

Improving the Structure of Deep Learning Algorithm in Image Processing Inspired by Representational Brain Dissimilarity Matrix

Zahra Heydaran Daroogheh Amnyieh¹, *PhD Student*, Seyed Mohammad Jalal Rastegar Fatemi¹,
Assistant Professor, Maryam Rastgarpour², *Assistant Professor*

1. Department of Electrical, College of Engineering, Saveh Branch, Islamic Azad University, Saveh, Iran
z.heidaran@iauda.ac.ir, m.rastgarpour@iau-saveh.ac.ir
2. Department of Computer, College of Engineering, Saveh Branch, Islamic Azad University, Saveh, Iran
jalal.pe77@gmail.com

Abstract:

Deep learning algorithms achieves some results at human level or even better in pattern recognition problems. Meanwhile they apply a different mechanism other than human brain. This paper describes a human-inspired segmentation and interpolation algorithm, which applies the retinal layer in the proposed model after the input layer. Following this retina, this layer encrypts the input image and transmits the input image to the second space, which try to change deep network structure inspired of the brain's visual path. Network feedback, recognition rate, and network energy level or the comprehensiveness of the trained network examined in subsets of the Caltech data set. In similar examples, deep learning algorithms require more data to learn other than human. In the difference between deep learning and human, there is a difference in the representation of information. In deep learning, weights improve in a way that optimizes the result in a particular experiment, but in millions of years of human evolution, the human brain has evolved optimally and effectively representation. Another point of contention is the deepening of deep learning layers. The number of these layers has multiplied compared to the brain that lead to more complexity and energy expenditure. However, in the brain it can make a diagnosis with less energy. The maximum recognition rate of the proposed model is 93% and the base model is close to 91%. Also, the proposed model is thinner and the rate of fire of neurons in the initial layers is lower and has a high stability to changes in light intensity. The Dissimilarity of the model layers has been higher and it has been able to show a better response in the face of noise images and record less recognition loss.

Keywords: Deep learning, convolutional system, recognition rate, network energy level, sparsity, representational dissimilarity matrix

Received: 20 March 2020

Revised: 15 September 2020

Accepted: 26 October 2020

Corresponding Author: Dr. Maryam Rastgarpour

بهبود ساختار الگوریتم یادگیری عمیق در پردازش تصویر با الهام از ماتریس تفکیک‌پذیری مغز

زهرا حیدران داروقه‌امنیه^۱، دانشجوی دکتری، سید محمد جلال رستگار فاطمی^۱، استادیار، مریم رستگار پور^۲،
استادیار

۱- گروه برق، دانشکده فنی و مهندسی، واحد ساوه، دانشگاه آزاد اسلامی، ساوه، ایران
z.heidaran@iauda.ac.ir, m.rastgarpour@iau-saveh.ac.ir
۲- گروه کامپیوتر، دانشکده فنی و مهندسی، واحد ساوه، دانشگاه آزاد اسلامی، ساوه، ایران
jalal.pe77@gmail.com

چکیده: الگوریتم‌های آموزش عمیق در بسیاری از مسائل بازشناسی الگو، نتایجی در سطح انسان و یا بهتر می‌توانند به ثبت رسانند. اما این نتایج با مکانیسمی متفاوت از مغز انسان به دست آمده است. مدل پیشنهادی در این مقاله یک الگوریتم تقسیم‌بندی و درون‌یابی با الهام از مغز انسان را توصیف نموده و بعد از لایه ورودی، لایه‌ی شبکه اعمال شده است که به پیروی از شبکه چشم، عمل رمزنگاری بر روی تصویر ورودی را انجام می‌دهد. سپس تصویر ورودی به فضای ثانی انتقال می‌یابد که تلاش برای تغییر ساختار شبکه عمیق با الهام از مسیر بینایی مغز خواهد بود. بازخورد شبکه، نرخ بازشناسی و سطح انرژی شبکه و با جامعیت شبکه‌ی آموزش دیده در زیرمجموعه‌هایی از مجموعه داده کلتک بررسی می‌گردد. در نمونه‌های مشابه الگوریتم‌های آموزش عمیق برای یادگیری نیاز به داده بیشتری در مقایسه با یادگیری انسان دارد. بعلاوه، اختلاف یادگیری عمیق و انسان، تفاوت در باز نمایش اطلاعات است. در یادگیری عمیق وزن‌ها در جهتی بهبود می‌یابند تا در یک آزمایش خاص نتیجه بهینه شود ولی در انسان با میلیون‌ها سال تکامل، مغز انسان به گونه‌ای تکامل یافته تا باز نمایش بهینه و مؤثر باشد. چالش مورد بررسی دیگر، عمیق‌تر شدن لایه‌های یادگیری عمیق است. تعداد این لایه‌ها نسبت به مغز چندین برابر گشته است و این مسئله منجر به پیچیدگی و صرف انرژی بیشتر می‌شود. اما در مغز با صرف انرژی کمتر می‌تواند تشخیص را انجام دهد. بیشینه نرخ بازشناسی مدل پیشنهادی به ۹۳ درصد می‌رسد و مدل پایه نزدیک به ۹۱ درصد است. همچنین مدل پیشنهادی تنک‌تر و نرخ آتش نوروها در لایه‌های ابتدایی کمتر و پایداری بالایی به تغییرات شدت روشنایی داشته، تفکیک‌پذیری در لایه‌های مدل بالاتر رفته و توانسته در مواجهه با تصاویر نویزی پاسخ بهتری نشان دهد و افت بازشناسی کمتری را ثبت کند.

کلمات کلیدی: یادگیری عمیق، سیستم کائولوشنی، نرخ بازشناسی، سطح انرژی شبکه، پراکندگی، ماتریس تفکیک‌پذیری مغز

تاریخ ارسال مقاله: ۱۳۹۹/۱/۱

تاریخ بازنگری مقاله: ۱۳۹۹/۶/۲۵

تاریخ پذیرش مقاله: ۱۳۹۹/۸/۵

نام نویسنده‌ی مسئول: دکتر مریم رستگار پور

نشانی نویسنده‌ی مسئول: ساوه- بلوار شهید بهشتی - دانشگاه آزاد اسلامی واحد ساوه- دانشکده مهندسی برق

۱- مقدمه

بازشناسی اشیا^۱ همواره از اهمیت بالایی برخوردار بوده است. تاکنون تحقیقات بسیار زیادی در این موضوع صورت گرفته است. مساله بازشناسی اشیا موضوع پیچیده‌ای است که سال‌ها مورد توجه و علاقه مهندسان کامپیوتر، الکترونیک و نیز محققان علوم عصب‌شناسی^۲ بوده است. به‌طور کلی، بازشناسی اشیا به معنای یافتن یک شی در یک صحنه یا تصویر است. انسان و بسیاری از پستان‌داران قادر هستند که بسیاری از اشیا پیرامون خود را بدون کوچک‌ترین مشکلی شناسایی و تفکیک نمایند، هر چند که ممکن است که اشیا موجود در یک صحنه در حالت‌های گوناگون و با زوایای دید مختلف و همچنین در اندازه‌های متفاوت باشند. اگر چه این امر به نظر برای انسان و پستان‌داران بسیار ساده و بدون تلاش و صرف زمان زیاد است، اما در نوع خود یک فرآیند محاسباتی بسیار مشکل و پیچیده است. این موضوع یعنی حل و شناخت چگونگی شناسایی اشیا در مغز کلیدی ارزشمند برای پاسخ به بسیاری از سوالات علوم شناختی^۳ و عصبی است که می‌تواند منجر به ساخت ماشین‌های هوشمند در آینده گردد.

شبکه‌های عصبی^۴ بر اساس عملکرد الکتریکی سلول‌های عصبی^۵ مطرح شده و توسعه یافته‌اند. وظیفه این شبکه‌ها ارائه مدلی از فعالیت‌های عصبی انسان در فراگیری دانش و بکارگیری دانش فراگیری شده در موقعیت‌های مشابه است. مدل‌های متنوعی برای مدل‌سازی ساختارهای عصبی انسان ارائه شده است، اما در این بین، ساختار شبکه عصبی کانولوشنی^۶ (CNN) که نوعی تکامل یافته‌ی شبکه عصبی پرسپترون چند لایه^۷ است موفقیت چشمگیری در مدل‌سازی سامانه بینایی انسان در ارتباط با قشر بینایی مغز داشته است. کارآمدی این روش را می‌توان با اشاره به تعدد بکارگیری این نوع شبکه به خصوص در تحقیقات اخیر در حل مسائل بینایی ماشین^۸ به اثبات رساند. یکی از مهم‌ترین ویژگی‌های شبکه عصبی کانولوشنی (CNN) که مد نظر متخصصین علوم اعصاب^۹ به خصوص در حوزه بینایی است شباهت ساختاری و عملکردی این نوع شبکه‌ها با سامانه بینایی انسان است.

با گسترش و موفقیت شبکه‌های کانولوشنی، موضوع یادگیری عمیق مورد توجه قرار گرفت زیرا شبکه‌های کانولوشنی شامل لایه‌های زیادی هستند و بنابراین یادگیری بهینه‌ی لایه‌های شبکه اهمیت بالایی دارد. مبحث یادگیری عمیق به شبکه‌های گسترش یافته با لایه‌های بالا (اصطلاحاً عمیق) اطلاق می‌شود. البته امروزه تقریباً یادگیری عمیق فقط برای شبکه‌های عصبی کانولوشنی استفاده می‌شود.

مدل‌های بسیاری با الهام از قشر بینایی مغز طراحی شده‌اند. به عنوان مثال سه مدل نئوکوگنیترون^{۱۰}، مدل سلسله مراتبی^{۱۱} و شبکه عصبی کانولوشنی که این سه مدل تقریباً از الگوی مشترکی الهام گرفته‌اند [۱،۲].

نئوکوگنیترون، مدل سلسله مراتبی و شبکه‌های کانولوشنی نوعی شبکه عصبی هستند که بر اساس تحقیقات هابل و ویزل بر روی سیستم بینایی ارائه شده و مدلی بر پایه سلول‌های ساده و پیچیده‌ی مغز در مسیر گیجگاهی پستان‌داران، معرفی شده‌اند. به عبارتی این مدل‌ها تنها به شناخت اشیا می‌پردازند و به مکان اشیا توجهی ندارند. در پیش از این در مدل معروف مدل سلسله مراتبی مدل‌سازی بخش‌های بسیاری از مسیر بینایی انجام شده که جای خالی آنها در مدل‌های یادگیری عمیق احساس می‌شود [۳]. برای مثال کدینگ رنگی در شبکه چشم که در مدل سلسله مراتبی پیاده‌سازی شد و نرخ بازشناسی را در آزمایش علائم راهنمایی و رانندگی افزایش داد [۴]. در آزمایشی دیگر ناحیه غدومی تالاموسی به نام Ign در مدل سلسله مراتبی پیاده شد و نتیجه مثبت در بازشناسی داشت [۵].

همان‌طور که بیان شد، هنوز بخش‌های بسیاری در مسیر بینایی هستند که در آموزش عمیق مدل نشده‌اند. برای مثال مدل‌سازی نگاشت رنگی در شبکه و یا مدل‌سازی زمانی تالاموس در مسیر بینایی و یا پیاده‌سازی روش یادگیری زیستی در یادگیری عمیق. مدل کردن کامل مسیر بینایی کمک شایانی به انطباق بیشتر به کارکرد مغز دارد. از این رو درک و تحلیل کارکرد بینایی در مغز را آسان‌تر می‌کند، همچنین می‌توان نقاط ضعف شبکه‌های عمیق را تا حد امکان برطرف کرد. مغز انسان همیشه الهام‌بخش برای محققان در حوزه هوش مصنوعی بوده است.

الگوریتم‌های یادگیری عمیق زیرمجموعه‌ای از الگوریتم‌های یادگیری ماشین هستند که هدف آنها کشف چندین سطح از بازنمودهای (نمایش) توزیع شده از داده ورودی است [۶،۷]. طی سال‌های اخیر، یادگیری عمیق به‌صورت گسترده در حوزه بینایی کامپیوتر مورد مطالعه قرار گرفته است [۸،۹] و به همین دلیل، تعداد زیادی از روش‌های مرتبط با آن بوجود آمده است [۱۰]. الگوریتم‌های آموزش عمیق در بسیاری از مسائل بازشناسی الگو توانسته‌اند نتایجی در سطح انسان و یا بهتر به ثبت برسانند

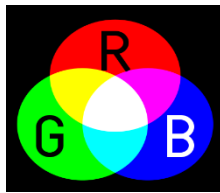
ولی این نتایج با مکانیسمی مختلف از مکانیسم مغز به دست آمده‌اند [۱۰]. در این مقاله یک الگوریتم تقسیم‌بندی و درون‌یابی الهام گرفته از انسان را توصیف نموده که در مدل پیشنهادی بعد از لایه ورودی، لایه‌ی شبکه‌ی اعمال شده است. این لایه به پیروی از شبکه‌ی چشم، عمل رمزنگاری بر روی تصویر ورودی را انجام می‌دهد و تصویر ورودی به فضای ثانی انتقال می‌یابد که تلاش برای تغییر ساختار شبکه عمیق با الهام از مسیر بینایی مغز خواهد بود و بازخورد شبکه و نرخ بازشناسی و سطح انرژی شبکه و یا جامعیت داشتن شبکه‌ی آموزش دیده در زیرمجموعه‌هایی از مجموعه داده کلتک بررسی خواهد گردید. برای مثال در نمونه‌های مشابه الگوریتم‌های آموزش عمیق برای یادگیری نیاز به تعداد بسیاری از داده دارند ولی انسان با داده نمونه بسیار کمتر یاد می‌گیرد [۱۱، ۱۰]. در اختلاف یادگیری عمیق و انسان، تفاوت در بازنمایش اطلاعات است. در یادگیری عمیق وزن‌ها در جهتی بهبود می‌یابند تا در یک آزمایش خاص نتیجه بهینه شود ولی در انسان میلیون‌ها سال تکامل، مغز انسان به گونه‌ای تکامل یافته تا بازنمایش به بهترین وجه بهینه و مؤثر باشد [۱۲]. یکی دیگر از موارد قابل بحث، عمیق‌تر شدن لایه‌های یادگیری عمیق است. تعداد این لایه‌ها نسبت به مغز چندین برابر گشته است و این مسئله منجر به پیچیدگی و صرف انرژی بیشتر شده است. اما در مغز با صرف انرژی بسیار کمتر می‌تواند تشخیص را انجام دهد [۱۳، ۱۴]. بیشینه نرخ بازشناسی مدل پیشنهادی بهبود یافته و به ۹۳ درصد می‌رسد و مدل پایه نزدیک به ۹۱ درصد است. همچنین مدل پیشنهادی تنک‌تر و نرخ آتش نورون‌ها در لایه‌های ابتدایی کمتر و پایداری بالایی به تغییرات شدت روشنایی داشته، تفکیک‌پذیری در لایه‌های مدل بالاتر رفته و توانسته در مواجهه با تصاویر نویزی پاسخ بهتری نشان دهد و افت بازشناسی کمتری را ثبت کند. ساختار مقاله به این شرح است: در قسمت دوم مدل پیشنهادی بیان می‌شود. در قسمت سوم به نتایج شبیه‌سازی و تحلیل مدل پیشنهادی پرداخته خواهد شد و در پایان، نتیجه‌گیری این پژوهش در قسمت چهارم بررسی خواهد شد.

۲- مدل پیشنهادی

با توجه به فقدان مدل‌سازی شبکه‌ی در شبکه‌های عمیق در این قسمت مدلی از رمزنگاری رنگی شبکه‌ی به همراه الگوریتم تقسیم‌بندی و درون‌یابی ارائه می‌شود. شناخت رفتار شبکه‌ی و چشم انسان بسیار اهمیت دارد تا مدل‌سازی صحیحی صورت گیرد. در ادامه یکی از مطرح‌ترین مدل‌های ارائه شده در مورد رمزنگاری رنگی شبکه‌ی بیان می‌شود. رمزنگاری به همراه شبیه‌سازی و معادلات ریاضی نوشته خواهد شد. همچنین مدلی از شبکه عمیق هشت لایه طراحی می‌شود. شبکه طراحی شده مبتنی بر شبکه کانولوشنی است که انطباق بیشتری به مسیر بینایی مغز دارد. شبکه طراحی شده از استانداردهای پایه و اولیه شبکه کانولوشنی پیروی دارد. مدل رمزنگاری شبکه‌ی پس از لایه ورودی در شبکه عمیق قرار می‌گیرد و تصاویر پس از ورود به شبکه، رمزنگاری می‌شوند. شبیه‌سازی بر روی دو کلاس تصاویر از مجموعه داده کلتک آزمایش می‌شود و نتایج با مدل پایه شبکه عمیق کانولوشنی مقایسه می‌شود.

۲-۱- شبیه‌سازی رنگی شبکه‌ی چشم

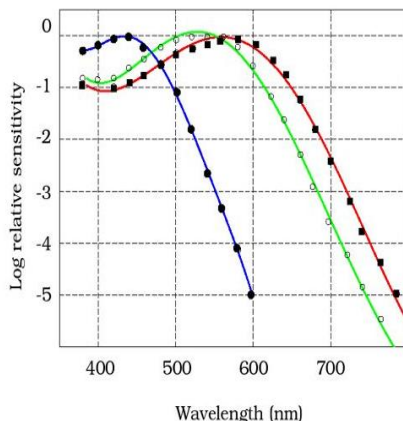
در ساختار چشم شبکه‌ی در انتهای کره چشم قرار دارد. شبکه‌ی پرده‌ای از سلول‌های حساس به نور است که در مواجهه با نور واکنش نشان می‌دهد. شبکه‌ی شامل سلول‌های مختلفی از جمله استوانه‌ای و مخروطی است [۱۵]. سلول‌های استوانه‌ای حساس به شدت نور و سلول‌های مخروطی در سه مدل حساس به سه طیف قرمز، سبز و آبی هستند. در شکل (۱) سه رنگ اصلی و ترکیب آنها نشان داده شده است. سلول‌های مخروطی تقریباً حساس به سه طیف نور اصلی هستند. بنابراین تصویر با چهار ویژگی شدت نور، قرمز، سبز و آبی کد می‌شود. در شکل (۲) دامنه طیف حساس به نور سلول‌های مختلف مخروطی نشان داده شده است [۱۶-۱۸].



شکل (۱): سه رنگ اصلی و ترکیب آنها

Figure (1): The three main colors and their combinations

Cone spectral sensitivities



شکل (۲): طیف موجی سه سلول مخروطی. شدت جذب نورها نرمالیزه شده‌اند درحالی‌که در واقعیت بسیار متفاوت هستند [۱۶-۱۸]
Figure (2): Wave spectrum of three cone cells. The intensity of light absorption is normalized while in reality they are very different [22-24]

در مدل پیشنهادی و ورودی شبکه ارتباط بین سلول‌ها بگونه‌ای است که داده به صورت خام به مسیر خروجی ارسال نمی‌شود و یک نوع لبه‌یابی صورت می‌گیرد [۱۹]. لبه یک تغییر ناگهانی در شدت روشنایی تصویر است [۱۵]. این عمل فقط لبه‌های تصویر را حفظ می‌کند و باعث کاهش انرژی فعالیت نورون‌ها می‌گردد. البته این لبه‌یابی بین رنگی است. فیلتر لبه مرسوم DOG^{11} است. معادله مربوط به فیلتر DOG در معادله (۱) نشان داده شده است [۲۰]. عمل فیلتر کردن تفاوتی با فیلترهای مرسوم دارد که هر یک از گوسی‌ها بر روی رنگی متفاوت اعمال می‌شوند. علامت منفی نشان‌دهنده عمل بازدارندگی یا فیلتر گوسی منفی است و علامت مثبت معرف تحریکی و یا فیلتر گوسی مثبت است. چنین عملی باعث استخراج لبه‌های بین رنگی می‌شود و از فعالیت نورانی زائد می‌کاهد و منجر به تنگی 12 می‌گردد. سه نوع لبه‌یابی انجام می‌شود، اولین لبه‌یابی در خروجی سلول‌های راد است. دومین لبه‌یابی بین خروجی سلول‌های کان قرمز و سبز و سومین لبه‌یابی بر روی خروجی سلول‌های آبی و زرد است. البته زرد از ترکیب سبز و قرمز ساخته می‌شود. این سه لبه‌یابی بیشترین رمزنگاری در شبکه را انجام می‌دهند. در شکل (۳) نمونه‌ای از رمزنگاری دو لایه زرد و آبی نشان داده شده است.

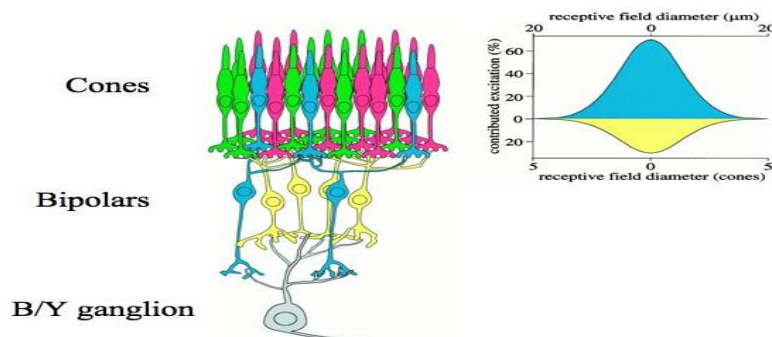
$$DoG(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} - \frac{1}{2\pi k^2\sigma^2} e^{-\frac{x^2+y^2}{2k^2\sigma^2}} \quad (1)$$

یکی از معایب مجموعه داده Caltech110 وجود عکس‌های مشابه از یک نمونه در هر گروه است که منجر به تضعیف نتیجه در پاسخ به نمونه‌های تنها (تک) می‌شود. ماتریس تفکیک‌پذیری عدم تشابه از یک لایه شبکه کانولوشنی نشان‌دهنده این ایراد در نمونه داده Caltech 110 است. در شکل (۴) وجود مربع‌های کوچک در ناحیه گروه‌ها نشان از وجود زیرگروه‌های مشابه در هر گروه است.

ماتریس تفکیک‌پذیری عدم تشابه 14 (RDM) به‌ازای همبستگی خروجی لایه‌ها برای تمام ورودی‌های مدل استخراج می‌شود و یک روش مؤثر برای نشان دادن عدم شباهت بین الگوهای پاسخ حاصل از دو تصویر است. بدینگونه که مقادیر خروجی هر لایه به صورت بردار در آمده و با مقادیر خروجی لایه متناظر به ازای ورودی متفاوت شباهت سنجی می‌شود. از توابع مختلفی چه

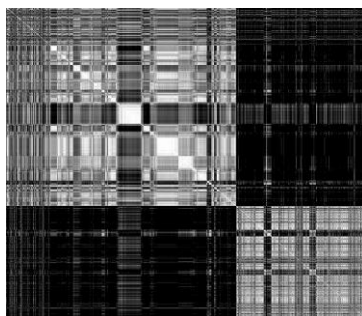
سنجش شباهت استفاده می‌شود. رابطه همبستگی در هر لایه بین خروجی لایه از تصاویر مختلف گرفته می‌شود. برای رسم ماتریس تفکیک‌پذیری عدم تشابه از جعبه ابزار تحلیل شباهت بازنمایی^{۱۵} (RSA) در متلب استفاده شد [۲۱].

Blue/yellow pathway



شکل (۳): لبه یابی بین دو لبه زرد و آبی و سپس انتقال داده از طریق سلولهای گنگلیون به ناحیه تالاموس [۱۸-۱۶]

Figure (3): Edging between the yellow and blue edges and then transferring data through the ganglion cells to the thalamus [24] - [22]



شکل (۴): ماتریس RDM نمونه‌های آزمایش برای یکی از لایه‌های مدل پایه [۲۱]

Figure (4): RDM matrix of test specimens for one of the base model layers [14]

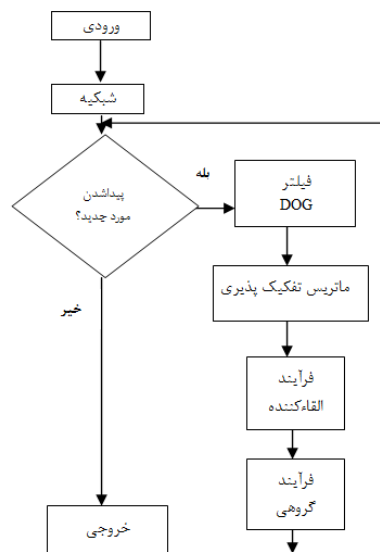


شکل (۵): فرآیند گروهی و القاءکننده [۱۲]

Figure (5): Group and inductive process [10]

بخش‌بندی تصویر پردازشی است که نواحی جالب (لبه‌ها) در تصویر را از سایر نواحی جدا می‌کند [۲۲]. در شکل (۵) فرآیند گروهی و القاءکننده، یک الگوریتم دو مرحله‌ای است که مجدداً تصویر در پردازش شرکت و تصاویر را انتخاب و تقسیم‌بندی می‌کند. الگوهای فیلتر القاءکننده‌ی درون‌یابی چهار لبه را در هر پیکسل تعریف می‌کنند: بالا، پایین، چپ و راست. این خطوط لبه برای شناسایی تصویر جهت کاهش فعالیت نورانی آن استفاده می‌شود [۱۲]. تصویر ورودی توسط الگوهای حساس به جهت گیری فیلتر می‌شود که لبه‌های درخشان‌تری را مجدداً جهت کاهش بیشتر فعالیت نورانی در هر یک شناسایی می‌نموده و تا هنگامی که هیچ مورد جدیدی پیدا نمی‌شود، سیستم به جستجو ادامه می‌دهد. شکل (۶) بلوک دیاگرام مدل پیشنهادی را نمایش می‌دهد.

اکنون بسیاری از کمپانی‌های عرصه‌ی هوش مصنوعی ساختاری برای شبکه‌های عمیق [۲۳، ۲۴]. ارائه داده‌اند که تقریباً همه آنها از ساختار اولیه شبکه عمیق کانولوشنی پیروی می‌کنند. در این کار از شبکه کانولوشنی عمیق اولیه هشت لایه استفاده شده است و تلاش می‌شود تاثیر اضافه شدن شبکه بر روی شبکه عمیق ارزیابی شود [۲۵].



شکل (۶): بلوک دیاگرام فلوجارت پیشنهادی
Figure (6): Block diagram of the proposed model

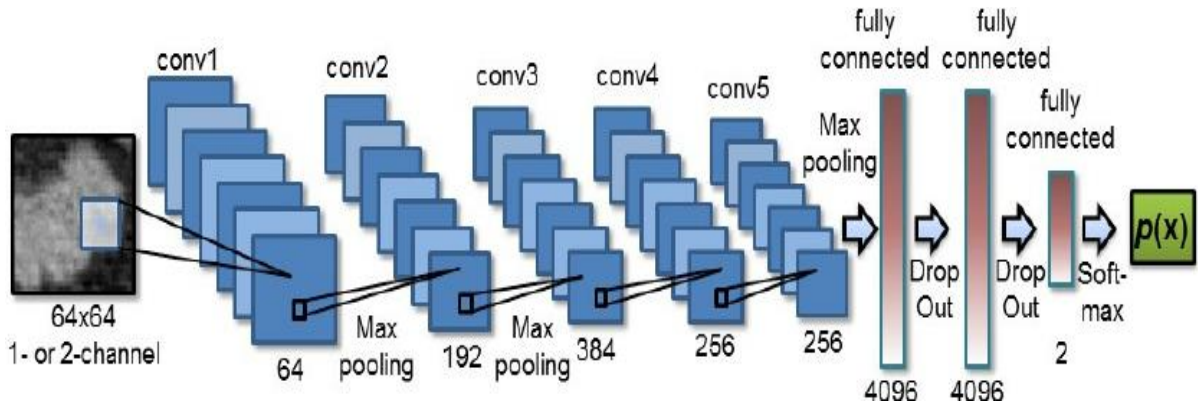
در شکل (الف-۷) ساختار شبکه عمیق کانولوشنی مدل پایه نشان داده شده است [۱۲]. در این ساختار بعد از هر لایه کانولوشنی یک لایه بیشینه‌گیری قرار داده شده که وظیفه کاهش بعد داده را دارند. در انتها سه لایه پرسپکتیونی تمام متصل قرار داده شده است و در نقش طبقه‌بند هستند. ورودی یک بار با تصاویر RGB^{۱۶} و بار دیگر با تصاویر خروجی شبکه در شبکه عمیق آزمایش می‌شود و بدین صورت یک لایه به نام شبکه به شبکه عمیق اضافه می‌گردد.

در شکل (ب-۷) ساختار مدل پیشنهادی مشخص است. در شبکه‌های عمیق رنگی پیشین به ویژه کانولوشنی، از تصاویر در دستگاه رنگی RGB برای آموزش شبکه‌ها استفاده می‌شده است و علیرغم ادعاهای زیستی بودن مدل‌ها، ولی هیچ شبیه‌سازی از شبکه چشم انجام نشده است. از این رو باید تاثیر شبکه چشم و الگوریتم تقسیم‌بندی و درون‌یابی با الهام از مغز انسان بر روی کارایی مدل‌های عمیق مورد ارزیابی قرار گیرد. جهت جلوگیری از افزایش بی‌رویه بعد تصاویر تنها از سه صفحه استفاده شده است. اولین مورد صفحه‌ی مربوط به شدت نور بوده و فیلتر DoG بر روی آن اعمال شده است. صفحه دوم دو رنگ قرمز و سبز و صفحه سوم دو رنگ زرد و آبی را رمز می‌کنند. در شکل (۸) یک نمونه از نتیجه رمزنگاری بر روی تصویر برگ درخت در لایه‌های مختلف نشان داده شده است. مشاهده می‌شود که انرژی تصویر به مقدار بسیار زیادی کاهش یافته است ولی ماهیت تصویر حفظ شده است. اگر انرژی یک تصویر از رابطه (۲) به‌دست بیاید مشاهده می‌شود که انرژی تصویر رمزنگاری شده در شبکه به میزان ۸۳ درصد کاهش یافته است [۲۶، ۲۷]. این کاهش انرژی باعث می‌شود تا نوران‌های کمتری در شبکه عمیق فعالیت داشته باشند و به عبارتی شبکه را تنک می‌کنند.

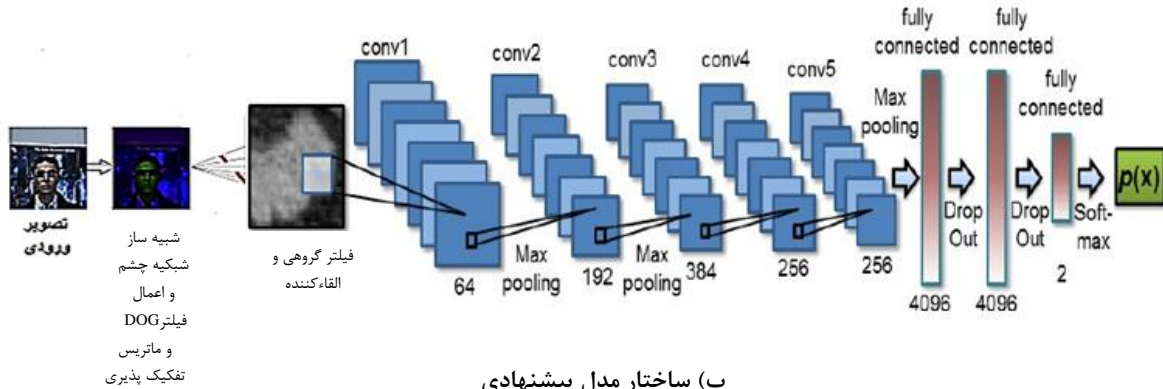
$$E = \sqrt{x_1^2 + \dots + x_n^2} \quad (2)$$

۳- شبیه‌سازی و تحلیل مدل پیشنهادی

جهت ارزیابی مدل‌ها از نمونه داده کلتک و نرم افزار متلب استفاده شده است. مدل پیشنهادی دوکلاسه است زیرا زمان محاسبات و پیچیدگی ارزیابی کاهش یابد. طبق نظریه اصلی در شبکه رمزنگاری رنگی بر روی تصاویر انجام می‌گیرد و ماهیت ریخت تصاویر تغییر نکرده و حتی به دلیل اعمال فیلتر گوسی وضوح شکل کاهش می‌یابد. مطابق این اصل مدل را از دو منظر ریخت و رنگ مورد ارزیابی قرار می‌گیرد. یعنی دو گروه تصاویری که انتخاب می‌شود یکبار مشخصه رنگ متفاوت داشته باند و در بار دوم مشخصه ریخت متفاوت داشته باشند.



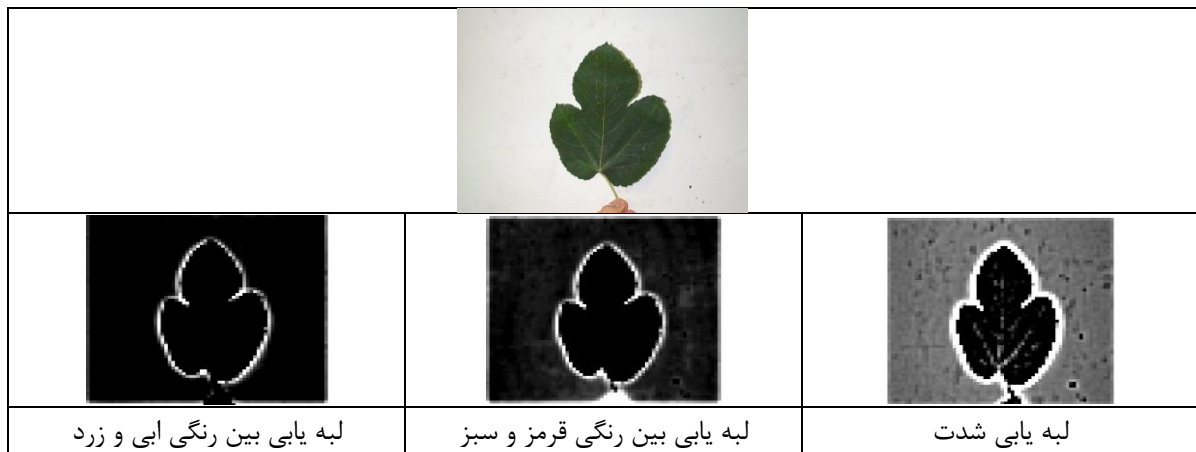
الف) ساختار شبکه عصبی عمیق هشت لایه مدل پایه [۱۲]



ب) ساختار مدل پیشنهادی

شکل (۷): الف) ساختار شبکه عصبی عمیق هشت لایه مدل پایه ب) ساختار مدل پیشنهادی

Figure (7): a) The structure of the deep neural network of eight layers of the base model
b) The structure of the proposed model

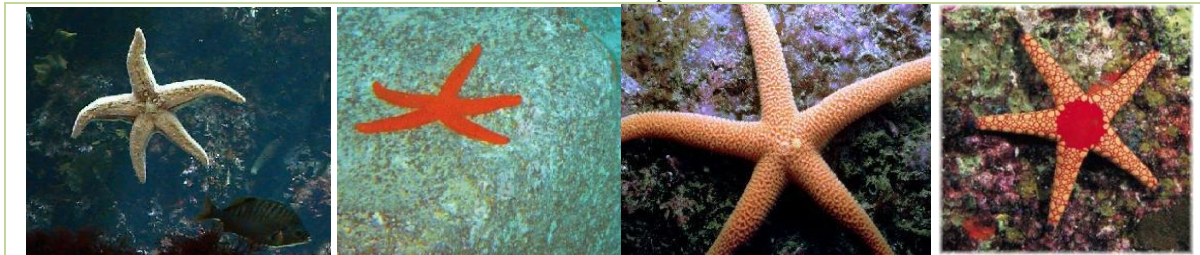


شکل (۸): نتیجه رمز نگاری شبکه چشم بر روی تصویر برگ درخت

Figure (8): The result of retina cryptography on a tree leaf image

در داده کلکت دو کلاس ستاره دریایی و برگ دارای تقریباً ریخت یکسان ولی رنگ متفاوت هستند. از این رو این دو جهت ارزیابی رنگی مدل انتخاب می‌شوند. در آزمایش دوم از دو گروه صورت و اتومبیل استفاده می‌شود که ریخت متفاوت و رنگ متنوع دارند. یکی از مهم‌ترین پارامترها در آموزش شبکه‌های کانولوشنی طول اندازه-دسته است. منظور از اندازه-دسته تعداد نمونه‌های مورد نیاز جهت اصلاح وزن‌ها در هر تکرار است. اندازه دسته ۵ انتخاب می‌شود. این عدد متداول بین نمونه کارهای مشابه است.

ستاره دریایی ۸۶ نمونه
Starfish 86 samples



برگ ۱۸۶ نمونه
Leaf 186 Sample

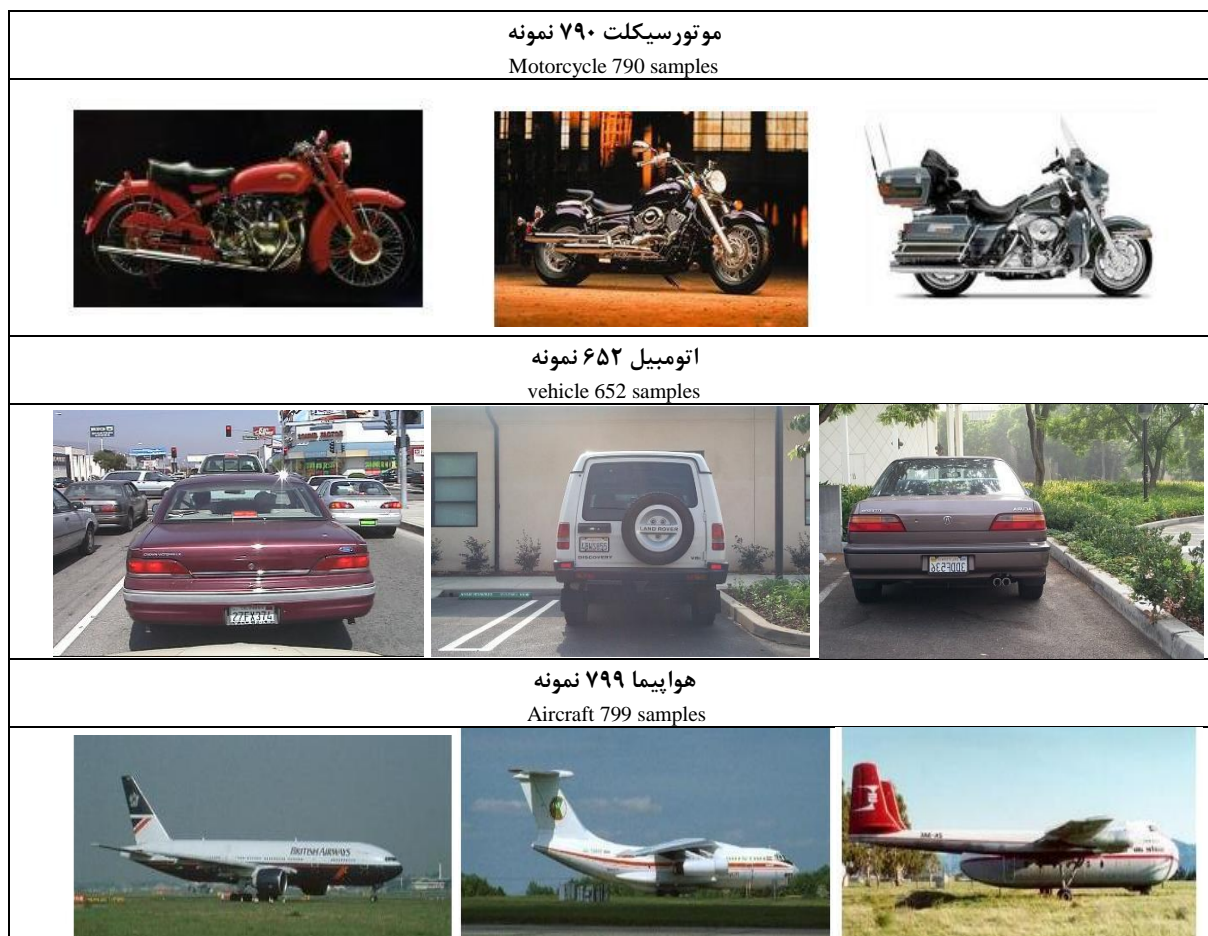


شکل (۹): تعدادی نمونه های برگ و ستاره دریایی و تعداد نمونه ها
Figure (9): Number of leaf and starfish samples and number of samples

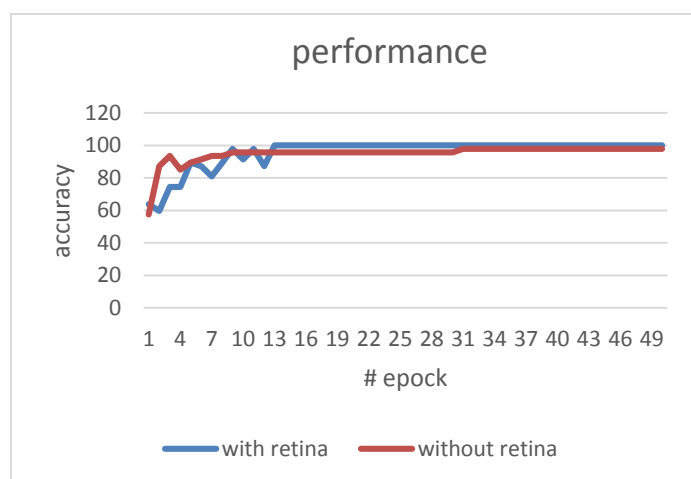
در شکل (۹) نمونه‌های از تصاویر گروه‌های ستاره دریایی و برگ نشان داده شده است. تصاویر از زاویه ثابت گرفته شده‌اند و تنها از فواصل مختلف تا جسم عکاسی شده است. مشاهده می‌شود که دو گروه از نظر ریخت بسیار شبیه هستند و مشخصه تفکیک-کنندگی غالب رنگ است. ستاره‌های دریایی معمولاً به رنگ گرم و برگ درختان به رنگ سبز هستند. در شکل (۱۰) نمونه تصاویر سه گروه وسیله نقلیه نشان داده شده است. تعداد نمونه‌ها نسبت به آزمایش قبلی بیشتر شده است. مشاهده می‌شود که تصاویر از نظم رنگی مشخصی پیروی نمی‌کنند. بنابراین مشخصه تفکیک‌پذیر ریخت تصاویر است. در آزمایش‌های ۲۵ درصد تصاویر برای آزمایش و بقیه برای آموزش شبکه استفاده می‌شوند. البته این آزمایشات به تعداد ۵ مرتبه تکرار شده‌اند و نتایج میانگین هستند. در آزمایش اول از تصاویر کلاس ستاره دریایی و برگ جهت آموزش مدل استفاده شد. نرخ بازشناسی به صورت درصد نمونه‌ی درست به کل نمونه‌های به‌ازای ۵۰ چرخه در شکل (۱۱) نشان داده شده است. با توجه به نمودار در صورت استفاده از شبکه نتیجه در چرخه ۱۰ و زودتر از حالت بدون شبکه پایدار می‌شود و همچنین نرخ بازشناسی در حدود ۳ درصد بالاتر دارد.

در ادامه آزمایش با تغییر جزئی مدل انرا برای سه کلاس خودرو، هواپیما و موتورسیکلت آماده شده است. در شکل (۱۲) نتیجه بازشناسی به‌ازای ۵۰ چرخه (ایپاک) نشان داده شده است. مشاهده می‌شود که مدل با شبکه ناپایداری بسیار بالاتری دارد. اما نرخ بازشناسی کمی بهتر ثبت کرده است. البته ناپایداری مدل پیشنهادی مشهود است. این آزمایشات از ۲۵ درصد نمونه‌ها جهت آزمایش استفاده شده است و کل فرایند آزمایش ۵ مرتبه به صورت تصادفی تکرار می‌شود. نمودار رسم شده میانگین کل نتایج حاصل از تکرار آزمایشات است.

باتوجه به نتایج در دو آزمایش می‌توان این نتیجه‌گیری را کرد که رمزنگاری رنگی زمانی تاثیر مثبت دارد که مشخصه غالب تفکیک‌پذیری کلاس‌ها رنگ باشد. در آزمایش اول دو کلاس بسیار به هم از نظر ریخت شبیه هستند ولی غالباً رنگ متفاوتی دارند و مدل پیشنهادی توانسته کارایی بهتری ثبت کند. ولی در حالت دوم یعنی بین سه کلاس وسیله نقلیه که مشخصه ریخت متفاوت دارند مدل نرخ بازشناسی بهتری ثبت نکرده البته در حالت پایدار کمی بهتر بوده است. در شکل (۱۲) مشخص است که در حالت اضافه شدن شبکه مینم‌های محلی افزایش یافته است و ممکن است مدل در یکی از آن‌ها گرفتار شود. جهت ارزیابی بیشتر مدل و نتیجه‌گیری بهتر لازم دانسته تا مدل بر روی داده‌های بیشتر آزمایش شود. از اینرو نرخ بازشناسی برای زیرمجموعه‌های دوکلاس از داده کلکت ۱۱۰ استخراج شد. در این آزمایش‌ها به این نکته توجه می‌شود که مشخصه رنگ یا شکل چه تاثیری در بازشناسی دارد. در شکل (۱۳) نرخ بازشناسی به‌ازای زیرمجموعه‌های مختلف رسم شده است.

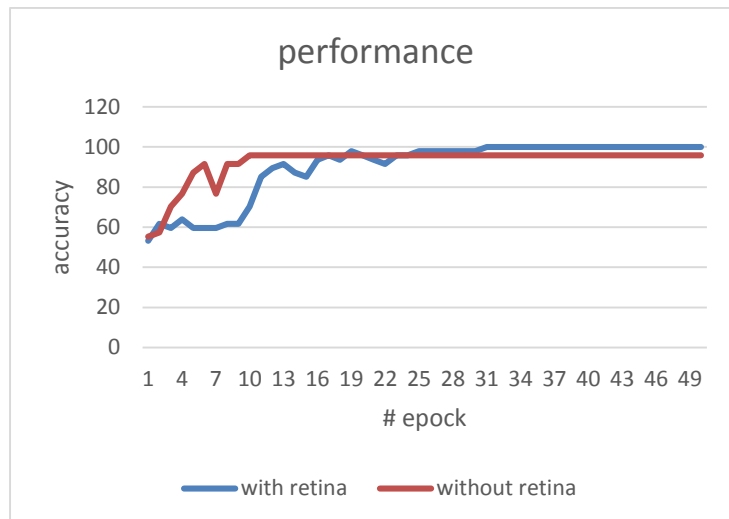


شکل (۱۰): نمونه تصاویر از سه گروه وسیله نقلیه اتومبیل، هواپیما و موتورسیکلت به همراه تعداد نمونه های آنها
 Figure (10): Sample images of three groups of vehicle, aircraft, and motorcycles with the number of samples



شکل (۱۱): نرخ بازشناسی مدل استاندارد قبل و بعد از الحاق شبکه برای داده دو گروه ستاره دریایی و برگ درختان از مجموعه داده کلکت

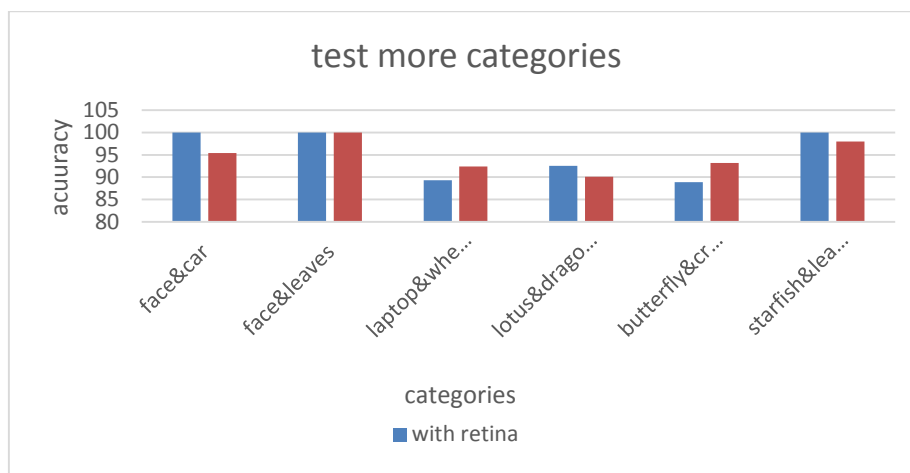
Figure (11): Recognition rate of standard model before and after retinal annexation for two-group data of starfish and tree leaves from Coltech data set



شکل (۱۲): نرخ بازشناسی مدل استاندارد قبل و بعد از الحاق شبکه برای داده سه گروه اتومبیل، هواپیما و موتورسیکلت از مجموعه داده کلتک

Figure (12): Recognition rate of the standard model before and after retinal annexation for three groups of cars, aircraft and motorcycles from the Caltech data set

در دسته‌هایی که رنگ مشخصه تفکیک غالب است نرخ بازشناسی مدل با شبکه بهتر بوده است. مانند تشخیص صورت از ماشین زیرا رنگ صورت انسان تقریباً در یک مقدار است. همچنین در مواردی که رنگ مشخصه مناسبی برای تفکیک نیست و ریخت و شکل بسیار غالب هستند، مدل پایه توانسته بهتر عمل کند. برای مثال در تشخیص پروانه از خرچنگ، مدل پایه بهتر بوده زیرا هر دو دسته پروانه و خرچنگ رنگ‌های گوناگونی دارند.



نمودار (۱۳): نرخ بازشناسی در زیرمجموعه‌هایی از مجموعه داده کلتک

Figure (13): Recognition rates in subsets of the Caltech database

۳-۱- آزمایش مدل بر روی مجموعه داده حیوان-غیرحیوان

یکی از آزمایش‌های رایج در بازشناسی تشخیص حیوان از غیر حیوان است. مجموعه داده از دو گروه تصاویر تشکیل شده است که در گروه اول حیوان وجود دارد و در گروه دوم حیوانی وجود ندارد. در شکل (۱۴) نمونه‌هایی از تصاویر مجموعه داده مورد آزمایش به همراه تعداد هر گروه نشان داده شده است. در شکل (۱۵) نرخ تشخیص برای تصاویر حیوانات به ازای تکرار اصلاح وزنی ۵۰ رسم شده است. همان‌طور که در شکل‌های (۱۱)، (۱۲)، (۱۳) و (۱۵) مشاهده می‌گردد، مدل پیشنهادی توانست نرخ

بازشناسی بهتری را ثبت کند. بیشینه نرخ بازشناسی مدل پیشنهادی به ۹۳ درصد می‌رسد و مدل پایه نزدیک به ۹۱ درصد است. البته مدل پایه زودتر به نقطه بهینه می‌رسد.

تصاویر بدون حیوان - ۶۰۰ نمونه

Non Animal images - 600 samples

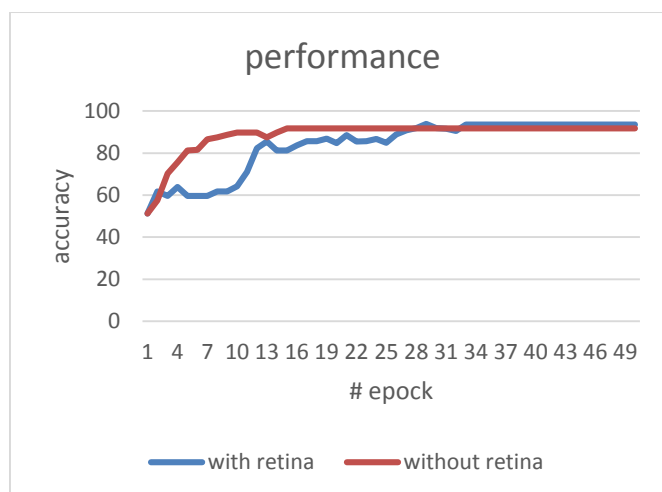


تصاویر حیوان - ۶۰۰ نمونه

Animal images - 600 samples



شکل (۱۴): نمونه‌ای از تصاویر حیوان-غیرحیوان
Figure (14): An example of animal- non animal images



شکل (۱۵): نرخ تشخیص حیوان در تصاویر
Figure (15): Animal detection rates in images

۲-۳- آزمایش میزان تنکی مدل

در این آزمایش میزان فعالیت نورون‌ها مورد بررسی قرار گرفت. هدف از اینکار بررسی میزان تنکی یا اسپارسنس مدل است. همان‌طور که قبلاً اشاره شد مغز انسان با کمترین فعالیت نورانی می‌تواند عمل شناسایی را انجام دهد. در مقاله‌ای از اندرو ان جی [۲۸] و سه عامل برای تنکی مدل بیان کرده است: (۱) هر کلاس تعداد محدودی ویژگی غالب داشته باشد (۲، ۱^۷، هر ویژگی منحصر به یک کلاس منحصر به فرد باشد^{۱۸} و (۳) توزیع ویژگی‌ها برای هر کلاس واحد باشد^{۱۹}.

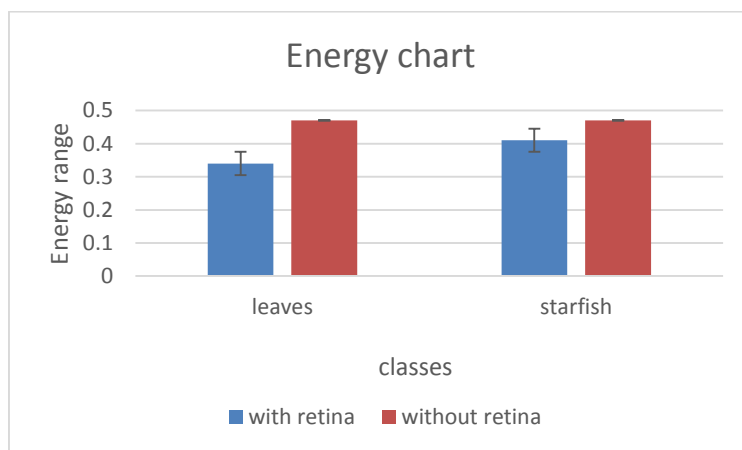
پی بردن به هر سه ویژگی کار بسیار سختی هست ولی در آزمایش‌های مختلف می‌توان تا حدودی هر ویژگی را به گونه جداگانه بررسی کرد. در ویژگی اول با انرژی مدل می‌توان تقریباً این مورد را بررسی کرد. در محاسبه انرژی از [۲۱،۲۰] استفاده شده و از رابطه $E = \sqrt{N_1^2 + N_2^2 + \dots + N_n^2}$ به دست می‌آید. در این رابطه N میزان خروجی هر نوران پس از فعال شدن است. هر چه این مقدار کمتر باشد نشان می‌دهد فعالیت نورانی در بازنشاسایی کاهش یافته است. در اینجا جهت مقایسه بهتر، رابطه انرژی نرمال می‌شود و با توجه به اینکه تابع فعال‌ساز هر نورون تابع تانسینگ هست بنابراین انرژی نرمال شده از رابطه (۳) به دست می‌آید:

$$E = \frac{\sqrt{N_1^2 + N_2^2 + \dots + N_n^2}}{\sqrt{n}} \quad (3)$$

نتیجه به صورت میانگین و انحراف معیار برای هر کلاس از داده در آزمایش ستاره دریایی و برگ در قالب نمودار میله‌ای در شکل (۱۶) رسم شده است. این نتیجه نشان می‌دهد که در حالت استفاده از شبکه میزان فعالیت نورون‌ها کاهش می‌یابد و این می‌تواند در شباهت بیشتر مدل به عملکرد مسیر بینایی تاثیرگذار باشد. در نمودار میزان واریانس به‌ازای تصاویر مختلف هر کلاس نشان داده شده است.

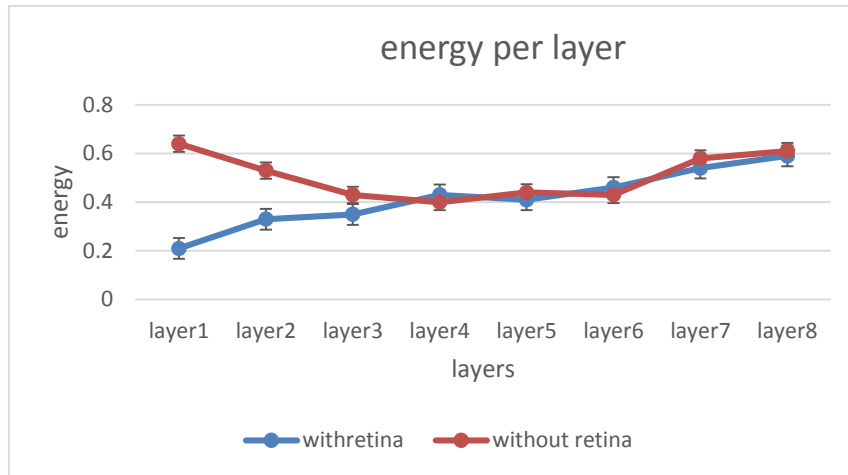
در شکل (۱۷) تحلیل انرژی در لایه‌های مختلف مدل‌ها انجام گردید و مشاهده شد که اختلاف اساسی در لایه‌های ابتدایی است. در لایه‌های ابتدایی مدل پیشنهادی با انرژی بسیار کمتری شروع می‌کند و به مراتب افزایش می‌یابد. در مدل پایه انرژی لایه‌های اولیه بسیار زیاد هستند و آهسته به مقدار مشخصی میرا می‌شوند. در لایه‌های خروجی دو مدل کارکردی تقریباً مشابه دارند. علت افزایش انرژی زیاد در مدل پایه را در استفاده از تصاویر خام است. زیرا مقدار پیکسل تصاویر خام بسیار زیاد است ولی در صورت استفاده از شبکه مقدار پیکسلی تصاویر کاهش می‌یابد، اصطلاحاً انرژی تصاویر بسیار کمتر می‌شود.

شرط دوم برای تنگی از نظر اندرو ان جی انحصار هر ویژگی به کلاسی خاص است. تحقیق این ویژگی از طریق استخراج ماتریس تفکیک‌پذیری (RDM) قابل بررسی است. ماتریس RDM به‌ازای همبستگی خروجی لایه‌ها به‌ازای تمام ورودی‌های مدل استخراج می‌شود. بدینگونه که مقادیر خروجی هر لایه به صورت بردار در آمده و با مقادیر خروجی لایه متناظر به‌ازای ورودی متفاوت شباهت سنجی می‌شود [۲۱]. از توابع مختلفی در سنجش شباهت استفاده می‌شود. در اینجا از رابطه (۴) استفاده شده است. رابطه همبستگی در هر لایه بین خروجی لایه از تصاویر مختلف گرفته می‌شود. در یک شبکه مناسب میزان شباهت برای خروجی با ورودی‌هایی از یک کلاس باید بیشتر باشد.

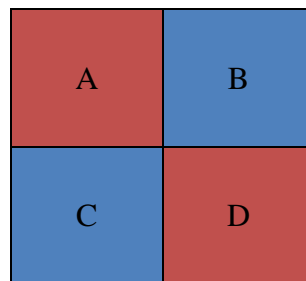


شکل (۱۶): میزان انرژی مدل یا به عبارتی فعالیت نورون‌ها برای دو کلاس برگ و ستاره دریایی نشان می‌دهد که در مدل پیشنهادی فعالیت نورانی کاهش محسوسی داشته است.

Figure (16): The energy level of the model, or in other words, the activity of neurons for the two classes of leaves and starfish, shows that in the proposed model, the luminous activity has significantly decreased.



شکل (۱۷): انرژی لایه‌های مختلف مدل‌ها به‌ازای تمامی تصاویر
Figure (17): Energy of different layers of models for all images



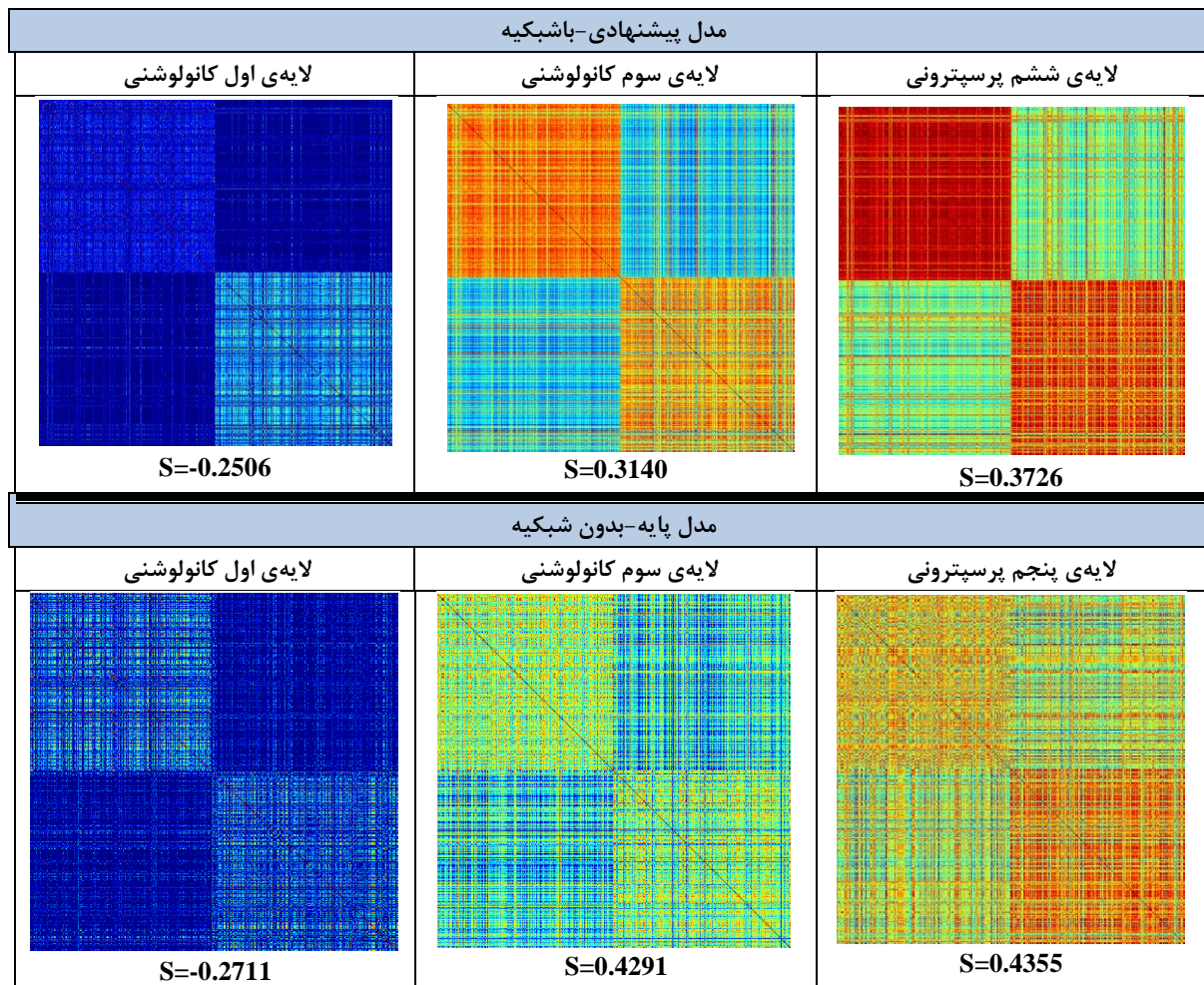
شکل (۱۸): فرم ماتریس تفکیک پذیری برای دو کلاس
Figure (18): Resolution matrix form for two classes

$$\text{corr} = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left[\sum_m \sum_n (A_{mn} - \bar{A})^2 \right] \left[\sum_m \sum_n (B_{mn} - \bar{B})^2 \right]}} \quad (4)$$

$$s = \frac{\text{sum}(C) + \text{sum}(B)}{\text{sum}(A) + \text{sum}(D)} \quad (5)$$

که A و B درایه‌های فضا و m سطر m درایه‌ی فضای A و n سطر n درایه‌ی فضای B است. اگر فرض شود شکل (۱۸) شماتیک یک ماتریس RDM دو گروه باشد آنگاه اگر درایه‌های فضای A و D بیشینه و مقادیر فضاهای B و C کمینه باشند، آنگاه تفکیک‌پذیری مدل در شرایط مطلوبی بوده است. بنابراین کمیت s طبق معادله (۵) به‌گونه‌ای تعریف می‌شود که به‌ازای ماتریس RDM مطلوب، مقدار حداقلی را داشته باشد و به عبارت دیگر مقدار کمتر برای کمیت s نشان از تفکیک‌پذیری بهتر دارد.

در ماتریس تفکیک‌پذیری اگر اختلاف سطوح شباهت و غیر شباهت هرچه بیشتر باشد نشان می‌دهد که تفکیک‌پذیری ویژگی‌ها بیشتر بوده است. در شکل (۱۹) ماتریس‌های تفکیک‌پذیری برای سه لایه از شبکه نشان داده شده است. مقادیر قرمز رنگ مقادیر شباهت بالا و مقادیر آبی رنگ میزان شباهت کم را نشان می‌دهد. شهودی مشخص است که در مدل پیشنهادی میزان شباهت در کلاس همنام بیشتر است و در عوض میزان غیر شباهت میان دو کلاس ناهمنام نیز در مدل پیشنهادی بیشتر است. البته با مقایسه کمیت s نیز نتیجه‌گیری می‌شود که در مدل پیشنهادی تفکیک‌پذیری در لایه‌های مدل بالاتر رفته است. از این ماتریس نتیجه گرفته می‌شود که انحصار ویژگی‌ها به کلاس خاص افزایش یافته است زیرا در تعداد ویژگی‌هایی که به هر دو کلاس پاسخ یک جور دهند کاهش یافته است.



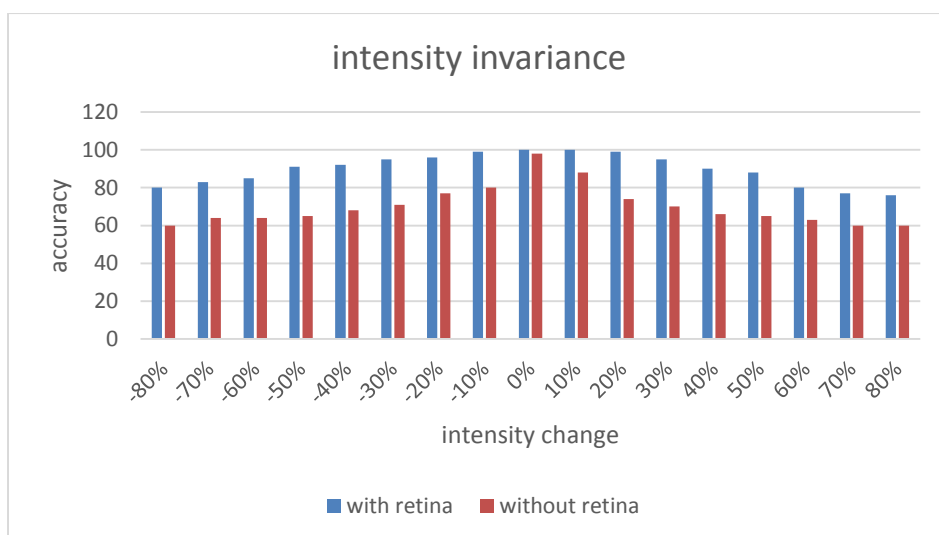
شکل (۱۹): ماتریس تفکیک‌پذیری برای سه لایه از دو مدل پیشنهادی و پایه
Figure (19): Representational Dissimilarity Matrix for three layers of the two proposed and basic models

۳-۳- پایداری مدل در برابر تغییرات رنگی

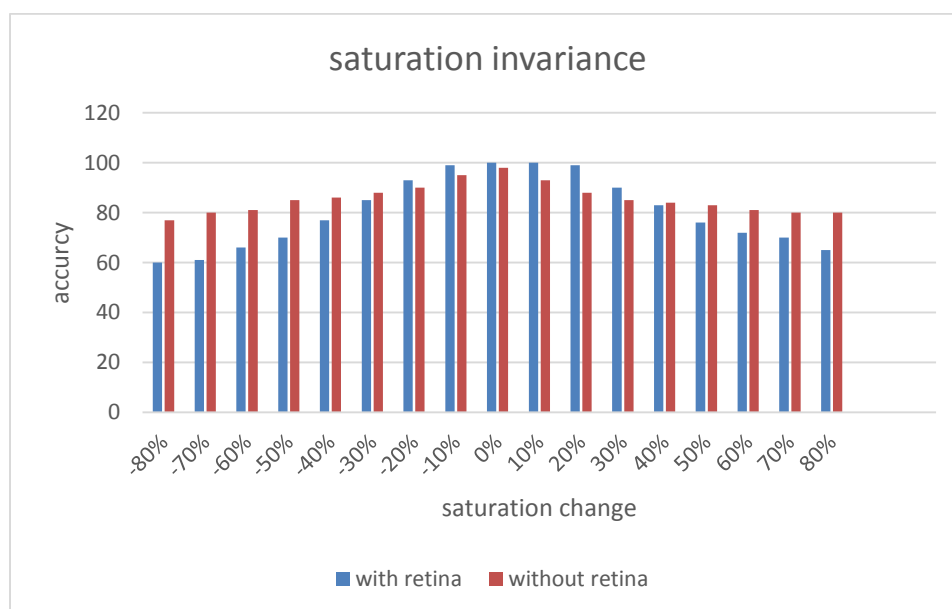
در بازشناسی تصاویر دو مشخصه اصلی جهت تفکیک بیشترین نقش را دارند. اولین مشخصه بافت و شکل است. مشخصه دوم رنگ است. رنگ تعاریف مختلفی دارد و از فاکتورهای مختلفی تشکیل شده است. مثلاً در سیستم RGB سه فاکتور قرمز، سبز و آبی هستند. در سیستم HIS^{۲۰} که کاربرد بالایی در تحلیل دقیق از مولفه‌های رنگی دارد سه فاکتور رنگدانه، اشباع شدگی و شدت روشنایی جهت تحلیل رنگ بیان می‌شود. مدهای مبتنی به تصاویر رنگی، پایداری مدل در برابر تغییرات فاکتورهای رنگی می‌تواند نقش مهمی داشته باشد. دو فاکتور شدت روشنایی و اشباع رنگ بسیار وابسته به محیط اطراف و شرایط عکس‌برداری هستند و یک عامل منفی در بازشناسی صحیح هستند. اما در مورد رنگدانه وضوح کمی متفاوت است. در بسیاری از مسائل که رنگدانه عامل مهم در بازشناسی است بنابراین پایداری به رنگدانه عاملی منفی تلقی می‌شود.

در آزمایش اول میزان شدت روشنایی تصاویر تغییر یافت و میزان دقت بازشناسی بررسی گردید. در شکل (۲۰) میزان دقت بازشناسی به‌ازای تغییرات شدت روشنایی نشان داده شده است. نمودار میله‌ای قرمز رنگ مدل پایه و میله‌های آبی مدل پیشنهادی است و با توجه به نمودار ملاحظه می‌شود که مدل پیشنهادی افت بسیار کمی نسبت به مدل پایه دارد. این یکی از مهم‌ترین شایستگی مدل پیشنهادی است و ثابت می‌کند که وابستگی مدل به شدت روشنایی بسیار کم است. استقلال مدل نسبت به روشنایی بسیار مهم است زیرا در زمان عکس‌برداری، شدت روشنایی عاملی تغییر پذیر است و میزان کردن آن امری دشوار است. علت اصلی این پایداری در اعمال فیلترهای لبه‌یاب بین رنگی است که میزان شدت روشنایی تصویر را بی‌تاثیر می‌کند.

در آزمایش بعدی میزان اشباع شدگی رنگی تصاویر تغییر یافت و میزان نرخ بازشناسی در شکل (۲۱) نشان داده شده است. در این نمودار ملاحظه می‌شود مدل پیشنهادی بعد تغییر ۲۰ درصد وابستگی بیشتری نسبت به مدل پایه از خود نشان می‌دهد و ناگهان افت زیادی را تجربه می‌کند. این حساسیت نسبت اشباع رنگ می‌تواند در بعضی از مسائل که رنگ کلیدی است نقش مثبت داشته باشد.

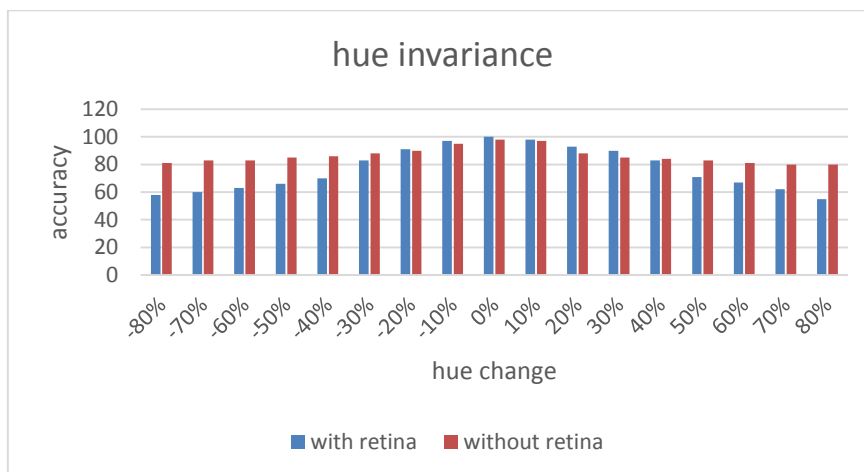


شکل (۲۰): مقدار نرخ بازشناسی به ازای تغییرات شدت روشنایی
Figure (20): Recognition rate value for changes in light intensity



شکل (۲۱): مقدار نرخ بازشناسی به ازای تغییرات اشباع رنگی
Figure (21): Recognition rate value for color saturation changes

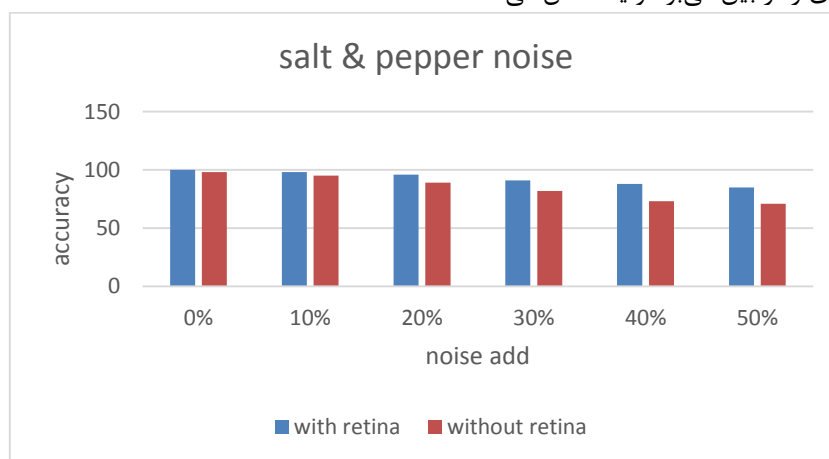
در آزمایش رنگدانه مقدار رنگدانه در فضای HSI تغییر کرد و میزان دقت مدل سنجیده شد. مشاهده شد که به ازای تغییرات رنگدانه میزان بازشناسی مدل پیشنهادی به شدت کاهش می‌یابد. با تحلیل شکل (۲۲) نتیجه‌گیری می‌شود که مدل پیشنهادی وابستگی بالاتری به رنگ دارد. این وابستگی در مسائلی که رنگ عامل مهم در تفکیک هست، بسیار مهم می‌شود. بنابراین از این منظر حساسیت مدل به رنگ یک عامل مثبت تلقی شده و نقشی مثبت در تفکیک رنگی تصاویر دارد.



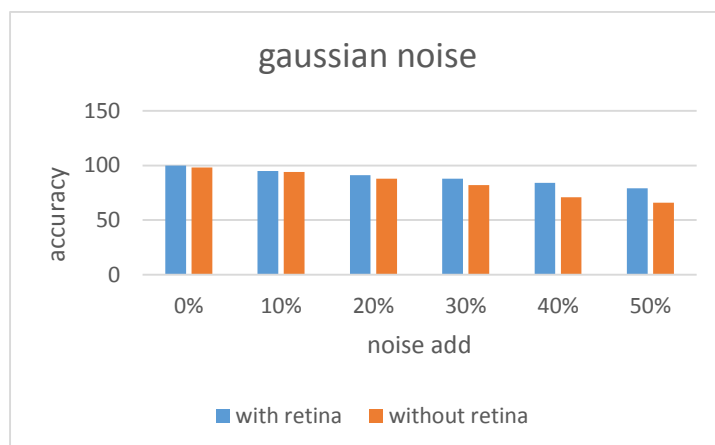
نمودار (۲۲): مقدار نرخ بازشناسی به ازای تغییرات رنگ
Figure (22): Recognition rate value for color changes

۳-۴- پایداری در برابر نویز

نویز یکی دیگر از عوامل مهم در شناسایی تصاویر هستند. نویز نمک و فلفل یکی از حملاتی حوزه مکانی به شمار می‌آید که پس از اعمال بر روی یک تصویر بر اساس درصد نویز اعمال شده، تعدادی از پیکسل‌های تصویر را به رنگ سفید و سیاه تبدیل می‌کند و باعث خراب شدن تصویر می‌گردد [۲۹]. هرچه نویز تصاویر افزایش یابد قطعاً تشخیص مشکل‌تر می‌گردد. نویزهای فراوانی در حوزه تصویر وجود دارند. در اینجا فقط دو نویز گوسی و فلفل نمکی مورد بررسی قرار می‌گیرند. در آزمایش‌های انجام‌شده به تصاویر آزمایش ورودی نویزهای فلفل نمکی و گوسی در مقادیر مختلف اضافه شد و پایداری مدل‌ها در برابر این نویزها سنجیده شد و مشاهدات در شکل‌های (۲۳) و (۲۴) به ازای تغییرات نرخ بازشناسی نشان داده شده است. در بعضی نمونه‌ها حتی تشخیص با چشم نیز امکان‌پذیر نیست. اما مدل‌ها تا حدودی کارایی خود را حفظ کرده‌اند. مدل پیشنهادی کمترین افت را در بازشناسی دارد و توانسته بهتر عمل کند. علت را می‌توان در اعمال فیلترهای گوسی بین رنگی دانست، زیرا این فیلتر گوسی تاثیر پیکسل‌های نویزی را از بین می‌برد و یا کاهش می‌دهد.



شکل (۲۳): میزان تغییرات نرخ بازشناسی به ازای افزایش نویز فلفل نمکی
Figure (23): The rate of change of the recognition rate in exchange for the increase of salt pepper noise



شکل (۲۴): میزان تغییرات نرخ بازشناسی به ازای افزایش نویز گوسی
 Figure (24): The rate of change of the recognition rate in exchange for the increase of Gaussian noise

۴- نتیجه‌گیری

در طول تاریخ مغز انسان به عنوان یکی از الهام بخش‌ترین پدیده در عرصه هوش مصنوعی بوده است. در این مقاله تلاش شد تا با الهام از مسیر بینایی سیستم عصبی انسان و یک الگوریتم تقسیم‌بندی و درون‌یابی الهام گرفته از انسان، تغییری بهینه در مدل‌های عمیق ایجاد گردد. این تبدیل رنگی که در شبکه انجام می‌گیرد را در دل شبکه عمیق کانولوشنی الحاق شد. طبق پیش‌بینی اعمال کدینگ رنگی ویژگی رنگ را در شبکه با ارزش‌تر قرار می‌دهد و بنابراین در شبکه‌هایی که رمزنگاری رنگی اعمال شده بود مشخصه رنگ جهت تفکیک دسته‌ها ارجح‌تر از دیگر مشخصه‌ها مانند ریخت و شکل شده بود. از این‌رو در مسائلی که رنگ عامل مهمی در بازشناسی بود، همان‌طور که در شکل‌های (۱۱)، (۱۲)، (۱۳) و (۱۵) مشاهده می‌گردد، مدل پیشنهادی توانست نرخ بازشناسی بهتری را ثبت کند. در ادامه جهت ارزیابی بیشتر مدل شرایط تنگی سنجیده شد و همان‌طور که در شکل (۱۶) مشخص است مدل پیشنهادی تنک‌تر و یا به‌عبارتی اسپارستر است. نرخ آتش نوروها در لایه‌های مختلف هر دو مدل مورد بررسی قرار گرفت و نشان داده شد که مدل پیشنهادی چنانچه در شکل (۱۷) قابل مشاهده است در لایه‌های ابتدایی بسیار انرژی کمتری دارد و به‌عبارتی نوروهای کمتری به نقطه فعال یا آتش می‌رسند. در مدل پیشنهادی در شکل (۱۹) دیده می‌شود که تفکیک‌پذیری در لایه‌های مدل بالاتر رفته است. در انتها پایداری مدل‌ها در برابر تغییرات کمیت‌های رنگی و نویز مورد ارزیابی قرار گرفت. مدل پیشنهادی در شکل (۲۰) مشخص است که پایداری بالایی به تغییرات شدت روشنایی دارد ولی حساسیت بیشتری در شکل (۲۱) به غلظت رنگ و در شکل (۲۲) به تغییرات رنگدانه دارد. این نکته در مواردی که رنگ عمل اصلی بازشناسی باشد نقش مثبت ایفا می‌کند. همچنین مدل پیشنهادی چنانچه در شکل‌های (۲۳) و (۲۴) ملاحظه می‌گردد توانست در مواجهه با تصاویر نویزی پاسخ بهتری نشان دهد و افت بازشناسی کمتری را ثبت کند، علت این امر در وجود فیلترهای گوسی بین رنگی است که تاثیر نویز را تا حدودی کم اثرتر می‌کنند.

References

مراجع

- [1] M. Riesenhuber, T. Poggio, "Hierarchical models of object recognition in cortex", *Nature Neuroscience*, vol. 2, no. 11, pp. 1019–25, Nov. 1999 (doi: 10.1038/14819).
- [2] K. Fukushima, "Artificial vision by multi-layered neural networks: neocognitron and its advances", *Neural Network*, vol. 37, no. 2013, pp. 103–19, Jan. 2013 (doi: 10.1016/j.neunet.2012.09.016)
- [3] Y. Li, W. Wu, B. Zhang, F. Li, "Enhanced HMAX model with feedforward feature learning for multiclass categorization", *Frontiers in Computational Neuroscience*, vol. 9, p. 123, Oct. 2015 (doi: 10.3389/fncom.2015.00123).
- [4] B. Yang, L. Zhou, Z. Deng, "C-HMAX: Artificial cognitive model inspired by the color vision mechanism of the human brain", *Tsinghua Science Technology*, vol. 18, no. 1, pp. 51–56, Feb. 2013 (doi: 10.1109/TST.2-013.6449407).

- [5] 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*, vol. 101, pp. 82–93, Aug. 2014 (doi: 10.1016/j.visres.2014.05.006).
- [6] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks", in *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [7] D. Ciregan, U. Meier, J. Schmidhuber, "Multi-column deep neural networks for image classification", *Proceeding of the IEEE/CVPR*, pp. 3642–3649, Providence, RI, USA, June 2012 (doi: 10.1109/CVPR.2012.6248110).
- [8] D. Berman, A. Buczak, J. Chavis, C. Corbett, "A survey of deep learning methods for cyber security", *Information*, vol. 10, no. 4, Article Number: 122, April 2019 (doi:10.3390/info10040122).
- [9] A. Gupta, H. K. Thakur, R. Shrivastava, P. Kumar, S. Nag, "A big data analysis framework using apache spark and deep learning", *Proceeding of the IEEE/ICDMW*, pp. 9–16, New Orleans, LA, USA, Nov. 2017 (doi:10.1109/ICDMW.2017.9)
- [10] L. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning", *APSIPA Trans. Signal Inf. Process.*, vol. 3, 2014. (doi: 10.1017/atsip.2013.9)
- [11] S. Helie, F. G. Ashby, "Learning and transfer of category knowledge in an indirect categorization task", *Psychological Research*, vol. 76, no. 3, pp. 292–303, 2012 (doi: 10.1007/s00426-011-0348-1).
- [12] M. Stettler, G. Francis, "Using a model of human visual perception to improve deep learning", *Neural Networks*, vol. 104, pp. 40–49, 2018 (doi: 10.1016/j.neunet.2018.04.005)
- [13] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", *Proceedings of the IEEE/CVPR*, pp. 770–778, Las Vegas, NV, USA, June 2016 (doi: 10.1109/CVPR.2016.90)
- [14] R. C. Fong, W. J. Scheirer, D. D. Cox, "Using human brain activity to guide machine learning", *Scientific Reports*, vol. 8, no. 1, Article Number: 5397, 2018 (doi: 10.1038/s41598-018-23618-6)
- [15] H. Mahdavi-Nasab, M. S. Memarzadeh, P. Moallem, "Automatic Persian license plate recognition by edge detection using hopfield neural network", *Journal of Intelligent Procedures in Electrical Technology*, vol. 1, no. 4, pp. 41–48, Winter 2011.
- [16] S. Chatterjee, G. D. Field, G. D. Horwitz, E. N. Johnson, K. Koida, K. Mancuso, "Advances in color science: from retina to behavior", *Journal Neuroscience*, vol. 30, no. 45, pp. 14955–63, Nov. 2010 (doi: 10.1523/JNEUROSCI.4348-10.2010).
- [17] M. E. Rudd, "Edge integration in achromatic color perception and the lightness–darkness asymmetry through retinex theory", *Journal of Vision*, vol. 13, pp. 1–30, Dec. 2013 (doi: 10.1167/13.14.18)
- [18] A. Lu, G. Xu, H. Jin, L. Mo, J. Zhang, J. X. Zhang, "Electrophysiological evidence for effects of color knowledge in object recognition", *Neuroscience Letters*, vol. 469, no. 3, pp. 405–10, Jan. 2010 (doi: 10.1016/j.neulet.2009.12.039).
- [19] R. Kozik, M. Chora's, M. Ficco, F. Palmieri, "A scalable distributed machine learning approach for attack detection in edge computing environments", *Journal of Parallel and Distributed Computing*, vol. 119, pp. 18–26, Sept. 2018 (doi: 10.1016/j.jpdc.2018.03.006).
- [20] W. Fu, M. Johnston, M. Zhang, "Automatic construction of Gaussian-based edge detectors using genetic programmings", *Applications of Evolutionary Computation*, pp. 336–375, 2011 (doi: 10.1016/j.patrec.2005.07.024).
- [21] Nili, Hamed, et al. "A toolbox for representational similarity analysis", *PLoS Computational Biology*, vol. 10, no. 4, April 2014 (doi: 10.1371/journal.pcbi.1003553).
- [22] A. R. Karimian, M. Torabian, M. R. Yazdchi, "Improvement of industrial radiography for defect detection of oil and gas pipelines in weld regions by image processing", *Journal of Intelligent Procedures in Electrical Technology*, vol. 1, no. 2, pp. 23–30, Summer 2010.
- [23] I. F. Salazar-reque, S. G. Huamán, G. Kemper, J. Telles, D. Diaz, "An algorithm for plant disease visual symptom detection in digital images based on superpixels", *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 1, pp. 194–203, 2019 (doi: 10.18517/ijaseit.9.1.5322)
- [24] J. G. A. Barbedo, "A new automatic method for disease symptom segmentation in digital photographs of plant leaves", *European Journal of Plant Pathology*, vol. 147, no. 2, pp. 349–364, 2017 (doi: 10.1007/s10658-016-1007-6).
- [25] G. Marcus, "Deep Learning: A Critical Appraisal", *arXiv Prepr. arXiv:1801.00631*, 2018.
- [26] J. Ngiam, Z. Chen, P. W. Koh, A. Y. Ng, "Learning deep energy models", *Proceedings of the ICML* pp. 1105–1112, June 2011.
- [27] K. He, X. Zhang, S. Ren, J. Sun, "Identity mappings in deep residual networks", *Computer Vision and Pattern Recognition*, pp. 630–645, March 2016.
- [28] A. Ng, "Sparse autoencoder", *CS294A Lecture Notes*, vol. 72, no. 2011, pp. 1–19, 2011.

[29] M. S. Goli, A. Nsghsh, "Robust digital image watermarking against cropping and salt & pepper noise using two-step sudoku", Journal of Intelligent Procedures in Electrical Technology, vol. 8, no. 31, pp. 21-32, Autumn 2017.

زیرنویس‌ها:

1. Object recognition
2. Neurologists
3. Cognitive science
4. Neural networks (NN)
5. Neuron
6. Convolutional neural network (CNN)
7. Multi-layer perceptron (MLP)
8. Machine vision
9. Neuroscientists
10. Neocognitron
11. Hierarchical model and X
12. Differential of Gaussian
13. Sparsity
14. Representational dissimilarity matrix
15. Representational similarity analysis
16. RGB: red-green-blue
17. Population sparsity
18. Lifetime sparsity
19. High disoersal
20. HIS: saturation-intensity-color name