Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: jscit.nit.ac.ir
DOI: 10.22034/jscit.2025.515576.2118
Volume 14, Number 2, Summer 2025, pp. 83-96
Received:04/07/2025, Revised:07/21/2025, Accepted: 07/23/2025

Improving Facial Emotional Expression Recognition for Individuals with Visual Agnosia Based on Early Biological Inspired Features Model

Elham Askari*
*Department of Computer Engineering, FSh.C., Islamic Azad University, Fouman, Iran.
askary.elham@iau.ir

Corresponding author's address: Elham Askari, Fouman and Shaft Branch, Islamic Azad University, Fouman, Iran.

Abstract- Agnosia is a neurological condition that results in the inability to recognize and extract meaning from the visual and auditory environment. Individuals with visual agnosia have difficulty in the communication between the mind and the brain and, as a result, in perceiving the images seen. In this paper, a model based on the visual pathway is proposed that will help these individuals with image and face recognition using fuzzy perceptual mapping. Also, an early biologically inspired feature model will be used to identify facial emotional expressions. First, appropriate features are extracted and then, using fuzzy perceptual mapping, the concept of the seen image is perceived. If the detected image is a face, the early biologically inspired model will be used to extract effective facial features and finally, the facial emotional states are recognized with a support vector machine. The results obtained show that the proposed model performs well with an accuracy of 98.1% in object recognition and an average accuracy of 99.01% in facial emotional states recognition.

Keywords- Visual Agnosia, Fuzzy Cognitive Map, Vision Model, Early Biological Model

مجله علمی رایانش نرم و فناوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: jscit.nit.ac.ir
شناسه دیجیتال: 10.22034/jscit.2025.515576.2118
جلد ۱۴، شماره ۲، تابستان ۱۴۰۴، صفحه ۸۳-۹۶
دریافت: ۱۴۰۴/۰۱/۱۸، بازنگری: ۱۴۰۴/۰۴/۳۰، پذیرش: ۱۴۰۴/۰۵/۰۱

بهبود بازشناسی حالت‌های هیجانی چهره در افراد مبتلا به ادراک پریشی بصری مبتنی بر ویژگی‌های الهام گرفته از مدل زیست‌شناختی زودهنگام

الهام عسکری *
*گروه کامپیوتر،واحد فومن و شفت،دانشگاه آزاد اسلامی، فومن، ایران.
askary.elham@iau.ir

* الهام عسکری نویسنده مسئول، فومن، خیابان انتظام، دانشگاه آزاد اسلامی واحد فومن و شفت.

چکیده- ادراک پریشی یک شرایط عصبی است که منجر به ناتوانی در شناسایی و استخراج معنا از محیط بصری و سمعی می‌شود. افراد مبتلا به ادراک پریشی بصری در ارتباط بین ذهن و مغز و درنتیجه ادراک تصاویر دیده شده دچار مشکل هستند. در این مقاله، مدلی پیشنهاد می‌شود که مبتنی بر مسیر بینایی است و با استفاده از نگاشت ادراکی فازی با بازشناسی تصاویر و چهره، به این افراد کمک خواهدکرد. همچنین درجهت شناسایی حالات هیجانی چهره از مدل ویژگی‌های الهام‌گرفته از زیست‌شناختی زودهنگام استفاده‌خواهد شد. ابتدا، ویژگی‌های مناسب استخراج شده و سپس با استفاده از نگاشت ادراکی فازی، مفهوم تصویر دیده شده ادراک می‌شود. درصورتیکه تصویر تشخیص داده شده چهره باشد از مدل الهام‌گرفته از زیست شناختی زودهنگام برای استخراج ویژگی‌های موثر چهره استفاده خواهد شد و در نهایت با ماشین بردار پشتیبان، حالات هیجانی چهره بازشناسی می‌شود. نتایج بدست‌آمده نشان می‌دهد که مدل پیشنهادی، بادقت ٪۹۸/۱، در بازشناسی اشیا و با میانگین دقت ٪۹۹/۰۱، در تشخیص حالات هیجانی چهره، عملکرد مناسبی دارد.

واژه‌های کلیدی: ادراک پریشی بصری، نگاشت ادراکی فازی، مدل بینایی، مدل زیست شناختی زودهنگام

۱- مقدمه

اگنوزیا یا ادراک پریشی، به شرایط عصبی اطلاق می‌شود که منجر به ناتوانی در دانستن، نامگذاری، شناسایی و استخراج معنا از محیط بصری، سمعی و حسی می‌گردد. این بیماری سبب ناتوانی در دریافت اطلاعات از طریق یکی از حواس می‌شود علیرغم این واقعیت که اندام گیرنده بی نقص است [۱و۲]. ادراک پریشی بصری[۱]، متداول ترین نوع این اختلال می‌باشد. اگنوزیای بصری یا اختلال در بازشناسی اشیا شرایطی است که در آن، بیمار توانایی تشخیص اشیاء را به صورت بصری و تصویری ندارد اگر چه نقش و عملکرد حس بینایی کاملاً نرمال و طبیعی است یعنی اشیاء مشاهده می‌شوند اما فرد قادر به درک معنای آن‌ها نیست و شیء را نمی‌تواند تعریف نماید [۳].

هاینریش لیشر درسال ۱۸۹۰، استدلال کرد که دو حالت وجود دارد که در آن اختلال تشخیص شی میتواند رخ دهد. یکی از راه‌هایی که می‌تواند در آن اختلال رخ دهد این است که آسیب در پردازش ادراکی اولیه باشد و یا اختلال در نمایش شیء واقعی وجود داشته باشد [۴].

ادراک پریشی بصری پس از آسیب به قشر ارتباط بصری و بخش‌هایی ازجریان شکمی ایجاد می‌شود چرا که این مسیر اجازه

می‌دهد فرد اشیا را بشناسد و اجازه می‌دهد رفتار فرد به‌سوی بینایی نسبتا نرمال هدایت شود. ادراک پریشی بصری اغلب ناشی از آسیب به قشر پس‌سری-گیجگاهی مغز است [۵]. غشای مغز از سلول‌های عصبی تشکیل شده که لایه خارجی مغز را می‌سازد. غشای بینایی به نواحی مختلفی تقسیم شده است.V1[۲] و V2 از بزرگترین ناحیه‌ها هستند. اطلاعات رسیده از شبکیه، از طریق LGN[۳] از تالاموس به V1 در دو مسیر جداگانه قرار می‌گیرند و بطور مستقل پردازش می‌گردند. کار مسیر تشخیص شی، شناسایی ویژگی اشیاء (رنگ، شکل و...) است و در مسیر V1، V2، V4،[۴] AIT و PIT[۵] قسمت پس‌سری قرار دارد. درجریان شکمی چیزی که تجسم می‌شود خروجی همان چیزیست که آگاهانه می توان دید وتوصیف کرد. این جریان تفسیرذهنی از جهان اطراف ما ایجاد می کند. مسیر تشخیص مکان شی، مشخصات فضایی صحنه (جهت حرکت و..) است و در مسیر V1، V2، MT،V3 و MST[۶] قرار دارد.

عده ای از پژوهشگران با کمک علوم پزشکی وعلوم مهندسی سعی در تشخیص و درمان بیماری ادراک پریشی داشته‌اند و البته بیشتر این پژوهش‌ها در زمینه رفتاری بوده است زیرا بیشتر افراد مبتلا به این اختلال تصور می‌کنند که دچار نوعی توهم و گیجی شده اند. تشخیص ادراک پریشی بصری شامل ارزیابی، سنجش و آزمایش میدان دید فرد، درک رنگ، تصویر، مهارت خواندن ، شناخت چهره‌ها، ترسیم و شناخت اشیاء واقعی و خطوط ترسیمی می‌باشد [۶].

آندریا سرینو و همکاران در تحقیق خود تشخیص بیماری ادراک پریشی درکی را به‌صورت یک گزارش موردی تشریح کردند. در ادراک پریشی درکی باتوجه به ماهیت عوامل آسیب، اختلال بصری اولیه وجود خواهد داشت [۷]. اریکسون و همکاران در مطالعه‌ای یک مورد از ادراک پریشی بصری شدید را در یک کودک با الگوی الکتروفیزیولوژیک از خروجی، درمناطق پس‌سری-گیجگاهی مورد بررسی قرار دادند. پس از تحلیل نقشه‌ها اختلال در پس‌سری-گیجگاهی سمت چپ تشخیص داده شد و علائمی از ادراک پریشی ادراکی دیده شد [۸ و ۹] .

بارتون و همکاران، بر روی بیمارانی که در تشخیص شی (افرادمبتلا به ادراک پریشی بصری) مشکل داشتند مطالعه کردند و تفاوت آنها را با افراد سالم با در نظر گرفتن مسیر حرکت چشمشان وقتی به تصویر نگاه می کردند نشان دادند. تفسیر آنان بدین صورت بود که حرکات چشم می‌تواند به عنوان انتخاب نقطه‌های برجسته مدل شود و نشان می دهد افراد آگنوزیک یک اتکای افزایشی روی قسمتهای برجسته بصری مانند روشنایی وکنتراست دارند [۱۰].

مدل‌های محاسباتی بسیاری برای شناخت ساختارهای ذهنی ارائه شده است. مورن و بالکینوس مدلی محاسباتی برای توصیف ساختارهایی از مغز که در توجه و ادراک دخالت دارد ارائه دادند. به کمک این مدل‌ها می‌توان به بیماران ذهنی ازجمله ادراک پریشی، آلزایمر و ... کمک کرد. مسیر درک ورودی‌های حسی بینایی را باتوجه به ساختار این مدل‌ها نیز می‌توان تفسیرنمود [۱۱]. این مدل‌ها، یک مدل محاسباتی از یادگیری هیجانی در آمیگدال هستند زیرا آمیگدال مکررا در واکنش‌های عاطفی و یادگیری و محرک‌های علائم احساسی جدید دخالت می‌کند و بخش مهمی از موتور یادگیری و نیز توجه را تشکیل می‌دهد.

مدل محاسباتی درک تصویرها یک مدل مهندسی است که این مدل را می‌توان به سه بخش تقسیم نمود:

۱-بینایی اولیه[۷] ۲- بینایی میانی[۸] و ۳- بینایی سطح بالا[۹]. در لایه بینایی اولیه درک قانونمندی‌های تصویر شامل بافت تصویر، لبه، شیب و یا تقارن انجام می‌شود. در بینایی لایه‌ی میانی، تلاش‌هایی که در زمینه تحلیل تصویر برای تقسیم‌کردن تصویر به اشیایی که کم و بیش قابل نام‌گذاری باشند می‌باشد. خروجی لایه‌ی میانی مجموعه‌ای از ویژگی‌هاست که از تصویر اولیه استخراج شده‌اند. این ویژگی ها باید حاوی اطلاعاتی در مورد ساختار اصلی جسم بوده و برای نمونه‌های مختلف از یک کلاس تا حد امکان ثابت باشند و همچنین به ازای کلاس‌های متفاوت با یکدیگر اختلاف داشته باشند. در لایه بینایی سطح بالا نتیجه مرحله قبل به صورت مجموعه‌ای از ویژگی‌های کمی که جسم را به صورت خاصی نمایش می‌دهد وارد می‌شود و پس از کم کردن از میزان داده‌ها در مورد هرجسم در حالی که اطلاعات اصلی ومهم حفظ شوند طبقه‌بندی صحیح اطلاعات انجام می‌شود [۱۱و۱۲].

مدل BEL[۱۰]، یک مدل هیجانی مغز است که مطابق با سیستم لیمبیک مغز انسان می‌باشد. در مدل BEL، شبکه در حین آموزش پارامترهایی که برای رسیدن به جواب مورد نیاز است یاد می‌گیرد که اینکار را با استفاده از پاداش یک تابع هزینه که بصورت تقویتی است محاسبه کند [۱۳-۱۵]. مدل BEL، در سیستم‌های مهندسی می‌تواند سبب بالا بردن درجه آزادی و ظرفیت کنترل و قابلیت اطمینان و پایداری شود. همچنین عملکرد کنترلگر BEL روی سیستم‌های غیرخطی گوناگون، سازگاری پایدار و بالایی را نشان داده است.

یکی از مدل‌هایی که اخیرا ارائه شده مدل الهام گرفته زیست شناختی (BIM[۱۱])، بوده است. این مدل به‌گونه‌ای طراحی شده که بتواند مانند مغز انسان عمل کند. واحدهای موجود در لایه‌های آن

با سلول‌های قشر بینایی به سادگی قابل تطبیق هستند. این واحدها ورودی‌های اولیه را با استفاده از یک گروه از فیلتر گابور که هر کدام از آن‌ها محصول یک بیضی گاوسی و یک صفحه مختلط است ترکیب می‌کنند. از مزایای فیلتر گابور این است که مانند چشم انسان عمل می‌کند اما این نوع ورودی متراکم منجر به هزینه محاسباتی سنگین در واحدهای آن می‌شود چون هر تصویر با فیلترهای گابور با پارامترهای مختلف، درهم‌تنیده خواهد شد [۱۶]. BIM، به دلیل نبود یک مرحله بازخورد، به طور کورکورانه ویژگی‌ها را انتخاب می‌کند تا نشان دهد کدام یک از ویژگی‌ها مهم هستند. سئرری و همکارانش تلاش کردند شکاف بین دید کامپیوتر و علوم اعصاب را با مدل استانداردی که برای بازشناسی جسم در دنیایی واقعی ارائه دادند از بین ببرند [۱۳].

در سال‌های اخیر، پژوهشگران از مدل‌های یادگیری عمیق و شبکه‌های عصبی برای تحلیل تصاویر مغزی و شبیه‌سازی فعالیت‌های مغزی در افراد مبتلا به ادراک پریشی استفاده کرده‌اند. مدل‌های AI قادر به شناسایی الگوهای پیچیده در داده‌های تصویری مغز هستند که ممکن است برای انسان‌ها پیچیده و زمان‌بر باشد. در پژوهشی در سال ۲۰۲۰، الگوریتم‌های شبکه عصبی پیچشی[۱۲] برای تحلیل تصاویر مغزی بیماران مبتلا به ادراک پریشی استفاده کرده و توانسته است الگوهای مربوط به آسیب‌های مغزی را شناسایی کند [۱۷]. در سال ۲۰۲۱، به کارگیری شبکه‌های عصبی برای شبیه‌سازی نحوه کارکرد مناطق مغزی مختلف در تشخیص چهره‌ها (که اغلب در اختلالات ادراک پریشی مشاهده می‌شود) پرداخته شده است [۱۸]. مدل‌های یادگیری عمیق می‌توانند نقش کلیدی در توسعه برنامه‌های توانبخشی مبتنی بر واقعیت مجازی برای بیماران مبتلا به ادراک پریشی ایفا کنند [۱۹].

سیکری و همکاران در تحقیقی، تشخیص بیماری آگنوزیای درکی را به صورت یک گزارش موردی تشریح کردند. در آگنوزیای درکی باتوجه به ماهیت عوامل آسیب، اختلال بصری اولیه وجود خواهد داشت. هوش فضایی پایین مانع از تشخیص ویژگی‌های منحصر به فرد محرک‌های بصری شده بود بطوری‌که جهت خطوط، تشخیص لبه‌های یک محرک بصری سطح اولیه از تجزیه وتحلیل لازم برای درک شکل در آن‌ها وجود نداشت [۲۰].

اریکسون و همکاران در تحقیقی، آگنوزیای بصری شدید را در یک کودک با الگوی الکتروفیزیولوژیک ازخروجی، با موج پیوسته درطول خواب آرام درمناطق پس سری-گیجگاهی مورد بررسی قراردادند. کودک باتشنج پراکنده دچار ناتوانی درتشخیص اشیا بدون اختلال در بینایی بود و علایم در قسمت پس سری-گیجگاهی مشاهده شده است [۲۱].

فولشم و همکاران، از بیمارانی که در تشخیص شی عمیق (افرادمبتلا به آگنوزیای بصری) مشکل داشتند گزارشی تهیه کردند و تفاوت آنها را با افراد سالم با در نظر گرفتن مسیر حرکت چشمشان وقتی به تصویر نگاه می کردند و در تصاویر کنترل می شدند نشان دادند. آنان دریافتند که حرکات چشم می‌تواند به عنوان انتخاب نقطه‌ها، مدل شود و نشان می‌دهد که افرادآگنوزیک دارای مشکل اساسی بر روی روشنایی وکنتراست به عنوان ویژگی‌های سطح پایین محرک دارند. [۲۲].

عملاً درمان مستقیمی برای بیماران ادراک پریشی وجود ندارد. گفتار درمانی یا کار درمانی ممکن است به جبران بیماری کمک کند. از دیدگاه سیستمی، انسان‌ها از ماشین‌ها بهتر عمل می‌کنند و اینکه بتوان یک دیدگاه سیستمی شبیه‌سازی کرد که بتواند بازشناسی جسم را به‌خوبی شناسایی کند یکی از رویاهای دانشمندان علوم اعصاب و مهندسین کامپیوتر بوده است ولی در علوم مهندسی سعی شده است تا با مدل‌سازی ذهن افراد در جهت تشخیص و درمان هوشمند بیماری‌های مغزی-ذهنی گامی برداشته شود. درواقع مدل ادراکی، سعی در ارائه ادراک افراد دارد تا بتواند بر اساس این مدل‌ها به افرادی که در ادراک دچار مشکل هستند کمک کند [۲۳] . با توجه به اینکه به بیماری ادراک پریشی در علوم مهندسی کمتر پرداخته شده است و برخلاف بیشتر پژوهش‌های پیشین که ادراک پریشی را از منظر علوم اعصاب و روان‌پزشکی بررسی کرده‌اند، این مقاله رویکردی مهندسی اتخاذ کرده و با مدل‌سازی ذهن، راهکاری محاسباتی برای کمک به بیماران مبتلا به ادراک پریشی بصری ارائه داده است. در این مقاله به ارائه مدلی برای کمک به بیماران ادراک پریشی پرداخته شده است تا بتوان با استفاده از مزایای مدل محاسباتی ذهن و ترکیب آن با مسیر ادراک بینایی، مشکل پردازش و تشخیص تصویر در این افراد بهبود داده شود. در این مقاله مدل‌های هوشمندی ارائه و استفاده شده که علاوه براینکه دقت بهتری در طبقه‌بندی نسبت به روش‌های گذشته دارد و عملکرد ذهن را بهتر نشان می‌دهد حتی می‌توان با این مدل‌ها مسیرهای بیولوژیک جدیدی را مدل‌سازی کرد. ابتدا مسیر بینایی افراد طراحی شده سپس جهت ادراک تصویر از نگاشت ادراکی فازی(FCM[۱۳]) استفاده می‌شود [۲۴]. در FCM، اشیا، فرایندها و ارتباطاتشان ارزش‌دهی می‌شود و در واقع برای مدل‌سازی سیستم پیچیده ادراک تصویر، مناسب است و منطبق برمنطق فازی و شبکه‌های عصبی می‌باشد. حال اگر تصویر ادراک شده چهره باشد باز هم در ادراک حالت هیجانی چهره، این افراد دچار مشکل هستند بنابراین این مساله نیز با مدل زیست شناختی EBIF[۱۴]، (ویژگی‌های الهام‌گرفته

از زیست شناختی زودهنگام) قابل حل می‌باشد. علت انتخاب این مدل این است که الهام‌گرفته از مدل سیستم عصبی و مغز انسان می‌باشد و برای استخراج ویژگی چهره جهت بازشناسی حالت هیجانی چهره مناسب است در واقع استفاده از فیلتر گابور در این مدل که خاصیت تفکیک‌پذیری خوب و چندگانه در حوزه مکان و فرکانس دارد و یک ابزار قوی برای تحلیل بافت محسوب می‌شود سبب بهبود نرخ بازشناسی می‌شود [۲۵].

مدل پیشنهادی براساس مسیر بینایی و درک تصویر در افراد نرمال طراحی شده است بنابراین در این قسمت به تشریح مسیر بینایی و بازشناسی اشیا در افراد نرمال پرداخته می‌شود.

۱-۱- مسیر بازشناسی اشیا در مغز

در کلیه افراد مراحل پردازش اطلاعات برای تشخیص اشیا دارای سه مرحله است:

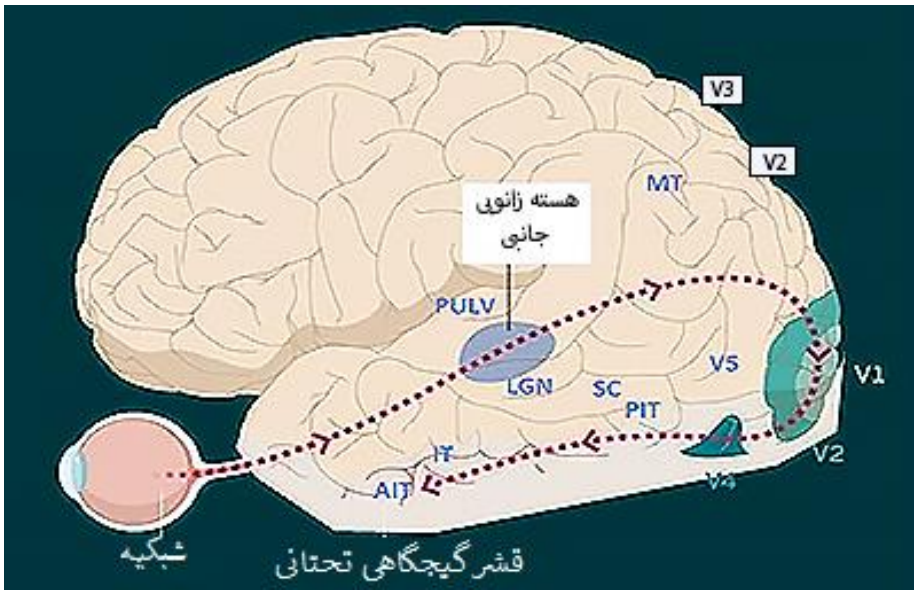
مرحله اول: دریافت محرک بینایی؛ مرحله دوم: درک شی؛ مرحله سوم: تشخیص شی. بعد از دریافت محرک، افکار فرد، همه اطلاعات را بصورت تصویر در می‌آورد و به عبارتی اطلاعات به شکل تصویر نمایان می‌شود و سپس افراد این تصویر را به آن چیزی که از گذشته می‌دانستند ارتباط داده و معنی آن را درک می‌کنند. انسان قادر است بسیاری از اشیای پیرامون خود را بدون کوچکترین مشکلی بازشناسی کند هر چند ممکن است این اشیا در حالت‌های گوناگون و با زوایای دید مختلف و همچنین در اندازه های متفاوت باشند. حتی انسان قادراست اشیا را درحالتی که بخش‌هایی از آنها را نمی‌بیند یا شی دیگری در مسیر دیدش قرار گرفته را نیز بازشناسی کند. اگرچه این امر برای انسان و پستانداران بسیار ساده و عملی است اما در نوع خود یک فرایند محاسباتی بسیار مشکل و پیچیده است [۲۶].

غشای مغز از سلول‌های عصبی تشکیل شده که لایه خارجی مغز را می‌سازد. این غشا فعالیت‌های پیچیده‌ای را مانند حافظه، یادگیری، حل مسئله، طرح‌ریزی، بینایی ، شنوایی وحرکت را کنترل می کند. غشای مخ به دو نیم کره تقسیم شده که هر کدام به چهار ناحیه‌ی گیجگاهی، پیشانی، آهیانه و پس‌سری تقسیم شده است. بخش پس‌سری، در قسمت عقب قرار دارد و اطلاعات بینایی را پردازش می‌کند.

غشای بینایی به نواحی مختلفی تقسیم شده است.V1 [۱۵]و V2 بزرگترین ناحیه هستند و مساحت هرکدام حدود ۱۱۰۰ تا۱۲۰۰ mmاست. اطلاعات رسیده از شبکیه، از طریق LGN از تالاموس به V1 در دو مسیر جداگانه قرار می‌گیرند و بطور مستقل پردازش می‌گردند. شکل ۱، مسیر بازشناسی اشیا را بر روی مغز نشان می‌دهد.

۱-۱-۱- مسیر تشخیص شی (شاخه شکمی)

کار این مسیر شناسایی ویژگی اشیاء (رنگ، شکل و...) است و در مسیر V1، V2، V4، AIT و PIT قسمت پس‌سری قرار دارد. درجریان شکمی چیزی که تجسم می‌شود خروجی همان چیزیست که آگاهانه می توان دید وتوصیف کرد. این جریان تفسیرذهنی از جهان اطراف ما ایجاد می کند. رنگ، اندازه نسبی، بافت و شکل، همه در این منطقه پردازش می‌شوند. هنگامی که افراد یک شی را



شکل ۱: مسیر بازشناسی اشیا

مشاهده می کنند یک بازنمایی ذهنی از آن شی ایجاد می‌شود و این بازنمایی حداقل به صورت نیمه دائم به آنچه که به عنوان حافظه به آن مراجعه می‌شود تبدیل می‌گردد [۲۷].

۱-۱-۲- مسیر تشخیص مکان شی (شاخه خلفی)

کار این مسیر شناسایی مشخصات فضایی صحنه (جهت حرکت و..) است و در مسیر V1، V2، V3، MT[۱۶] و MST قرار دارد. جریان خلفی، ورودی بصری را پردازش می‌کند بنابراین اغلب به عنوان موتور بینایی نامیده می‌شود. جریان پشتی برخلاف شکمی خاطرات دائمی را شکل نمی‌دهد بلکه بروزرسانی مداوم اطلاعات ضبط شده را به مدت دو ثانیه ارائه می‌دهد [۲۸].

محققین در ترسیم مسیر بینایی، اظهار داشت که نواحی بصری شیاردار نه تنها به‌صورت آناتومیک بلکه به‌صورت عملکردی هم به خوبی می‌تواند مجزا شود [۲۸]، که شامل مسیر تکتوپولوینار[۱۷] (مسیری برای کنترل موتور بصری)، مسیر زیر قشری شامل فرآیندهای sc و pulv برخی جنبه‌های درک ناخودآگاه، مسیر ژنیکولوستریا[۱۸] (مسیری برای ادراک بصری) می‌باشد. بیمارانی که دارای صدمات لوب راست یا چپ گیجگاهی هستند در تمایز قایل شدن و تشخیص تصاویر و حتی در مهارت‌هایی که با تشکیل اشکال هندسی ارتباط دارد با مشکل مواجه‌اند. بیماران تحت صدمه لوب راست یا چپ گیجگاهی، در یادآوری و شناسایی اشیا و به خاطرسپاری مکان آن‌ها، با مشکلاتی روبرو هستند.

پردازش اطلاعات بینایی در منطقه V4، با الگوهای حافظه ذخیره شده ادغام می شود. آسیب این مسیر می‌تواند منجر به بیماری ادراک پریشی بصری شود [۲۸].

در نهایت، هدف اصلی مقاله، جبران اختلال در درک تصویری این بیماران از طریق شبیه‌سازی مسیرهای بیولوژیکی مغز است. ابتدا تصویر با FCM شناسایی شده و در صورت تشخیص چهره، بازشناسی هیجانات با مدل EBIF انجام می‌شود. نتایج نشان داده‌اند که این مدل‌ها علاوه بر افزایش دقت طبقه‌بندی، عملکرد مغز را در حوزه ادراک بهتر شبیه‌سازی می‌کنند. این مقاله بدین صورت سازماندهی شده است که پس از بیان مقدمه، در بخش دوم، مواد و روش پیشنهادی ارائه می‌شود. در بخش سوم به تحلیل نتایج پرداخته می‌شود و در بخش چهارم نتیجه‌گیری بیان خواهد شد.

۲-مواد و روش‌ها

بر خلاف اغلب پژوهش‌های پیشین که بیماری ادراک پریشی را از دیدگاه علوم اعصاب یا روان‌پزشکی بررسی کرده‌اند، در این مقاله سعی شده است از منظر مهندسی و با استفاده از مدل‌سازی ذهن انسان، راهکاری محاسباتی برای کمک به این بیماران ارائه شود. این مقاله دو مدل محاسباتی را به‌صورت تلفیقی پیشنهاد داده است:

- نگاشت ادراکی فازی (FCM) : برای شبیه‌سازی مسیر بینایی و ادراک تصویری در بیماران که یک گراف جهت‌دار وزن‌دار است که در آن گره‌ها (یا مفاهیم) نمایانگر فاکتورها، متغیرها یا اجزای سیستم هستند و یال‌ها روابط علّی بین این مفاهیم را نشان می‌دهند. وزن یال‌ها (بین ۱- و ۱) میزان و نوع تأثیر یک مفهوم بر دیگری را نشان می‌دهد (مثبت، منفی یا خنثی) [۲۴].

- مدل EBIF : برای شناسایی و بازشناسی حالت‌های هیجانی چهره که مبتنی بر ویژگی‌های الهام‌گرفته از زیست شناختی مغز و استفاده از فیلتر گابور است. این مدل، به ویژگی‌هایی اشاره دارند که بر اساس نحوه پردازش اطلاعات بصری در مراحل ابتدایی سیستم بینایی مغز، به‌ویژه در ناحیه V1 قشر بینایی، طراحی شده‌اند. این ویژگی‌ها معمولاً شامل پاسخ به لبه‌ها، بافت‌ها، فرکانس‌های فضایی و جهت‌های خاص هستند و اغلب با استفاده از فیلترهایی مانند گابور مدل‌سازی می‌شوند. این فیلترها عملکرد نورون‌های بینایی اولیه را تقلید کرده و به استخراج اطلاعات ساختاری مهم از تصاویر کمک می‌کنند [۲۵].

مقاله با استفاده از این مدل‌ها سعی در مدل‌سازی مسیرهای بیولوژیکی و ذهنی بازشناسی تصویری دارد که در مطالعات گذشته یا مطرح نشده یا به صورت کامل به آن پرداخته نشده است. هدف این مقاله ابتدا بهبود بازشناسی تصاویر و همچنین بازشناسی حالات هیجانی چهره در ذهن افراد مبتلا به ادراک پریشی بصری است. در واقع، در این مقاله با توجه به مسیر درک تصویر و در نظر گرفتن قسمت دچار مشکل افراد اگنوزیک، به جبران مشکل عدم پردازش صحیح تصاویر در مغز و ذهن این افراد پرداخته می‌شود. در افراد نرمال، پس از مشاهده یک تصویر، اطلاعات تصویر در مسیر مربوط به بینایی شروع به بالارفتن از طبقات مختلف قشر بینایی مغز می‌کند تا به مرحله ادراک تصویر برسد. از آنجا که بیماران مبتلا به ادراک پریشی معنا و ویژگی‌های اشیا را نمی‌توانند تشخیص دهند

بنابراین، در این مقاله برای جبران این کمبود، ابتدا با استفاده از نقشه ادراکی فازی، تصویر تشخیص داده می‌شود و درصورتیکه چهره، تشخیص داده شود سپس با استفاده از ویژگی‌های الهام‌گرفته از مدل زیست شناختی زودهنگام اقدام به بازشناسی حالت هیجانی چهره خواهد شد. نتایج به‌دست‌آمده در این پژوهش نشان داده است

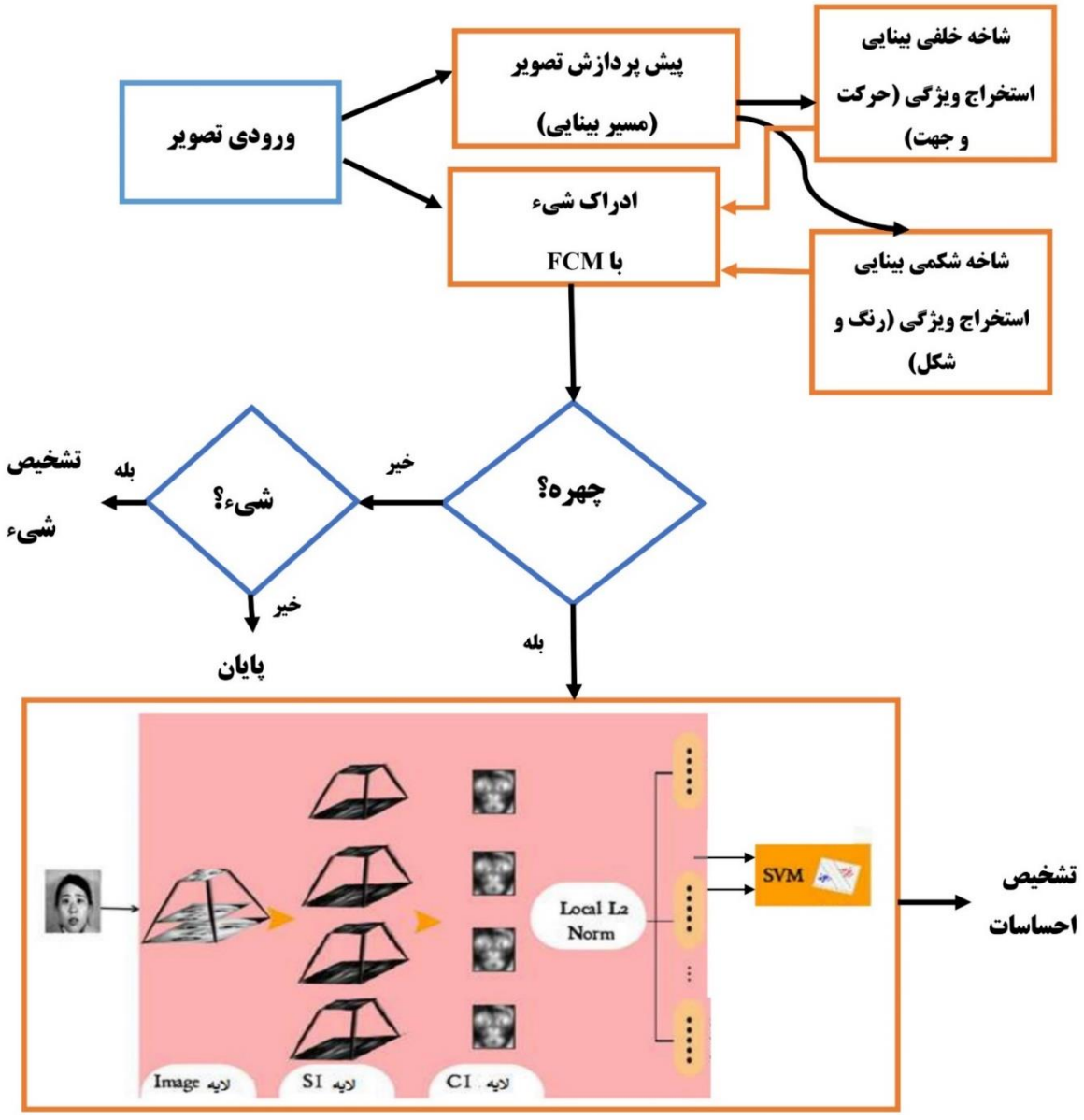
که مدل‌های پیشنهادی نه تنها دقت طبقه‌بندی را نسبت به روش‌های پیشین افزایش می‌دهند، بلکه به شبیه‌سازی دقیق‌تر عملکرد مغز انسان در زمینه ادراک کمک می‌کنند. در ادامه به بررسی جزئیات مدل پیشنهادی پرداخته خواهد شد.

۲-۱-مدل پیشنهادی

با توجه به عدم درک اشیا و چهره در این افراد، می‌بایست مدلی ارائه شود که بتواند هر دو مشکل را حل نماید. مشکل ادراک تصویر با FCM حل خواهد شد و مشکل شناسایی حالات چهره نیز با EBIF مرتفع می‌شود. ساختار مدل پیشنهادی بصورت شکل ۲ است.

همانطور که بیان شد مدل پیشنهادی مبتنی بر مسیر بینایی از دیدن تصویر تا درک آن در مغز بیان شده است. ابتدا تصویر بعنوان ورودی خوانده می‌شود. تصویر ورودی در ماژول Tectopulvinar، پیش‌پردازش (حذف نویز، بهبود رزلوشن و ...) خواهد شد. در ماژول‌های شاخه شکمی و شاخه خلفی، ویژگی‌های تصویر استخراج شده و به ماژول Geniculostriate، جهت ادراک وارد خواهد شد.

مسیر Geniculostriate یکی از اصلی‌ترین مسیرهای پردازش بینایی در مغز انسان است. این مسیر از شبکیه آغاز می‌شود، سیگنال‌های بینایی از طریق عصب بینایی به هسته زانویی جانبی (LGN) در تالاموس منتقل می‌شوند و از آن‌جا به قشر بینایی اولیه V1 هدایت می‌گردند.



شکل ۲: ساختار مدل پیشنهادی

حال اگر تصویر مشخص شده چهره باشد برای شناسایی حالات چهره وارد مدل EBIF شده تا حالات چهره ادراک و شناسایی شود.

مهم‌ترین نوآوری مدل پیشنهادی استفاده و مدلسازی مسیر بینایی در مغز از مشاهده تا ادراک جهت تشخیص شیء و حالات هیجانی چهره است. در واقع، در مدل ارائه شده، استفاده از ابزارهای هوشمند FCM و EBIF، سبب تشخیص بهینه و منطبق‌تر بر ادراک افراد سالم می‌شوند. در زیر گام‌های مدل پیشنهادی به ترتیب بیان شده است و در ادامه هر یک از اجزای مدل پیشنهادی به‌تشریح بیان خواهد شد.

جدول ۱: گام‌های مدل پیشنهادی

گام ۱:ورود تصویر
گام ۲: پیش پردازش تصویر: حذف نویز، تغییر سایز
گام ۳: استخراج ویژگی شامل: رنگ، شکل، جهت
گام ۴: ورود به FCM برای تشخیص
گام ۵: اگر تصویر چهره است
گام ۵-۱:ورود به EBIF و تشخیص حالت هیجانی چهره.
گام ۶:اگر تصویر شیء است، تشخیص داده شود.

۲-۱-۱- استخراج ویژگی

پس از انتخاب تصاویر، باید اقدام به استخراج ویژگی کرد. در اینجا ویژگی‌هایی که بیشترین اهمیت را در جداسازی کلاس‌های الگوهای تصویر دارند استخراج می‌شود. ویژگی‌های استفاده شده در زیر توضیح داده شده است.

الف- لبه‌یابی: برای لبه‌یابی از فیلتر سوبل بر مبنای تقریب مشتق سوبل، استفاده شده است. با استخراج لبه می‌توان نواحی مهم در تصاویر را استخراج کرد. با استفاده از فیلتر سوبل برای استخراج محدوده چهره و اجزای مهم آن چون چشم‌ها و لب‌ها، گام نخست شناسایی چهره انجام می‌شود. در فیلتر سوبل، دو ماسک عمودی

و افقی وجود دارد .ماسک سوبل افقی، بیشتر لبه های افقی را مشخص می کند. ماسک سوبل عمودی، لبه‌های عمودی را مشخص می‌کند.

ب-بافت: ماتریس‌های هم‌رخدادی[۱۹] جهت استخراج ویژگی از بافت استفاده می‌شوند. ویژگی‌های که از این طریق بدست می‌آیند کنتراست، انرژی آنتروپی و همگنی هستند.

- کنتراست

کنتراست، بافت‌ها را اندازه‌گیری می‌کند. در واقع، کنتراست به عنوان یکی از ویژگی‌های آماری تصویر، می‌تواند اطلاعاتی درباره‌ی ساختار فضایی و تغییرات شدت روشنایی در ناحیه‌ای از تصویر که معرف بافت است، فراهم کند و برای این کار از فرمول زیر استفاده می شود [۲۹].

$$C=\sum_i\sum_j (i-j)^2 C(i,j) \qquad (۱)$$

C، کنتراست تصویر است و c(i,j)، شدت نور هر پیکسل را نشان می‌دهد بطوریکه i و j، سطر و ستون پیکسل‌های تصویر می‌باشد.

- انرژی

انرژی، E، یکرنگی بافتی تصاویر را اندازه گیری می‌کند و برای این کار از فرمول زیر استفاده می‌شود [۲۹].

$$E=\sum_i\sum_j C^2(i,j) \qquad (۲)$$

- آنتروپی

آنتروپی، H، درجه بی نظمی را اندازه گیری می کندو برای این کار از فرمول زیر استفاده می‌شود [۳۰].

$$H=\sum_i\sum_j C(i,j)\log\big(C(i,j)\big) \qquad (۳)$$

- همگنی

همگنی، HO، میزان توزیع شدگی المان ها را اندازه گیری می‌کند و برای این کار از فرمول زیر استفاده می‌شود.

$$HO=\sum_i\sum_j \frac{C(i,j)}{1+|i-j|} \qquad (۴)$$

ج) گشتاور رنگ

گشتاورهای رنگ که در آن تشخیص توزیع رنگ در یک تصویر به همان روش که گشتاورهای منحصر به فرد مرکزی یک توزیع احتمال را توصیف می‌کنند اندازه گیری می‌شود. گشتاورهای رنگ به‌طور عمده برای مقاصد شاخص‌گذاری رنگ به عنوان ویژگی هایی در کاربردهای بازیابی تصویر به‌منظور مقایسه چگونگی شباهت دو

عکس براساس رنگ به‌کاربرده می‌شوند [۳۱] .

- میانه

میانه، E_i،که به عنوان اولین گشتاور رنگ است, می‌تواند به عنوان رنگ میانگین در تصویر تفسیر شود و میتوان آن را از رابطه زیرمحاسبه کرد.

$$E_i = \sum_{j=1}^{N} \frac{1}{N} Pij \qquad (۵)$$

که درآن N تعداد پیکسل ها و j شمارنده پیکسل‌ها در تصویر و Pij ارزش پیکسل iام از تصویر در کانال رنگ iام است.

- انحراف معیار

دومین گشتاور رنگ انحراف معیار، σ، استاندارد است، که با در نظر گرفتن ریشه دوم واریانس توزیع رنگ به دست آمده است.

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^{N} \left(p_{ij} - E_i \right)^2 \right)} \qquad (۶)$$

که در آن Ei ارزش میانه یا اولین گشتاور رنگ، برای کانال رنگ iام تصویر است.

- چولگی

گشتاور رنگ سوم چولگی، S_i، است. آن چگونگی توزیع رنگ را می‌سنجد و در نتیجه اطلاعاتی در مورد شکل توزیع رنگ می دهد. چولگی می تواند ازرابطه زیرمحاسبه شود.

$$s_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^{N} \left(p_{ij} - E_i \right)^3 \right)} \qquad (۷)$$

- شاخص گذاری رنگ

گشتاور رنگ، d، را می توان به منظور مقایسه چگونگی تشابه دو تصویر بکار برد. این یک رویکرد نسبتا جدید برای شاخص‌گذاری رنگ است. شاخص‌گذاری رنگ کاربرد اصلی گشتاور رنگ است. تصاویر می‌توانند شاخص گذاری شوند و شاخص شامل محاسبه گشتاور رنگ خواهد بود. بنابراین اگر کسی یک تصویر دارد و می‌خواهد تصاویر مشابه آن را در پایگاه داده پیدا کند، گشتاورهای رنگ تصویر مورد نظر محاسبه می‌شود. پس از آن از تابع زیر به منظور محاسبه نمره شباهت بین تصویر مورد نظر و تصاویر پایگاه داده استفاده خواهد شد [۳۲] .

$$\begin{aligned} d_{mom}(H,I) &= \sum_{i=1}^{r} \omega_{i1} \left| E_i^1 - E_i^2 \right| \\ &+ \omega_{i2} \left| \sigma_i^1 - \sigma_i^2 \right| + \omega_{i3} \left| s_i^1 - s_i^2 \right| \end{aligned} \qquad (۸)$$

E^1 وE^2، بترتیب تصاویر مورد نظر و تصاویر پایگاه داده هستند. ω، بسامد زاویه‌ای است.

۲-۱-۲-طراحی نگاشت ادراکی فازی (FCM)

FCM، یک روش محاسباتی نرم برای مدل کردن سیستم‌ها است و به‌طور گسترده‌ای برای تجزیه و تحلیل سیستم‌های پیچیده و تصمیم‌گیری استفاده می‌شود. FCM یک گراف جهت‌دار برای بازنمایی ارتباطات علی میان چند مفهوم است و کاربرد زیادی در حوزه های مختلف برای پشتیبانی تصمیم‌گیری و طبقه‌بندی وظایف داشته است. نگاشت ادراکی فازی روشی برای ارائه دانش کارآمد و استدلال است که براساس تجربه و دانش انسان می‌باشد و دربرگیرنده‌ی نظرات خبرگان در مورد یک واقعیت ذهنی است و نیاز به ورود داده‌ها و آموزش دارد [۳۳].

ویژگی‌های سودمند FCMها مانند سادگی، پشتیبانی از دانش متناقض، حوزه رابطه بین علت ومعلول برای مدل‌سازی دانش واستنتاج وهمچنین قابلیت یادگیری، آنها را برای بسیاری از حوزه‌های متنوع علمی مدل‌سازی دانش، پیش بینی وتصمیم‌گیری قابل اجرا می‌سازد. در واقع، FCM حوزه‌های خاص را با استفاده از گره‌ها، مفاهیم (متغیرها، وضعیت‌ها، ورودی‌ها و خروجی‌ها)، ارتباطات علی و روابط فازی علامتدار بین آنها که می‌تواند مثبت یا منفی باشد توصیف می‌کند. از روش‌های یادگیری برای آموزش FCMها استفاده می‌شود که شامل بروزرسانی وزن‌های روابط علی می‌باشد. نگاشت ادراکی فازی (FCM) به صورت زیرتفسیر می‌شود.

اگریک FCM، باتعداد N گره C_i باشد مقدار هرگره در هر تکرار بصورت رابطه زیر محاسبه شود.

$$A_i^{t+1} = F\left(A_i^t + \sum_{j=1}^{n} A_j^t w_{ji} \right) \qquad (۹)$$

که در آن A_i^t مقدارمفهوم C_i در زمان t+1، A_j^t مقدار مفهوم C_i در زمان t، A_i^t متناظراست با W، وزن فازی بین دو گره و F تابع آستانه است که نتیجه ی ضرب را تبدیل به عددی در بازه [۱-۰] می‌کند. تابعی که در این مقاله استفاده شده است تابع لجستیک است که در رابطه زیر آمده است.

$$F(x) = \frac{1}{1 + e^{-\lambda \chi}} \qquad (۱۰)$$

λ، میزان شیب منحنی، x نقطه میانه سیگموید و هرگره یک مجموعه فازی است که می تواند از ۰ تا ۱۰۰ درصد برانگیخته شود. FCM، اشیا و فرایندها را به ارزش‌ها ارتباط می‌دهد. روش مدل‌سازی FCM، منطبق برمنطق فازی و شبکه های عصبی است که بصورت پسخور رشد می‌کنند. ابتدا یک FCM ارزش‌گذاری اولیه می‌شود سپس سطح فعال‌سازی هرگره از سیستم مقدار خاصی می‌گیرد و مفاهیم مختلف در اثرگذاری متقابل آزاد هستند. فعال‌سازی یک گره روی گره‌های دیگر که به آن متصل است اثر می‌گذارد. این کار تا زمانی که سیستم به نقطه ثابت تعادل یا یک دور محدود و یا یک رفتار آشوب برسد ادامه می‌یابد [۳۴].

در این مقاله، یک FCM، با ۱۰۰ گره و ۵۰ مفهوم در نظر گرفته شده است. مقدار هر گره، ویژگی‌های استخراج شده از هر تصویر است. در مرحله اول میزان وزن بین گره‌ها بصورت تصادفی بین ۰ تا ۰,۵ است. در ادامه با توجه به ورودی‌ها و ارتباط بین گره‌ها که گره‌های مشابه وزن‌های بیشتری بین آنها قطعا به وجود خواهد آمد. ضمن اینکه فاصله بین مقادیر با استفاده از فاصله اقلیدسی محاسبه می‌شود. مقدار گره‌ها با هر ورودی تغییر می‌کند. رویکرد تغییر وزن و عملیات درون شبکه براساس روش پسخور می‌باشد.

۲-۱-۳- مدل EBIF

مدل EBIF بهبود یافته مدل BIM است که در اینجا جهت استخراج ویژگی برای بالا بردن نرخ بازشناسی حالات هیجانی از آن استفاده شده است. این مدل شامل چهار لایه Image Layer، S1Layer، C1 Layer، Local L2-norm Layer می‌باشد که وظایف هر کدام در زیرشرح داده می‌شود [۳۵].

الف- لایه تصویر

در لایه تصویر، ابتدا تصاویر به خاکستری (gray scale) تبدیل می‌شوند و سپس بصورت هرمی در N مقیاس (در این پژوهش N=6) از بزرگ به کوچک بدست آورده می‌شوند. هر مقیاس از تصویر بالایی، فاکتوری از $\sqrt{2}$ کوچکتر از تصویر جاری می‌باشد. هرم‌ها به صورت مقیاس K تا K+1 هستند بطوریکه (k=1,2,…,N-1) می‌باشد. در واقع تصاویر بعد از خاکستری شدن به یک هرم در شش مقیاس تبدیل می‌شوند. تصاویر پایگاه داده زمانی که به صورت مقیاس‌های K تا K+1 ایجاد شدند وارد لایه S1 می‌شوند.

ب- لایه S1

در لایه S1 از بانک فیلتر گابور استفاده می‌شود. فیلترهای گابور برای استخراج ویژگی از نواحی صورت استفاده می‌شود. مهم‌ترین دلیل استفاده از فیلترهای گابور مقاومت در مقابل چرخش و مقیاس است. علاوه‌براین در برابر اختلالات فتومتریک مانند روشنایی، نویز تصاویر و غیره نیز مقابله می‌نماید. فیلترهای گابور درصورتی که به طور مناسب و دقیق تنظیم شوند، عملکرد بسیار مناسبی در تشخیص ویژگی های بافت و لبه بافت دارند. ویژگی مهم دیگر فیلترهای گابور درجه تفکیک مشترک بالای آنهاست. این بدان معنی است که پاسخ آنها هم در حوزه مکان و هم در حوزه فرکانس کاملا محلی و قابل تنظیم کردن است. فیلتر گابور با M جهت‌گیری، تمام قسمت‌های تصاویر هرمی را فیلتر می‌کند. هسته کانولوشن فیلتر گابور حاصلضرب یک تابع نمایی مختلط و گوسین است. فرمول استفاده شده در فیلترگابور در مدل پیشنهادی در زیر آمده است.

$$G(x,y)=\exp(-\frac{X^2+\gamma^2Y^2}{2\sigma^2})\cos(\frac{2\pi}{\lambda}X) \qquad (۱۱)$$

در اینجا $X= x\cos\Theta+y\sin\Theta$ و $Y=-x\sin\Theta+y\cos\Theta$ می‌باشد. بطوریکه Θ، جهت‌گیری‌های فیلتر را کنترل می‌کند. بانک فیلتر شامل M جهت‌گیری است که در اینجا M=9 می‌باشد. جهت‌گیری‌ها به صورت مساوی در محدوده $[\pi,0)$ توزیع می‌شود. سایز هر فیلتر ۵×۵ می‌باشد و x و y بین ۲ و ۲- متغیر است. پارامترهای نسبت ابعاد، عرض موثر و طول موج به ترتیب ۱، ۳ و ۵ تنظیم شده است. در معادله (۱۲) پاسخ فیلتر گابور به یک تصویر آمده است.

$$R=G\times I \qquad (۱۲)$$

پاسخ های فیلتر گابور می‌توانند شی را محلی‌سازی کند و ساختار آن را متمایز و با جهت‌گیری‌های مختلف بکند.

ج- لایه C1

لایه C1 مطابق سلول‌های پیچیده قشر بصری عمل می‌کند. در این قسمت هرم گابور بطور مساوی مطابق با یک جهت خاص تقسیم شده و زیرهرم‌ها تنظیم می‌شود. در پایین، سایز هرم‌ها ۴×۴ و در بالا سایز ۳×۳ می‌باشد که مطابق با مدل پایه BIM است. در این لایه توسط MaxPooling کاهش بعد انجام می‌شود. در واقع از ترکیب هرمی استفاده می‌شود و از میانگین و انحراف استاندارد به جای بیشترین پاسخ برای نمایش هر زیرهرم استفاده می‌شود. مزیت استفاده از لایه C1 این است که هر عنصر در این لایه، با ترکیب پاسخ فیلتر گابور محلی که مبنای آشکارسازی لبه و موقعیت کمی و مقیاس قابل تحمل بر روی موقعیت همسایه و جهت گیری‌های متفاوت است بدست می‌آید.

- کاهش ابعاد توسط Pooling

کارکرد Pooling کاهش ابعاد بصورت عمقی می‌باشد. به این صورت که اندازه مکانی (عرض و ارتفاع) تصویر را به جهت کاهش تعداد پارامترها و محاسبات در داخل شبکه کم می‌کند. لایه Pooling بصورت مستقل روی هر برش عمقی از توده ورودی عمل کرده و آن را با استفاده از عملیات Max، از لحاظ مکانی تغییر اندازه می‌دهد. رایج‌ترین شکل استفاده از این لایه به صورت استفاده این لایه با فیلترهایی با اندازه ۲×۲ به همراه گام ۲ (S=2) است که هر برش عمقی در ورودی را با حذف دو عنصر از عرض و دو عنصر از ارتفاع کاهش داده و باعث حذف ۷۵٪ مقادیر موجود در آن برش عمقی می‌شود. هر عملیات Max در اینجا ماکسیمم بین ۴ عدد (یک ناحیه ۲×۲ در برش عمقی) را بدست می‌آورد. بطور کلی لایه Pooling، یک توده با اندازه D1×H1×W1 را بعنوان ورودی دریافت می‌کند. که W نشانگر عرض، H نشانگر ارتفاع و D نشانگر عمق آن می‌باشد. Pooling نیازمند دو فراپارامتر است:

- اندازه وسعت مکانی (اندازه (x و y فیلترها) ناحیه ادراکی) F
- اندازه گام یا S

سپس یک توده خروجی با اندازه D2×H2×2 تولید می‌کند که:

$$W2 = \frac{(W1 - F + 2P)}{(S+1)} \qquad (۱۳)$$

$$H2 = \frac{(H1 - F + 2P)}{(S+1)} \qquad (۱۴)$$

روابط بالا نشان می‌دهد که عرض و ارتفاع هر دو بطور مساوی بصورت متقارن محاسبه می‌شوند. Max Pooling همچنین کار حذف نویز (Noise Suppressant) را انجام می‌دهد. MaxPooling، به علت همگرایی سریعتر و تعمیم بهتر و انتخاب ویژگی‌های بهینه، بسیار خوب عمل می‌کند.

د- لایه نرمال محلی L2

در لایه نرمال محلیL2، واریانس روشنایی که ویژگی بسیار مهمی برای بازشناسی چهره می‌باشد محاسبه می‌شود. برای داشتن یک مدل EBIF قدرتمند در برابر تغییرات نور ابتدا باید نرمال‌سازی شود. ابعاد ویژگی بردار f از هر مکان لایه C1 با L2-norm بصورت زیر محاسبه می‌شود.

$$\tilde{f} = \frac{f}{\sqrt{\|f\|^2 + \xi}} \qquad (۱۵)$$

در اینجا ξ یک ثابت است. ویژگی‌های بردار، در موقعیت‌های متفاوت از لایه C1 از فرم ویژگی زیرمجموعه‌ها برای پاسخ شی‌های مقیاس k، نرمال‌سازی می‌شود. همه N-1ها، زیرمجموعه ویژگی‌های چند مقیاسی هستند که حاوی اطلاعات لبه و چند جهت‌گیری نرمال شده محلی هستند که موقعیت کمی و مقیاس قابل تحمل از پاسخ‌دهی ویژگی EBIF نهایی می‌باشند.

۲-۲-دسته‌بندی ماشین بردار پشتیبان

ماشین بردار پشتیبان (SVM) یک الگوریتم طبقه بند است که یکی از بهترین تکنیک‌های دسته‌بندی بشمار می‌آید. در SVM، انتخاب کرنل‌های غیرخطی به ما اجازه‌ی ساخت جدا کننده‌های خطی در فضای ویژگی را می‌دهد در صورتی‌که آنها در فضای اصلی غیرخطی هستند. همچنین در حل مسائل محاسباتی که دارای ابعاد زیادی هستند بسیار خوب عمل می‌کنند. بعلاوه استفاده از ابعاد نامتناهی را ممکن می‌سازد و از لحاظ زمانی و حافظه کارآمد است[۳۵]. در این مقاله در SVM، از کرنل چند جمله‌ای که بهترین نتیجه را حاصل کرد، استفاده شده است و معادله آن در زیر آمده است.

$$K(x_i, x_j) = (\mathbf{1} + x_i^T x_j)^{\mathbf{d}} \qquad (۱۶)$$

بطوریکه x، داده‌های آموزشی و d=۲ ، درجه تابع برای کرنل چند جمله‌ای است.

۳-نتایج آزمایش‌ها

در این آزمون، از پایگاه داده کورل 5K (شامل پنج کلاس آفریقا، ساحل، بناها، اتوبوس و دایناسور) و پایگاه داده چهره JAFFE (ژاپنی) استفاده شده است. که شامل ۲۱۳ تصویر چهره‌های احساسی زنان ژاپنی است. شکل ۳، تصاویر پایگاه داده‌های مورد استفاده، نشان داده شده است. بطور کلی از ۵۰۰ تصویر که تصاویری از دو پایگاه داده می‌باشد در آموزش و آزمایش مدل پیشنهادی استفاده شده است. ۳۰۰ تصویر از پایگاه داده کورل و ۲۰۰ تصویر از پایگاه داده JAFFE مورد استفاده قرار گرفته است. در فاز دوم مدل پیشنهادی که تشخیص چهره می‌باشد فقط از تصاویر پایگاه داده JAFFE استفاده می‌شود.

بعد از فراخوانی تصاویر مجموعه داده، از تصاویر استخراج ویژگی شده و عمل آموزش انجام می‌شود. در مرحله تست، نوع و کلاس آن تصویر تشخیص داده خواهد شد. برای اعتبارسنجی مدل پیشنهادی از k-fold و k=10 استفاده شده است.



(الف) (ب)

شکل ۳: تصاویر پایگاه داده‌های مورد استفاده. الف) تصاویر پایگاه داده JAFFE، ب) پایگاه داده کورل 5K

در آزمایش‌های انجام شده مشاهده شده است که هنگامی که k=10 است مدل پیشنهادی از راندمان بیشتری برخوردار است. این مقدار از طریق تجربی و با آزمایش حاصل می‌شود. در این نوع تقسیم، داده‌ها به k-1 دسته آموزش و ۱ دسته آزمایش تقسیم می‌شوند. در این نوع اعتبارسنجی داده‌ها به K زیرمجموعه افراز می‌شوند. از این Kزیرمجموعه، هر بار یکی برای اعتبارسنجی و K-1 تای دیگر برای آموزش بکار می‌روند. این روال K بار تکرار می شود و همه داده‌ها دقیقا یکبار برای آموزش و یکبار برای اعتبارسنجی بکار می‌روند.

برای تعیین کارایی مدل پیشنهادی از معیارهای دقت[۲۰]، صحت[۲۱] و بازیابی[۲۲] استفاده شده است. روابط زیر نحوه محاسبه آنها را نشان می دهد.

$$Accuracy = \frac{T_p + T_n}{T_p + F_p + F_n + T_n} \quad (۱۷)$$

$$\Pr ecision = \frac{T_p}{T_p + F_p} \quad (۱۸)$$

$$\mathrm{Re}call = \frac{T_p}{T_p + F_n} \quad (۱۹)$$

که Fn، تصاویر مرتبط بازیابی نشده، Tp تصاویر مرتبط بازیابی شده، Tn تصاویر نامرتبط بازیابی نشده و Fp تصاویر نامرتبط بازیابی شده است. جدول ۱ مقایسه میزان دقت روش پیشنهادی برای بازیابی انواع مختلف تصاویر در تعداد تصاویر مختلف با روش‌های SVM و شبکه عصبی را نشان می‌دهد.

همانطور که مشاهده می‌شود مدل پیشنهادی برای انواع تصاویر در تعداد تصاویر بیشتر بهتر عمل می‌کند و با دقت ٪۹۸/۱، موفق به بازشناسی اشیا شده است در صورتی‌که SVM، با ٪۹۱/۴ و شبکه عصبی با ٪۸۱/۴، به عمل بازشناسی پرداخته است. کارایی مدل پیشنهادی با تعدادی مدل رقیب نیز مقایسه شد و این مقادیر در جدول ۲ گزارش شده است.

در ادامه، پس از تشخیص نوع تصویر، اگر تصویر شناسایی شده چهره باشد می‌بایست حالت آن چهره با مدل ارائه شده شناسایی شود. همانطور که در جدول ۲ مشاهده می‌شود مدل پیشنهادی به ترتیب با دقت، صحت و بازیابی٪۹۴/۶ ، ٪۹۳/۸ و ٪۹۴/۲، عملکرد بهتری در هر مرحله از خود نشان داده است. همانطورکه مشاهده می شود مدل پیشنهادی بدلیل استفاده از منطق فازی توانسته است نسبت به دیگر مدل‌ها از دقت بالاتری برخوردار باشد. در مدل پیشنهادی مبتنی بر FCM، گره‌ها نشان‌دهنده مفاهیم (مثلاً ویژگی‌های تصویری یا حالات هیجانی) هستند و یال‌ها روابط علت-معلولی بین این مفاهیم را مدل‌سازی می‌کنند و این روش برخلاف Naive Bayes است که فرض استقلال ویژگی‌ها را دارد و در مدل BEL نیز وابستگی‌ها بر پایه احتمال شرطی است اما قدرت نمایش روابط دینامیک و چرخه‌ای در آن محدودتر از FCM است بنابراین روش پیشنهادی بهتر عمل نموده است.

جدول ۱: مقایسه میزان دقت بازشناسی انواع تصاویر با تعداد تصاویر مختلف در مدل پیشنهادی و دیگر روش‌ها با استفاده از تصاویر دو پایگاه داده

تعداد تصاویر	100	200	300	400	500
مدل پیشنهادی	82٪	85.30٪	86.90٪	94.2٪	98.1٪
SVM [۳۵]	67٪	69.4٪	73.1٪	86.3٪	91.4٪
شبکه عصبی [۳۵]	42٪	58.2٪	63.5٪	78٪	81.4٪

جدول ۲: مقایسه کارایی مدل پیشنهادی با مدل‌های رقیب برای بازشناسی انواع تصاویر با استفاده از تصاویر دو پایگاه داده

معیارارزیابی	مدل پیشنهادی	یادگیری Zero-shot [۲۶]	H-MAX [۲۳]	Naive Bayes [۳۷]	مدل بل [۱۳]
دقت	94.6٪	93.1%	92.8%	83.3٪	91.2٪
صحت	93.8%	91.5%	92.1%	82.1%	90.6%
بازیابی	94.2%	92.9%	91.7%	84.3%	92.1%

جدول ۳: مقایسه میانگین دقت مدل پیشنهادی با دو مدل رقیب در حالات مختلف هیجانی با استفاد ه از پایگاه داده JAFEE

حالت	[۱۳] BEL (m±SD)	[۳۶] Moren (m±SD)	EBIF (m±SD)	p-value
ترس	96.96±0.012	96.75±0.015	98.32±0.014	0.02
غم	98.2±0.032	94.4±0.011	99.1±0.012	0.04
خشم	97.3±0.018	94.3±0.030	99.4±0.017	0.02
شادی	97.9±0.022	97.1±0.022	99.2±0.002	0.01
تعجب	96.3±0.002	95.5±0.011	99.03±0.011	0.02

مدل پیشنهادی با حالات مختلف با دو مدل رقیب، Moren و BEL، مورد مقایسه قرار گرفته و نتایج در جدول ۳ نشان داده شده است. در این جدول میانگین (m) و انحراف معیار (SD) دقت در هر حالت، مبتنی بر آزمون تی می‌باشد. $p\text{-value}<0.05$، به عنوان مقدار معتبر و قابل قبول در نظر گرفته می‌شود.

همانطور که مشاهده می‌شود مدل پیشنهادی با دقت۹۸/۳۲درصد، در بازشناسی ترس در رتبه اول و مدلBEL با دقت۹۶/۹۶ درصد، در رتبه دوم و مدل Moren در رتبه سوم با دقت ۹۶/۵۷درصد می‌باشد. همچنین مدل پیشنهادی با دقت ۹۹/۱درصد، در بازشناسی غم در رتبه نخست و مدل BEL با دقت۹۸/۲درصد، در رتبه دوم و مدل Moren در رتبه سوم با دقت۹۴/۴درصد، است. مدل پیشنهادی در هنگام تشخیص خشم، شادی و تعجب نیز عملکرد بهتری از خود نشان داده است. همانطور که بیان شد مدلEBIF، که در مدل پیشنهادی استفاده شده است به‌طور خاص ویژگی‌هایی را از تصویر استخراج می‌کند که با احساسات انسانی مرتبط هستند (مانند شدت رنگ، کنتراست، بافت، ترکیب‌بندی تصویر و محتوای معنایی). در مقابل، مدلMoren ، بیشتر مبتنی بر مدلی نورونی و انتزاعی از سیستم پردازش هیجان مغز است و ویژگی‌های تصویر را به‌طور مستقیم بررسی نمی‌کند. به علاوه مدل BEL، نیز بیشتر بر مبنای یادگیری احتمالاتی از داده‌ها کار می‌کند، اما فاقد تمرکز اختصاصی بر ویژگی‌های تصویری هیجانی است لذا مدل پیشنهادی موفق‌تر عمل کرده است.

در شکل ۴، تعدادی از تصاویری که مدل پیشنهادی صحیح شناسایی نموده و تصاویری که به اشتباه ادراک کرده است نشان داده شده است.



(الف) (ب)

شکل ۴: الف)تصاویر دیده شده ب) تصاویر درک شده

تصاویر الف، تصاویری هستند که توسط فرد دیده شده (تصویر ورودی) اما تصاویر ب، تشخیص و ادراک شده است (تصویر خروجی). دیده می‌شود که تصاویری که از لحاظ شکل، رنگ و بافت مشابه هستند مدل را دچار چالش می‌کنند.

۴-نتیجه‌گیری

در این مقاله مدلی پیشنهاد شد که در جهت رفع مشکل تشخیص اشیا و بازشناسی حالات هیجانی چهره در افراد مبتلا به ادراک پریشی بصری برآمد. در مدل ارائه شده از سیستم بینایی انسان بر اساس ساختار قشر بینایی مغز الهام گرفته شد تا تصویر دیده شناسایی شود. در این قسمت از مدل، با استفاده از FCM، بصورت هوشمند به بهبود ادراک تصاویر در ذهن این افراد کمک شد. با استخراج ویژگی‌های لبه، کنتراست، انرژی، آنتروپی، ناهمگنی، گشتاور رنگ، میانه و انحراف معیار از تصاویر و با تشخیص معنا و روابط بین اشیای تصویر دیده شده عمل ادراک انجام شد. در نهایت اگر تصویر شناسایی شده چهره می‌بود با استفاده از مدل ویژگی‌های الهام گرفته زیست شناختی زود هنگام به استخراج ویژگی‌های حالت چهره و بازشناسی آن با SVM پرداخته شد. مدل پیشنهادی، مشکل درک و بازشناسی اشیا و همچنین حتی حالات هیجانی چهره را در بیماران مبتلا به ادراک پریشی بصری کاهش داد. معیار دقت در مدل پیشنهادی و دیگر مدل‌ها مقایسه شد و مشاهده گردید که مدل پیشنهادی در بازشناسی اشیا با دقت ٪۹۸/۱ و در بازشناسی حالات چهره با میانگین دقت ٪۹۹/۰۱ کارایی بهتری در مقایسه با دیگر روش‌ها دارد.

در روش پیشنهادی، اگرچه مدل EBIF عملکرد مناسبی روی پایگاه داده‌های استاندارد نشان داده، اما ممکن است در شرایط بالینی واقعی، به دلیل تنوع بالاتر در کیفیت تصاویر و حرکات غیرارادی چهره در افراد مبتلا به ادراک پریشی، دقت کاهش یابد. به دلیل محدودیت دسترسی به داده‌های واقعی از افراد دارای ادراک پریشی، ارزیابی مدل به‌طور کامل روی داده‌های شبیه‌سازی شده یا عمومی صورت گرفته است. لذا، تعمیم‌پذیری نتایج به جمعیت بالینی باید با احتیاط انجام شود. ساختار پیشنهادی با وجود دقت بالا، از نظر محاسباتی سنگین‌تر از برخی مدل‌های پایه است که می‌تواند اجرای آن در دستگاه‌های کلینیکی محدود باشد.

مراجع

[1] M. J. Yang, T. J. Nail, and J. Winer, "Left parietal tumors presenting with smartphone icon visual agnosia: Two cases of a modern presentation of Gerstmann syndrome," *World Neurosurgery*, vol. 142, pp. 233–238, 2020.

[2] N. Seijdel, H. S. Scholte, and E. H. F. de Haan, "Visual features drive the category-specific impairments on categorization tasks in a patient with object agnosia," *Neuropsychologia*, vol. 161, 2021.

[3] C. Gerlach and R. Robotham, "Chapter 9 - Object recognition and visual object agnosia," in *Handbook of Clinical Neurology*, Elsevier, vol. 178, pp. 155–173, 2021.

[4] H. J. Peel and P. A. Chouinard, "A review of the impairments, preserved visual functions, and neuropathology in 21 patients with visual form agnosia – A unique defect with line drawings," *Neuropsychologia*, vol. 190, 2023.

[5] J. H. Missimer, E. Abela, M. Pastore-Wapp, R. Wiest, and B. J. Weder, "Distinguishing transient from persistent tactile agnosia after partial anterior circulation infarcts—Behavioral and neuroimaging evidence for white matter disconnection," *NeuroImage: Clinical*, vol. 36, 2022.

[6] R. Ptak, F. Turri, and N. Doganci, "Object recognition and visual agnosia," in *Encyclopedia of Behavioral Neuroscience*, 2nd ed., pp. 587–596, 2022.

[7] B. Fernández, A. Moreira, H. Gomes, J. M. Justo, and V. Pimenta, "Mirror self-image agnosia: A case report," *Journal of the Neurological Sciences*, vol. 455, 2023.

[8] L. Iterson, L. T. L. Sie, P. B. Augustijn, A. C. S. Rooze, and F. E. Jansen, "Acquired visual agnosia as an uncommon presentation of epileptic encephalopathy in a 6-year-old boy with CSWS," *Epilepsy & Behavior Reports*, vol. 16, 2021.

[9] O. Martinaud, "Visual agnosia and focal brain injury," *Revue Neurologique*, vol. 173, no. 7–8, pp. 451–460, 2017.

[10] X. Duanmu, "Image retrieval using color moment invariant," in *Seventh International Conference on Information Technology*, pp. 200–203, 2010.

[11] R. Ptak and F. Lazeyras, "Functional connectivity and the failure to retrieve meaning from shape in visual object agnosia," *Brain and Cognition*, vol. 131, pp. 94–101, 2019.

[12] O. Martinaud, "Visual agnosia and focal brain injury," *Revue Neurologique*, vol. 173, no. 7–8, pp. 451–460, 2017.

[۱۳] ف. یاقوتی، س. معتمد، «بازشناسی حالات هیجانی چهره با استفاده از مدل یادگیری هیجانی مغز»، *فصلنامه تازه‌های علوم شناختی*، سال ۲۰، شماره ۴، صفحات ۶۱-۴۶، ۱۳۹۷.

[14] J. Morén, *Emotion and learning—a computational model of the amygdala*, Ph.D. dissertation, Lund University, Lund, Sweden, 2002.

[15] E. Lotfi, "Mathematical modeling of emotional brain for classification problems," in *Proceedings of IAM*, vol. 2, no. 1, pp. 60–71, 2013.

[16] H. Yongzhen, H. Kaiqi, W. Liangsheng, T. Dacheng, T. Tieniu, and L. Xuelong, "Enhanced biologically inspired model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, USA, pp. 23–28, 2008.

[17] H. Lechner and J. Zimmermann, "Application of deep learning algorithms to detect agnosia from MRI brain images," *Journal of Neuroscience Methods*, vol. 330, 2020.

[18] K. L. Ng and E. T. Wong, "A machine learning approach for classifying agnosia types based on behavioral assessments," *Frontiers in Neurology*, vol. 12, p. 689832, 2021.

[19] Y. Dauphin and H. Cho, "Personalized treatment plans for agnosia: The role of artificial intelligence in rehabilitation," *Journal of Cognitive Neuroscience*, vol. 34, no. 8, pp. 1293–1305, 2022.

[20] R. Cecere, N. Dundon, and C. Bertini, "When apperceptive agnosia is explained by a deficit of primary visual processing," *Elsevier Ltd.*, pp. 12–27, 2014.

[21] K. Eriksson, A. inen, K. Hirvonen, P. Nieminen, and M. Koivikko, "Visual agnosia in a child with non-lesional occipito-temporal CSWS," *Brain & Development*, pp. 262–267, 2003.

[22] T. Foulsham and J. Barton, "Modeling eye movements in visual agnosia with a saliency map approach: Bottom–up guidance or top–down strategy?," *Neural Networks*, vol. 24, pp. 665–677, 2011.

[۲۳] ف. یاقوتی و س. معتمد، «بازشناسی چهره با استفاده از مدل بهبودیافته HMAX»، *فصلنامه تازه‌های علوم شناختی*، سال ۲۱، شماره ۱، صفحات ۲۹–۴۴، ۲۰۱۹.

[24] N. Safari and S. Motamed, "Face recognition based on hierarchical model and X (HMAX)," *Advances in Cognitive Science*, vol. 21, no. 1, pp. 29–44, 2019.

[25] E. Askari, S. K. Setaredan, A. Sheikhani, M. R. Mohammadi, and M. Teshnelab, "Modeling the connections of brain regions in children with autism using cellular neural networks and electroencephalography analysis," *Journal of Artificial Intelligence in Medicine*, vol. 89, pp. 40–50, 2018.

[26] P. Zhao, H. Xue, X. Ji, H. Liu, and L. Han, "Zero-shot learning via visual feature enhancement and dual classifier learning for image recognition," *Information Sciences*, vol. 642, 2023.

[27] H. Lissauer, "A case of visual agnosia with a contribution to theory," *Cognition Neuropsychology*, vol. 5, pp. 153–192, 1988.

[28] B. L. Finlay, "Focusing the eyes and recognizing objects: evo-devo and the sensitive period," *Current Opinion in Behavioral Sciences*, vol. 36, pp. 36–41, 2020.

[29] H. Sufikarimi and K. Mohammadi, "Role of the secondary visual cortex in HMAX model for object recognition," *Cognitive Systems Research*, vol. 64, pp. 15–28, 2020.

[30] Y. Wu, "Fusion-based modeling of an intelligent algorithm for enhanced object detection using a deep learning approach on radar and camera data," *Information Fusion*, vol. 113, 2025.

[31] S. Etemad and M. Amirmazlaghani, "Color texture image retrieval based on Copula multivariate modeling in the Shearlet domain," *Engineering Applications of Artificial Intelligence*, vol. 102, 2021.

[32] C. Singh, E. Walia, and K. PreetKaur, "Color texture description with novel local binary patterns for effective image retrieval," *Pattern Recognition*, vol. 76, pp. 50–68, 2018.

[33] H. Zoubir, M. Rguig, M. E. Aroussi, R. Saadane, and A. Chehri, "Pixel-level concrete bridge crack detection using Convolutional Neural Networks, Gabor filters, and attention mechanisms," *Engineering Structures*, vol. 314, 2024.

[34] J. L. Salmeron and I. Arévalo, "Concurrent vertical and horizontal federated learning with fuzzy cognitive maps," *Future Generation Computer Systems*, vol. 162, 2025.

[35] S. M. Tabatabaii, T. M. Nazeri, and M. Dastorani, "Performance comparison of GP, ANN, BCSD and SVM models for temperature simulation," *Journal of Meteorology and Atmospheric Sciences*, vol. 1, no. 1, pp. 53–64, 2018.

[36] V. Kumar, V. Kumar, and T. Chan, "Bayes and Naive Bayes Classifier," *arXiv*, 2014.

[37] J. Morén, T. Shibata, and K. Doya, "The mechanism of saccade motor pattern generation investigated by a large-scale spiking neuron model of the superior colliculus," *PloS One*, vol. 8, 2013.

پاورقی‌ها:

[1] Visual Agnosia
[2] Primary visual cortex
[3] Lateral Geniculate Nucleus
[4] Anterior Inferotemporal Cortex
[5] Posterior Inferotemporal Cortex
[6] Medial Superior Temporal area
[7] Primary Vision
[8] Intermediate Vision
[9] High Level Vision
[10] Brain's Emotional Learning Model
[11] Biologically Inspired Features Model
[12] Convolutional Neural Network
[13] Fuzzy Cognitive Map
[14] Early Biologically Inspired Features
[15] Primary visual cortex
[16] Middle Temporal Area
[17] Tectopulvinar
[18] Geniculostriate
[19] GLCM
[20] Accuracy
[21] Precision
[22] Recall