

# تعیین میزان رضایت‌مندی در حالات چهره به کمک شبکه‌ی عصبی

احسان تمنایی (دانشجوی کارشناسی ارشد)

علی مقداری\* (استاد)

دانشکده‌ی هندی مکنابنک، دانشگاه صنعتی شریف

مینو عالمی (دانشیار)

دانشکده‌ی علوم انسانی، دانشگاه آزاد اسلامی، واحد تهران غرب

مهندسی مکانیک شریف، پاییز ۱۴۰۰ (۱۳-۲، پژوهشی)  
دردی ۳-۳، شماره ۲، ص. ۱۳-۲۰، (پژوهشی)

امروزه ربات‌ها نقش مهمی در زندگی روزانه‌ی افراد کم‌توان و ناتوان و حتی انسان‌های عادی ایفا می‌کنند به نحوی که تقریباً در تمامی حوزه‌های درمانی و کمک‌درمانی، آموزشی و توان‌بخشی، بازی و سرگرمی شاهد حضور انواع و اقسام ربات‌های اجتماعی هستیم. یکی از معضلاتی که ممکن است در تعاملات بین انسان و ربات وجود داشته باشد، عدم درک متقابل عاطفی است؛ بدین معنا که معمولاً ربات هیچ درک عاطفی از حالات روحی انسان ندارد و گاه‌ها به همین دلیل، کیفیت تعاملات کاهش می‌یابد. شاید بتوان درک میزان رضایت‌مندی افراد را به‌عنوان یک پارامتر اصلی در تعاملات بین انسان و ربات در نظر گرفت؛ بدین معنا که همواره ایجاد تعامل مناسب باعث افزایش میزان رضایت‌مندی در انسان می‌شود و از طرفی دیگر افراد با ابراز ناراضایتی می‌توانند عدم تمایل خود را برای تداوم یک تعامل بیان کنند. از این رو سعی شده است در این نوشتار با استفاده از مدل شبکه‌ی عصبی پیچشی<sup>۱</sup> میزان رضایت‌مندی<sup>۲</sup> فرد در هنگام مواجهه با یک سناریوی از پیش تعیین شده بررسی شود. برخلاف تحقیقات متعدد ارزشمندی که با استفاده از شبکه‌ی عصبی عمیق<sup>۳</sup> به دنبال تشخیص حالت چهره هستند و از تصویر خام فرد به‌عنوان ورودی شبکه استفاده می‌کنند، در این جا از بردار هیستوگرام شیب‌های جهت‌دار<sup>۴</sup> چهره به‌عنوان بردار مشخصه‌ی توصیف‌کننده‌ی میزان رضایت‌مندی و از یک مدل شبکه‌ی عصبی کوچک به‌عنوان رده‌بند<sup>۵</sup> استفاده شده است. مدل به دست آمده علاوه بر قدرت تشخیص میزان رضایت‌مندی و قابلیت تعمیم<sup>۶</sup> بالا، قابلیت تشخیص عواطف منفی را نیز دارد. کوچک بودن و هزینه‌ی پایین پردازش دو عنصر بسیار مهم در کارایی سیستم‌های منفک است که به‌عنوان دو قید اساسی در مدل لحاظ شده است؛ گاهی برای رسیدن به این دو مهم از پارامترهای دیگر چشم‌پوشی شده است.

**واژگان کلیدی:** میزان بروز رضایت‌مندی، شبکه‌ی عصبی، ربات اجتماعی، حالت چهره.

## ۱. مقدمه

کیفیتی همواره ثابت امور محوله را انجام دهند. اما در تعاملات اجتماعی بین ربات و انسان درک متقابل فی‌مابین آن‌ها در کیفیت تعاملات بسیار مؤثر است به نحوی که علی‌رغم توانایی بیشتر ربات‌ها در انجام امور محوله، نداشتن همین درک متقابل سبب کارایی به مراتب کم‌تر آن‌ها نسبت به انسان می‌شود. از همین رو در این مقاله به نحوه‌ی ایجاد و آموزش مدل‌های مبتنی بر شبکه‌ی عصبی به منظور درک میزان رضایت‌مندی نمایان شده در حالت چهره‌ی افراد خواهیم پرداخت.

قدمت کارهای انجام شده در زمینه‌ی تشخیص احساسات و دلیل بروز آن‌ها در افراد، به حدود ۴۰ سال پیش برمی‌گردد. چارلز داروین<sup>[۱]</sup> اولین فردی است که در حوزه‌ی عواطف و بروز احساسات دارای فرضیه است و ایشان بروز احساسات در

ربات‌های اجتماعی امروزه نقش مهمی در زندگی روزانه‌ی افراد کم‌توان و ناتوان و حتی انسان‌های عادی ایفا می‌کنند به نحوی که تقریباً در تمامی حوزه‌های درمانی و کمک‌درمانی، آموزشی و توان‌بخشی، بازی و سرگرمی شاهد حضور انواع و اقسام ربات‌های اجتماعی هستیم. از آن‌جا که فعالیت در این حوزه‌ها مستلزم داشتن روحیه‌ی مقاوم و انجام چندین و چندباره‌ی یک عمل خاص با کیفیت ثابت است، ربات‌ها می‌توانند به‌خوبی جای انسان را گرفته و بدون خستگی و دل‌زدگی و با

\* نویسنده مسئول

تاریخ: دریافت ۱۰/۱۳۹۹، اصلاحیه ۲۸/۱۴۰۰، پذیرش ۲۶/۲۰۲۰/۱۴۰۰.

DOI:10.24200/J40.2021.56649.1562

ehtamanaee@yahoo.com  
meghdari@sharif.edu  
minooalemi2000@yahoo.com

حالت ناراحتی به‌عنوان عدم رضایت‌مندی در قالب یک قابلیت تعمیم سیستم مورد بررسی قرار گرفته است.

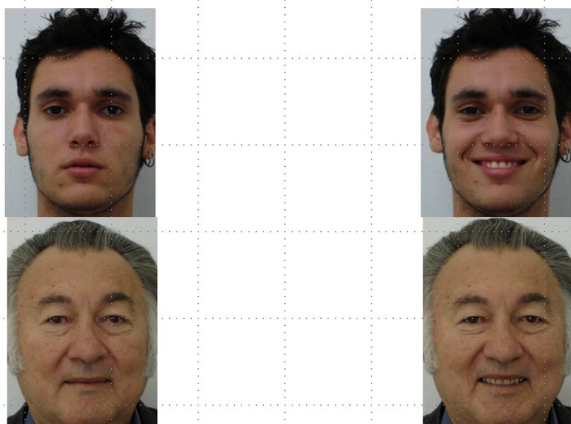
### ۳. استخراج بردار ویژگی چهره

در این نوشتار ابتدا به اصل تفاوت بین حالات چهره توجه شده است؛ به این معنی که اساساً افراد با دیدن چه خصوصیتی به حالت ابراز شده توسط فرد پی می‌برند؟ برای درک بهتر مطلب بیان شده به تصاویر نشان داده شده در شکل ۲ توجه کنید: چگونه با داشتن یک تصویر با چند خط ساده می‌توان احساسات بیان شده در شکل ۱ را به‌صورت کامل درک کرد؟ در واقع به نظر می‌رسد که درک بصری ما از حالت چهره با استفاده از اطلاعات بسیار ساده‌ی خطوط چهره اتفاق می‌افتد (شکل ۳).

به زوایای خطوط چهره در شکل ۳ دقت کنید، تقریباً زوایای ایجاد شده در تمامی خطوط چهره‌ی شاد بیشتر از چهره ناراحت است. به همین سادگی می‌توان حالت چهره را فهمید بدون توجه به رنگ و اندازه‌گیری فاصله‌ی نسبی بین نقاط کلیدی چهره و ... گویی ذهن انسان در مسیر تکامل خود، برای هر حالت چهره الگویی از خطوط چهره را اختصاص داده است و با همین الگوی ساده در باره‌ی عواطف ابراز شده توسط دیگران قضاوت می‌کند. اما چگونه می‌توان این تفاوت در خطوط چهره را به‌صورت کمی بیان کرد؟ بدین منظور از «بردار هیستوگرام شیب‌های جهت‌دار» استفاده می‌شود. این اصطلاح رایج در علم بینایی ماشین<sup>۷</sup> در واقع چگالی زاویه‌ی خطوط تصویر در نقاط مختلف آن را بازگو می‌کند که این امر همان بردار جادویی است که به نظر می‌رسد می‌توان از آن برای تشخیص حالت چهره استفاده کرد. در ادامه به تاریخچه‌ی پیدایش هیستوگرام شیب‌های جهت‌دار اشاره خواهد شد. در سال ۱۹۸۶ میلادی مک‌کونل<sup>۸</sup> هیستوگرام شیب‌های جهت‌دار را معرفی



شکل ۲. بیان عواطف با رسم چند خط ساده.



شکل ۳. زوایای خطوط در حین ابراز عواطف.

یک فرد را مستقیماً زاده‌ی احساس فرد نسبت به محیط می‌داند و معتقد است که دامنه و شدت بروز عواطف ریشه در خصوصیات ژنتیکی افراد و تعاملات محیطی دارد. تشخیص عواطف چهره نیز دارای سابقه‌ی طولانی است به نحوی که در سال ۱۹۷۸ دانشمندی به نام سوا در مقاله‌ی این موضوع را مطرح و تحلیل خود را ارائه کرده است.<sup>۹</sup> دانشمندان زیادی در اواخر قرن بیستم روی موضوع تشخیص عواطف و احساسات انسان کار کرده‌اند<sup>۱۰-۱۲</sup> اما هنوز هم بعد از حدود چهل سال، تحقیقات در این حوزه ادامه دارد و پروژه‌هایی در مقاطع مختلف تحصیلی در این حوزه تعریف می‌شود<sup>۱۳</sup> که گاهی در آن‌ها به منظور تشخیص عواطف از مدل‌های شبکه‌ی عصبی استفاده شده است.<sup>۱۴</sup> دو دلیل اصلی توجه دوباره به تشخیص عواطف عبارت است از:

- نیاز صنعت و کاربردهای تجاری؛
- پیشرفت در حوزه‌ی علوم پردازش تصویر و یادگیری ماشین و ...

### ۲. توصیف و نمایش احساسات

به‌صورت کلی دو روش برای توصیف عواطف چهره وجود دارد، در روش اول حالت چهره به‌صورت درصد عضویت آن در کلاس‌های گسسته از پیش تعیین شده بیان می‌شود.<sup>۱۱</sup> رایج‌ترین و معتبرترین کلاس‌های گسسته‌ی عواطف چهره توسط اکمن پیشنهاد شده است: خوشحالی، ناراحتی، تعجب، ترس، عصبانیت و چندان.

اکمن بیان کرده است که این روش یکسان برای شناخت و بیان همه‌ی انسان‌ها مستقل از فرهنگشان تکامل یافته است. مطالعات اکمن در این حوزه نشان داده است که احساسات پایه بیان شده بدون در نظر گرفتن فرهنگ قابل درک است؛ در شکل ۱ شش حالت اصلی صورت در این احساسات نشان داده شده است.

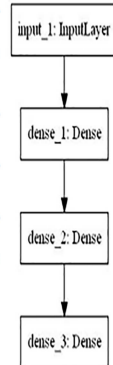
احساسات پایه در شناخت خودکار احساسات بسیار مؤثرند و تاکنون اکثر پژوهش‌ها به‌طور خاص به شناخت این احساسات اختصاص داشته است.<sup>۱۲</sup> روش دیگر برای تعریف احساسات استفاده از طبقه‌بندی ابعادی است<sup>۱۳</sup> که در آن «حالت عاطفی» به‌عنوان نقطه‌ی در فضای چندبعدی مشخص شده است که در این حالت هر بعد نشان‌دهنده‌ی یک مفهوم کیفی، نظیر شدت ابراز عواطف یا ارزش‌مند بودن آن حالت عاطفی یا هر مفهوم کیفی قابل تعریف دیگر است. این ابعاد به نوعی تفاوت‌ها و شباهت‌های عواطف را در کیفیت ابرازشان بررسی می‌کند.<sup>۱۴</sup> مثال‌هایی از این ابعاد احساسی عبارت است از: ارزشی (دل‌پذیری در مقابل ناخوشایندی)، قدرت (حس کنترل، تسلط در مقابل تسلیم)، فعال بودن (آرامش بخش در مقابل محرک بودن) و ارزش (پیش‌بینی و ارزیابی نوآوری و غیرقابل پیش‌بینی بودن).

طبقه‌بندی ابعادی در مقایسه با طبقه‌بندی دسته‌ی اعطاف بیشتری برای تجزیه و تحلیل احساسات در اختیار ما می‌گذارد. در این مقاله رضایت‌مندی به‌عنوان یک بعد ارزشی تعریف شده است و حالت شاد به‌عنوان غایت رضایت‌مندی و حالت معمولی به‌عنوان غایت عدم رضایت تعریف شده است و توانایی سیستم در تشخیص



شکل ۱. شش حالت اصلی بیان عواطف: خوشحالی، ناراحتی، تعجب، ترس، عصبانیت و چندان.

Layer (type)	Output Shape	Param #
input_3 (InputLayer)	(None, 648)	0
dense_7 (Dense)	(None, 51)	33099
dense_8 (Dense)	(None, 25)	1300
dense_9 (Dense)	(None, 2)	52
Total params: 34,451		
Trainable params: 34,451		
Non-trainable params: 0		



شکل ۵. لایه‌ها و پارامترهای مدل بهینه.

دسته <sup>۱۲</sup> بزرگ استفاده شده است:

$$batch\_size = \text{int}(\circ, 5 * HOG - size)$$

مدل شبکه‌ی عصبی طراحی شده از یک لایه‌ی ورودی و دو لایه‌ی پنهان با تعداد نرون متناسب با اندازه‌ی ورودی تشکیل شده است:

$$dense1 =$$

$$layers.Dense(\text{int}(\circ, 7 * HOG - size), activation) = 'relu'$$

$$dense2 =$$

$$layers.Dense(\text{int}(\circ, 4 * HOG - size), activation) = 'relu'$$

در شکل ۵ چکیده‌ی مدل نشان داده شده است.

این مدل تنها دارای ۳۴۴۵۱ پارامتر قابل آموزش است، هرچند رسیدن به پاسخ نسبی با تعداد بسیار کم‌تری (زیر ۱۰ هزار پارامتر) هم امکان‌پذیر است ولی خود همین بهینه‌سازی تعداد پارامترها می‌تواند یک چالش جذاب و قابل پیگیری باشد. این پروژه برای آموزش از مجموعه داده‌های برجسب‌دار استاندارد اف‌ای‌آی (FEI) <sup>۱۳</sup> استفاده شده است، که شامل دویست عکس در حالت عادی و دویست عکس در حالت نهایت رضایت‌مندی (خنده) افراد است. داده‌ها ۱۶۰ داده‌ی آموزشی، برای شروع مرحله‌ی بعدی آموزش، داده‌ها به صورت تصادفی در دو گروه فوق قرار می‌گیرند.

روند آموزش و اعتبارسنجی مدل فوق روی داده‌های برجسب‌دار استاندارد FEI Face Database در شکل ۶ نشان داده شده است.

همان‌طور که مشاهده می‌کنید مدل شبکه‌ی عصبی به یک دقت یکسان برای داده‌های آموزش و اعتبارسنجی رسیده است که نوید یک قابلیت تعمیم عالی را می‌دهد.

برای آزمایش مدل از داده‌های استاندارد CK+ استفاده شده است تا قابلیت کلی‌نگری مدل به خوبی مورد ارزیابی قرارگیرد (آموزش و اعتبارسنجی روی یک سری داده و آزمایش روی داده‌هایی از دسته‌ی دیگر انجام شده است).

## ۵. آزمون نظریه

برای آزمون نظریه‌ی مدل شبکه‌ی عصبی به دست آمده از داده‌های استاندارد CK+ استفاده شده است و آزمون با دو هدف رده‌بندی و تشخیص روند انجام شده

(۸۰، ۸۰) = اندازه پنجره  
 (۴۰، ۴۰) = اندازه بلوک  
 (۲۰، ۲۰) = لغزش سلول  
 (۲۰، ۲۰) = اندازه سلول  
 ۱۸ = تعداد بین



شکل ۴. مشخصات پنجره‌ی بلوک و سلول در مدل بهینه.

کرد و در سال ۱۹۹۴ مؤسسه‌ی تحقیقاتی شرکت میتسوبیشی <sup>۹</sup> آن را مورد استفاده قرار داد. استفاده از این روش در سال ۲۰۰۵ گسترده‌تر شد، و کار تکمیلی محققان مؤسسه‌ی ایرینا <sup>۱۰</sup> در این زمینه، در کنفرانس بینایی ماشین و الگوی تشخیص (CVPR) <sup>۱۱</sup> مطرح شد. در این کار آن‌ها بر تصویر عابر پیاده در تصاویر ثابت متمرکز شدند اگرچه از آن پس کارشان را بر تصاویر ویدیویی متحرک و همچنین بر حیوانات و وسایل نقلیه در تصاویر ثابت نیز بسط دادند. <sup>[۱۷]</sup> این توصیف‌گر در تکه‌های ۱۲۸ × ۶۴ پیکسلی از تصویر محاسبه شده است، اگرچه تصویر می‌تواند هر اندازه‌ی داشته باشد. معمولاً تکه‌های عکس در مقیاس‌های متعدد و مکان‌های متفاوتی از تصویر تحلیل خواهد شد. <sup>[۱۸، ۱۹]</sup>

## ۱.۳. بردار ویژگی‌های مناسب برای تشخیص حالت چهره

از آن‌جا که هدف اصلی بردار ویژگی‌ها استخراج چگالی زاویه‌ی خطوط در صورت است، تنها فید ما در کاهش سایز بردار ویژگی‌ها عدم حذف نواحی مهم صورت است. ممکن است در نگاه اول این پیشنهاد به ذهن خطور کند که ابتدا چهره را به نواحی مهمی چون چشم‌ها، ابروها، دهان، گونه و پیشانی و ... تقسیم کرده و تنها چگالی خطوط در نواحی مهم بررسی شود. این کار بسیار مفید و کارساز خواهد بود اگرچه پردازش اولیه برای تشخیص نواحی چهره را بر روند تشخیص حالت اضافه می‌کند که خود این عملیات پردازشی می‌تواند زمان کلی تشخیص سیستم را زیاد کند. این امر برای سیستم‌های برخط مطلوب نیست؛ لذا کلیت چهره به عنوان داده‌ی ورودی در نظر گرفته شده و سعی شده است از آن یک بردار ویژگی کمینه استخراج شود. بدین منظور اندازه‌ی پنجره ۸۰ × ۸۰ در نظر گرفته شده است و صورت فرد با چهار بلوک کناری و ۵ بلوک میانی (لغزش بلوک‌های کناری به‌ازای گام لغزش ۲۰ پیکسل) ۴۰ × ۴۰ ناحیه‌بندی شده است و چگالی زوایای خطوط در سلول‌های ۲۰ × ۲۰ محاسبه شده است (شکل ۴).

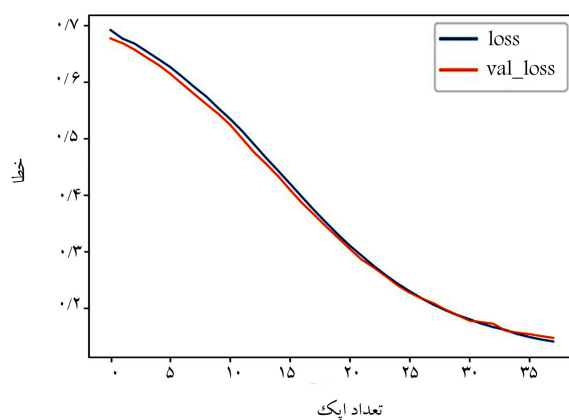
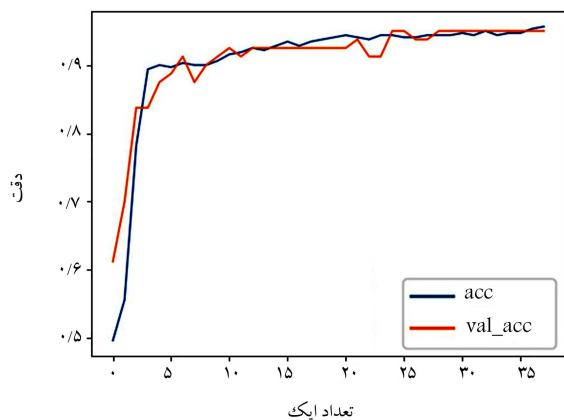
با توجه به مشخصات اشاره شده، اندازه‌ی بردار ویژگی چهره عبارت خواهد بود از:

$$HOG - size =$$

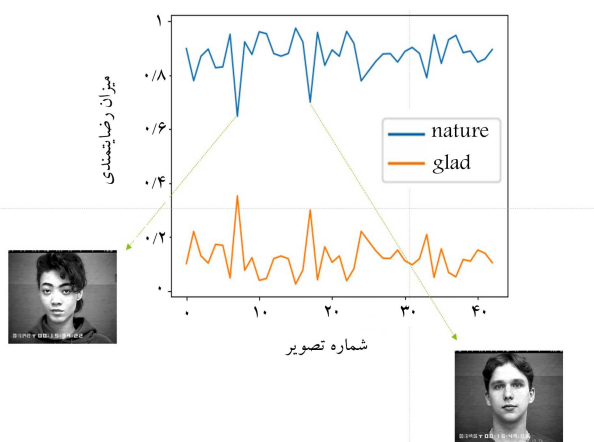
$$3 * 3(blocks) * 4(cells in each block) * 18(bins) = 648$$

## ۴. آموزش مدل شبکه‌ی عصبی

برای بیان رضایت‌مندی فرد از یک مدل شبکه‌ی عصبی پیچشی استفاده شده است که در واقع رضایت‌مندی را به صورت درصد عضویت در دسته‌ی کاملاً راضی و کاملاً ناراضی بیان می‌کند. برای آن که مدل به مرکز خوشه‌ها همگرا شود از اندازه



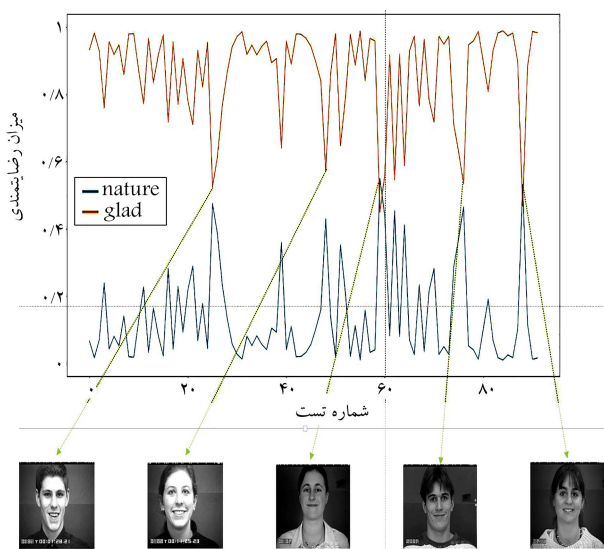
شکل ۶. روند آموزش مدل.



شکل ۸. خروجی مدل برای حالت عادی تصاویر.



شکل ۷. نمونه داده آموزشی CK+.



شکل ۹. خروجی مدل برای حالت خندان تصاویر.

( $TP = 103, FN = 2$ ) که نشان‌دهنده قدرت بالای قابلیت تعمیم سیستم است. اما آنچه که باعث نگرانی می‌شود، قله‌های نزدیک به هم برای دست‌کم پنج تصویر نشان داده شده است؛ این امکان وجود دارد که با یک مدل غیربهمینه یا با

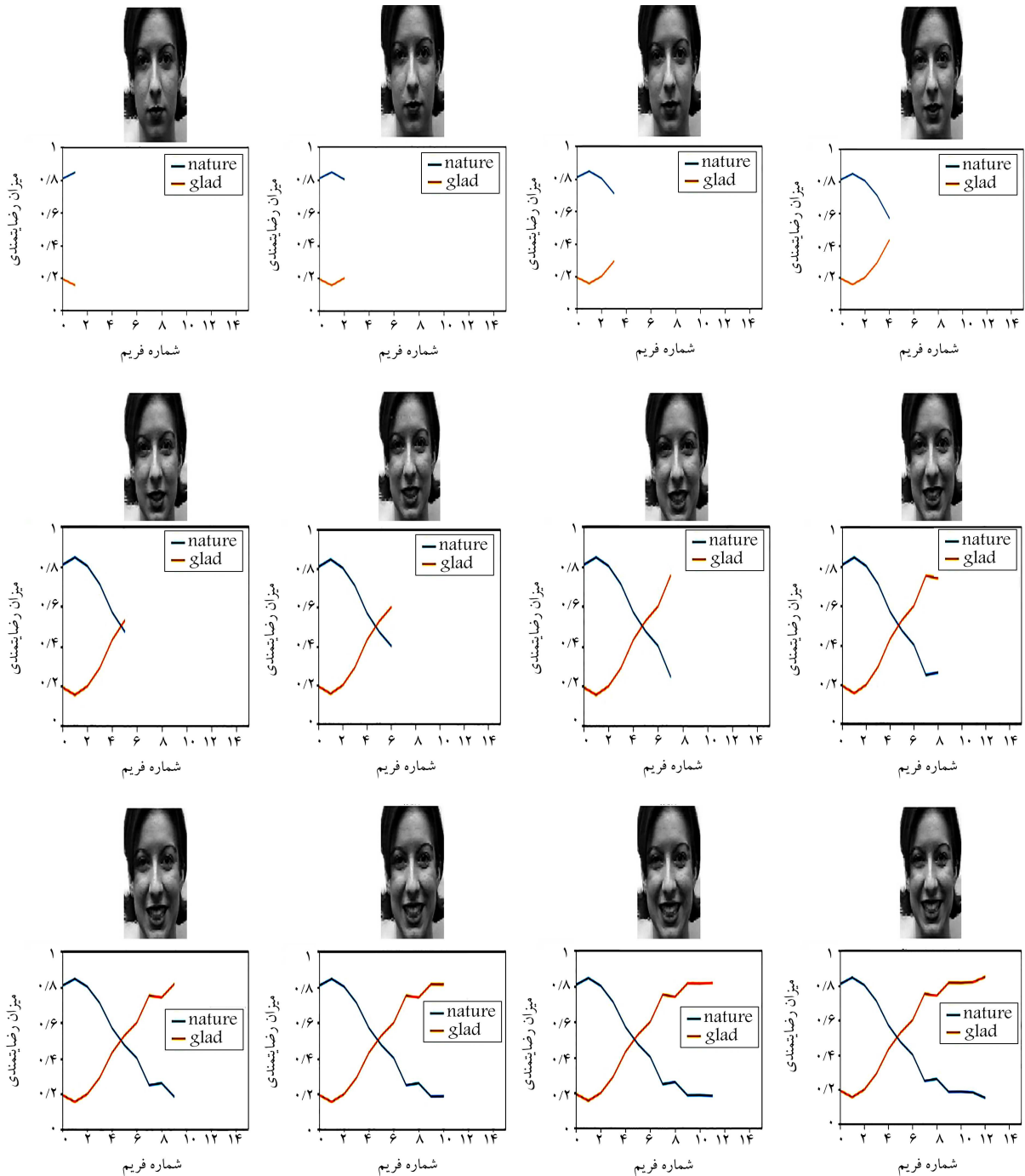
است. چنان که در شکل ۷ مشاهده می‌کنید تصاویر موجود در داده‌های استاندارد CK+ به صورت تدریجی از حالت معمولی (بدون هیچ‌گونه ابراز رضایت‌مندی) به حالت خندان (رضایت‌مندی کامل) تغییر وضعیت داده‌اند. لذا اولین تصویر هر دسته به عنوان تصویر عادی و آخرین تصویر ذخیره شده را به عنوان تصویر خندان در نظر گرفته و به عنوان ورودی آزمون به مدل وارد می‌شوند (شکل ۷).

در شکل ۸، خروجی مدل برای چهل تصویر عادی موجود در داده‌های CK+ نشان داده شده است که خود موید این مطلب است که مدل بدون نقص حالت عادی را برچسب می‌زند و در واقع با دقت ۱۰۰ درصد ( $FP = 0, TN = 40$ ) حالت عادی را تشخیص می‌دهد.

قله‌های مقادیر به دست آمده در خروجی مدل (بدترین مقادیر منتصب شده به تصاویر) کاملاً از هم فاصله دارند و این بسیار عالی است و نشان می‌دهد که مدل بهمینه حالت عادی را با قاطعیت مناسبی (عضویت بالا در دسته طبیعی و عضویت ناچیز در دسته خوشحال) تشخیص می‌دهد.

در مرحله بعد، تصاویر حالت خوشحال به عنوان ورودی به مدل معرفی شده‌اند. از آنجا که دقت پاسخ‌دهی مدل برای تصاویر خوشحال، مانند تصاویر عادی نبود تمام ۱۰۵ داده‌ی ممکن مورد بررسی قرار گرفته‌اند. خروجی مدل برای تصاویر خندان در شکل ۹ نشان داده شده است.

دقت برچسب‌زنی برای تصاویر خوشحال کم‌تر از تصاویر عادی است و چنان که مشاهده می‌کنید مدل به دقت تقریبی ۹۸ درصد رسیده است



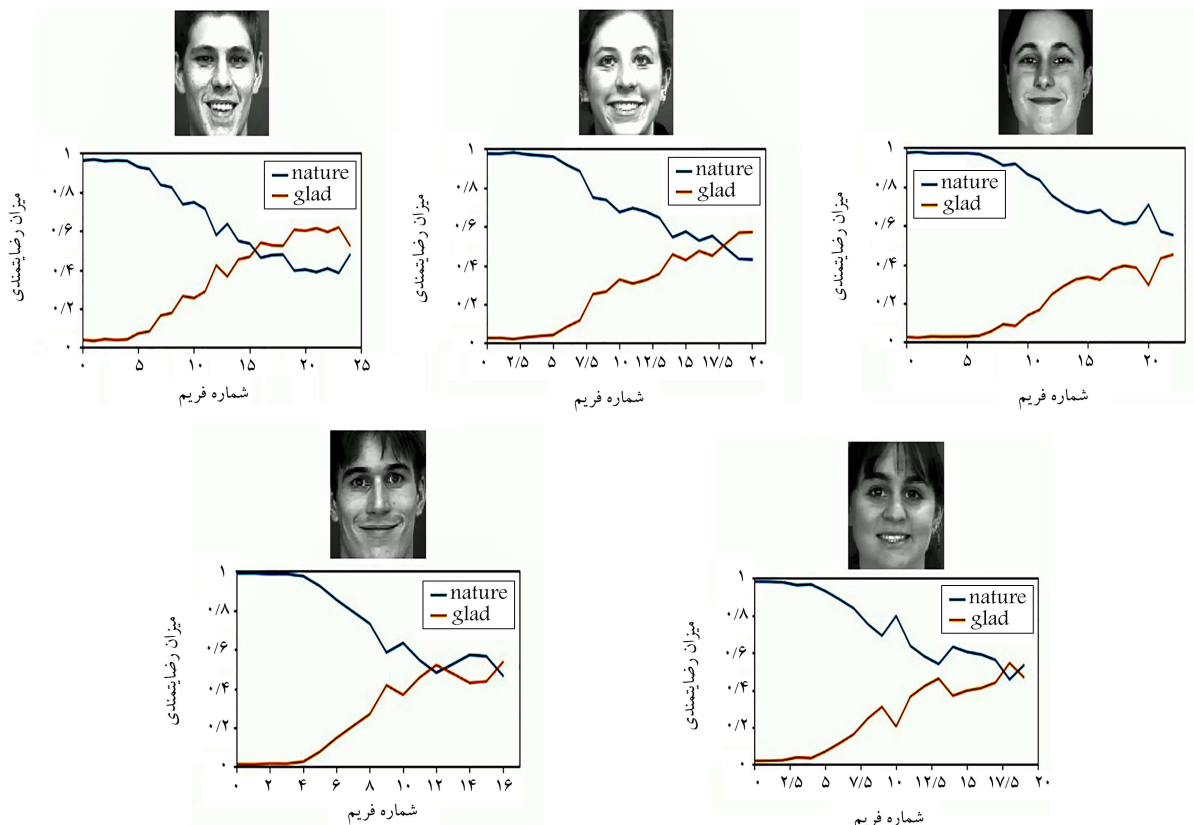
شکل ۱۰. تغییرات میزان رضایت مندی.

ایجاد شده در عضویت فرد در رده‌های خروجی باید روند تغییرات رضایت مندی را ثبت کند. مثلاً ۱۲ فریم اول از تصاویر نشان داده شده در شکل ۵ به عنوان ورودی به مدل اعمال شده که خروجی سیستم را در شکل ۱۰ مشاهده می‌کنید.

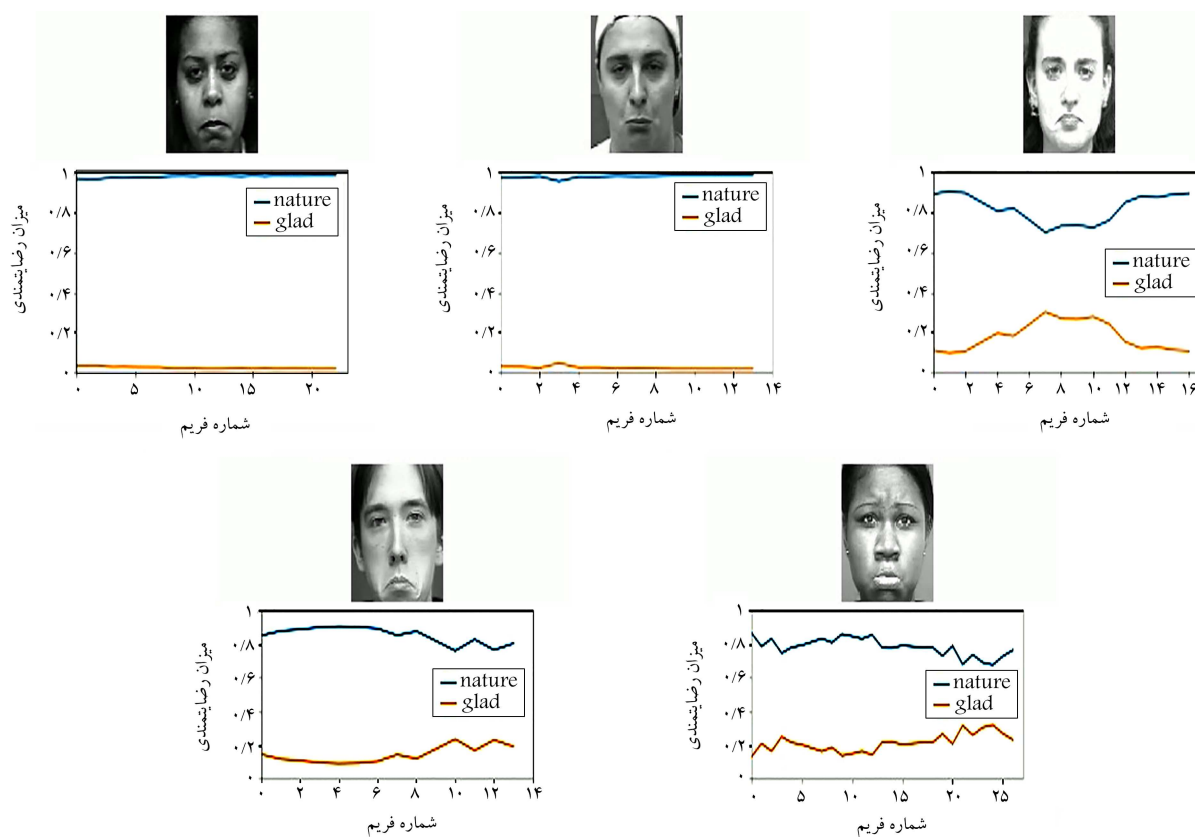
چنان که مشاهده می‌شود هر قدر بروز حس رضایت مندی در تصویر بیشتر می‌شود، درصد تعلق فرد به دسته‌ی رضایت مند نیز افزایش می‌یابد. برای نشان دادن توانایی مدل در بیان دینامیک تغییرات رضایت مندی نمودار نهایی به دست آمده از دسته تصاویری که مدل در رده بندی آن‌ها دچار مشکل بود، در شکل ۱۱ آورده شده است.

انجام دوباره‌ی مراحل آموزش، مدل به دست آمده نتواند برچسب‌زنی برای تصاویر خوشحال را با دقت مورد نظر انجام دهد (احتمال رسیدن به دقت ۹۵ درصد نیز وجود دارد).

اما از آن‌جا که هدف، قضاوت درباره‌ی دینامیک تغییر حالت افراد در مواجهه با یک محرک خارجی است (نه رده بندی لحظه‌ی) سناریوی آزمون را تغییر داده و در آزمایشی دیگر تغییرات میزان رضایت مندی در دسته‌ی تصاویر موجود در داده‌های استاندارد CK+ مورد تحقیق قرار گرفته است؛ به این ترتیب که تصاویر تغییر تدریجی ابراز رضایت مندی به ترتیب به مدل معرفی می‌شود و در نهایت مدل با توجه به تغییرات



شکل ۱۱. تشخیص روند تغییر رضایت‌مندی.



شکل ۱۲. تغییرات میزان رضایت‌مندی در حین ابراز ناراحتی.

ابراز ناراحتی، مدل عدم رضایت‌مندی را تشخیص می‌دهد. این توانایی بسیار مهمی است که ما با یک داده‌ی استاندارد (FEI) و با داشتن دو کلاس در انتها و مرکز بعد رضایت‌مندی، مدل را آموزش داده ایم و مدل در قضاوت تصاویر استاندارد دیگر (CK+) در کلاس متعلق به انتهای دیگر بعد رضایت‌مندی موفق عمل کرده است.

## ۷. نتیجه‌گیری

در نوشتار حاضر یک مدل کوچک از شبکه‌ی عصبی پیچشی ارائه شده است که با توجه به بردار هیسیتوگرام زوایای خطوط چهره درباره‌ی بعد ارزشی رضایت‌مندی افراد در مواجهه با یک سناریوی از پیش طراحی شده، قضاوت می‌کند. از آنجا که تعداد پارامترهای قابل آموزش این مدل کم است (۳۴۴۵۱ پارامتر قابل آموزش)، استفاده و آموزش تطبیقی آن در سیستم‌های منفک بسیار ساده‌تر از مدل‌های شبکه‌ی عصبی عمیق با تعداد پارامترهای بالاست. در این نوشتار نشان داده شد که مدل به دست آمده دارای دقت مناسب در رده‌بندی حالات مورد نظر بروز داده شده است، به نحوی که توانسته است حالت عادی افراد را با دقت ۱۰۰ درصد ( $FP = 0, TN = 40$ ) و حالت خندان را با دقت ۹۸ درصد ( $TP = 103, FN = 2$ ) برچسب‌گذاری کند. اما مهم‌ترین قابلیت این مدل در تشخیص روند تغییرات میزان رضایت‌مندی افراد، در حین مواجهه با یک سناریوی از پیش طراحی شده است به نحوی که توانسته است روند تغییرات میزان رضایت‌مندی را در ۱۰۵ پوشه تصاویر داده‌های استاندارد CK+ بدون خطا تشخیص دهد ( $F \setminus event = 100\%$ ).

قابلیت دیگر این مدل شبکه‌ی عصبی آن است که حتی با وجود آن که مدل با داشتن تصاویر خنده به‌عنوان غایت رضایت‌مندی و حالت طبیعی به‌عنوان عدم رضایت‌مندی آموزش دیده است، می‌تواند اگر سناریوی ایجاد شده باعث ناراحتی فرد شد (به‌عنوان یک حالت ناخواسته) این عدم رضایت‌مندی را تشخیص دهد.

چنان که مشاهده می‌کنید، تغییرات میزان رضایت‌مندی به خوبی در فریم‌های متوالی تشخیص داده شده است. دلیل این امر آن است که در واقع با بیان رضایت‌مندی نسبی فرد در فریم‌های مختلف در واقع به نوعی خروجی مدل برای هر فرد نسبت به چهره‌اش به‌هنگار شده است؛ بدین معنی که ابتدا فریم اول تصویر عادی فرد به مدل داده شده است و در فریم‌های بعدی در واقع خروجی با توجه به این ورودی اولیه در گراف قضاوت می‌شود.

## ۶. قابلیت تعمیم

«قابلیت تعمیم» توانایی مدل در قضاوت تصاویری است که نه در مرحله‌ی آموزش و نه در مرحله‌ی اعتبارسنجی به مدل معرفی شده‌اند. همان‌طور که مشاهده کردید، برای آموزش و اعتبارسنجی از تصاویر موجود در داده‌های استاندارد FEI استفاده شده است و مدل روی داده‌های استاندارد CK+ آزمایش شده است. بدین ترتیب قابلیت تعمیم‌دهی مدل اثبات شده است ولی در این قسمت سعی خواهد شد قابلیت کلی‌نگری مدل از زاویه‌ی دیگری بررسی شود.

در واقع از آنجا که هدف از محرک خندانندن فرد بوده است، حالت عادی به‌عنوان عدم رضایت‌مندی برچسب زده شده است و به نوعی عدم واکنش به این محرک، عدم رضایت‌مندی یا بی‌تفاوتی را نشان می‌دهد. حال فرض کنید فرد با دیدن محرک نه تنها رضایت‌مندی خود را با خندیدن نشان ندهد، بلکه به هر دلیلی ناراحت شود، آیا مدل می‌تواند این ناراحتی را به‌عنوان یک حالت ابراز عواطف ناخواسته تحلیل کند!

در عین تعجب مشاهده خواهید کرد که این توانایی در مدل وجود دارد زیرا بردار هیسیتوگرام شیب‌های جهت‌دار (HOG) چهره در حالت ناراحت به حالت عادی شبیه‌تر است. برای اثبات این ادعا به گراف‌های به دست آمده از ابراز ناراحتی تصاویر استاندارد CK+ توجه کنید (شکل ۱۲).

چنان که مشاهده می‌کنید، با توجه به گراف‌های به دست آمده از افراد در حین

## پانوشته‌ها

1. convolutional neural network
2. level of Pleasure
3. deep neural network
4. histogram of oriented gradients (HOG)
5. classifier
6. generalization
7. Machine Vision Science
8. R. McConnell
9. mitsubishi electric research laboratories

10. IRINA
11. computer vision and pattern recognition (CVPR)
12. batch size
13. face emotion onformation (FEI)

## منابع (References)

1. Darwin, C. and Prodger, P. "The expression of the emotions in man and animals", Oxford University Press (1998).

2. Motoi Suwa, Noboru Sugie, and Keisuke Fujimura. "A preliminary note on pattern recognition of human emotional expression", *In International Joint Conference on Pattern Recognition*, pages pp. 408-410 (1978).
3. Ekman, P. and Friesen, W. "Facial action coding system: a technique for the measurement of facial movement", Palo Alto: Consulting Psychologists (1978).
4. Ekman, P., Friesen, W.V., O'sullivan, M. and et al. "Universals and cultural differences in the judgments of facial expressions of emotion", *Journal of Personality and Social Psychology*, **53**(4), American Psychological Association, pp. 712-717 (1987).
5. Russell, J. and Mehrabian, A. "Evidence for a three-factor theory of emotions", *Journal of Research in Personality*, **11**(3), pp.273-294 (1977).
6. Plutchik, R. "The nature of emotions human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice", *American Scientist*, **89**(4), Sigma Xi, the Scientific Research Society, pp. 344-350 (2001).
7. Ashok Samal and Prasana A. Iyengar. "Automatic recognition and analysis of human faces and facial expressions: a survey", *Pattern Recognition*, **25**(1), pp. 65-77 (1992).
8. Noroozi, F. "Multimodal emotion recognition based human-robot interaction enhancement", Degree: PhD (2018).
9. Raut, Nitisha. "Facial emotion recognition using machine learning", Degree: San Jose State University ,MS (2018).
10. Hu, Jiaqi. "A personal facial expression monitoring system by deep learning", Degree: Texas A&M University, PhD (2017).
11. Ghayoumi, M. "Facial expression analysis using deep learning with partial integration to other modalities to detect emotion", Degree: Kent State University (PhD 2017).
12. Paul Ekman. "An argument for basic emotions", *Cognition and Emotion*, **6**(3), pp. 169-200 (1992).
13. Ekman, P. and Friesen, W. "Pictures of Facial Affect. Consulting Psychologists Press", (1976).
14. Zeng, Z., Pantic, M. and Huang. "A survey of affect recognition methods: audio, visual, and spontaneous expressions", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**(1), pp. 39-58 (2009).
15. Russell, J. and Mehrabian, A. "Evidence for a three-factor theory of emotions", *Journal of Research in Personality*, **11**(3), pp. 273-294 (1977).
16. James, A. "Russell and Albert Mehrabian. evidence for a three-factor theory of emotions", *Journal of Research in Personality*, **11**(3), pp. 273-294 (1977).
17. Dalal, N. and Triggs, B. "Histograms of oriented gradients for human detection. in: computer vision and pattern recognition (CVPR)", *IEEE Computer Society Conference on. IEEE*, **1**, pp. 886-93, (2005).
18. Aliasghari, P., Taheri, A., Meghdari, A. and et al. "Implementing a gaze control system on a social robot in multi-person interactions", *In: Springer Nature Applied Sciences: Socio-Cognitive Engineering*, **2**(6), pp. 1-13 (2020).
19. Esfandbod, A. Rokhi, Z., Taheri, A. and et al. *7<sup>th</sup> International Conference on Robotics and Mechatronics (ICRoM-IEEE)*, pp. 69-73 (2019).