



دانشکده مهندسی کامپیوتر

تشخیص خودکار حالت چهره توسط روش‌های یادگیری ماشین

پایان‌نامه برای دریافت درجه کارشناسی

در رشته مهندسی کامپیوتر - گرایش نرم‌افزار

یاسر سوری

استاد راهنما:

دکتر محسن سریانی

دکتر عادل ترکمان رحمانی

آبان ماه ۱۳۹۰



دانشکده مهندسی کامپیوتر

تشخیص خودکار حالت چهره توسط روش‌های یادگیری ماشین

پایان‌نامه برای دریافت درجه کارشناسی
در رشته مهندسی کامپیوتر گرایش نرم‌افزار

یاسر سوری

استاد راهنما:

دکتر محسن سریانی

دکتر عادل ترکمان رحمانی

آبان ماه ۱۳۹۰

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تأییدیه صحت و اصالت نتایج

باسمه تعالی

اینجانب یاسر سوری به شماره دانشجویی ۸۶۵۲۱۲۵۸ دانشجوی رشته کامپیوتر، گرایش نرم افزار در مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه‌ی نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه‌برداری شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. در صورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض در خصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسئولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذیصلاح (اعم از اداری و قضایی) به عهده اینجانب خواهد بود و دانشگاه هیچ‌گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی:

امضا و تاریخ:

تقدیم بہ ہمسرم

چکیده

در چند دهه‌ی اخیر تشخیص حالت چهره یکی از سریع‌ترین رشد‌ها را در بین شاخه‌های هوش مصنوعی داشته است. این رشد سریع به خاطر کاربردهای زیادی مثل بررسی حالات روحی، بازیابی تصاویر و واسطه‌های انسان و رایانه‌ی سرگرم‌کننده‌تر بوده است. در این پژوهش ما به بررسی ساختار کلی سامانه‌های تشخیص حالت چهره و معرفی روش‌های بهتر می‌پردازیم.

سپس سامانه‌ی تشخیص خودکار حالت چهره‌ی این پروژه معرفی خواهد شد. تصاویر ورودی ابتداً توسط یک تشخیص دهنده‌ی چهره پردازش می‌شوند. چهره‌های پیدا شده به ابعاد ثابتی تغییر اندازه داده می‌شوند، سپس سیستم رنگی آن‌ها به سیاه و سفید تغییر می‌کند و توسط مجموعه‌ای از فیلترهای گبور^۱ به فضای اندازه‌ی گبور^۲ برده می‌شوند. سپس توسط تحلیل مشخص کننده‌ی خطی^۳، تعدادی ویژگی انتخاب می‌شود و توسط ۴ ماشین بردار پشتیبان^۴، به صورت دودویی رده‌بندی می‌شود. پاسخ رده‌بندی در نهایت یکی از ۴ حالت: خوشحال، عصبی، شگفت‌زده و معمولی است. نتایج آزمایش نشان می‌دهد که به طور میانگین ۸۱/۲۵ درصد دقت توسط این سیستم بر روی مجموعه داده‌ها تولید شده بدست می‌آید. به خاطر عدم دسترسی به مجموعه داده‌های معروف، در این پژوهش مجموعه داده‌ای شامل ۳۲ عکس چهره از دو شخص تهیه شده است.

واژه‌های کلیدی: تشخیص حالت چهره - بینایی ماشین - یادگیری ماشین - پردازش تصویر

^۱ Gabor

^۲ Gabor Magnitude

^۳ Linear Discriminant Analysis

^۴ Support Vector Machine

فهرست مطالب

عنوان	صفحه
فصل ۱: مقدمه	۱
۱-۱- شرح مسأله	۲
۲-۱- انگیزه‌های پژوهش	۲
۳-۱- اهداف پژوهش	۲
۴-۱- ساختار پایان‌نامه	۳
فصل ۲: ساختار کلی سامانه‌های تشخیص حالت چهره	۴
۱-۲- مقدمه	۵
۲-۱-۱- دو رویکرد کلی برای تحلیل حالت چهره در تصاویر ثابت	۵
۲-۱-۲- روش‌های مبتنی بر هندسه و روش‌های مبتنی بر ظاهر	۶
۲-۲- ساختار کلی سامانه‌های تشخیص خودکار حالت چهره	۷
۲-۲-۲- تصویر ورودی	۷
۲-۲-۳- پیش پردازش	۸
۲-۲-۴- استخراج ویژگی	۸
۲-۲-۵- انتخاب ویژگی	۹
۲-۲-۶- رده بندی	۹
فصل ۳: طرح پیشنهادی	۱۰
۱-۳- مقدمه	۱۱
۲-۳- نرمال سازی تصاویر	۱۱
۳-۳- استخراج ویژگی	۱۲
۴-۳- انتخاب ویژگی	۱۳
۵-۳- رده بندی	۱۴
فصل ۴: ارزیابی طرح پیشنهادی	۱۵
۱-۴- مقدمه	۱۶
۲-۴- داده‌های آزمایشی	۱۶
۲-۲-۴- مجموعه داده‌های ساخته شده	۱۷
۳-۴- تشریح جزئیات ارزیابی	۱۸
۴-۴- نتایج ارزیابی	۱۸
۴-۴-۱- بررسی حالت‌های پاسخ اشتباه توسط سامانه	۱۹

۲۰	فصل ۵: نتیجه‌گیری و کارهای آینده
۲۱	۵-۱- نتیجه‌گیری
۲۱	۵-۲- کارهای آینده
۲۲	مراجع

فهرست شکل‌ها

عنوان	صفحه
شکل (۱-۲) ساختار کلی سامانه‌های تشخیص حالت چهره	۷
شکل (۱-۳) تصویر نمونه گرفته شده از دوربین - تصویر ورودی	۱۱
شکل (۲-۳) نمونه تصویر نرمال شده	۱۲
شکل (۳-۳) خروجی pca برای انتخاب دو ویژگی از بین ویژگی‌ها و رسم این ویژگی‌ها بر روی نمودار	۱۳
شکل (۴-۳) خروجی lda برای انتخاب دو ویژگی از بین ویژگی‌ها و رسم این ویژگی‌ها بر روی نمودار	۱۴
شکل (۱-۴) نمونه‌ای از تصاویر مجموعه داده‌ی ساخته شده	۱۶
شکل (۲-۴) تصویر حالت خندان که شگفت‌زده تشخیص داده شده است	۱۹
شکل (۳-۴) تصویر حالت معمولی که عصبانی تشخیص داده شده است	۱۹

فصل ۱:

مقدمه

۱-۱- شرح مسأله

حالت چهره به معنی بروز حالات روحی مثل خوشحالی یا ناراحتی در چهره است. حالات چهره را می‌توان در حالت کلی به ۷ حالت دسته بندی کرد [1]. این ۷ حالت عبارتند از: خوشحال، ناراحت، متعجب، عصبانی، ترسیده، متنفر و معمولی. در این پروژه هدف تشخیص این حالات کلی از روی تصویر چهره‌ی انسان است.

۱-۲- انگیزه‌های پژوهش

تشخیص خودکار حالت چهره مسئله ایست که جنبه‌های زیادی از جمله بینایی ماشین، یادگیری ماشین و پردازش تصویر دارد. برای حل این مسئله نیاز هست که در تمام این شاخه‌های مطالعه انجام شود. همین موضوع از اصلی‌ترین انگیزه‌های این پژوهش بوده است.

ساخت سیستمی که بتواند به صورت خودکار حالت چهره‌ی فرد را تشخیص دهد استفاده‌های بسیاری در علوم مختلف دارد. برای مثال در علوم پزشکی، فرودگاه‌ها و غیره برای تشخیص حالت روحی افراد از این روش‌ها استفاده می‌شود. همچنین چنین سیستمی می‌تواند به واسطه‌های ارتباط انسان و رایانه^۱ جنبه‌های اجتماعی ببخشد.

۱-۳- اهداف پژوهش

در این پژوهش هدف مطالعه‌ی کارهای علمی انجام شده برای حل این مسئله و همچنین پیاده‌سازی سیستمی است که بتواند تصویر چهره‌ی فرد را به یکی از حالت‌های خوشحال، عصبی، عادی و متعجب تبدیل کند است.

۱-۴- ساختار پایان نامه

در ادامه، در فصل ۲ ساختار کلی سامانه‌های تشخیص حالت چهره را بررسی می‌کنیم، سپس در فصل ۳ به بررسی طرح پیشنهادی این پایان‌نامه خواهیم پرداخت. در فصل ۴ نتایج بدست آمده از آزمایشات را بررسی خواهیم کرد.

فصل ۲:

ساختار کلی سامانه‌های تشخیص حالت

چهره

در مقالات علمی برای بررسی حالت چهره، دو رویکرد متفاوت در نظر گرفته شده است. هر کدام از این رویکردها روش‌های متفاوتی دارند. تقسیم چهره به واحدهای حرکتی متفاوت و یا در نظر گرفتن آن به صورت یک جزء برای پردازش‌های بعدی، مهم‌ترین تفاوت این رویکردهاست. در هر دوی این رویکردها، دو روش متفاوت به نام «مبتنی بر هندسه»^۱ و «مبتنی بر ظاهر»^۲ می‌تواند مورد استفاده قرار بگیرد. در ادامه به بررسی این دو روش خواهیم پرداخت. همچنین در این فصل به بررسی کارهای مرتبط انجام شده خواهیم پرداخت.

۲-۱-۱- دو رویکرد کلی برای تحلیل حالت چهره در تصاویر ثابت

دو رویکرد کلی را می‌تواند به صورت زیر خلاصه کرد:

۱. استفاده از کل تصویر تمام رخ چهره و پردازش آن و رده‌بندی^۳ این تصاویر در شش رده‌ی: تنفر، ترس، لذت، تعجب، ناراحتی و عصبانیت، کلیات این رویکرد را تشکیل می‌دهد. در این رویکرد فرض بر این است که هر کدام از حالت‌های چهره‌ی نام‌برده شده، دارای مشخصه‌های قابل مشاهده بر چهره هستند. اکمن^۴ و فرایزن^۵ [1] و ایزارد^۶ [3] این پیشنهادها را داده‌اند و بارتلت^۷ و لیتلورت^۸ [4] [5] از این روش برای ساخت سامانه‌های تشخیص کاملاً خودکار استفاده کرده‌اند [6].

۲. به جای استفاده از تصاویر چهره به صورت کلی، تقسیم کردن آن‌ها به زیر بخش‌هایی برای پردازش‌های بعدی، کلیات این رویکرد را نشان می‌دهد. به خاطر اینکه حالت‌های چهره بیش‌تر

^۱ Geometric-based

^۲ Appearance-based

^۳ Classification

^۴ Ekman

^۵ Friesen

^۶ Izard

^۷ Bartlett

^۸ Littlewort

به تغییرات اندک از ویژگی‌های جدا از هم چهره مثل چشم‌ها، ابروها و لب‌ها مرتبط است، این روش‌ها بهتر هستند. از این رویکرد همان «سامانه‌ی رمزگذاری حرکت چهره» است که در ابتدا توسط اکمن و فرایزن [6]، برای توصیف حالات چهره توسط ۴۴ واحد حرکتی بر چهره، توصیف شده است. بیش‌تر فعالیت‌های جدید علمی در این زمینه انجام می‌شود [6].

۲-۱-۲- روش‌های مبتنی بر هندسه و روش‌های مبتنی بر ظاهر

دو روش کلی برای حل مسئله‌ی تشخیص خودکار حالت چهره وجود دارد که در هر دو رویکردی که در بالا توضیح داده شد مورد استفاده قرار می‌گیرند.

۱. روش مبتنی بر هندسه یک روش قدیمی است که شامل دنبال کردن و پردازش حرکت بعضی نقاط خاص روی تصویر است.

۲. به جای دنبال کردن نقاط خاصی روی تصویر، رنگ بخش‌های خاصی برای تشکیل بردار ویژگی پردازش می‌شوند. ویژگی‌های مختلفی مثل گبور^۱ و هار^۲ به همراه روش‌های انتخاب ویژگی مثل pca ^۳، lda ^۴ و $adaboost$ ^۵ در این روش‌ها مورد استفاده قرار می‌گیرند.

^۱ Gabor

^۲ Haar

^۳ Principal Component Analysis

^۴ Linear Discriminant Analysis

^۵ Adaptive Boosting

۲-۲- ساختار کلی سامانه‌های تشخیص خودکار حالت چهره



شکل (۱-۲) ساختار کلی سامانه‌های تشخیص حالت چهره

سامانه‌های تشخیص خودکار حالت چهره معمولاً ساختار مشابهی دارند. در ادامه ساختار کلی این سامانه‌ها را بررسی خواهیم کرد.

۲-۲-۲- تصویر ورودی

معمولاً تصویر ورودی تصویر چهره‌ی یک انسان است که برای سادگی گاهی از دوربینی مثل دوربین رایانه‌های امروزی گرفته می‌شود. این تصویر معمولاً شامل یک چهره است. از این به بعد سامانه فقط با این تصویر کار دارد. برای موارد استفاده‌ی بی‌درنگ باید مراحل بعدی به حدی سریع باشد که حدود ۳۰ بار در

هر ثانیه این تصاویر ورودی پردازش شوند.

۲-۲-۳- پیش پردازش

تصویر ورودی می‌تواند بسیار تغییر بکند. برای مثال فاصله‌ی فرد از دوربین کم یا زیاد شود، جهت سر فرد نسبت به دوربین عوض شود یا اینکه رنگ تصویر عوض شود. برای خنثی کردن چنین تأثیراتی معمولاً یک مرحله پیش پردازش روی تصویر انجام می‌شود. در این مرحله معمولاً ابتدا چهره‌ی از داخل تصویر تشخیص داده می‌شود. سپس با توجه به محل چهره ویژگی‌های دیگر از جمله محل چشم‌ها پیدا می‌شوند. همچنین در این مرحله تأثیرات تابش نور، تأثیرات گردش سر و غیره حذف می‌شوند. خروجی این مرحله معمولاً تصویر چهره‌ی فرد است که اندازه‌ی آن نرمال شده و شامل اطلاعاتی مثل محل ویژگی‌های دیگر چهره مثل محل چشم‌هاست.

۲-۲-۴- استخراج ویژگی

هدف از مرحله‌ی استخراج ویژگی تبدیل اطلاعات نقاط^۱ تصویر به نمایش سطح بالاتری از شکل، حرکت، رنگ و بافت چهره است. این مرحله به صورت تبدیلی بر روی تصویر نرمال شده‌ی چهره است. از جمله تبدیلات مورد استفاده در این مرحله می‌توان به تبدیل هار و گبور [6] اشاره کرد. از مزایای اصلی تبدیل هار سرعت بسیار زیاد آن است. ولی در عمل در مقایسه با تبدیل فضای موجی گبور ضعیف‌تر است [8]. با این حال برای سامانه‌های بی‌درنگ از تبدیل هار هم استفاده می‌شود. عیب بزرگ تبدیل گبور سرعت کم آن و تولید ویژگی‌های زیاد آن است. سرعت کم گبور به خاطر محاسبات پیچیده‌ی آن است که بسته به تعداد فیلترهای مورد استفاده به صورت خطی افزایش پیدا می‌کند. خروجی تبدیل گبور مقادیر موهومی به ازای هر نقطه در تصویر و هر فیلتر است [9]. معمولاً مقدار اندازه‌ی این اعداد موهومی مورد استفاده قرار می‌گیرد.

۲-۲-۵- انتخاب ویژگی

بعد از تبدیلاتی که در مرحله‌ی قبل روی تصاویر انجام می‌شود، ابعاد فضای مسئله‌ی ما معمولاً زیاد است. برای مثال یک تصویر نرمال شده‌ی چهره با ابعاد 48×48 که توسط بانکی از ۸ جهت و ۹ فرکانس از فیلترهای گبور به نمایش مقدار گبور تبدیل شود شامل $9 \times 8 \times 48 \times 48 = 165888$ ویژگی است [10]. با این تعداد ابعاد از ویژگی یادگیری به خوبی انجام نمی‌شود و کارایی خوبی ندارد. ضمناً واضح است که این تعداد از ویژگی‌های نمی‌تواند واقعاً در رده‌بندی تأثیرگذار باشد. به همین خاطر توسط روش‌های کاهش ابعاد یا انتخاب ویژگی ابعاد مسئله کاهش داده می‌شود. از جمله روش‌های کاهش ابعاد می‌توان به lda ، pca و $adaboost$ اشاره کرد. تحقیقات انجام شده نشان می‌دهد که $adaboost$ بهتر از تمام روش‌ها برای انتخاب ویژگی عمل می‌کند [10].

۲-۲-۶- رده بندی

مرحله‌ی رده‌بندی مرحله‌ای است که از روی ویژگی‌های انتخاب شده از تصویر باید حالت چهره تشخیص داده شود. دو دسته‌ی کلی از رده‌ها در این مرحله مورد استفاده قرار می‌گیرند: واحدهای حرکتی [11] و حالت‌های چهره‌ی کلی که توسط اکمن [1] معرفی شدند.

۶ حالت چهره‌ی کلی عبارتند از خوشحال، ناراحت، متعجب، عصبانی، ترسیده و متنفر است. واحد حرکتی یکی از ۴۶ المان قابل رؤیت اصلی در چهره است. یک حالت چهره معمولاً حاصل ترکیب تعدادی از این واحدهای حرکتی است [1] [11]. واحدهای حرکتی در سامانه رمزگذاری حرکت چهره معرفی شده‌اند [6].

روش‌های معمول برای رده‌بندی عبارتند از svm ، $adaboost$ ، lda . این روش‌ها در [10] با هم مقایسه شده‌اند. نتایج نشان می‌دهد که اگر انتخاب ویژگی با $adaboost$ و رده‌بندی توسط svm با هسته‌ی خطی انجام شود بهترین نتایج از نظر دقت و سرعت حاصل می‌شود [10].

فصل ۳:

طرح پیشنهادی

۳-۱- مقدمه

با توجه به مطالعات انجام شده، زمان موجود و توانایی خود طرحی که در ادامه توضیح داده می‌شود را برای اجرا انتخاب کردم.

تصمیم بر این شد که چهار حالت کلی چهره شامل خوشحال، ناراحت، متعجب و عادی تشخیص داده شود. به طور مختصر در این طرح تصاویر ورودی از دوربین رایانه و تمام رخ هستند. در هر تصویر چهره‌ی فرد تشخیص داده می‌شود و تصویر چهره‌ی فرد به سیاه و سفید تبدیل می‌شود. سائز تصویر چهره هم در مرحله‌ی نرمال سازی به 128×128 تغییر داده می‌شود.

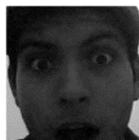
توسط بانکی از فیلترهای گبور شامل ۶ جهت و ۳ فرکانس تصویر به صورت مقادیر گبور نمایش داده می‌شود. سپس توسط روش lda دو ویژگی برای هر رده بند انتخاب می‌شود.

برای ۴ ویژگی که داریم تصمیم بر این شد که ۴ رده بند آموزش داده شود. هر رده بند باید تشخیص دهد که تصویر یکی از حالات را دارد یا خیر. در تصمیم نهایی بر اساس مقادیری که رده بندها اطمینان دارند می‌توان تشخیص داد که کدام یک از این چهار حالت انتخاب خواهد شد.

۳-۲- نرمال سازی تصاویر



شکل (۳-۱) تصویر نمونه گرفته شده از دوربین - تصویر ورودی



شکل (۲-۳) نمونه تصویر نرمال شده

برای هر تصویر ورودی ابتدا مکان چهره را توسط یک تشخیص دهنده‌ی چهره تشخیص می‌دهیم. برای این طرح از تشخیص دهنده‌ی تصویر ویولا^۱ و جونز^۲ استفاده شده است [12]. تصویر ورودی ابتدا ابعادش نصف می‌شود و رنگش به سیاه و سفید تغییر می‌کند. این مرحله به خاطر این است که تشخیص دهنده‌ی تصویر بر روی تصاویر کوچک‌تر سریع‌تر عمل می‌کند. سپس تصویر سیاه سفید و کوچک شده، توسط تشخیص دهنده‌ی چهره بررسی می‌شود تا مکان چهره مشخص شود. سپس قسمت مربوط به چهره که تشخیص داده شده است بریده شده و ابعادش به 128×128 تغییر داده می‌شود.

این مراحل بر روی تمام تصاویر آموزش و یا آزمایش انجام داده می‌شود.

۳-۳- استخراج ویژگی

همان‌گونه که توضیح داده شد، ویژگی‌ها توسط نمایش مقادیر گبور استخراج می‌شوند. این انتخاب به دلیل این است که فیلترهای گبور به صورت گسترده در تشخیص بافت مورد استفاده قرار گرفته‌اند [9]. همچنین خروجی آن‌ها بسیار به درک انسان نزدیک‌تر است. از فیلتر گبور پیاده‌سازی شده توسط [12] در این طرح استفاده شده است. بانک فیلترهای استفاده شده شامل ۶ جهت (۰، ۳۰، ۶۰، ۹۰، ۱۲۰ و ۱۵۰ درجه) و ۳ فرکانس (۲، ۴ و ۶) هستند. در مجموع ۱۸ فیلتر مورد استفاده قرار گرفته‌اند. در انتهای این مرحله برای هر تصویر ورودی، $128 \times 128 \times 18$ یا ۲۹۴۹۱۲ ویژگی استخراج شد.

^۱ Viola

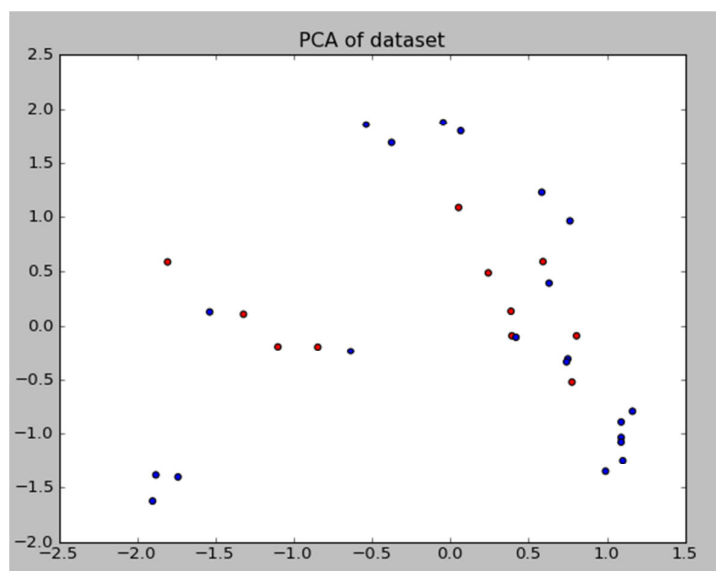
^۲ Jones

۳-۴- انتخاب ویژگی

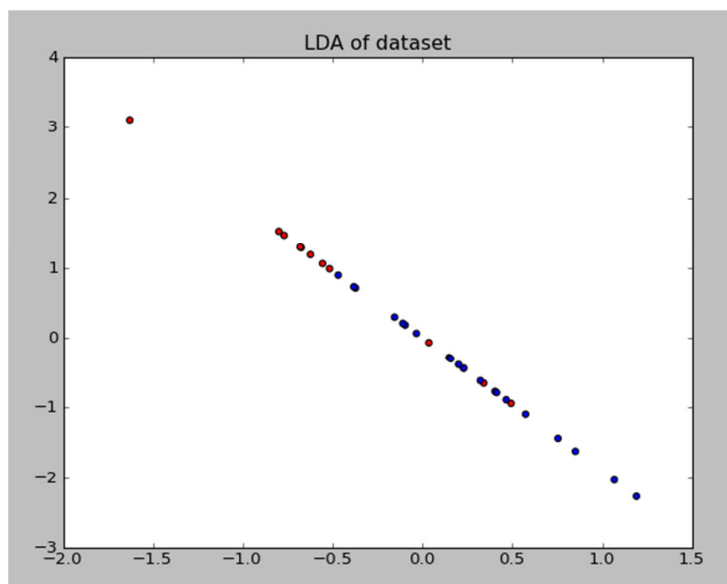
با توجه به تعداد زیاد ویژگی‌های استخراج شده در مرحله‌ی قبل لازم بود که تعداد این ویژگی‌ها کاهش یابد. زیرا این تعداد ویژگی قابل یادگیری نبودند. همچنین تعداد زیاد آن‌ها نشان از غیر مرتبط بودن آن‌ها داشت.

تصمیم گرفتم از بین pca و lda یکی را انتخاب کنم.

pca روشی است که از بین ویژگی‌ها آن‌هایی را انتخاب می‌کند که غیر خطی‌تر هستند نسبت به هم. در واقع pca سعی می‌کند که ابعاد کاذب را حذف بکند. lda عکس pca عمل می‌کند و آن ویژگی‌هایی را انتخاب می‌کند که نسبت به هم خطی‌تر هستند.



شکل (۳-۳) خروجی pca برای انتخاب دو ویژگی از بین ویژگی‌ها و رسم این ویژگی‌ها بر روی نمودار



شکل (۳-۴) خروجی lda برای انتخاب دو ویژگی از بین ویژگی‌ها و رسم این ویژگی‌ها بر روی نمودار

با توجه به نمایش داده‌های انتخاب شده تصمیم گرفتم که برای هر رده بند دو ویژگی را توسط lda انتخاب کنم. به نظر می‌رسد که خروجی lda دارای ویژگی‌های بهتری برای یادگیری باشد. آنچه در این تصاویر می‌بینید، مقادیر دو ویژگی انتخاب شده توسط روش مربوطه است که روی نمودار رسم شده‌اند. هر کدام از نقاط قرمز مربوط به یک عکس است که چهره‌ی خوشحال دارد و هر کدام از نقاط آبی مربوط به یک عکس است که حالتی غیر از خوشحال را دارد.

۳-۵- رده بندی

برای هر یکی از حالت‌های چهره یک رده بند آموزش داده شد. برای رده‌بندی هم از رده بند svm استفاده شد.

هسته‌ی رده بندها خطی با پارامتر خطای ۳ هستند.

فصل ۴:

ارزیابی طرح پیشنهادی

۴-۱- مقدمه

در این فصل به بررسی نتایج حاصل شده از آزمایشات توسط سیستم توصیف شده در فصل پیش خواهیم پرداخت.

با استفاده از داده‌هایی که در ادامه معرفی خواهند شد، این سامانه دقت متوسط ۸۱.۲۵ درصد را دارا می‌باشد.

۴-۲- داده‌های آزمایشی

برای ارزیابی سامانه‌های تشخیص حالت چهره معمولاً از مجموعه داده‌ی کوهن^۱ و کاناد^۲ [14] استفاده می‌شود [8] [10] [5] [6]. متأسفانه برای این پروژه مقدور نبود که از این مجموعه داده استفاده شود.^۳ به همین خاطر مجموعه داده‌ای شامل ۳۲ تصویر از دو فرد تهیه شد.



شکل (۴-۱) نمونه‌ای از تصاویر مجموعه داده‌ی ساخته شده

در این مجموعه از هر کدام از ۴ حالت: معمولی، خوشحال، متعجب و عصبی، برای هر کدام از افراد ۴ تصویر تهیه شد.

^۱ Cohn

^۲ Kanade

^۳ تقاضا برای دسترسی به این مجموعه داده، پر شد و ارسال شد ولی پاسخی دریافت نشد!

۴-۲-۲- مجموعه داده‌های ساخته شده

در ادامه تمامی مجموعه داده‌ها را خواهید دید.

عصبانی



خوشحال



معمولی





۴-۳- تشریح جزئیات ارزیابی

هر بار داده‌ها به دو بخش آموزشی و آزمایشی تقسیم می‌شد. به طوری که داده‌های آزمایشی ۸ عدد بودند. سپس، برای هر کدام از ۴ حالت کلی تصویر یک lda و یک svm با داده‌های آموزشی، آموزش داده شدند. سپس برای هر کدام از داده‌های آزمایشی، با استفاده از روش رده‌بندی دودویی [6] از بین svmهای آموزش داده شده، آن رده‌ای را که با اطمینان بیش‌تری انتخاب شده بود را انتخاب می‌کنیم.

۴-۴- نتایج ارزیابی

طی چند مرحله آزمایش بر روی ۸ عدد از تصاویر آزمایشی، به طور متوسط دقت ۸۱.۲۵ درصد به دست آمده است. انتخاب ۸ تصویر آزمایشی به این صورت بود که از هر کدام از اشخاص برای هر کدام از ۴ حالت چهره یک تصویر انتخاب شد.

در آزمایش بعدی مجموعه ۱۶ تصویر شامل تصاویر فقط یک شخص برای آموزش استفاده شد و تصاویر شخص دیگر برای آزمایش استفاده شد. این آزمایش با جا به جا کردن تصاویر آزمایشی و آموزشی بار دیگر تکرار شد. به طور میانگین دقت ۳۴/۷۲ درصد بود.

۴-۴-۱- بررسی حالت‌های پاسخ اشتباه توسط سامانه

برای نمونه دو مورد از پاسخ‌های اشتباه سامانه را بررسی خواهیم کرد:

- نمونه‌ی آزمایشی حالت خنده که توسط سامانه شگفت‌زده تشخیص داده شده است:



شکل (۴-۲) تصویر حالت خندان که شگفت‌زده تشخیص داده شده است

- نمونه‌ی آزمایشی حالت معمولی که عصبانی تشخیص داده شده است:



شکل (۴-۳) تصویر حالت معمولی که عصبانی تشخیص داده شده است

فصل ۵:

نتیجه‌گیری و کارهای آینده

۵-۱- نتیجه‌گیری

در این پژوهش ابتدا رویکردهای کلی «استفاده از کل تصویر چهره» و «تقسیم کردن تصویر چهره به زیر بخش‌های چهره» پرداختیم. سپس به بررسی روش‌های «مبتنی بر هندسه» و «مبتنی بر ظاهر» پرداختیم. بیش‌تر تحقیقات جدید با رویکرد تقسیم کردن چهره به زیر بخش‌ها و با روش‌های مبتنی بر ظاهر با استفاده از ابزارهایی مثل فیلترهای گبور و انتخاب کننده‌های ویژگی مثل adaboost است. ساختار کلی سامانه‌های تشخیص خودکار حالت چهره شامل قسمت‌های پیش پردازش، استخراج ویژگی، انتخاب ویژگی و رده‌بندی توضیح داده شد.

سپس سامانه‌ی تشخیص خودکار حالت چهره‌ی این پروژه معرفی شد. تصاویر ورودی ابتداً توسط یک تشخیص دهنده‌ی چهره پردازش می‌شوند. چهره‌های پیدا شده به ابعاد ثابتی تغییر اندازه داده می‌شوند، سپس سیستم رنگی آن‌ها به سیاه و سفید تغییر می‌کند و توسط مجموعه‌ای از فیلترهای گبور به فضای اندازه‌ی گبور برده می‌شوند. سپس توسط تحلیل مشخص کننده‌ی خطی، تعدادی ویژگی انتخاب می‌شود و توسط ۴ ماشین بردار پشتیبان، به صورت دودویی رده‌بندی می‌شود. پاسخ رده‌بندی در نهایت یکی از ۴ حالت: خوشحال، عصبی، شگفت‌زده و معمولی است. نتایج آزمایش نشان می‌دهد که به طور میانگین ۸۱/۲۵ درصد دقت توسط این سیستم بر روی مجموعه داده‌ها تولید شده بدست می‌آید. تمامی منابع، پیاده‌سازی‌ها و مجموعه تصاویر در [15] موجود و قابل دسترس همه است.

۵-۲- کارهای آینده

طرح پیشنهادی این پژوهش کاستی‌های زیادی داشت که در آینده می‌توان بر روی آن‌ها کار کرد. در مرحله‌ی پیش پردازش باید بتوان تأثیرات چرخش سر و تغییرات نورانی و ویژگی‌های مثل موهای چهره را حذف کرد.

همچنین به دلیل زمان‌گیر بودن مرحله‌ی تشخیص چهره و تبدیل گبور این سامانه به صورت بدون درنگ قابل استفاده نیست. ایجاد تغییراتی برای بدون درنگ کردن این سامانه می‌تواند از کارهای آینده باشد. ضمناً استفاده از ویژگی‌های بیش‌تر برای رده بندها و استفاده از انتخاب کننده‌های ویژگی بهتر مثل adaboost می‌تواند نتایج بهتری را به همراه داشته باشد.

مراجع

- [1] Ekman, P., *Emotion in the Human Face*. Cambridge University Press, 1982.
- [2] Ekman, P. and Friesen, W.V., *Pictures of Facial Affect*. Palo Alto, Calif.: Consulting Psychologist, 1976.
- [3] Izard, C.; Dougherty, L. and Hembree, E.A., "A System for Identifying Affect Expressions by Holistic Judgements," Univ. of Delaware unpublished manuscript, 1983.
- [4] Bartlett, M.; Hager, J.; Ekman, P. and Sejnowski T., "Measuring Facial Expressions by Computer Image Analysis," *Psychophysiology*, vol. 36, pp. 253-264, 1999.
- [5] Bartlett, M.S; Littlewort, G.; Lainscsek, C.; Fasel, I. and Movellan, J., "Machine Learning Methods for Fully Automatic Recognition of Facial Expressions and Facial Actions," in *IEEE International Conference on Systems, Men and Cybernetics*, 2004, pp. 592-597.
- [6] Cemre Z., "Facial Expression Recognition," Ms Dissertation, Department of Electronic Engineering, University of Surrey, Surrey, 2008.
- [7] Ekman, P. and Friesen, W.V., *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. San Francisco: Consulting Psychologists Press, 1978.
- [8] Fasel, B. and Luetttin, J., "Automatic facial expression analysis: a survey," *Elsevier Pattern Recognition*, 2003.
- [9] Manjunath, B.S. and Ma, W.Y., "Texture features for browsing and retrieval of image data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1996.
- [10] Bartlett, M.S.; Littlewort, G.; Frank, M.; Lainscsek, C.; Fasel, I. and Movellan, J., "Recognizing facial expression: machine learning and application to spontaneous behavior," *Computer Vision and Pattern Recognition*, 2005.
- [11] Donato, G.; Bartlett, M.S.; Hager, J.C. and Ekman, P., "Classifying Facial Actions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 974-989, 1999.
- [12] Viola, P. and Jones, M.J, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [13] Momma. (2011, Nov.) Gabor Filter. [Online]. <http://www.eml.ele.cst.nihon-u.ac.jp/~momma/wiki/wiki.cgi/OpenCV/Gabor%20Filter.html>
- [14] Kanade, T.; Cohn, J.F.; Tian, Yingli, "Comprehensive database for facial expression analysis," *IEEE Automatic Face and Gesture Recognition*, 2000.
- [15] Y. Sourì. (2011, Nov.) yassersouri/FER repository at github. [Online]. <http://github.com/yassersouri/FER>