یازدهمین کنگره ملی سراسری
فناوریهای نوین در حوزه توسعه پایدار ایران
11th National Congress of
the New Technologies in Sustainable Development of Iran
senaconf.ir

بازشناسی احساسات بر اساس تشخیص چهره

معصومه مشهدی زاده[۱]، رقیه نوروزیان[۲]، کبری باقری[۳]
[۱] گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران، masomehmashhadizadeh@gmail.com
[۲] گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران، norouzianroghayeh@gmail.com
[۳] گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران bagheri128@gmail.com

چکیده

احساسات در زندگی روزمره ما حضور بسیار پررنگی دارند. احساسات خود را از طریق واکنش‌های فیزیولوژیکی (تپش قلب، گرما و افزایش نبض)، واکنش‌های حرکتی (حالات ژست، بیانات صورت) و تغییر در صدا نشان می‌دهند. بیان صورت ابزار مهمی برای برقراری ارتباط احساسات یک فرد به صورت غیر کلامی و بازشناسی احساسات با استفاده از حالات چهره می‌باشد. پیچیدگی احساسات توجه محققان مختلفی را به خود جلب کرده است و تشخیص بیان چهره موضوع بسیاری از تحقیقات در سال‌های اخیر بوده است. در نتیجه بازنمایی‌های متفاوتی از احساسات به منظور ارائه مجموعه‌ای از نظریه‌های احساسات پیشنهاد شده است. تکنیک‌های مختلفی برای تشخیص چهره، استخراج ویژگی‌ها و طبقه‌بندی احساسات پیشنهاد شده است، زیرا تشخیص بر اساس ویژگی‌های بصری بیان متفاوت است. در این مقاله ابتدا به تعریف احساسات و سپس نظریه‌های که در این مورد پدید آمده‌اند و در ادامه به تشخیص چهره و مراحل آن پرداخته شده است.

واژه‌های کلیدی

بازشناسی احساسات، بیان صورت، تشخیص چهره، استخراج ویژگی، طبقه‌بندی احساسات

۱. مقدمه

در عصر توسعه سریع علم و فناوری، هوش مصنوعی به یکی از موضوعات اصلی تحقیقات تبدیل شده است [۱]. در نتیجه، تشخیص چهره در هوش مصنوعی توسط بسیاری از محققان مورد بررسی قرار گرفته است. در ارتباطات روزمره بین افراد معمولاً دو نوع ارتباط کلامی و غیر کلامی وجود دارند [۲]. مطالعات نشان داده است که حدود ۴۵ درصد از ارتباطات انسان کلامی است، در حالی که ۵۵ درصد باقی مانده غیرکلامی است که عمدتاً با احساسات صورت بیان می‌شود. بنابراین، حالت چهره یک وسیله غیرکلامی بسیار ضروری و مهم در ارتباطات بین‌فردی می‌باشد که عمدتا برای نشان دادن پیچیدگی افکار درونی متغیر افراد استفاده می‌شود. احساس به هفت دسته تقسیم می شود که عبارتند از: خنثی، عصبانی، تنفر، ترس، خوشحال، غمگین و متعجب. احساسات صورت سیگنال‌هایی را منتقل می کند که منعکس کننده هوشیاری و حالات ذهنی انسان است[۳]. داروین اولین کسی بود که ارتباط بین حالت چهره و احساس را در سال ۱۸۷۲ پیشنهاد کرد. بر اساس گفته‌های داروین، احساسات و بیان آن‌ها از نظر زیستی ذاتی می‌باشد و انطباق تکامل و شباهت به صورت فیلوژنتیکی می‌باشد [۴]. در سال ۱۹۷۲ اکمن، فریسن و السورث [۵] روی ایده داروین کار کردند و یافتند که هر حالت چهره فرهنگ خاص خود را دارد، مانند هر فرهنگی که زبان کلامی خاص خود را دارا می‌باشد، احساسات نیز زبان بیان چهره اختصاصی خود را دارند. مک کارتر و توکمین، در سال ۱۹۶۴ اولین مطالعه را انجام دادند و نشان دادند که حالات چهره با وضعیت عاطفی در ارتباط است [6]. مطالعات روانشناسی نشان داد که فرد می تواند احساسات و نگرش خود را ۷ درصد با صحبت کردن، ۳۸ درصد از طریق بیان صوتی و ۵۵ درصد از طریق حالات صورتش بیان کند [7-8].

از چند دهه گذشته، فناوری تشخیص چهره مورد توجه محققان به دلیل کاربردهای متعدد آن قرار گرفته است. تشخیص بیان چهره بخشی از تشخیص چهره است. روش‌های دیگر تشخیص چهره هم‌ترازی چهره، هم‌ترازی سر، تشخیص حالت چهره، احراز هویت چهره، سن و جنسیت یک شخص و غیره هستند. با توجه به دید انسان، تشخیص چهره یک فرد در هر شرایطی بسیار آسان است. اما برای دید کامپیوتری به دلیل تغییرات تصویر (به عنوان مثال جهت صورت، روشنایی تصویر، ژست، افزایش سن، انسداد، حالات چهره، مدل مو، آرایش و غیره) بسیار دشوار است [9]. از اوایل دهه ۹۰ میلادی، تشخیص خودکار حالت چهره موضوع مورد علاقه محققان بوده است. پیش از این تحقیقات کافی بر روی تشخیص چهره، ردیابی، استخراج ویژگی‌ها و تکنیک های طبقه‌بندی انجام شده بود. در این مقاله با استناد از مطالعات و پیشرفت‌های انجام شده در زمینه تشخیص چهره ، ویژگی‌های ایده آل سیستم مورد استفاده در بیان چهره بیان می شود (10)

چالش اصلی در زمینه تشخیص احساسات چهره، طبقه‌بندی خودکار تصاویر ثابت بر اساس هفت احساس اصلی می‌باشد. تجزیه و تحلیل و تشخیص خودکار حالات چهره موضوع بسیار مهم تحقیقات تشخیص احساسات گسسته و همچنین تشخیص واحد عمل FACS (AU) می‌باشد. پایگاه داده ای متعددی برای تشخیص بیان چهره بسته به سازگاری و استاندارد سازی وجود دارد. با این حال، مقایسه سیستم‌ها دشوار است اگر سیستم جزئیات کافی برای بازتولید جزئیات گزارش شده بر اساس پروتکل ارزیابی مشترک نداشته باشد، زیرا بر بهبود وضعیت تشخیص چهره تأثیر می‌گذارد. تجزیه و تحلیل و تشخیص حالت چهره این امکان را می‌دهد تا اهداف و چالش های جدید کشف شوند. این مقاله چالش‌هایی را در زمینه تشخیص خودکار چهره ارائه می‌دهد. این چالش‌ها شامل تشخیص AU و تشخیص احساسات گسسته هستند[11] .

در این مقاله پروتکل ارزیابی بیان می‌گردد و نتایج این دو چالش فرعی نیز مورد بحث قرار می‌گیرد [12]. برای تجزیه و تحلیل چهره و تشخیص کارآمد چهره، یک توصیفگر ویژگی محلی و الگوی عددی جهت‌دار محلی (LDN) به عنوان یک تکنیک جدید برای تشخیص حالت چهره پیشنهاد شده است. داده‌های جهت دار بافت صورت از طریق الگوی عددی جهت ار محلی به صورت خیلی فشرده کدگذاری می‌شوند. این روش در مقایسه با رویکردهای موجود، کد انتخابی تری ارائه می‌دهد. علاوه بر این، ساختار هر ریز الگو با کمک ماسک قطب‌نما محاسبه می‌شود که مسئول استخراج اطلاعات جهت‌دار

است. چنین داده‌هایی با استفاده از اعداد جهت‌دار کدگذاری می‌شوند و نشانه‌هایی که موجب تمایز الگوهای یکسان با شدت انتقال متفاوت می‌شوند. صورت به مناطق مختلفی به منظور استخراج ویژگی‌های عددی جهت‌دار محلی توزیع شده تقسیم می‌شود. این ویژگی‌ها بیشتر به عنوان بردارهای ویژگی به هم پیوسته و به عنوان یک توصیفگر چهره استفاده می‌شوند. آزمایش‌های متعددی انجام می‌شود تا توصیفگر بتواند تحت هر شرایطی آزمایشی مانند تغییرات در بیان، گذشت زمان و نویز به خوبی عمل کند. علاوه بر این، عملکرد این توصیفگرها با ماسک‌های مختلف تحت شرایط مختلف تحلیل چهره تحلیل می‌شوند. [13].

۲. بازشناسی احساسات

بازشناسی احساسات مطالعه شناخت هفت احساس جهانی (خشم، انزجار، لذت، ترس، شادی، غم و تعجب) می‌باشد. احساسات وضعیت ذهنی فرد را از طریق اعمال غیرعمدی او منعکس می نند که ممکن است فرازبانی یا غیر فرازبانی باشد. بنابراین، عواطف یک فرد با استفاده از ویژگی‌های رفتاری مانند صدا، دست خط، بیان صورت، سیگنال های مغزی، سیگنال-های قلبی و غیره شناسایی می‌شود. ویژگی‌های رفتاری نه تنها برای شناسایی یک شخص استفاده می‌گردد بلکه به شناخت احساسات نیز کمک می‌کند، بنابراین به عنوان بیومتری‌های نرم شناخته می‌شوند [14]. بیومتری‌های نرم به صورت صفات فیزیکی، صفات رفتاری و عادات انسانی طبقه‌بندی می‌شوند. قد، وزن، رنگ پوست، رنگ چشم، صفات فیزیکی هستند. صدا، طرز راه رفتن صفات، رفتاری هستند و سبک پوشیدن، استفاده از زیورآلات عادات انسانی هستند. علاوه بر بازشناسی احساسات، عوامل مختلف دیگری مانند ظرفیت، قطبیت و برانگیختگی نقش مهمی در تشخیص وضعیت ذهنی یک فرد ایفا می‌کنند [15]، نقشه‌برداری از مغز با استفاده از ظرفیت، قطبیت، برانگیختگی و بازشناسی احساسات، تحلیل احساسی نامیده می‌شود. تحلیل احساسی درک نظر و نگرش فرد نسبت به موضوع خاص با استفاده از رویکردهای محاسباتی مختلفی انجام می‌شود. بازشناسی احساسات در بسیاری از زمینه‌ها از جمله تعامل انسان با کامپیوتر، امنیت بیومتری و غیره کاربرد دارد و بینشی در مورد هوش مصنوعی یا هوش ماشینی ارائه م دهد که از انواع مختلف الگوریتم‌های یادگیری ماشینی تحت نظارت و بدون نظارت برای شبیه سازی مغز انسان استفاده می کند. مطالعه احساسات انسان، تفسیر، پردازش و انطباق آنها توسط ماشین ها به عنوان محاسبات عاطفی یا هوش مصنوعی عاطفی شناخته می‌شود [16]. حالت عاطفی انسان از حالات چهره، حرکات بدن، گفتار، نگارش، سیگنال‌های مغز یا قلب و غیره شناسایی می‌گردد. با استفاده از تکنیک های مختلف یادگیری ماشینی ویژگی‌ها یا الگوهای مورد نیاز از داده‌های جمع‌آوری شده استخراج می‌شود.

۳. طبقه بندی بیان صورت مبتنی بر احساسات

پیشرفت‌های اخیر در زمینه تشخیص بیان بر تشخیص بیان طبیعی‌تر چهره تأکید دارد. به طور کلی هفت دسته احساسات جهانی وجود دارد که از روی حالات چهره قضاوت می‌شوند و عبارتند از بیان خشم، ابراز انزجار، ابراز ترس، بیانات خنثی، احساسات شاد، احساسات غم انگیز و احساسات غافلگیر کننده [17]. پایگاه داده‌های AFEW و SFEW بیانات طبیعی-تری از این هفت دسته احساس را ثبت کرده اند. پایگاه داده AFEW حاوی تصاویر ثابت است در حالی که پایگاه داده SFEW-dataset حاوی ویدئوهایی از فیلم است. دلیل استفاده از صحنه‌های فیلم این است که احساسات در فیلم‌ها در مقایسه با بیانات تقلیدی پایگاه داده کنترل‌شده آزمایشگاهی، خودانگیخته‌تر هستند. در این مقاله یک الگوریتم تشخیص بیان جدید ارائه شده است که ساده‌تر و سریعتر از سایر الگوریتم‌های موجود می‌باشد[18].

هدف اصلی این الگوریتم، در دسترس قرار دادن آگاهی احساسی در برنامه‌های مبتنی بر زمان واقعی در حال اجرا در سطح ثانویه است. الگوریتم پیشنهادی در سه مرحله پیش پردازش، استخراج ویژگی و یادگیری ماشین انجام می‌شود. این الگوریتم استاندارد بر اساس استخراج ۱۹ علامت مشخصه چهره است که برای تعریف حرکات اصلی صورت استفاده می‌شود و همچنین برخی از AU های تعریف شده در FACS را تشخیص می‌دهد. علاوه بر این، مفهوم جدیدی از CAU (Combined Action Units)معرفی شده است. CAU به معنای تعداد واحدهای اقدام (AUs) است که با همدیگر به صورت یک واحد جداگانه گروه‌بندی شده اند. در اینجا طبقه‌بندی احساسات بر اساس قواعد منطقی انجام می‌شود. از این رو هیچ یادگیری یا آموزشی لازم نیست. پیاده سازی‌های اولیه بر روی پلتفرم موبایل انجام می‌شود. تشخیص‌های AU و CAU بر اساس یک الگوی خاص حاصل از مقایسه تصویر واقعی و خنثی به دست آمده است. این الگو بر اساس یک گروه از نشانه‌های چهره است. AUs و CAU بر اساس سه عامل یعنی فاصله بین نقاط، بردارهای حرکتی و مساحت محصور در مثلث‌ها تعریف می‌شود. سپس این سه عامل با مقادیر آستانه از پیش تعریف شده مقایسه می‌شوند که فعال‌سازی AU یا CAU را توجیه می‌کند [19] .

برای تشخیص دقیق احساسات، سیستم آموزش دیده است. این آموزش در سه مرحله ۱) یادگیری ویژگی، ۲) انتخاب ویژگی و ۳) مدل‌سازی طبقه بند (classifier) انجام می‌شود. مطالعات اکتشافی بسیاری باید برای جستجوی ترکیبی بهینه از ویژگی‌ها، مجموعه ویژگی‌ها و طبقه‌بندی‌کننده به منظور دستیابی به بهترین عملکرد انجام شود. روش جدید شبکه باور عمیق برای آموزش معماری حلقه یکپارچه در سه مرحله پیشنهاد شده است. شبکه باور عمیق پیشنهادی به یادگیری مجموعه خاصی از ویژگی‌ها کمک می‌کند که بیانات را بر اساس ظاهر یا شکل دسته‌بندی می‌کند. علاوه بر این، این آموزش به تشکیل طبقه بند باور قوی به روش آماری کمک می‌کند. با جلسات مستمر آموزش و یادگیری، طبقه‌بندی-کننده تقویت می‌شود و همچنین قابلیت تمایز بین بیانات مختلف نیز با طبقه‌بندی‌کننده کارآمد از طریق یک فرآیند دقیق در چارچوب شبکه باور عمیق بهبود می‌یابد [20].

۴.تشخیص احساسات بر اساس حالات چهره

در فرآیند ارتباط انسان‌ها با همدیگر نوسانات احساسات مختلف وجود دارد. زمانی که افراد احساسات اولیه را تجربه می‌کنند، در چهره آن‌ها الگوهای بیانی مختلفی ظاهر می‌شود که هر کدام مجموعه‌ای از ویژگی‌ها و مقیاس توزیع خاص خود را دارند. تشخیص حالت چهره بخش مهمی از تعامل انسان و رایانه است که به رایانه‌ها اجازه می دهد تا حالات چهره را بر اساس تفکر انسان درک کنند. فرآیند پردازش تشخیص چهره به سه فرآیند مهم یعنی تشخیص چهره، استخراج ویژگی و مدول طبقه‌بندی تقسیم می شود، فرآیند تشخیص چهره فناوری کلیدی در تشخیص چهره به شمار می‌رود [21-22] و پیشرفت سریعی داشته است و از تصویر چهره اصلی، ویژگی‌های عالی استخراج می‌شود و ویژگی‌های صحیح طبقه‌بندی از عوامل تاثیرگذار و مهم بر نتیجه تشخیص هستند. به عنوان مثال، گائو و ما (۲۰۲۰) [23]ویژگی‌های حالت چهره را از تصاویر چهره به دست آوردند تا با توجه به تغییرات حالت چهره، حالات عاطفی را پیش‌بینی کنند.

۵. نظریه‌های تشخیص حالت چهره

۵. ۱. تشخیص حالت چهره با استفاده از رویکرد مبتنی بر ویژگی‌های هندسی

مرحله مهم در رویکرد تشخیص چهره مبتنی بر ویژگی‌های هندسی، مکان‌یابی و ردیابی نقاط چهره با موقعیت متراکم می‌باشد. برای ردیابی نقاط متراکم چهره، تقریباً همه رویکردهای مبتنی بر ویژگی های هندسی از مدل ظاهری اکتیو و

تغییرات آن استفاده می‌کنند. استخراج ویژگی های حرکتی صورت و شکل اجزای صورت با کمک تنوع بیان با استفاده از مکان ویژگی به روش‌های مختلف انجام می شود. اخیر به منظور تعیین احساسات چهره یک رویکرد مبتنی بر مدل ظاهری اکتیو پیشنهاد شده است که الگوریتم‌های مختلف مدل ظاهری اکتیو را با هم مقایسه کرده و عملکرد آنها را ارزیابی می‌کند. تشخیص چهره از ویژگی‌های هندسی مانند ویژگی‌های چهره، رنگ پوست، زمینه پوست و ویژگی‌های هندسی چند گانه استفاده می‌کند [24]. از این ویژگی‌های هندسی برای تشخیص احساسات استفاده می‌شود. این روش از تشخیص دستی نقاط مشخصه صورت و ردیابی تغییر شکل حجمی قطعه قطعه بزیر برای ردیابی دستی نقاط چهره استفاده می شود. تکنیک‌های یادگیری ماشین متعددی مورد مطالعه و آزمایش قرار گرفته است و روش ساده نزدیکترین همسایه k بهترین نتایج را ارائه می‌دهد [25]. در مطالعه دیگری از مدل ظاهری اکتیو استریو به منظور بهبود ردیابی و برازش مدل ظاهری اکتیو استاندارد با کمک دوربین‌های متعدد به منظور مدل سازی شکل سه بعدی و پارامترهای حرکت صلب استفاده گردید. علاوه بر این، برای ترکیب شکل سه بعدی و ظاهر دو بعدی ثبت شده، از طبقه بند تحلیل تفکیک لایه عمومی استفاده گردید.

۵. ۲. تشخیص حالات چهره با استفاده از رویکرد مبتنی بر ظاهر

از ویژگی های مبتنی بر الگوهای ظاهری یک فرد در درک احساسات استفاده می شود. نمونه جهت نزدیک، الگوهای باینری محلی گابور، تحلیل تفکیک خطی، داده‌های بافت فیلتر گابور، تجزیه و تحلیل مؤلفه های اصلی، هیستوگرام جهت گرا و فاکتور سازی شبکه غیر منفی و غیره مواردی از الگوهای ظاهری هستند. توزیع، ماشین بردار پشتیبانی، شبکه عصبی، مدل پنهان مارکوف، طبقه بند ساده لوح بیز و ویژگی‌های مبتنی بر رویکرد نظری اطلاعات برخی از ویژگی های مبتنی بر ظاهر هستند[26].

نظریه فاکتورسازی شبکه غیر منفی مبتنی از رویکردهای امیدوارکننده به نظر می رسد. در این روش تشخیص چهره، از روش استخراج اطلاعات بافت موجک گابور برای تجزیه و تحلیل تاثیر انسداد جزئی بر حالت چهره استفاده می‌شود. در این روش، سایر الگوریتم های مورد استفاده روش مبتنی بر شکل و روش تجزیه تصویر تحت نظارت تفکیک فاکتورسازی شبکه غیر منفی می‌باشد. یکی دیگر از تکنیک‌های تشخیص چهره، فاکتورسازی شبکه غیر منفی نمودار پراکنده است که یک رویکرد کاهش ابعاد اثرات انسداد است. تصاویر چهره با ابعاد بالا را به یک زیرفضا با حفظ موقعیت تبدیل می کند. بسیاری از متخصصان از عملگر الگوی باینری محلی برای تجزیه و تحلیل و تشخیص بیان چهره استفاده می-کنند. نظریه‌ای در مورد تشخیص احساسات بر اساس عملگر الگوی باینری محلی ارائه شده است که از طبقه‌بند ماشین بردار پشتیبانی با ویژگی‌های الگوی باینری محلی تقویت‌شده با هدف دستیابی به بالاترین دقت در بیان چهره استفاده می‌کند. از ویژگی های مبتنی بر ظاهر نیز برای تشخیص حالت چهره استفاده می شود که قادر به تشخیص خودکار حالات چهره از ویدیوها می‌باشد. روش های مختلف یادگیری ماشین بررسی شده است. در این مقاله روش زیر مجموعه فیلترهای گابور با استفاده از AdaBoost و سپس آموزش ماشین بردار پشتیبانی بسته به خروجی فیلترهای انتخابی AdaBoost بهترین نتایج را ارائه می دهند [27-28] .

۶. مراحل تشخیص بیان چهره

روش تشخیص حالات چهره برای تصاویر دیجیتال، سیستم مرتبط و کاربردهای آن در اینجا ارائه شده است. این روش تجزیه و تحلیل تصویر دیجیتال به تعیین تمایز بین لبخند یا پلک زدن روی چهره یک شخص کمک می کند. طبقه بندی این تصاویر دیجیتال در تشخیص چهره یا تعیین موقعیت امکان استفاده از طبقه بند نسبی بسته به کاربردهای اختصاصی

را فراهم می کند. مراحل تشخیص چهره عبارتند از دریافت تصویر، پیش پردازش استخراج ویژگی، طبقه بندی و پس-پردازش [11] .

۶. ۱. دریافت تصویر

دستگاه حسگر مانند دوربین وظیفه ثبت چهره یک سوژه و تبدیل آن به شکل دیجیتال با کمک مبدل آنالوگ به دیجیتال را به عهده دارد. دریافت چهره اولین قدم پردازش تصویر چهره می‌باشد. هدف اصلی این مرحله تشخیص خودکار بخشی از صورت با استفاده از تصاویر یا فیلم می‌باشد. این تصویر از آشکارساز عبور می‌کند و هر گونه قالب چهره را تشخیص می‌دهد. به منظور شناسایی دقیق ناحیه صورت از روی صحنه و تشخیص لحظه‌ای، رویکردی جدید پیشنهاد شد که اصلاح شده اولین روش چارچوب تشخیص چهره یعنی "Viola-Jones" می‌باشد. این چارچوب تشخیص چهره «Viola-Jones» حاوی طبقه بندهای آماده است که توسط AdaBoost انجام می شود. علاوه بر این طبقه بند از فیلترهای تصویر یکپارچه مشابه توابع Haar Basis استفاده می کند و در هر مکان و مقیاس با سرعت بالا محاسبه می شود. ویژگی های تابع Haar Basis مستطیلی، دو لبه و خطی می باشد. در مرحله آبشاری، از "AdaBoost" بر اساس روش انتخاب ویژگی برای انتخاب مجموعه ویژگی استفاده می شود و به تشخیص چشم‌ها، لب‌ها و بینی کمک می‌کند و دو جزء دیگر یعنی ابروها و گوی چشم افزوده شده است. تمام اجزای فریم که با موفقیت شناسایی شده اند از طریق جعبه مرزی متمایز می شوند. علاوه بر این، نقطه مرکزی هر جعبه مرزی مربوط به اجزای صورت شناسایی شده است. [29].

۶. ۲. پیش‌پردازش چهره

مرحله دوم تشخیص چهره، پیش پردازش تصویر است. در این مرحله، تبدیل مقیاس خاکستری تصویر گرفته شده انجام می شود و حتی اندازه تصویر نیز کاهش می یابد [30]. فشرده سازی داده های نامربوط انجام می شود، در نتیجه ویژگی‌های مهم تصویر بهبود می یابد. هدف این مرحله دستیابی به تصاویر چهره غیر تقلبی با قدرت استاندارد و اندازه و شکل یکنواخت می‌باشد. به منظور دستیابی به عملکرد تشخیص با کیفیت بالا در زمینه بازشناسی احساسات، پیش پردازش چهره یک مرحله مهم است که باید دنبال شود. این مرحله به حذف نویزهای نامربوط کمک می کند و همه چهره‌ها دارای یک دامنه مشترک می شوند. از آنجایی که کارشناسان تصمیم گرفتند مدل شبکه عمیق را در FER قبل از تشخیص چهره آموزش دهند، اندازه چهره‌های مجموعه داده SFEW به اندازه ۴۸×۴۸ تغییر می‌کند و به مقیاس خاکستری به منظور تطبیق با داده های FER تبدیل می شود. علاوه بر این، اکولایزر هیستوگرام استاندارد برای استفاده از پیش پردازش تصاویر صورت حاصل از پایگاه داده‌های FER و SFEW حفظ می‌شود و سپس برازش صفحه خطی انجام می شود که نور نامتعادل را از بین می‌برد. در مرحله نهایی پس از برازش صفحه، عادی سازی پیکسل‌های تصویر تا رسیدن به مقدار میانگین صفر و بردار واریانس واحد انجام می‌شود [31].

۶. ۳. استخراج ویژگی

استخراج ویژگی سومین مرحله از تشخیص چهره است که ویژگی های مورد نیاز تصویر (چهره) را استخراج می کند. استخراج ویژگی متشکل از مراحل - ۱) کاهش ابعاد ۲) استخراج ویژگی و ۳) انتخاب ویژگی تشکیل می باشد. از این سه مرحله، کاهش بعد نقش محوری در تشخیص الگوی چهره/تصویر ایفا می کند. از SIFT برای تعیین مقدار نقاط کلیدی استفاده می کند.. [30] [31].

الف) تکنیک استخراج ویژگی مبتنی بر تبدیل کسینوس گسسته

در این روش تصویر ورودی برای تبدیل کسینوس گسسته اعمال می شود. ضرایب تبدیل کسینوس گسسته با روش زیگزاگی مرتب شده اند و به منظور اجرای تبدیل ماتریس تصویر دو بعدی تبدیل کسینوس گسسته به بردارهای ویژگی پیاده سازی شده اند. اجزای حوزه فرکانس در نقاط اولیه بردار قرار می گیرند. ویژگی های تصویری صورت مانند چشم‌ها، لب‌ها و بینی از تصویری که به‌عنوان ورودی تغذیه می‌شود، استخراج می‌شوند که در روش تبدیل کسینوس گسسته تصویر فرعی بیشتر مورد استفاده قرار می گیرد. در نهایت، ضرایب تبدیل کسینوس گسسته برای چشم، لب و بینی به دست آمده و با بردارهای ویژگی ترکیب می شود. طبقه‌بندی «AdaBoost» برای تشخیص بیشتر احساسات استفاده می‌شود که به طور کلی نرخ شناسایی ۷۵/۹۴ درصد به دست می‌آید[32].

ب) تکنیک استخراج ویژگی مبتنی بر فیلتر گابور

در این روش ابتدا چهره ها در تصاویر تحت پیش پردازش مبتنی بر تبدیل افین به منظور عادی سازی تصاویر قرار می گیرند. ارزیابی تصاویر پیش پردازشی و نرمال با استفاده از فیلترهای مختلف گابور انجام می شود که برای تمایز بین احساسات مختلف استفاده می شود. تکنیک های کاهش بعد مورد استفاده برای تشخیص امکان سنجی، دو نوع متا طبقه بند یعنی تحلیل تشخیصی خطی و تحلیل مؤلفه‌های اصلی هستند. طبقه بند تابع تشخیصی را مانند تابع خطی داده های ویژگی فرض می کند. ویژگی بردار حاصل شامل داده های این سناریو است. به گفته نویسنده، انتخاب فیلتر گابور پیشنهادی به کاهش پیچیدگی محاسبات و همچنین بعد فضای ویژگی کمک می کند [33].

پ) تکنیک استخراج ویژگی مبتنی بر تحلیل مولفه اصلی

از روش، WPCA (Weighted PCA) [34] برای ادغام چندین ویژگی استفاده می شود. این ویژگی های مکانی خود مرتبط با ابعاد بالا از تصاویر چهره استخراج می شوند و در ادامه به مناطق مختلف تقسیم می شوند. در مرحله بعد، Weighted PCA برای کاهش ابعاد ویژگی ها اعمال می شود. وزن ها بر روی پایه سریعتر با کمک سیستم کدگذاری عملکرد چهره تشخیص داده می شوند که در آن از تابع پایه شعاعی نیز استفاده شده است. احساسات و ویژگی های چهره با استفاده از الگوریتم ماشین بردار پشتیبانی طبقه بندی می شوند. در نهایت برای دستیابی به شباهت بین الگوها، فاصله اقلیدسی تخمین زده می شود و سپس تشخیص حالت چهره با استفاده از نزدیکترین الگوریتم انجام می شود. در این تکنیک نرخ تشخیص ۸۸/۲۵ درصد می باشد.

ت) تکنیک استخراج ویژگی مبتنی بر تحلیل اجزای مستقل

در این روش [30]، مؤلفه‌های تحلیل مؤلفه‌های مستقل از قانون انتقال داده‌های ایده‌آل از طریق نورون های سینوسی حرکت می کنند. تحلیل مؤلفه‌های مستقل بر روی تصاویر چهره حاصل از پایگاه داده FERET بر اساس دو مدل معماری اعمال می شود. مدل اول تصویر را به عنوان یک متغیر تصادفی و پیکسل را به عنوان یک آزمایش تصادفی فرض می کند. از این رو تحلیل مؤلفه‌های مستقل تصاویر را از نظر آماری تصاویر مستقل در نظر می گیرد. ویژگی های چهره شبیه فضایی این تصاویر پراکنده و محلی است. مدل دوم پیکسل را به صورت متغیر تصادفی و چهره را به صورت آزمایش تصادفی فرض می کند. ضرایب تصویر مستقل هستند و از این رو کد فاکتوریل چهره را نشان می دهد.

۶. ۴ . طبقه بندی

این آخرین مرحله یا یکی مانده به آخرین مرحله در پردازش تشخیص چهره بسته به نوع کاربرد و نیاز سیستم است. طبقه بندی بر روی ویژگی های حاصل از نتیجه استخراج ویژگی انجام می شود. در ادامه طبقه بندهای مختلفی برای بازشناسی دقیق چهره بسته به ویژگی های بهینه، کاربرد و آموزش سیستم ارائه شده است [30] [31] .

الف) مدل پنهان مارکوف

این مدل برای طبقه بندی بیان با سطح بالاتر پیشنهاد شده است. که برای تشخیص همراه با طبقه بند عادی، ضروری هستند. این بیانات سطح بالاتر عبارتند از دلسرد کننده، تشویق کننده، مخالف، علاقه مند و نامطمئن در حالی که سطح پایین تر بیانات عبارتند از شادی، تعجب، غم و بیانات خنثی. یک مدل نمایه سازی عاطفی وضعیت عاطفی را درک می‌کند، در نتیجه توابع بر اساس پایگاه داده عمل می‌کند [35].

ب) شبکه‌های عصبی

در این مدل ترکیبی از دو استراتژی استخراج ویژگی و سازمان عصبی و دو مرحله برای تشخیص چهره و شخصیت‌پردازی وجود دارد. تصاویر به منظور کاهش زمان و بالا بردن ماهیت تصویر از قبل به کار گرفته شده است. طبقه بند مبتنی بر سازمان عصبی برای گروه‌بندی بردار مؤلفه استفاده می‌شود [36].

پ). ماشین بردار پشتیبانی

در این روش ابتدا اسکن هر دنباله فریم ویدیویی در زمان واقعی برای تشخیص جلوی چهره انجام می‌شود. سپس چهره‌ها با تبدیل پچ‌های تصویری با اندازه یکسان به بانک فیلتر انرژی گابور مقیاس گذاری می‌شود [38]. خروجی این فیلتر به عنوان ورودی تشخیص طبقه بند در نظر گرفته می‌شود که کدگذاری بیانات را به ابعاد متفاوت انجام می‌دهد. ویژگی های چهره از فیلترهای گابور به کمک AdaBoost انتخاب می شود که با استفاده از ماشین بردار پشتیبانی آموزش دیده اند. در ماشین بردار پشتیبانی پیشنهادی میزان تشخیص ۹۳ درصد بدست آمد [37].

ت) طبقه بند AdaBoost

در این روش تصاویر جلوی چهره بسته به نوع سطح بیان طبقه بندی می شوند که برای تشخیص با طبقه بند عادی حیاتی است. بیانات سطح بالاتر شامل بیانات دلسرد کننده، مشوق، مخالف، علاقه مند و نامطمئن در حالی که بیانات سطح پایین تر شامل بیانات شادی، تعجب، غم و خنثی است. سیستم تشخیص هیچ گونه بلوک مشخصه ای ندارد. یکی از روش های تشخیص رنگ پوست چهره روش RGB، HSI و YCbCr می باشد که بر روی تصاویر چهره اعمال می شود. بنابراین، بخشی از تصویر که رنگ پوستی دارد از بخش رنگ غیر پوستی با کمک آستانه کران پایین و بالایی جدا می شود. بنابراین، مقادیر طیف رنگ پوست بین ۳ تا ۳۸ در نظر گرفته می شوند. علاوه بر این، ویژگی‌های چهره با استفاده از تبدیل فضای رنگ شناسایی می‌شوند. تأیید تشخیص مردمک ناحیه چهره مانند تقسیم بندی صورت، ارتفاع و عرض ناحیه صورت و غیره با تکنیک برچسب‌گذاری مولفه متصل انجام می شود. این به تشخیص نشانه های چهره مانند - ابرو، چشم، لب و بینی کمک می کند. در نهایت، ویژگی های چهره استخراج می شود و ویژگی‌ها با جابجایی ویژگی‌های شاخص چهره با استفاده از الگوریتم طبقه‌بندی پیشنهادی AdaBoost از هم جدا می‌شوند که دقت آن ۹۰ درصد می باشد [39].

۶. ۵. پس پردازش

آخرین مرحله تشخیص چهره پس پردازش است. تصویر حاصل معمولاً پس از طبقه‌بندی موفق تصویر، مراحل پس از پردازش را طی می‌کنند. در این مرحله پس پردازش، از بانک خطی فیلترها برای رمزگذاری تغییر و تغییر ناپذیری روشنایی استفاده می شود. هدف اصلی پس پردازش افزایش دقت بازشناسی با بهره گیری از دانش دامنه برای اصلاح خطاهای طبقه بندی است. خطاهای طبقه بندی را نیز می توان با جفت کردن چندین سطوح سلسله مراتب طبقه بندی اصلاح کرد. یک مزیت استفاده از رویکرد پس پردازش تغییر ناپذیری تصاویر می باشد زیرا احتمال دقت تشخیص را افزایش می دهد. یکی مطالعه نشان داد که پس پردازش تصویر ردیابی شده به ویژه در شرایط نوری قابل اعتماد مفید است. در مرحله پس پردازش ، پارامتر صافی تصویر نیز بدون نرمال سازی روی پیش فرض تنظیم شده است. [40].

۷. نتیجه گیری

در این مقاله معروف ترین نظریه‌ها و بازنمایی احساسات ارائه گردید. رویکردهای استخراج مشخصه به سه دسته (هندسی، ظاهری و ترکیبی) و همچنین روش‌هایی با استفاده از انتخاب توصیفگر معرفی شدند. تشخیص خودکار حالات چهره بر اساس ویژگی های گذرا یک موضوع تحقیقاتی جدید است. نتایج به دست آمده ثابت کرد که رویکرد پیشنهادی می تواند مبنای پذیرفته شده ای برای تشخیص حالات چهره باشد.

منابع

[1] Michael, R I., Sam, E . A., 2018. Survey on Human Face Expression Recognition Techniques, Journal of King Saud University Computer & Information Sciences.
[2] Ullman, S., 2019. Using neuroscience to develop artificial intelligence, Science, 363(6428):692– 693.
[3]Tian, Y. L. , Kanade, T. , Cohn, J. F. , Li, S. Z. , & Jain, A. K. , 2005. Facial expression analysis. Handbook of Face Recognition, Springer London, 247–275.

[4] Darwin, C. , 1872. The Expression of the Emotions in Man and Animal, J. Murray., London.
[5] Ekman, P. and Friesen, W. V., 1976. Pictures of Facial Affect, Consulting Psychologists Press, Palo Alto, CA.
[6] Tomkins, S. S., and McCarter, R., 1964. What and where are the primary affects? Some evidence for a theory. Perceptual and Motor Skills.
[7] Majumdar, J., and Avabhrith, R., 2014. Human Face Expression Recognition", in the proceeding of IJETAE, 4, 7.
[8] De Silva, L. C., Miyasato, T, and Nakatsu, R , 1997. Facial Emotion Recognition Using Multimodal Information. In Proc. IEEE Int. Conf. on Information', Communications and Signal Processing (ICICS'97), Singapore, 397- 401.
[9] Mandal, B., Lim, R.Y., Dai, P., Sayed, M.R., Li, L. and Lim, J.H., 2016. Trends in machine and human face recognition. In Advances in Face Detection and Facial Image Analysis , Springer, Cham, 145-187.
[10] Bettadapura, V., 2012. Face expression recognition and analysis: the state of the art. arXiv preprint arXiv,1-27.
[11] Nathawat, P., Chaplot, V., 2020. A review of ficial expression recognition. European Journal of Molecular & Clinical Medicine, 7 (4), 1232-1253.
[12] Savran, A., Sankur, B. and Bilge, M.T., 2012. Regression-based intensity estimation of facial action units. Image and Vision Computing, 30(10), 774-784.
[13] Rivera, A. R., Jorge R. C., Oksam Oksam C., 2013 .Local directional number pattern for face analysis: Face and expression recognition." IEEE transactions on image processing, 22, 1740-1752.
[14] Nandakuma, K.,. Jain, AK ., 2021. Soft biometrics, *Encyclopedia of Biometrics*. International Journal of Applied Engineering Research , 16(2),103-110.
[15] Deng, D., Jing, L., Yu, J., Sun, S., & Ng, M.K., 2019. Sentiment Lexicon Construction With Hierarchical Supervision Topic Model. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27, 704-718.
[16] Gong, C., Lin, F., Zhou, X., Lü, X, Amygdala-inspired affective computing: To realize personalize intracranial emotions with accurately observed external emotions. China Communications, 16, (8), 115-129.
[17] Mollahosseini, A., Hasani, B. and Mahoor, M.H., 2017. Affectnet: A database for facial expression, valence, and arousal computing in the wild, arXiv preprint arXiv:1708.03985.
[18] Li, Y., Wang, S., Zhao, Y. and Ji, Q., 2013. Simultaneous facial feature tracking and facial expression recognition. IEEE Transactions on Image Processing, 22(7), 2559-2573.
[19] Eckert, M., 2016. Fast facial expression recognition for emotion awareness disposal. Consumer Electronics-Berlin (ICCE-Berlin), IEEE, 6th International Conference on. IEEE, 2016.
[20] Liu, P., Han, S., Meng, Z., Tong ,Y., 2014. Facial Expression Recognition via a Boosted Deep Belief Network. IEEE Conference on Computer Vision and Pattern Recognition,1805-1812.
[21] Adjabi, I., Ouahabi A., Benzaoui, A., Taleb-Ahmed, A. ,2020. Past, present, and future of face recognition: a review. Electronics, 9,1188.
[22] Zhang, L., Sun, L., Yu, L., Dong, X., Chen, J., Cai, W., 2021. ARFace: attention-aware and regularization for face recognition with reinforcement learning, in IEEE

Transactions on Biometrics, Behavior, and Identity Science, (Piscataway, NJ: IEEE;), 1-10.
[23] Gao, H., Ma B., 2020. A robust improved network for facial expression recognition. Front. Signal Process, 4 (4),81-87.
[24] N., 2020. A Review of Facial Expression Recognition. European Journal of Molecular & Clinical Medicine, 7(4), 1232-1253.
[25] Sandbach, G.; Zafeiriou, S.; Pantic, M.; Rueckert, D., 2012. Recognition of 3D facial expressions dynamics, Image Vision Comput, 762–773.
[26] Zhao, X.; Zhang, S. 2011. Facial expression recognition based on local binary patterns and kernel discriminant isomap. Sensors, 11, 9573–9588.
[27] Moore, S.; Bowden, R., 2011Local binary patterns for multi-view facial expression recognition. Comput. Vision Image Understanding, 115, 541–558.
[28] Zhao, G.; Huang, X.; Taini, M.; Li, S.Z.; Pietikäinen, M. 2011. Facial expression recognition from nearinfrared videos. Image Vision Computing, 29 (9), 607-619,
[29] Thorat, B., Ganesh M., PravinY. , 2015. Automatic Classification of Facial Expressions from Video Stream using Decision Tree.” International Journal of Computer Applications, 121, 22-36.
[30] Hemalatha,G., Sumathi, C.P., 2014. A Study of Techniques for Facial Detection and Expression Classification. International Journal of Computer Science and Engineering Survey, 5 (2), 27-37.
[31] Sharan, A., Kumar, S., 2017. A Hybrid Approach of Facial Emotion Detection using Genetic Algorithm along with Artificial Neural Network, International Journal of Computer Applications, 175. 1-6.
[32] Owusu, E., Zhan, Y. and Mao, Q.R., 2014. A neural-AdaBoost based facial expression recognition system. Expert Systems with Applications, 41(7), 3383-3390.
[33] Bashyal, S., Venayagamoorthy, G.K., 2008. Recognition of facial expressions using Gabor wavelets and learning vector quantization. Engineering Applications of Artificial Intelligence, 21(7), 1056-1064.
[34] Hegde, G.P., Seetha, M., Hegde, N., 2016. Kernel locality preserving symmetrical weighted Fisher discriminant analysis-based subspace approach for expression recognition. Engineering science and technology, an international journal, 19(3), 1321-1333.
[35] Meng, H., Bianchi-Berthouze, N., 2011. Naturalistic affective expression classification by a multi-stage approach based on hidden markov models. In Affective computing and intelligent interaction . Springer, Berlin, Heidelberg, 378-387.
[36] Saudagare, P.V., Chaudhari, D.S., 2012. Facial expression recognition using neural network–an overview. International Journal of Soft Computing and Engineering (IJSCE), 2(1), 224-227
[37] Luo, Y., Wu, C.M., Zhang, Y., 2013. Facial expression recognition based on fusion feature of PCA and LBP with SVM. Optik-International Journal for Light and Electron Optics, 124 (17),.2767-2770.
[38] Ruan, J., Yin, J., Chen, Q. and Chen, G., 2014. Facial expression recognition based on gabor wavelet transform and relevance vector machine. Journal of Information & Computional Science, 11 (1), 295-302.
[39] Owusu, E., Zhan, Y., Mao, Q.R., 2014. A neural-AdaBoost based facial expression recognition system. Expert Systems with Applications, 41(7), 3383-3390.

[40] Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., Tapp, R., 2014. Featural processing in recognition of emotional facial expressions. Cognition & emotion, 28(3), 416-432