Published by: Iranian Society for Biomedical Engineering / www.isbme.ir

Iranian Journal of Biomedical Engineering

www.ijbme.org / P-ISSN: 2008-5869 / E-ISSN: 8006-9685

Volume 14, Issue 1, Spring 2020, 69 - 79

A Temporal Computational Model for Object Recognition inspired by Human Visual System

Sadeghnejad, Naser [1] / Ezoji, Mehdi [2*] / Ebrahimpour, Reza [3]

[1] - Ph.D. Student, Electronic Department, Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

[2] - Assistant Professor, Electronic Department, Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

[3] - Professor, Computer Department, Faculty of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

ARTICLE INFO

DOI: 10.22041/IJBME.2020.119227.1548

Received: 27 December 2019 Revised: 6/4/2020-28/4/2020 Accepted: 30 April 2020

KEYWORDS

Object Recognition
Computational Model
Deep Neural Network
Decision Making Model
Basic Level Categorization

ABSTRACT

Object recognition is one of the main cognitive abilities of human and animals. Human visual system, as a fast and accurate system can be a source of inspiration for the computational models of object recognition. Studies on the human visual system have emphasized its processing over time, whereas it is not considered in the conventional computational models of object recognition. In this paper, we attempt to present a time-based multilevel model for object recognition. In the first layer of the model, the input image information is sent to the next layer in a temporal representation. In the middle layer of the model, a deep neural network is used as a feature extractor. Finally, in contrast to the popular computational models for object recognition, a decision-making model such as drift-diffusion model is proposed based on the neuronal decision-making mechanisms in the brain. In other words, adaption to the human visual system has been considered in all of three layers. Several experiments have been conducted to evaluate the performance of the proposed computational model in object recognition. The experimental results show that as the input image becomes more complicated, noise increases, or occlusion occurs, the performance/reaction time of the model decreases/increases, which is consistent with the behavior of human visual system. The performance of the model for object recognition and base-level categorization is also investigated for application of the original images and the inverted images. The results show the difference between the processes of the object recognition and base-level categorization, which is consistent with the behavior of human visual system reported in the referenced papers.

[*]Corresponding Author

Address	Electronic Department, Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran		
Postal Code	4714871167	**Tel**	*+98-11-35501435*
E-Mail	*m.ezoji@nit.ac.ir*	**Fax**	*+98-11-32320570*

ناشر: انجمن مهندسی پزشکی ایران / www.isbme.ir

مجله‌ی مهندسی پزشکی زیستی

شاپای چاپی: ۵۸۶۹-۲۰۰۸ / شاپای الکترونیکی: ۹۶۸۵-۸۰۰۶ / www.ijbme.org

دوره: ۱۴، شماره: ۱، بهار ۱۳۹۹، ۶۹ - ۷۹

ارائه‌ی یک مدل محاسباتی بازشناسی اشیا مبتنی بر زمان با الهام از سامانه‌ی بینایی انسان

صادق‌نژاد، ناصر [۱] / ازوجی، مهدی [۲*] / ابراهیم‌پور، رضا [۳]

[۱] – دانشجوی دکتری مهندسی برق، گروه الکترونیک، دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل، ایران

[۲] – استادیار، گروه الکترونیک، دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل، ایران

[۳] – استاد، دانشکده‌ی مهندسی کامپیوتر، دانشگاه تربیت دبیر شهید رجائی، تهران، ایران

مشخصات مقاله

شناسه‌ی دیجیتال: 10.22041/IJBME.2020.119227.1548

ثبت در سامانه: ۶ دی ۱۳۹۸ بازنگری: ۱۳۹۹/۱/۱۸-۱۳۹۹/۲/۹ پذیرش: ۱۱ اردیبهشت ۱۳۹۹

واژه‌های کلیدی

بازشناسی اشیا

مدل محاسباتی

شبکه‌ی عصبی عمیق

مدل تصمیم‌گیری

طبقه‌بندی سطح پایه

چکیده

یکی از اصلی‌ترین توانایی‌های شناختی انسان و جانوران، بازشناسی اشیا است. سامانه‌ی بینایی انسان به عنوان یک سامانه‌ی سریع و دقیق می‌تواند منبع الهام مناسبی برای ارائه‌ی مدل‌های محاسباتی بازشناسی اشیا باشد. پژوهش‌های پیشین که به بررسی رفتار سامانه‌ی بینایی انسان در بازشناسی اشیا پرداخته‌اند، بر پردازش طی گام‌های زمانی در این سامانه تاکید کرده‌اند، در حالی که در مدل‌های محاسباتی موجود برای بازشناسی اشیا، چنین چیزی مورد توجه قرار نمی‌گیرد. در این مقاله سعی شده است تا یک مدل چندلایه‌ی مبتنی بر زمان برای بازشناسی اشیا ارائه شود. در لایه‌ی نخست مدل، اطلاعات تصویر ورودی در یک بازنمایی زمانی به لایه‌های بعدی ارسال می‌شود. در لایه‌ی میانی مدل، از یک شبکه‌ی عصبی عمیق به عنوان استخراج کننده‌ی ویژگی استفاده شده است. در پایان، برخلاف مدل‌های محاسباتی موجود برای بازشناسی اشیا، پیشنهاد شده است که برای طبقه‌بندی ویژگی‌های استخراج شده از مدل‌های تصمیم‌گیری مبتنی بر سازوکار نورونی تصمیم‌گیری در مغز مانند مدل رانشی-انتشار استفاده شود. به بیان دیگر، در هر یک از این سه لایه تلاش شده است تا تطبیق مناسبی با سازوکار سامانه‌ی بینایی انسان ایجاد شود. برای ارزیابی کارایی مدل محاسباتی پیشنهادی در بازشناسی اشیا، آزمون‌های متعددی انجام شده است. نتایج به دست آمده از بررسی مدل پیشنهادی نشان می‌دهد که با دشوارتر شدن تصاویر، افزودن نویز یا بروز انسداد، کارایی مدل در بازشناسی اشیا کاهش یافته و زمان پاسخ‌دهی آن افزایش می‌یابد که این روند با شواهد رفتاری انسانی مطابقت دارد. هم‌چنین عمل‌کرد مدل برای تشخیص شی و طبقه‌بندی سطح پایه در دو حالت تصاویر اصلی و تصاویر وارونه بررسی شده است. نتایج به دست آمده گویای تفاوت بین پردازش تشخیص شی با طبقه‌بندی سطح پایه است که این نتایج با آزمایش‌های رفتاری گزارش شده در مقاله‌های مرجع هم‌خوانی دارد.

[*]نویسنده‌ی مسئول

نشانی گروه الکترونیک، دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل، ایران

کد پستی ۴۷۱۴۸۷۱۱۶۷ **تلفن** ۳۵۵۰۱۴۳۵-۱۱-۹۸+

پست الکترونیک *m.ezoji@nit.ac.ir* **دورنگار** ۳۲۳۲۰۵۷۰-۱۱-۹۸+

۱- مقدمه

بازشناسی اشیا یکی از اصلی‌ترین توانایی‌های شناختی انسان و جانوران است که در زندگی روزمره نقش مهمی دارد [۱]. توانایی بازشناسی اشیا و دسته‌بندی آن‌ها در سامانه‌ی بینایی انسان با سرعت و دقت بالایی انجام می‌شود [۲]. انسان و بسیاری از پستان‌داران می‌توانند اشیای پیرامون خود را علی‌رغم بروز انسداد و تغییر در حالت‌ها و اندازه‌ها، از زوایای دید گوناگون بازشناسی و تفکیک کنند. اگر چه این کار برای انسان و پستان‌داران بسیار ساده بوده و در زمان کوتاهی انجام می‌شود، اما در نوع خود یک فرایند محاسباتی پیچیده است [۳]. شناخت چگونگی بازشناسی اشیا در مغز، کلیدی ارزشمند برای پاسخ به پرسش‌های علوم شناختی و اعصاب است که می‌تواند به ساخت ماشین‌های هوشمند کارامدتر در آینده بیانجامد [۴].

فرایند بازشناسی اشیا در مغز شامل سه گام اساسی بازنمایی زمانی، استخراج ویژگی و تصمیم‌گیری است. سامانه‌ی بینایی در ابتدا اطلاعات تصویر ورودی را در یک بازنمایی زمانی به لایه‌های بالاتر ارسال کرده، سپس در نواحی قشر بینایی اولیه[۱] و سلول‌های فروگیج‌گاهی[۲] به استخراج ویژگی در طول زمان پرداخته و سرانجام در نواحی مربوط به تصمیم‌گیری که شامل قشر پیش‌پیشانی[۳] است با استفاده از ویژگی‌های استخراج شده در طول زمان، درباره‌ی تصویر شی تصمیم می‌گیرد [۵].

سامانه‌ی بینایی انسان با توجه به توانایی‌های منحصر به فرد، همواره الهام‌بخش مدل‌های بینایی زیادی بوده است. طراحی یک مدل بازشناسی اشیا مبتنی بر سامانه‌ی بینایی انسان، علاوه بر کمک به حل مسائل حوزه‌ی بینایی ماشین، به شناخت بیش‌تر سازوکار قشر بینایی مغز در مساله‌ی بازشناسی اشیا نیز منجر می‌شود [۶، ۷]. درک چگونگی بازشناسی اشیا در انسان می‌تواند به ارائه‌ی مدل‌های محاسباتی یا ماشین‌های یادگیر کارامدتر، حل مساله‌های بینایی ماشین و گمانه‌زنی‌ها درباره‌ی چگونگی بروز برخی اختلالات شناختی در مغز کمک کند [۸].

هر کدام از مدل‌های محاسباتی پیشین به نوبه‌ی خود سعی در شبیه‌سازی بخشی از ویژگی‌های سامانه‌ی بینایی انسان داشته‌اند، اما هنوز مدل جامعی که گویای تمام یافته‌های بازشناسی اشیا در انسان باشد، ارائه نشده است. با وجود این که اخیرا در مدل‌های محاسباتی بازشناسی اشیا، زمان در پردازش مورد توجه قرار نمی‌گیرد، اما در مطالعات مربوط به سیستم بینایی انسان روی نقش پردازش زمانی در سامانه‌ی بینایی در بازشناسی اشیا تاکید شده است [۹، ۱۰]. از جمله مدل‌های مطرحی که پایه‌ی زیستی دارند می‌توان به مدل‌های HMAX [۱۱] و شبکه‌های عصبی پیچشی [۱۲] اشاره کرد. تمام این مدل‌ها ساختاری چندلایه دارند که از مطالعات فیزیولوژیکی و مدل توصیفی هابل و ویزل [۱۳] به دست آمده است. در این مدل‌ها میدان دریافت[۴] واحدهای هر لایه از تجمیع[۵] و ترکیب خروجی‌های واحدهای لایه‌ی پیشین ایجاد می‌شود. پس از چند مرحله‌ی پردازش، با ترکیب میدان دریافت‌های کوچک (دارای حساسیت بالا به محرک ساده)، میدان‌های دریافت بزرگ‌تری (حساس به محرک‌های پیچیده‌تر) ایجاد می‌شوند [۱۴-۱۶].

مدل HMAX در سال ۱۹۹۹ با الهام از کارکرد سلول‌های ساده و پیچیده ارائه شده است [۱۱]. طی آزمایش‌های مختلف مبتنی بر ثبت نورونی، نشان داده شده است که استفاده از عمل‌گر بیشینه‌گیر برای مدل‌سازی عمل‌کرد سلول‌های پیچیده در ناحیه‌ی V1 (و لایه‌های بالاتر) به توصیف بهتر رفتار سلول‌های پیچیده منجر می‌شود [۵]. در ساده‌ترین حالت، مدل استاندارد HMAX شامل چهار لایه از واحدهای محاسباتی بوده که با واحدهای ساده‌ی S شروع شده و با واحدهای پیچیده‌ی C ادامه می‌یابد. در لایه‌ی آخر، یک بیشینه‌گیری کلی انجام شده و بردار حاصل از آن به یک طبقه‌بند مانند ماشین بردار پشتیبان (SVM) یا K-همسایه‌ی نزدیک (KNN) داده می‌شود. مدل HMAX نظیر سامانه‌ی بینایی انسان در برابر تغییراتی چون جابه‌جایی و تغییر اندازه‌ی اشیا کارایی بالایی دارد اما به دلیل استفاده از یک طبقه‌بند کلاسیک در لایه‌ی آخر و عدم پردازش در طول زمان، تطابقی با رفتار انسان در بازشناسی اشیا ندارد.

شبکه‌های عصبی پیچشی که داده‌هایی با اطلاعات مکانی را پردازش می‌کنند، ویرایشی از پرسپترون چندلایه هستند که با الهام از پردازش‌های زیستی مغز انسان ایجاد شده‌اند. در شبکه‌های پیچشی، سلول‌های ساده و پیچیده به ترتیب با مجموعه‌ای از فیلترها در لایه‌های پیچشی و تجمعی مدل‌سازی می‌شوند [۱۷]. با کنار هم قرار دادن لایه‌های متفاوت پیچشی، تجمعی و لایه‌ی اتصال کامل می‌توان شبکه‌های متفاوتی ساخت و برای حل مسایل متفاوت از آن‌ها استفاده نمود. از جمله رایج‌ترین معماری‌های ارائه شده در این حوزه می‌توان به شبکه‌های AlexNet [۱۸]، LeNet [۱۹]، ZfNet [۲۰]، VGGNet [۲۱]، GoogleNet [۲۲] و ResNet [۲۳] اشاره کرد. در لایه‌های پایانی شبکه‌های عصبی پیچشی اغلب از لایه‌ی

[۱] Primary Visual Cortex
[۲] Inferotemporal
[۳] Prefrontal Cortex
[۴] Receptive Field
[۵] Accumulation

اتصال کامل استفاده شده که پایه‌ی زیستی نداشته و در بستر پردازش در طی زمان تعریف نشده است.

مطالعه در حوزه‌ی تصمیم‌گیری در شاخه‌هایی مانند بیولوژی، علوم کامپیوتر، اقتصاد، علوم سیاسی و روان‌شناسی صورت گرفته است [۲۴]. بررسی سازوکار نورونی واحد تصمیم‌گیرنده در مغز مورد توجه دانشمندان علوم اعصاب شناختی قرار گرفته که طی آن با ترکیب روش‌های فیزیولوژیکی و روان-فیزیکی به ارائه‌ی مدل‌های محاسباتی-شناختی منطبق بر شواهد بیولوژی پرداخته شده است [۲۵]. از رایج‌ترین مدل‌های تصمیم‌گیریِ ارائه شده می‌توان به مدل رانشی انتشار[۱] (DDM) [۲۶، ۲۷]، مدل رقابتی[۲] [۲۸] و مدل‌های مبتنی بر بیولوژی [۲۹] مانند مدل مهاری پیش‌رو[۳] (FFI)، مدلِ مهاری دوطرفه یا جانبی[۴] و مدل شبکه‌ی عصبی بازگشتی ونگ[۵] [۳۰] اشاره کرد.

در این مقاله سعی شده است تا با الهام از سامانه‌ی بینایی انسان، یک مدل محاسباتیِ چندلایه‌ی مبتنی بر پردازش در طی زمان و تصمیم‌گیری مبتنی بر DDM، برای بازشناسی اشیا ارائه شود. برای ارزیابی مدل پیشنهادی، تطبیق رفتاری آن با سامانه‌ی بینایی انسان طی سنجش کارایی مدل و زمان واکنش آن به ازای اعمال ورودی‌های گوناگون بررسی شده است.

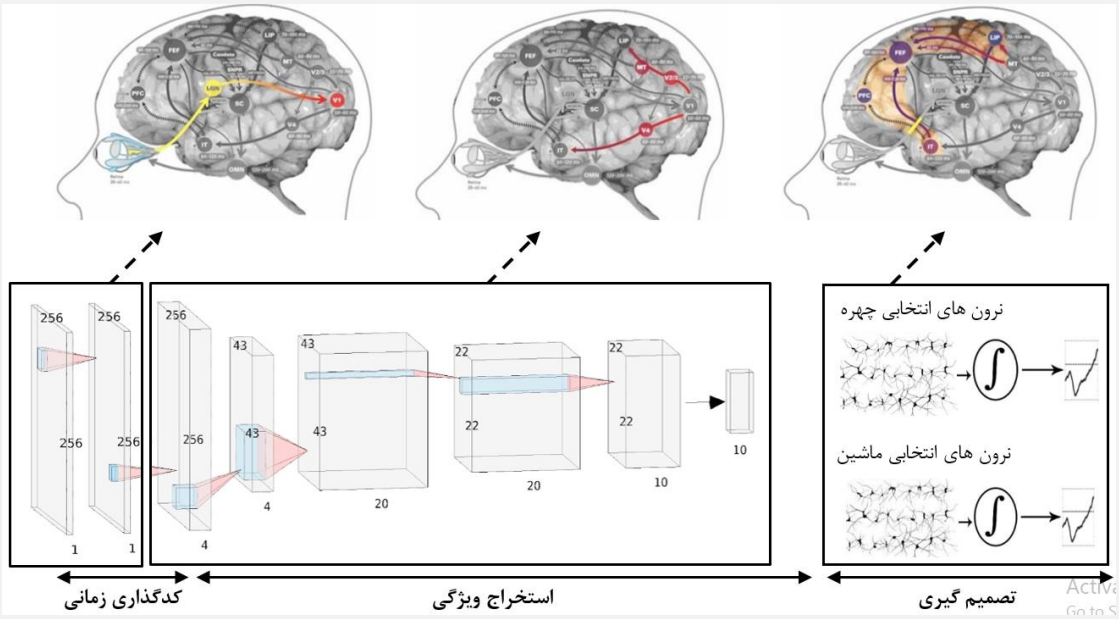
در ادامه، در بخش ۲ ساختار پیشنهادی شرح داده شده، در بخش ۳ روش پیشنهادی طی آزمایش‌های گوناگون ارزیابی شده و در پایان به نتیجه‌گیری و جمع‌بندی پرداخته شده است.

۲- مواد و روش‌ها

هنگامی که یک تصویر به سامانه‌ی بینایی انسان عرضه می‌شود، مولفه‌های مختلف آن مانند لبه‌های برجسته، ریزه‌کاری‌ها، جزئیات و کلیات تصویر هم‌زمان پردازش نشده و این مولفه‌های اطلاعاتی طی بازه‌های زمانی مختلفی مورد پردازش قرار می‌گیرند. به این صورت در سامانه‌ی بینایی انسان، درباره تصاویر آسان زودتر تصمیم‌گیری می‌شود. در حالی که در مدل‌هایی مانند شبکه‌های عصبی، کل تصویر ورودی به صورت یک‌جا پردازش شده و شبکه تفاوتی میان تصویرهای آسان و سخت قائل نمی‌شود. در این مقاله به ارائه‌ی یک مدل محاسباتی پرداخته شده که از این نظر تطابق بیش‌تری با رفتار سامانه‌ی بینایی انسان داشته باشد. ساختار مدل پیشنهادی و چگونگی پیاده‌سازی آن در ادامه توضیح داده شده است.

۲-۱- ساختار مدل پیشنهادی

مدل پیشنهادی که روند نمای کلی آن در شکل (۱) نشان داده شده است، مشابه سامانه‌ی بازشناسی اشیا در انسان دارای سه گام کدگذاری/بازنمایی زمانی تصویر ورودی، استخراج ویژگی در طی زمان و طبقه‌بندی یا تصمیم‌گیری مبتنی بر زمان می‌باشد. این مراحل در ادامه توضیح داده شده و مکان انجام هر یک از آن‌ها در قشر مغز نیز در نیمه‌ی بالایی شکل (۱) نشان داده شده است.



شکل (۱)- ساختار کلی مدل پیشنهادی بازشناسی اشیا

[۱] Drift Diffusion Model
[۲] Competitive Model
[۳] Feed Forward Inhibition
[۴] Mutual/Lateral Inhibition Model
[۵] Wang Recurrent Neural Networks

۲-۱-۱- کدگذاری زمانی تصاویر ورودی

در مرحله‌ی کدگذاری زمانی، باید رفتار و ویژگی‌های سلول‌های گانگلیونی[۱] شبکیه تقریب زده شده و در ادامه یک کدگذاریِ مشابه لایه‌ی V_1 در مسیر بینایی انسان حاصل شود [۳۱]. بدین منظور در این مقاله از یک فیلتر DoG استفاده شده است.

۲-۱-۲- استخراج ویژگی با شبکه‌ی عصبی پیچشی عمیق

در روش پیشنهادی از یک شبکه‌ی عصبی پیچشی عمیق[۲] (DCNN) برای استخراج ویژگی‌های مناسب استفاده شده است که از طریق یادگیری با سرپرست آموزش می‌بینند.

در روش آموزش با سرپرست، N تصویر ورودی و N بردار خروجیِ مطلوب متناظر با آن وجود دارد. اگر x_n n-امین تصویر در مجموعه‌ی آموزش، d_n n-امین بردار خروجی مطلوب متناظر با x_n و y_n خروجی واقعی باشد، تابع خطا به صورت زیر تعریف می‌شود که در آن C بیان‌گر تعداد کلاس است.

$$E(W) = -\sum_{n=1}^{N}\sum_{c=1}^{C}(y_n^c \log(d_n^c)) \qquad (۱)$$

طبق با قانون دلتا، در صورت تغییر وزن‌های شبکه‌ی عصبی در خلاف جهت بردار گرادیان تابع خطا، بیش‌ترین کاهش در اندازه‌ی تابع خطا رخ خواهد داد. بنابراین طبق رابطه‌ی (۲) وزن‌های شبکه‌ی عصبی را تغییر داده تا خطا کمینه شود.

$$W^l\left(t+۱\right) = W^l(t) - \eta_t \frac{\partial E(t)}{\partial W^l} \qquad (۲)$$

در رابطه‌ی (۲)، l شماره‌ی لایه‌ی مدل، η_t نرخ یادگیری و t گام زمانی است. تعداد لایه‌ها، اندازه‌ی فیلترها و پارامترهای یادگیری باید برای عمل تشخیص مطلوب، بهینه‌سازی شوند.

۲-۱-۳- طبقه‌بندی/تصمیم‌گیری زمانی

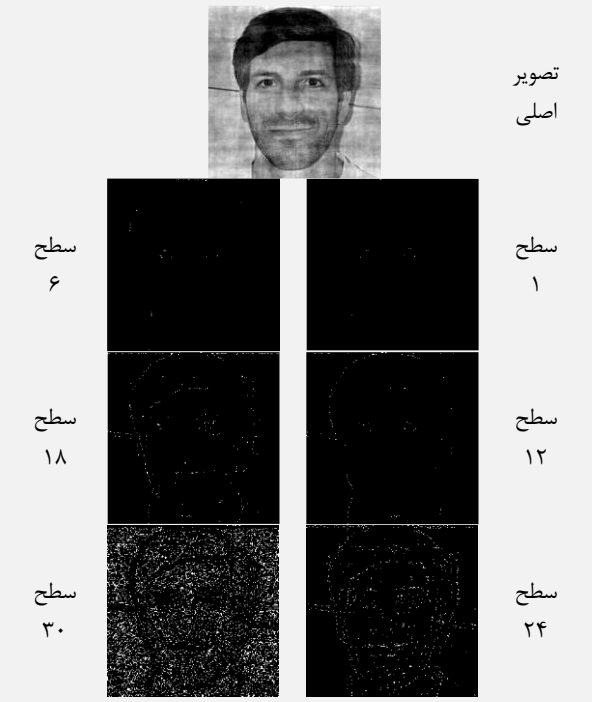
پس از دو گام کدگذاری زمانی تصاویر ورودی و استخراج ویژگی، در گام سوم از یک طبقه‌بند استفاده شده است. طبقه‌بندهای کلاسیک مانند SVM[۳] به بعد زمان وابستگی نداشته و به کارگیری آن‌ها باعث از بین رفتن پردازش در گام‌های زمانی در مراحل پیشین می‌شود. بنابراین در این مقاله برای بالا بردن تطابق رفتار مدل پیشنهادی با مغز انسان و حفظ زمان در گام، پیشنهاد می‌شود که مطابق شکل (۱)، از یک مدل تصمیم‌گیری مانند مدل تصمیم‌گیری رانشی انتشار استفاده شود [۲۶]. در این مدل، شواهد برای انتخاب‌های مختلف جمع‌آوری و اختلاف بین شواهد جمع‌آوری شده محاسبه می‌شود. زمانی که این اختلاف به آستانه‌ی از پیش تنظیم شده برسد، تصمیم به نفع انتخاب با بیش‌ترین شواهد گرفته می‌شود.

۲-۲- چگونگی پیاده‌سازی مدل پیشنهادی

در ادامه به چگونگی پیاده‌سازی گام‌های گوناگون ساختار پیشنهادی یعنی کدگذاری زمانی تصاویر ورودی، استخراج ویژگی و نیز تصمیم‌گیری پرداخته شده است.

۲-۲-۱- کدگذاری زمانی تصویر ورودی

در این مقاله، در پیاده‌سازی‌های این گام از یک فیلتر DoG با اندازه‌ی ۷×۷ پیکسل و انحراف معیارهای ۱ و ۲ استفاده شده است. با این روش از یک تصویر ورودی، یک تصویر لبه‌یابی شده‌ی نرمالیزه (با مقادیر بین ۰ تا ۲۵۵) استخراج شده است. مقادیر بیش‌تر (نزدیک به ۲۵۵) شامل لبه‌های تیزتر و مقادیر کم‌تر با لبه‌های نرم‌تر مرتبط هستند. سپس خروجی فیلتر DoG بر اساس مقادیر پیکسل‌های آن در ۳۰ سطح[۴]، تفکیک شده (تعیین آستانه‌های متناظر با هر سطح، طی آستانه‌گذاری یک‌نواخت بر بازه‌ی مقادیر تصویر ورودی) و در هر گام زمانی، یکی از این ۳۰ سطح به عنوان ورودی به مدل اعمال شده که تعدادی از این سطح‌ها در شکل (۲) نشان داده شده است.



شکل (۲)- نمونه‌ای از تصویر اصلی و برخی از سطح‌های تفکیک شده‌ی خروجی لایه‌ی DoG

[۱] Ganglion Cell
[۲] Deep Convolutional Neural Network
[۳] Support Vector Machine
[۴] Level

۲-۲-۲- استخراج ویژگی

در این گام، یک شبکه‌ی عصبی عمیق پیچشی با سه لایه‌ی پیچشی و سه لایه‌ی تجمعی ارائه شده است. لایه‌های ۱، ۲ و ۳ پیچشی به ترتیب دارای ۴، ۲۰ و ۱۰ نقشه‌ی عصبی (فیلتر) با اندازه‌ی ۵×۵، ۴×۱۶×۱۶ و ۲۰×۵×۵ بوده و اندازه‌ی فیلترهای تجمعی لایه‌های ۱ و ۲ به ترتیب برابر با ۷×۷ و ۲×۲ با اندازه‌ی گام‌های ۶ و ۲ است. لایه‌ی ۳ تجمعی یک پردازش سراسری بیشینه‌گیری انجام داده که با توجه به وجود ۱۰ فیلتر در لایه‌ی ۳ پیچشی، دارای ۱۰ خروجی است. وزن‌های اولیه‌ی فیلترهای پیچشی به صورت تصادفی انتخاب شده و یادگیری در آن‌ها به صورت با سرپرست و با الگوریتم پس انتشار انجام شده است.

۲-۲-۳- به‌کارگیری DDM در تصمیم‌گیری

در این گام از مدل تصمیم‌گیری رانشی-انتشار استفاده شده است. ابتدا نورون‌های حساس به چهره و ماشین با آزمایشی در خروجی ۱۰ نورونی لایه‌ی استخراج ویژگی تعیین شده و میانگین خروجی نورون‌های این دو مجموعه به دو نورون مدل تصمیم‌گیری داده شده است. این مدل اطلاعات مراحل قبل را طی گام‌های زمانی (۳۰ سطح آستانه‌گذاری شده در مرحله‌ی کدگذاری زمانی) جمع کرده، در هر لحظه آن را با آستانه‌ی تصمیم‌گیری مقایسه کرده و هر زمان که یکی از مجموعه‌ها به آستانه برسد، دسته‌بندی تصویر ورودی انجام می‌شود.

۳- آزمایش‌ها و ارزیابی

برای ارزیابی مدل پیشنهادی و تطابق کارکرد آن با رفتار انسان، دو نوع ارزیابی طراحی و با آزمایش‌های مختلفی اجرا شده است. در ارزیابی اول، رفتار مدل در روبه‌رو شدن با تصاویری با درجه‌ی سختی متفاوت (اضافه شدن نویز فرکانسی از کلاس مقابل)، تصاویر نویزی (حضور نویز گوسی با واریانس گوناگون) و تصاویر با درجه‌ی انسداد مختلف بررسی شده و در ارزیابی دوم کارکرد مدل در تشخیص و دسته‌بندی شی در سطح پایه سنجیده شده است. برای انجام این آزمایش‌ها از مجموعه‌ی داده‌ی Caltech [۳۲] استفاده شده است. در ارزیابی اول از مجموعه‌ی داده‌ی چهره و ماشین و در ارزیابی دوم از ۹ کلاس مختلف از این مجموعه استفاده شده که در ادامه شرح داده شده است.

۳-۱- ارزیابی کارایی مدل در تشخیص اشیا

هدف از این آزمایش‌ها، بررسی تطبیق رفتار مدل پیشنهادی از لحاظ کارایی و زمان پاسخ‌گویی با رفتار سامانه‌ی بینایی انسان است. برای این منظور، کارکرد مدل از دید سرعت و دقت به حضور نویز در فاز، نویز گوسی و هم‌چنین اثر بروز برش و انسداد در تصاویر ورودی مورد بررسی قرار گرفته است.

۳-۱-۱- بررسی کارایی در برابر تصاویر با درجه‌ی سختی متفاوت

در این آزمایش، تصاویر چهره از مجموعه‌ی داده‌ی Caltech [۳۲] و تصاویر ماشین از اینترنت گردآوری شده که نمونه‌ای از این تصاویر در شکل (۳) نشان داده شده است.

شکل (۳)- چند نمونه از تصاویر چهره‌ی Caltech و ماشین [۳۲]

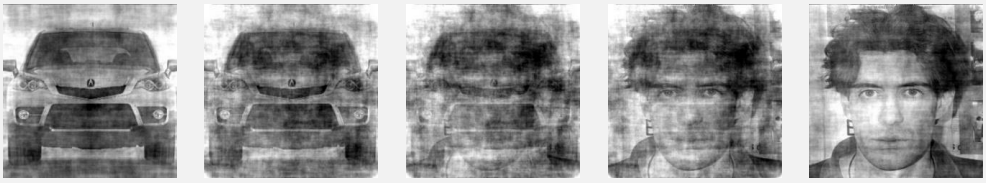
کل مجموعه‌ی داده شامل ۳۵۰ تصویر چهره و ۳۵۰ تصویر ماشین با اندازه‌ی ۲۵۶×۲۵۶ پیکسل است. برای بررسی حساسیت مدل پیشنهادی به تصاویر با نویز در فاز، ۱۰ تصویر از هر مجموعه برای آزمون کنار گذاشته شده که با استفاده از آن‌ها و طبق الگوریتم (۱)، تصاویر ترکیبی گوناگون ایجاد شده است. بخشی از تصاویر آموزشی برای تنظیم پارامترهای شبکه به عنوان داده‌های اعتبارسنجی در نظر گرفته شده است.

الگوریتم (۱)- ایجاد تصاویری با درجه‌ی سختی متفاوت از دید ادارکی با افزودن نویز فاز از کلاس مقابل

۱- اعمال تبدیل فوریه به تمامی تصاویر
۲- محاسبه‌ی میانگین دامنه‌ی تبدیل فوریه‌ی تمام تصاویر
۳- ضرب ماتریس فاز هر تصویر چهره (ماشین) در عدد r (۱٪ تا ۱۰۰٪) و ماتریس فاز هر تصویر ماشین (چهره) در عدد r-۱۰۰
۴- محاسبه‌ی ماتریس فاز ترکیبی از جمع دو ماتریس فاز حاصل
۵- تبدیل فوریه‌ی معکوس از ماتریس فاز ترکیبی و میانگین دامنه‌ها، به تصویر تلفیقی با درجه‌ی سختی نسبی |r-۵۰|۲-۱۰۰ می‌انجامد

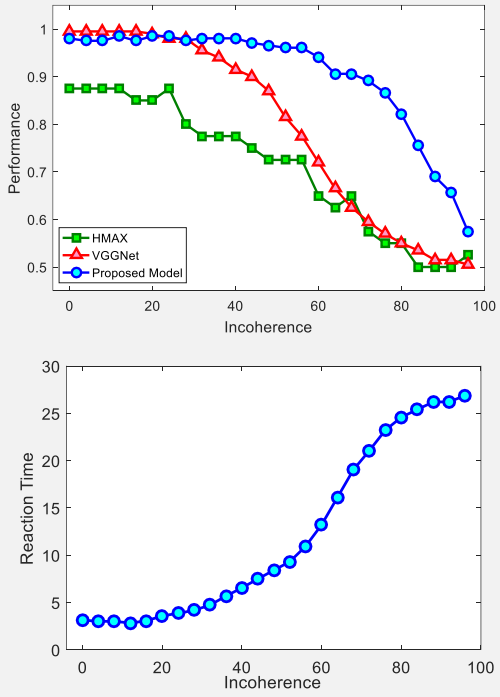
چند نمونه از تصاویر ترکیبی حاصل با درجه‌های سختی مختلف در شکل (۴) ارائه شده است. مشاهده می‌شود که به ازای r=۵۰، به کارگیری الگوریتم (۱) به سخت‌ترین تصویر برای بازشناسی (تصویری با ناهم‌سازی[۱] ۱۰۰) منجر شده است. با داشتن ۱۰ تصویر چهره و ۱۰ تصویر ماشین، ۱۰۰۰۰ تصویر با درجه‌های سختی مختلف ایجاد شده که در ۳۰ سطح به مدل داده شده و تعدادی از آن‌ها برای آزمایش شبکه انتخاب شده است.

[۱] Incoherence



شکل (۴)- چند نمونه از محرک بینایی مبتنی بر الگوریتم (۱) (با تلفیق r درصد چهره و ۱۰۰-r درصد ماشین)، از راست به چپ: با درجه‌های سختی نسبی برابر با ۰٪، ۶۰٪، ۱۰۰٪ ۶۰٪ و ۰٪

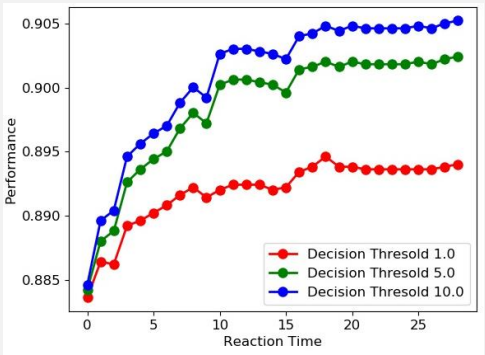
دقت بازشناسی (کارایی) مدل پیشنهادی در مقایسه با مدل HMAX و VGGNet بر حسب سختی‌های گوناگون تصاویر ورودی در شکل (۵-الف) ارائه شده که در آن محور افقی بیان‌گر سطوح گوناگون سختی (مبتنی بر الگوریتم ۱) و محور عمودی بیان‌گر کارایی مدل بر حسب درصد است. مشاهده می‌شود که با افزایش سطوح سختی، کارایی مدل کاهش یافته است.



شکل (۵)- ارزیابی مدل پیشنهادی در برابر تصاویر با سختی‌های گوناگون، بالا) دقت بازشناسی، پایین) زمان واکنش

زمان واکنش مدل (تعداد گام زمانی لازم برای رسیدن به آستانه‌ی تصمیم‌گیری در DDM) بر حسب سختی‌های مختلف تصاویر ورودی در شکل (۵-ب) ارائه شده که نشان دهنده‌ی افزایش زمان واکنش با سخت شدن تصویر ورودی است.

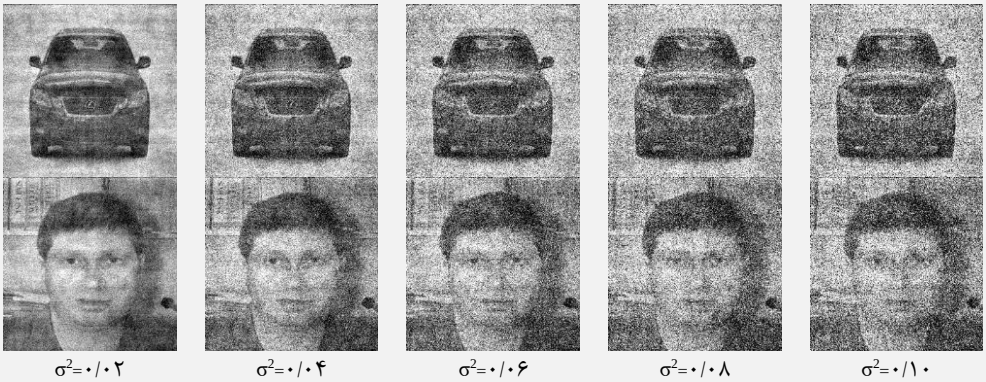
شکل (۶) رابطه‌ی بین زمان واکنش و کارایی مدل برای تصاویر با درجه‌ی سختی ۰٪ در آستانه‌های تصمیم‌گیری مختلف را نشان می‌دهد. مشاهده می‌شود که با افزایش زمان واکنش، دقت مدل افزایش یافته که بیان‌گر تعادل بین دقت و سرعت است.



شکل (۶)- رابطه‌ی بین زمان واکنش و کارایی مدل با آستانه‌های تصمیم‌گیری گوناگون

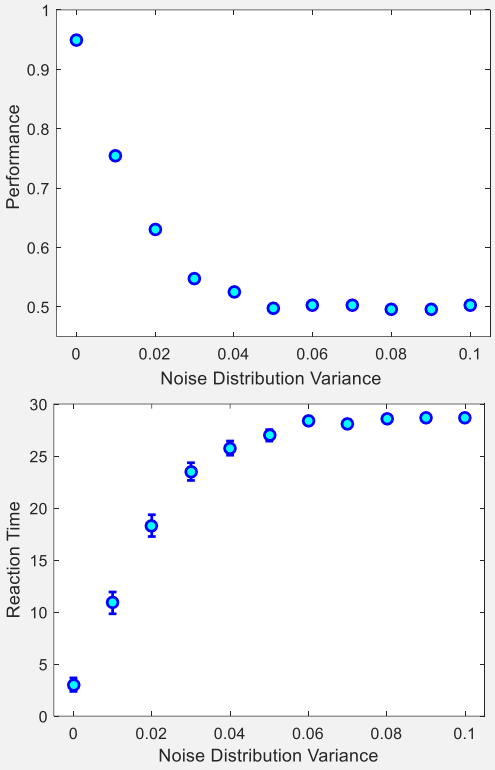
۳-۱-۲- بررسی کارایی مدل پیشنهادی در برابر نویز

برای بررسی تاثیر نویز بر زمان واکنش و کارایی مدل، ۲۰ تصویر از هر مجموعه‌ی ماشین و چهره برای آزمون کنار گذاشته شده و به این تصاویر نویز گوسی با میانگین صفر و واریانس متفاوت افزوده شده که چند نمونه از آن‌ها در شکل (۷) ارائه شده است.



شکل (۷)- چند نمونه از تصاویر آزمونِ نویزی شده با نویز گوسی با میانگین صفر و واریانس‌های مختلف

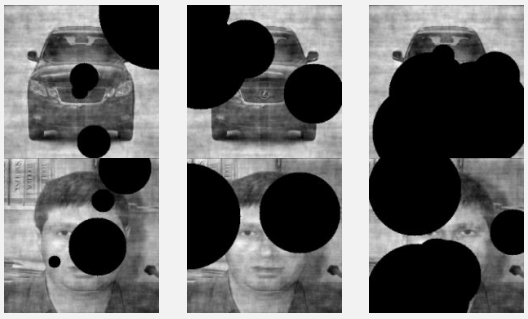
اثر نویز گوسی با اندازه‌های متفاوت بر دقت بازشناسی (کارایی) و زمان واکنشِ مدل در شکل (۸) نشان داده شده است.



شکل (۸)- اثر نویز گوسی بر بالا) دقت بازشناسی (کارایی) و پایین) زمان واکنش مدل، به ازای نویزهای گوناگون

۳-۱-۳- بررسی کارایی مدل پیشنهادی در برابر انسداد

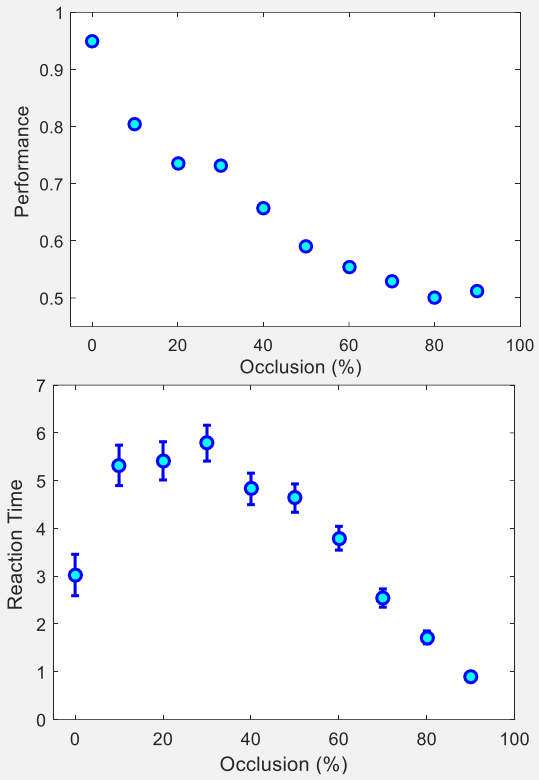
برای بررسی اثر انسداد تصاویر بر کارکرد مدل، ۲۰ تصویر از هر مجموعه برای آزمون کنار گذاشته شده، بروز انسدادهای مختلف در این تصاویر با استفاده از دایره‌های مشکی با اندازه و محل تصادفی ایجاد شده و تصاویری با درصدهای انسداد ۰، ۱۰، ۲۰، ۳۰، ۴۰، ۵۰، ۶۰، ۷۰، ۸۰ و ۹۰ انتخاب شده که چند نمونه از آن‌ها با درجه‌های انسداد گوناگون در شکل (۹) ارائه شده است.



٪۲۰ ٪۴۰ ٪۸۰

شکل (۹)- چند نمونه از تصاویر آزمونِ نویزی شده با انسدادهای گوناگون

اثر انسداد تصاویر بر کارایی و زمان واکنش مدل در شکل (۱۰) نشان داده شده است.



شکل (۱۰)- نمودار کارایی و زمان واکنش مدل بر حسب بروز انسدادهای گوناگون در تصویر، بالا) دقت بازشناسی (کارایی)، پایین) زمان واکنش (تعداد سطح) مدل

۳-۲- ارزیابی مدل پیشنهادی در طبقه‌بندی

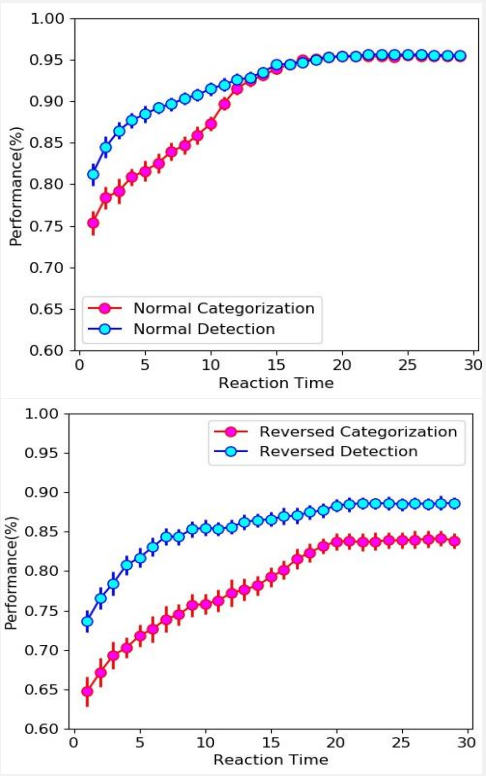
هدف این ارزیابی، مقایسه‌ی زمان واکنش مدل در تشخیص شی و طبقه‌بندی آن در سطح پایه و بررسی تطابق آن با رفتار انسان است. لذا مانند مرجع [۳۳] از مجموعه‌ی داده‌ی Caltech [۳۲] استفاده شده و الگوهای بدون شی به طور تصادفی از مربع‌های ۸×۸ پیکسلی تصاویر طبیعی ایجاد شده است (شکل ۱۱).

در آزمایش تشخیص شی، تصاویر هدف شامل یک شی از هر دسته‌ی سطح پایه و تصاویر مزاحم شامل الگوهای بدون شی است. در آزمایش طبقه‌بندی شی، تصویر هدف شامل یک شی از یکی از دسته‌های سطح پایه (سگ، صندلی و ماشین) و تصاویر مزاحم از دیگر دسته‌های سطح پایه در همان گروه (پرنده، ماهی، تخت، میز، کشتی و هواپیما) است.

مدل ارائه شده در این مقاله برای بازشناسی اشیا مبتنی بر زمان است، لذا مقایسه‌ی زمان واکنش آن در تشخیص شی، طبقه‌بندی آن در سطح پایه و بررسی و تطابق آن با رفتار انسانی بسیار مهم می‌باشد. در مقاله‌ی [۳۳] در هر آزمایش رفتاری، تصویر ورودی برای مدت ۱۷، ۳۳، ۵۰، ۶۸ و ۱۶۷ میلی‌ثانیه به افراد نمایش داده شده است. در ارزیابی مدل پیشنهادی نیز طی روندی مشابه، هر بار یکی از ۳۰ سطح تفکیک شده‌ی خروجی DoG به مدل داده شده و عمل‌کرد مدل بررسی شده است.

شکل (۱۱)- نمونه‌هایی از تصاویر استفاده شده در ارزیابی دوم، چپ) نمونه‌هایی از محرک‌های استفاده شده برای آزمایش تشخیص اشیا شامل تصاویر هدف (تصاویر اشیا) و تصاویر بدون شی، راست) نمونه‌هایی از تصاویر مورد استفاده برای آزمایش طبقه‌بندی سطح پایه شامل تصاویر هدف (سگ، صندلی و ماشین) و تصاویر مزاحم (پرنده، ماهی، تخت، میز، کشتی و هواپیما)

عمل‌کرد مدل پیشنهادی در دو آزمایش تشخیص شی و طبقه‌بندی سطح پایه با اعمال تصاویر اصلی تعیین شده و برای حالتی که تصاویر به صورت وارونه اعمال شود، مورد بررسی قرار گرفته است. دقت تصمیم‌گیری مدل بر حسب تعداد گام زمانی مورد نیاز برای رسیدن به آستانه‌ی تصمیم‌گیری در DDM، در شکل (۱۲) نشان داده شده است.



شکل (۱۲)- کارایی مدل پیشنهادی بر حسب زمان واکنش برای دو آزمایش تشخیص شی و طبقه‌بندی شی در سطح پایه، الف) ارائه‌ی تصاویر اصلی، ب) ارائه‌ی تصاویر وارونه

۴- یافته‌ها و بحث

در این مقاله، یک مدل محاسباتی برای بازشناسی اشیا مبتنی بر ارائه‌ی تدریجی تصویر ورودی و طبقه‌بندی مبتنی بر انباشت شواهد طی گام‌های زمانی و با الهام از سامانه‌ی بینایی انسان پیشنهاد شده است. کارایی مدل پیشنهادی طی آزمایش‌های متعددی برای سنجش تطابق کارکرد آن با رفتار انسان مورد ارزیابی قرار گرفته که در این بخش به بحث و تفسیر نتایج به دست آمده از آزمایش‌ها پرداخته شده است.

در ابتدا نتایج ارزیابی نخست که به بررسی رفتار مدل در روبه‌رو شدن با تصاویری با درجه‌ی سختی متفاوت، تصاویر نویزی و تصاویر با درجه‌ی انسداد گوناگون پرداخته، بیان شده است. همان‌طور که در شکل (۵) مشاهده می‌شود، با افزایش سطح سختی تصویر ورودی، دقت مدل در تشخیص کاهش و زمان واکنش آن افزایش یافته است. افزایش زمان واکنش مدل با افزایش سطح سختی تصاویر ورودی، بر اساس بررسی‌های انجام شده در مقاله‌ی [۳۴]، مطابق با رفتار انسانی است.

اثر تغییرات آستانه‌ی تصمیم‌گیری بر تغییرات بین زمان واکنش و کارایی مدل در شکل (۶) نشان داده شده است. مشاهده می‌شود که با افزایش آستانه‌ی تصمیم‌گیری در DDM، کارایی و زمان واکنش مدل افزایش یافته است. تنظیم این پارامتر می‌تواند به تطابق بیش‌تر کارکرد مدل و رفتار انسان منجر شود. با توجه به شکل (۸) مشاهده می‌شود که با افزایش توان نویز در تصاویر ورودی و در نتیجه سخت‌تر شدن مساله، کارایی مدل از لحاظ دقت و سرعت کاهش یافته است. این رفتار در بررسی کارکرد مدل در برابر انسداد نیز نمایان می‌شود. در شکل (۱۰) مشاهده می‌شود که با افزایش میزان انسداد در تصاویر ورودی،

زمان واکنش ابتدا افزایش و پس از انسداد ۴۰٪ کاهش یافته است. به نظر می‌رسد که در انسدادهای پایین، مدل پیشنهادی برای جبران اطلاعات از دست رفته به دلیل انسداد، نیازمند پردازش سطوح بیش‌تری از تصویر ورودی برای رسیدن به حد آستانه‌ی مدل در لایه‌ی تصمیم‌گیری است. اما در انسدادهای بالاتر، کافی نبودن اطلاعات موجود در سطوح گوناگون تصویر ورودی به تصمیم‌گیری زودهنگام و در نتیجه کاهش کارایی مدل منجر می‌شود. این کارکرد مدل با رفتار انسانی تطابق دارد، زیرا انسان نیز در تصمیم‌گیری درباره‌ی تصاویر با انسداد بالا، برای دریافت اطلاعات ناموجود تصویر، بیش‌تر صبر نمی‌کند.

در ارزیابی دوم، عمل‌کرد مدل پیشنهادی در تشخیص شی و طبقه‌بندی سطح پایه در دو حالت تصاویر اصلی و وارونه بررسی شده است. نمودار کارایی بر حسب زمان واکنش مربوط به این آزمایش در این دو حالت در شکل (۱۲) نشان داده شده است. مشاهده می‌شود که در حالت تصاویر اصلی، نمودارهای تشخیص شی و طبقه‌بندی سطح پایه از لحاظ روند کارایی و زمان واکنش بسیار به هم نزدیک هستند، در حالی که در حالت تصاویر وارونه، کارایی مدل مربوط به هر دو آزمایش کاهش یافته و این کاهش در طبقه‌بندی شی در سطح پایه بیش‌تر از تشخیص شی است. این اختلاف بین این دو نمودار در حالت تصاویر وارونه افزایش یافته که گویای پردازش زمانی متفاوت در تشخیص شی و طبقه‌بندی سطح پایه است. همان‌طور که در شکل (۱۲) مشاهده می‌شود، یافته‌های به دست آمده از مدل پیشنهادی با نتایج حاصل از آزمایش رفتاری گزارش شده در مقاله‌ی [۳۳] مطابقت دارد.

۵- نتیجه‌گیری

در این مقاله یک مدل محاسباتی برای بازشناسی اشیا ارائه شده است که تصاویر ورودی را در چندلایه و در گام‌های زمانی پردازش می‌کند. در لایه‌ی اول، اطلاعات تصویر ورودی برای ارسال به لایه‌های بعدی در گام‌هایی زمانی بازنمایی شده، در لایه‌ی میانی، از یک شبکه‌ی عصبی عمیق به عنوان مدل پایه برای استخراج ویژگی استفاده شده و در لایه‌ی آخر، مشابه با سازوکار نورونی تصمیم‌گیری مغز از مدل تصمیم‌گیری DDM برای طبقه‌بندی ویژگی‌های استخراجی استفاده شده است.

با افزایش سطوح سختی تصاویر، کارایی مدل در بازشناسی اشیا کاهش و زمان پاسخ‌دهی آن افزایش یافته که این نتایج با شواهد رفتاری انسانی سازگار است. هم‌چنین کارایی مدل بر حسب زمانِ واکنش برای دو آزمایش تشخیص شی و طبقه‌بندی شی در سطح پایه بررسی شده است. این مدل بین پردازش تشخیص شی با طبقه‌بندی سطح پایه در دقت عمل‌کرد بر حسب زمان واکنش تفاوت قائل شده که این نتایج با آزمایش‌های رفتاری گزارش شده در مقاله‌های مرجع هم‌خوانی دارد.

از محدودیت‌های این پژوهش می‌توان به بررسی مساله‌ی بازشناسی دوکلاسی، به کارگیری شبکه‌ی عصبی عمیق پیش‌رو و تطبیق کلی کارایی شبکه با رفتار انسانی اشاره کرد.

پیاده‌سازی‌های مبتنی بر ایده‌ی پیشنهادی نشان می‌دهد که با گزینش ساختاری مناسب برای مدل، به کارگیری ساختارهایی دارای تطابق بیش‌تر با سامانه‌ی بینایی انسان و تنظیم پارامترها می‌توان مدل‌های دقیق‌تر، سازگارتر و کاراتری را در آینده ارائه کرد. طراحی و انجام آزمایش‌های روان-فیزیک متناسب با شیوه‌ی ارائه‌ی الگوها در این مقاله، می‌تواند به تنظیم مناسب‌تر پارامترهای مدل برای تطابق بیش‌تر کارکرد آن با رفتار انسانی منجر شود. ارائه‌ی مدلی که ضمن سازگاری با کارکرد سامانه‌ی بینایی انسان بتواند در برابر عوامل کنترل نشده‌ی صحنه مقاوم باشد، در ادامه‌ی این پژوهش می‌تواند مورد بررسی قرار گیرد.

۶- مراجع

[1] E. Contini, S. Wardle, T. Carlson, “Decoding the time-course of object recognition in the human brain: From visual features to categorical decisions”, Neuropsychologia, 2017.

[2] M. Dehaqani, A. Vahabie, R. Kiani, M. Ahmadabadi, B. Araabi, H. Esteky, “Temporal dynamics of visual category representation in the macaque inferior temporal cortex”, Journal of Neurophysiology, 116:587-601, 2016.

[3] K. Rajaei, Y. Mohsenzadeh, R. Ebrahimpour, S. Khaligh-Razavi, “Beyond Core Object Recognition: Recurrent processes account for object recognition under occlusion”, PLOS Computational Biology, l 15(5), 2019.

[4] H. Fujiyoshi, T. Hirakawa, T. Yamashita, “Deep learning-based image recognition for autonomous driving”, IATSS Research, 2019.

[5] James DiCarlo, D. Zoccolan, N. Rust, “How does the brain solve visual object recognition?” Neuron, Vol 73 PP 415-434, 2012.

[6] H. Sufikarimi, K. Mohammadi, “Feature extraction for object recognition inspired by human visual system”, Iranian Journal of Biomedical Engineering, 11(4): 337-349, 2018.

[7] M. Jazlaeiyan, H. S. Shahhoseini, “Optimal Feature Selection in Biologically Inspired Model for Object Recognition Using Mutual Information Maximisation”, Iranian Journal of Biomedical Engineering, 8: 371-383, 2015.

[8] S. Khaligh-Razavi, S. Habibi, M. Sadeghi, H. Marefat, M. Khanbagi, S. Nabavi, E. Sadeghi, C. Kalafatis, “Integrated Cognitive Assessment: Speed and Accuracy of Visual Processing as a Reliable Proxy to Cognitive Performance.”, Sci Rep vol. 9, pp: 1102, 2019.

[9] A. Mirzaei, S. M. Khaligh-Razavi, M. Ghodrati, S. Zabbah, R. Ebrahimpour, "Predicting the human reaction time based on natural image statistics in a rapid categorization task", Vision Research, 81: 36-44, 2013.
[10] S. Zabbah, K. Rajaei, A. Mirzaei, R. Ebrahimpour, S.M. Khaligh-Razavi, "The impact of the lateral geniculate nucleus and corticogeniculate interactions on efficient coding and higher-order visual object processing", Vision Research,101: 82-93, 2014.
[11] M. Riesenhuber, T. Poggio, "Hierarchical models of object recognition in cortex", Nat Neurosci vol. 2, pp: 1019–1025, 1999
[12] B. Le Cun, J. Denker, D. Henderson, R. Howard, W. Hubbard, L. Jackel, "Handwritten digit recognition with a back-propagation network," in Advances in neural information processing systems, 1990.
[13] D. Hubel, T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," J. Physiol., vol. 160, no. 1, pp: 106-154, 1962.
[14] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biol. Cybern., vol. 36, no. 4, pp: 193-202, 1980.
[15] K. Fukushima, "Training multi-layered neural network neocognitron," Neural Networks, vol. 40, pp: 18-31, 2013.
[16] K. Fukushima, "Neocognitron for handwritten digit recognition," Neurocomputing, vol. 51, pp: 161-180, 2003.
[17] S. Khaligh-Razavi, N. Kriegeskorte, "Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation", PLoS Comput Biol 10(11), 2014
[18] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet classification with deep convolutional neural networks", Communications of the ACM, vol. 60, no. 6, pp: 84-90, 2017
[19] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE 86(11): 2278–2324, 1998.
[20] M. Zeiler, R. Fergus, "Visualizing and Understanding Convolutional Networks", European Conference on Computer Vision (ECCV), pp 818-833, 2014.
[21] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", Computer Vision and Pattern Recognition, ICLR, 2015.
[22] C. Szegedy, C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich "Going deeper with convolutions", Conference on Computer Vision and Pattern Recognition(CVPR), Boston, MA, pp. 1-92015.
[23] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition", Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas, NV, pp. 770-778, 2016.
[24] H. Timothy, C. Summerfield, "Perceptual Decision Making in Rodents, Monkeys, and Humans", Neuron, vol. 93-1, pp: 15-31, 2017
[25] J. Gold, M. Shadlen, "The neural basis of decision making", Annu Rev Neurosci. 30(1): 535-74, 2007.
[26] R. Ratcliff, J. Rouder, "Modeling response times for two-choice decisions", Psychological Science, 9(5):347-35, 1998.
[27] R. Ratcliff, G. McKoon, "The diffusion decision model: theory and data for two-choice decision tasks", Neural computation, 20(4):873-922, 2008.
[28] D. Vickers, "Evidence for an accumulator model of psychophysical discrimination." Ergonomics, 13(1):37-58, 1970.
[29] X. Wang, "Probabilistic decision making by slow reverberation in cortical circuits." Neuron, 36(5):955-968, 2002.
[30] K. Wong, X. Wang, "A Recurrent Network Mechanism of Time Integration in Perceptual Decisions." The Journal of Neuroscience, 26(4):1314 –1328, 2006.
[31] S. Thorpe, A. Delorme, R. Van Rullen, "Spike-based strategies for rapid processing." Neural Netw, 14(6-7):715-25, 2001
[32] G. Griffin, A. Holub, P. Perona, "Caltech-256 object category dataset.", Technical Report 7694, California Institute of Technology, 2007.
[33] M. Mack, I. Gauthier, J. Sadr, T. Palmeri, "Object detection and basic-level categorization: Sometimes you know it is there before you know what it is", Psychonomic Bulletin & Review, 15(1), 28-35, 2008.
[34] A. Diaz, F. Queirazza and G. Philiastides, "Perceptual learning alters post-sensory processing in human decision-making", Nature Human Behaviour, vol. 1, no. 0035, 2017.