. مطابق شکل ۱، این ابر صفحه بهینه دقیقاً در بین کلاسها قرار گرفته و از هر یک از آنها دارای فاصله ρ_0 می گردد. در این شرایط داده‌هایی که در فاصله ρ_0 از ابر صفحه قرار گرفته اند به نام بردارهای پشتیبان خوانده می شوند.



شکل (۱): وضعیت بردارهای پشتیبان نسبت به ابر صفحه بهینه

اکنون مجموعه داده‌های آموزش $\{(\mathbf{X}_i, d_i)\}_{i=1}^N$ در نظر گرفته می شوند که \mathbf{X}_i بیانگر i امین داده ورودی و d_i نشان دهنده کلاس تعلق داده مربوطه است. در صورتی که داده ورودی به کلاس اول متعلق باشد، مقدار d_i +۱ و در صورتی که داده ورودی به کلاس دوم متعلق باشد، مقدار d_i -۱ را اختیار می نماید. با فرض اینکه داده‌ها به صورت خطی جداپذیر باشند می توان معادله ابر صفحه جدا کننده را به صورت رابطه ۶ نوشت:

$$\mathbf{w}^T \mathbf{x} + b = 0 \quad (۶)$$

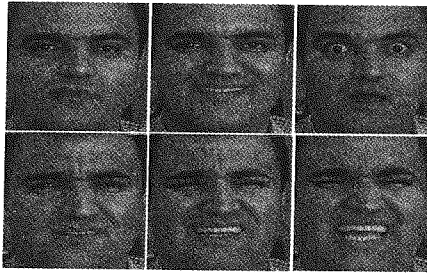
که در آن \mathbf{x} نشان دهنده متغیر فضای داده‌ها، b یک مقدار بایاس ثابت و \mathbf{w} بیانگر بردار ضرایب تعریف کننده ابر صفحه می باشد. بر اساس معادله ۶ می توان نوشت:

$$\mathbf{w}_0^T \mathbf{x}_i + b_0 \geq 0 \quad (\text{برای } \mathbf{x}_i \text{ با } d_i + 1) \quad (۷)$$

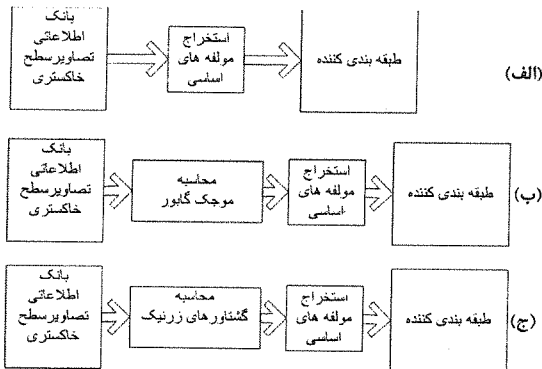
$$\mathbf{w}_0^T \mathbf{x}_i + b_0 \leq 0 \quad (\text{برای } \mathbf{x}_i \text{ با } d_i - 1)$$

می توان نشان داد که بر اساس شرط موجود در رابطه ۷، ابر صفحه بهینه از طریق کمینه کردن مقدار $\|\mathbf{w}_0\|$ به دست می آید. بر این اساس تابع هدف مناسبی می توان به صورت رابطه ۸ تعریف نمود.

آزمایش‌ها از همه تصاویر بانک اطلاعاتی استفاده شده است.



شکل (۲): نمونه‌ای از تصاویر جایابی شده بانک اطلاعاتی.



شکل (۳): بلوک دیاگرام اجرای کار. الف) استخراج مولفه‌های اساسی از تصاویر خام سطح خاکستری. ب) استفاده از موجک گابور برای استخراج ویژگی. ج) استفاده از گشتاورهای زرنیک برای استخراج ویژگی.

از سه ابزار مولفه‌های اساسی، تبدیل موجک گابور و گشتاورهای زرنیک برای استخراج ویژگی استفاده شده است. ردیفهای الف، ب و ج از شکل ۳ به ترتیب نشان دهنده بلوک دیاگرام به کار رفته برای این سه روش می‌باشند.

در روش منطبق بر موجک گابور، این تبدیل در چهار جهت به زوایای صفر، ۴۵، ۹۰ و ۱۳۵ درجه و سه سطح به تصاویر اعمال می‌گردد. مقدار σ_x در هر سطح به ترتیب برابر با مقدار $\pi/15$ ، $\pi/7/5$ و $\pi/2/7/5$ ، همچنین $\sigma_y = 0/5 \sigma_x$ و فرکانس مدولاسیون فیلترها در هر سطح به ترتیب برابر با $\pi/2$ ، $\pi/5$ و $\pi/12/5$ انتخاب شده است. از کنار هم قرار دادن خروجی تبدیل گابور، بردار ویژگی اولیه حاصل می‌شود در حالی که این بردار طولی بزرگ داشته و از مقادیر مختلط تشکیل می‌گردد. در این تحقیق از مقادیر "دامنه"، "فاز"، ترکیب آنها و همچنین "گذر از صفر ضرایب تبدیل گابور" برای استخراج ویژگی استفاده شده است. با توجه به بعد زیاد بردارهای تولیدی، مسئله کاهش بعد ویژگی‌ها اهمیت زیادی پیدا می‌نماید. برای مثال با اعمال تبدیل گابور در ۴ جهت و سه سطح به یک تصویر 64×64 ، یک بردار ویژگی به طول

$$J(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^N \alpha_i [(w^T x_i + b) d_i - 1] \quad (۸)$$

که ضرایب غیر منفی α_i ، ضرایب لاگرانژ نامیده می‌شوند. جوابهای عمومی معادله بالا تشکیل یک رویه می‌دهند که جواب بهینه، نقطه زینی در رویه تولیدی است. این نقطه بهینه از طریق کمینه سازی $J(w, b, \alpha)$ نسبت به w و b و بیشینه سازی نسبت به α محاسبه می‌شود.

در حالت عمومی، مرز بین دو کلاس ورودی لزوماً یک خط راست نبوده و همواره نمی‌توان آنها را به وسیله یک ابرصفحه جدا نمود. در این شرایط، یک تابع غیر خطی مناسب $\phi(\cdot)$ پیدا می‌شود که بتواند داده‌های ورودی را به یک فضای جدید نگاشت نماید، در این فضای جدید، وضعیت داده‌های ورودی می‌تواند به صورتی در آید که دو کلاس به صورت خطی و به وسیله یک ابر صفحه از یکدیگر متمایز گردند. در این شرایط، روش SVM از نوع مبتنی بر هسته غیر خطی نامیده می‌شود. برای اطلاعات بیشتر در زمینه ماشین بردار پشتیبان می‌توان به [۸] و [۱۷] مراجعه نمود.

ع- بانک اطلاعاتی

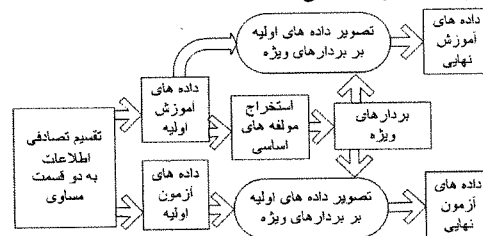
بانک اطلاعاتی به کار رفته، مشتمل بر ۴۷۶ تصویر سطح خاکستری از ۳۰ نفر می‌باشد. این بانک اطلاعاتی در دو مرحله تهیه شده و در هر مرحله از هر نفر درخواست گردیده تا حالتهای احساسی "خنثی"، "خوشحالی"، "تعجب"، "عصبانیت" و همچنین حالت "درد" در سه درجه مختلف را تولید نماید. در ۱۰ نفر از افراد بانک اطلاعاتی، از محرکهای دردآور استفاده شده تا تصاویر بدست آمده به واقعیت نزدیکتر باشند. تصاویر بدست آمده توسط دو فرد مجرب و همچنین یک نفر روانشناس برچسب زده شده اند. درصد فراوانی کلاس‌های درد شدید، متوسط و کم به کل نمونه‌ها برابر با ۰/۰۹، ۰/۱۴ و ۰/۱۸ می‌باشد. فرض بر این است که محل چهره در تصاویر مورد استفاده، توسط یک الگوریتم مناسب [۲۷] جایابی شده است. خطای روش مورد استفاده برای جایابی چهره، برابر با یک درصد بوده که درحالت خطا، مکان چهره به صورت دستی تعیین شده است. ناحیه چهره در تصاویر جایابی شده به اندازه 64×64 نرمالیزه می‌شوند. شکل ۲ نمونه‌ای از تصاویر مورد استفاده در بانک اطلاعاتی را نشان می‌دهد.

۵- الگوریتم‌ها

الگوریتمهای مورد استفاده از دو قسمت اصلی شامل "استخراج ویژگی" و "طبقه بندی ویژگی‌ها" تشکیل می‌گردد. بلوک دیاگرام روش‌ها در شکل ۳ ارائه شده است. در تمامی

۶۴×۶۴×۴×۳=۴۹۱۵۲ ایجاد می‌شود. به منظور کاهش بعد ویژگی‌ها، از تبدیل مولفه‌های اساسی استفاده شده است. در آزمایش‌هایی که مقادیر "فاز" و "اندازه" به صورت همزمان برای ساختن بردار ویژگی مورد استفاده قرار گرفته‌اند، تبدیل مولفه‌های اساسی به صورت جداگانه به هر یک از این مقادیر اعمال شده و نهایتاً بردار ویژگی هر یک از آنها در کنار یکدیگر قرار می‌گیرد. از تمام بردارهای ویژه که توسط تبدیل مولفه‌های اساسی ایجاد می‌شوند بهره برده شد. به وسیله تصویر نمودن داده‌های اولیه آموزش و آزمون بر این بردارهای ویژه، داده‌های نهایی آموزش و آزمون، با ابعاد بسیار کمتر ایجاد می‌گردند.

برای با گشتاورهای زرنیکی از ۳۰ جمله اول استفاده شده است. در این حالت اگرچه مشکل بزرگی ابعاد بردارهای ویژگی وجود ندارد اما به جهت کامل نبودن فضای ویژگی، دوباره باید از تبدیل مولفه‌های اساسی استفاده شود تا رتبه فضا کامل گردد. در روش مولفه‌های اساسی، ابتدا مقادیر ویژه و بردارهای ویژه از تصاویر آموزشی استخراج می‌گردند و سپس تصاویر آموزشی و همچنین تصاویر آزمون بر روی این بردارهای ویژه تصویر می‌گردند تا بردارهای ویژگی جهت آموزش و آزمون طبقه بندی کننده فراهم گردد. بلوک دیاگرام این روش در شکل ۴ نشان داده شده است.



شکل (۴): بلوک دیاگرام مرحله اعمال مولفه‌های اساسی و استخراج داده‌های آموزش و آزمون برای طبقه بندی کننده.

در صورتی که از تصاویر خام به عنوان اطلاعات ورودی به بخش مولفه‌های اساسی استفاده شود، بردار ورودی بعدی برابر با ۶۴×۶۴ دارد. نیمی از تصاویر بانک اطلاعاتی به عنوان نمونه‌های آموزشی استفاده شده و از آنها مولفه‌های اساسی استخراج می‌گردد. فضای جدیدی که تولید می‌شود بعدی نزدیک به ۲۰۰ پیدا می‌کند. این بعد مقداری ثابت نبوده و وابستگی به خصوصیت اطلاعات ورودی دارد. اما در حالت کلی، کمتر یا مساوی با تعداد نمونه‌های آموزش است.

۶- نتایج پیاده سازی‌ها

در تمامی موارد، داده‌های آموزش و آزمون به صورت تصادفی و به تعداد مساوی انتخاب شده، به طوری که هیچگونه

داده یکسانی بین دادگان آموزش و آزمون موجود نبود. آزمایش‌ها به صورت Cross-Validation [۸] و به تعداد ۵۰ تکرار انجام شده است. در هر تکرار، هریک از کلاسهای احساسی موجود در بانک اطلاعاتی به صورت تصادفی به دو قسمت مساوی تقسیم شده تا دادگان آموزش و آزمون ساخته شود. به عبارتی در هر تکرار، بلوک دیاگرام نمایش داده شده در شکل ۴ یکبار اجرا می‌شود.

در ادامه، منحنی مشخصه عملکرد گیرنده (ROC: Receiver Operating Characteristic) [۱۳]، [۲] برای ارزیابی نتایج حاصل از عملکرد قسمت طبقه بندی کننده بر روی داده های آزمون، رسم می‌گردد. در تکرار بعد، این تقسیم تصادفی مجدداً انجام می‌پذیرد. در پایان بین منحنی های بدست آمده از تکرارهای مختلف، متوسط گیری می‌شود.

بر اساس نتایج مربوط به شبکه عصبی، در عموم موارد به اجبار، باید از یک شبکه دو لایه پنهان استفاده نمود. تعداد نورون‌های لایه پنهان بین ۰/۴۰ تا ۰/۷۵ تعداد گره‌های ورودی انتخاب شده است. مشاهده گردید که با توجه به محدودیت تعداد نمونه‌های آموزش، هرچه بعد بردار ویژگی بزرگتر باشد، باید تعداد نورون‌های لایه پنهان نیز بیشتر باشد.

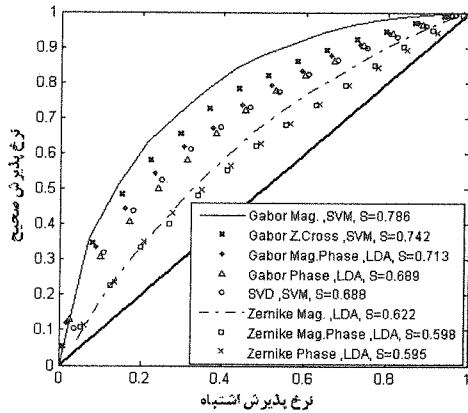
هر یک از روشهای استخراج ویژگی در کنار سه طبقه بندی کننده شبکه عصبی پس انتشارخطا، LDA و SVM مورد استفاده قرار گرفته است که به ازای هر روش استخراج ویژگی، بهترین نتیجه به دست آمده از این سه طبقه بندی کننده، انتخاب شده و در شکل‌های ۵ الی ۷ نشان داده شده است. راهنمای استفاده از شکل‌ها در جدول ۱ ارائه شده است.

هرچه منحنی ROC به گوشه بالا-چپ نزدیکتر باشد، سطح زیر منحنی بزرگتر و عملکرد طبقه بندی کننده بهتر است. سطح زیر منحنی‌های ROC بر روی شکل‌ها ذکر شده است.

محل تقاطع قطر فرعی مربع واحد با منحنی ROC تعیین کننده نقطه (EAR: Equal Acceptance Rate) است. در این نقطه دقت طبقه بندی دو کلاس مورد نظر با یکدیگر برابر است. بزرگی این مقدار نیز می‌تواند معیاری از عملکرد طبقه بندی کننده باشد.

جدول (۱): راهنمای کلمات مخفف استفاده شده در نتایج آتی .

P3	حالت درد با درجه شدید
P2	حالت درد با درجه متوسط
P1	حالت درد با درجه کم
S	سطح زیر منحنی ROC
SVD	اعمال SVD برای استخراج مولفه‌های اساسی
Mag.	استفاده از مقدار دامنه برای استخراج ویژگی
Phase	استفاده از مقدار فاز برای استخراج ویژگی
Gabor Z.Cross	استفاده از گذر از صفر تبدیل Gabor
Mag.Phase	استفاده همزمان از مقادیر فاز و اندازه برای استخراج ویژگی
NN	طبقه بندی کننده شبکه عصبی



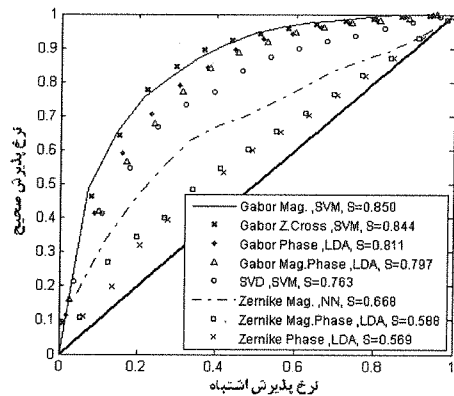
شکل (۷): بهترین نتایج برای طبقه بندی حالت درد کم. از هر نفر تقریباً دو تصویر برای هر حالت احساسی وجود دارد.

پایه‌سازی‌های قبل برای شکل ساده شده‌ای از بانک اطلاعاتی نیز مورد استفاده قرار گرفته است. در این حالت داده‌های بانک اطلاعاتی دستچین شده و با توجه به آنکه بعضی از افراد در ابراز بعضی از تظاهرات احساسی با مشکل روبرو بوده‌اند، از این‌رو حداکثر یک تصویر به ازای هر حالت احساسی از هر نفر وجود دارد. تمامی این شرایط معادل با حالتی است که از فرد تحت آزمایش، تصویری جهت آموزش طبقه بندی کننده وجود ندارد. با توجه به کاهش تعداد نمونه های بانک اطلاعاتی، در هر بار آموزش، از سه چهارم داده ها برای آموزش طبقه بندی کننده و از یک چهارم باقیمانده برای آزمون نتایج استفاده شده است.

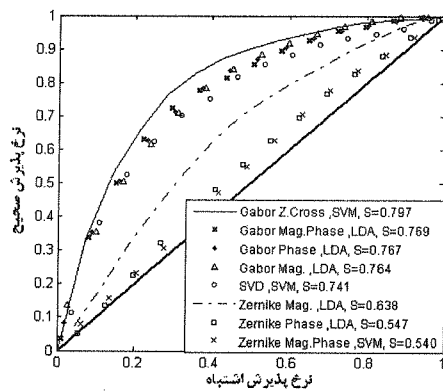
جدول (۲): بهترین نتایج بدست آمده برای طبقه بندی سه درجه درد. در هر خانه جدول، سطر اول درصد بدست آمده در نقطه EAR، سطر دوم روش مورد استفاده برای استخراج ویژگی و سطر سوم نوع طبقه بندی کننده را نشان می‌دهند.

	درد کم	درد متوسط	درد شدید
بانک اطلاعات اصلی پروژه	۷۱/۱۵ % Gabor Mag. SVM	۷۳/۷۱ % Gabor Z.Cross SVM	۷۷/۴۴ % Gabor Z.Cross SVM
بانک اطلاعاتی ساده شده	۵۹/۳۱ % Zenike Mag SVM	۶۱/۸۸ % Gabor Mag.Phase SVM	۷۵/۱۱ % Gabor Mag. SVM

بیشترین مقدار EAR بدست آمده برای این قسمت در کنار بهترین نتایج به دست آمده از مرحله قبل، در جدول ۲ ارائه شده است. قابل ذکر است که در نتایج به دست آمده از بانک اطلاعاتی ساده شده، درصد شناسایی درد کم برای روش‌های مختلف استخراج ویژگی بسیار نزدیک به یکدیگر می‌باشد.



شکل (۵): بهترین نتایج برای طبقه بندی حالت درد شدید. از هر نفر تقریباً دو تصویر برای هر حالت احساسی وجود دارد.



شکل (۶): بهترین نتایج برای طبقه بندی حالت درد متوسط. از هر نفر تقریباً دو تصویر برای هر حالت احساسی وجود دارد.

در کنار طبقه بندی کننده‌های LDA, SVM و شبکه‌های عصبی پس انتشار خطا، از سه روش کلی استخراج ویژگی شامل مولفه‌های اساسی، گشتاورهای زرنیکی و موجک گابور برای پیاده سازی یک سیستم تشخیص دهنده تظاهرات احساسی درد در چهره استفاده شده است.

نتایج به دست آمده نشان می‌دهند که جداسازی حالت درد کم (P1) نسبت به جداسازی دو حالت دیگر مشکل تر است. در صورتی که از فرد مورد آزمایش، داده آموزشی وجود نداشته باشد، دقت استخراج درد کم، بسیار کاهش می‌یابد.

ویژگی اندازه تبدیل گابور در اکثر موارد موفق ترین ویژگی میدان بوده و پس از آن با اختلافی کم، ویژگی گذر از صفر ضرایب گابور قرار گرفته است.

یکی از نکات قابل توجهی که در خلال آزمایش‌ها مشاهده شده، مربوط به عملکرد ویژگی فاز تبدیل گابور در کنار طبقه بندی کننده‌های LDA و SVM است. مشاهده شده که دقت نتایج حاصل از طبقه بندی کننده LDA در تمامی موارد بسیار بهتر از طبقه بندی کننده SVM می باشد این در حالی است که برای سایر ویژگی‌ها، نتایج حاصل از این دو طبقه بندی کننده به یکدیگر نزدیک می‌باشد.

با شدت گرفتن درد از حالت متوسط به شدید، توقع می‌رود که نتایج طبقه بندی بهتر شوند. اما چنین پدیده‌ای اتفاق نمی‌افتد. بررسی نتایج نشان می‌دهد که علت این امر نمایان شدن بیش از حد دندان‌ها در حالت درد شدید و شباهت این کلاس به کلاس خوشحالی است.

وجود شباهت‌های ناحیه‌ای بین تظاهرات احساسی مختلف باعث بروز خطا در نتایج می‌گردد. نمونه‌ای از این خطا نمایان شدن دندان‌ها در دو حالت درد شدید و همچنین خندیدن نسبتاً شدید است. دور شدن لب‌های بالا و پائین از یکدیگر و همچنین بالا رفتن ابروها نیز می‌توانند به عنوان یک تغییر حالت مشترک بین حالت درد و حالت تعجب تلقی گردند. جمع شدن لب‌ها در بعضی از تظاهرات درد، باعث تداخل بین کلاس عصبانیت و درد می‌گردد. در کنار واحدهای عملکردی مشترک بین حالت درد و سایر حالات احساسی، خطاهایی که در جایابی چهره و هم‌اندازه کردن تصاویر ورودی رخ می‌دهند نیز می‌توانند باعث خطای طبقه بندی کننده گردند. انطباق کامل تصاویر بر روی یکدیگر، اهمیت زیادی در کیفیت نتایج دارد. در صورتی که جایابی چهره به گونه‌ای باشد که اجزای اصلی چهره دقیقاً بر روی یکدیگر منطبق نگردند، خطای طبقه بندی ممکن است رخ دهد. با توجه به تنوع چهره افراد، این مورد تا حدی به وسیله

تنوع افراد بانک اطلاعاتی برطرف می‌شود ولی به صورت کامل رفع نمی‌گردد. نکته مهم دیگر، وجود چرخش‌های جزئی در تصاویر اخذ شده است. باید دقت شود که در هنگام بروز تظاهرات احساسی، چرخش‌هایی به صورت ناخودآگاه و در زوایایی مختلف در سر اتفاق می‌افتند، که این امر تنوع دادگان بانک اطلاعاتی را افزایش داده و همچنین باعث بروز کمی خطا در الگوریتم مورد استفاده برای جایابی می‌گردد. این مورد نیز می‌تواند در انطباق کامل تصاویر بر روی یکدیگر، تولید خطا کرده و در نتیجه باعث ایجاد خطا در نتایج خروجی گردد.

طبقه بندی کننده SVM با هسته‌های مختلف غیر خطی تحت آزمایش قرار گرفته و مشاهده شده که هسته خطی جواب بهتری نسبت به هسته‌های غیر خطی تولید می‌نماید.

یکی دیگر از ویژگی‌های مورد استفاده، گشتاورهای زرنیک می‌باشد که برای اولین بار در رابطه با طبقه بندی تظاهرات احساسی چهره به کار گرفته شده است. با وجود جواب‌های قابل توجهی که از این ویژگی در کاربرد تشخیص هویت به وسیله چهره گزارش شده است [۱۵]، عملکرد این روش در طبقه بندی تظاهرات احساسی چهره، پائین ترین امتیاز را در بین سایر روش‌های پیاده سازی شده به دست آورده است.

ویژگی گذر از صفر ضرایب گابور نیز نه تنها برای اولین بار است که در زمینه تظاهرات احساسی مورد استفاده قرار گرفته، بلکه در زمینه شناسایی ساختارهای تصاویر نیز برای اولین بار مورد استفاده قرار گرفته است. تنها کاربرد این روش در گذشته، مقایسه بافت تصاویر بوده است برای مثال می‌توان به کاربرد آن در شناسایی هویت به وسیله عنبیه و مقایسه بافت تصاویر هم اندازه عنبیه اشاره نمود [۶]. نتایج حاصل از گذر از صفر ضرایب گابور نشان می‌دهند که می‌توان از این روش برای استخراج ویژگی‌های نسبتاً قوی استفاده نمود در حالی که بردار ویژگی به شکل باینری دو بیتی بوده و پیچیدگی آن در مقابل بردارهای ویژگی که از مقادیر حقیقی تشکیل می‌شوند، بسیار پائین تر می‌باشد. این خصوصیت باعث افزایش سرعت محاسبات به حد بسیار بالایی می‌گردد.

طبقه بندی تظاهرات احساسی چهره از مباحث مهم در علوم پزشکی، روانشناسی و سایبرنتیک می‌باشد. از آنجا که حالت درد با تعدادی از تظاهرات اصلی چهره دارای مشترکاتی است، لذا از دیدگاه روانشناسی برای حالت درد یک طبقه مجزا اختصاص داده نشده است. طبقه بندی تظاهرات درد در چهره از دیدگاه پزشکی و مراقبتی می‌تواند بسیار مهم و سودمند باشد. از نگاه پزشکی می‌توان درد را در یک کلاس جداگانه فرض نموده و سایر حالات را در کلاسی دیگر دسته بندی کرد. نوآوری اصلی که در این تحقیق ارائه شده پیاده‌سازی این

Pub Co Inc, 2004.

Haykin, S.; *Neural networks: A comprehensive foundation, second ed.*, Prentice Hall, 1999. [۱۷]

Hong, H.; Neven, H.; Malsburg, C. v. d.; "Online facial expression recognition based on personalized gallery", Intl. Conference on Automatic Face and Gesture Recognition, Nara, Japan, pp. 354-359, IEEE Comp. Soc., April 14-16 1998. [۱۸]

Huang, L. L.; Shimizu, A.; Kobatake, H.; "Robust face detection using gabor filter features", Pattern Recognition Letters, vol. 26, no. 11, pp. 1641-1649, 2005. [۱۹]

Khotanzad, A.; Hong, Y. H.; "Invariant image recognition by zernike moments.", IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 5, pp. 489-497, 1990. [۲۰]

Kimura, S.; Yachida, M.; "Facial expression recognition and its degree estimation", IEEE Computer Vision and Pattern Recognition or CVPR, pp. 295-300, 1997. [۲۱]

Kobayashi, H.; Hara, F.; "Facial interaction between animated 3d face robot and human beings.", Int'l Conf. Systems, Man, Cybernetics, pp. 732-737, 1997. [۲۲]

Lyons, M. J.; Budynek, J.; Akamatsu, S.; "Automatic classification of single facial images", IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 12, pp. 1357-1362, 1999. [۲۳]

Ma, W.; Manjunath, B.; "Texture features for browsing and retrieval of image data", IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 8, pp. 837-842, 2006. [۲۴]

Monwar, M. M.; Rezaei, S.; "Appearance-based pain recognition from video sequences", IEEE International Joint Conference on Neural Networks-IJCNN, pp. 2429-2434, 2006. [۲۵]

Monwar, M. M.; Rezaei, S.; "Pain recognition using artificial neural network", IEEE International Symposium on Signal Processing and Information Technology, pp. 28-33, 2006. [۲۶]

Noruzi, M.; Vafadoost, M.; "Face detection by gabor transform", Accepted for publication in AKUJOURNAL, no. 67. [۲۷]

Otsuka, T.; Ohya, J.; "Recognition of facial expressions using hmm with continuous output probabilities.", Proc. Int'l Workshop Robot and Human Comm., pp. 323-328, 1996. [۲۸]

Pichler, O.; Teuner, A.; Hosticka, B. J.; "A comparison of texture feature extraction using adaptive gabor filtering, pyramidal and tree structured wavelet transforms", Pattern Recognition, vol. 29, no. 5, pp. 733-742, 1996. [۲۹]

Seyedarabi, H.; Aghagolzadeh, A.; Khanmohammadi, S.; "Facial expressions animation and lip tracking using facial characteristic points and deformable model", International Journal of Information Technology, vol. 1, no. 4, 2004. [۳۰]

دیدگاه متفاوت است در حالی که نشان داده شده می‌توان
تظاهرات احساسی درد را در سه شدت مختلف درجه بندی
نمود.

۸- مراجع

Belkasim, S.; Hassan, E.; Obeidi, T.; "Radial zernike moment invariants.", CIT, pp. 790-795, 2004. [۱]

Bradley, A. P.; "Use of the area under the roc curve in the evaluation of machine learning algorithms", Pattern Recognition, vol. 30, no. 7, pp. 1145-1159, 1997. [۲]

Brahnam, S.; Chuang, C. F.; Shih, F. Y.; Slack, M. R.; "Machine recognition and representation of neonatal facial displays of acute pain", Artificial Intelligence in Medicine, vol. 36, no. 3, pp. 211-222, 2006. [۳]

Clausi, D.; Jernigan, M.; "Designing gabor filters for optimal texture separability", Pattern Recognition, vol. 33, no. 11, pp. 1835-1849, November 2000. [۴]

Cohen, A.; *Biomedical Signal Processing*, CRC Press, Florida, third edition, 1998. [۵]

Daugman, J. G.; "High confidence visual recognition of persons by a test of statistical independence", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 11, pp. 1148-1161, 1993. [۶]

C-Williams, A. C. d.; "Facial expression of pain: An evolutionary account", Behavioral and Brain Sciences, vol. 25, no. 4, pp. 439-488, 2002. [۷]

Duda, R.; Hart, P.; *Pattern Classification and Scene Analysis*, Wiley Interscience, New York, USA, third edition, 1996. [۸]

Edwards, G. J.; Cootes, T. F.; Taylor, C. J.; "Face recognition using active appearance models", European Conf. Computer Vision, pp. 582-595, 1998. [۹]

Ekman, P.; Friesen, W.; *Facial Action Coding System (FACS): Manual*, Consulting Psychologists Press, Calif., 1978. [۱۰]

Fasel, B.; "Automatic facial expression analysis: A survey", Pattern Recognition, , no. 36, 2003. [۱۱]

Fausett, L., Ed.; *Fundamentals of neural networks: architectures, algorithms, and applications*, Prentice-Hall Inc., Upper Saddle River, NJ, USA, 1994. [۱۲]

Fawcett, T.; "Roc graphs: Notes and practical considerations for researchers", Tech. Rep. HPL-2003-4, HP Labs, 2003. [۱۳]

Grigorescu, S.; Petkov, N.; Kruizinga, P.; "Comparison of texture features based on gabor filters", International Conference on Image Analysis and Processing, pp. 142-147, 1999. [۱۴]

Haddadnia, J.; Faez, K.; Ahmadi, M.; "An efficient human face recognition system using pseudo zernike moment invariant and radial basis function neural network", IJPRAI, vol. 17, no. 1, pp. 41-62, 2003. [۱۵]

Hancock, P. J. B.; Frowd, C. D.; Niven, C.; "Recognition of pain expressions", Neural Computation and Psychology Workshop on Modelling of Language, Cognition and Action. pp. 339-348, World Scientific [۱۶]