



تسریع همگرایی شبکه های عصبی پیچشی با ماتریس عدم تشابه بازنمایی مغز در پردازش تصویر

زهرا حیدران داروقه امنیه*^(۱) گلناز آقایی قزوینی^(۲)

(۱) گروه مهندسی برق، واحد دولت آباد، دانشگاه آزاد اسلامی، اصفهان، ایران*

(۲) گروه مهندسی کامپیوتر، واحد دولت آباد، دانشگاه آزاد اسلامی، اصفهان، ایران

تاریخ دریافت: ۱۴۰۲/۰۲/۲۱ تاریخ پذیرش: ۱۴۰۲/۰۵/۰۴

چکیده

امروزه موضوع یادگیری عمیق با پیشرفت موفقیت آمیز شبکه‌های پیچشی و یادگیری بهینه‌ی لایه‌های شبکه، مورد اهمیت واقع شده است. در این مقاله یک الگوریتم پیش‌پردازش با الهام از ماتریس عدم تشابه بازنمایی مغز برگرفته از مسیر بینایی ارائه گردیده که با تعریف تابع *RDM* و ایجاد تنکی کمک می‌کند تا وزنهای *CNN* سریع‌تر همگرا شود. این پارامتر جدید شتاب دهنده یا کاتالیزور فرایند آموزش *CNN* با افزایش تعداد صفرهای ماتریس عدم تشابه بازنمایی می‌باشد. میزان نرخ بازشناسی برای داده‌های دو گروهه و سه گروهه و میزان نرخ بازشناسی در زیرمجموعه‌هایی از مجموعه داده کلتک ۲۵۶ مورد تحلیل و بررسی قرار گرفت که میزان بهبود مدل پیشنهادی ۴ درصد نسبت به مدل پایه ارزیابی گردید. همچنین مقایسه مدل پیش پردازش پیشنهادی با مدل‌های مرجع در حیطه‌ی پارامترهای ارزیابی بررسی گردید.

کلمات کلیدی: سیستم پیچشی، تابع هزینه، همگرایی، نرخ بازشناسی، ماتریس عدم تشابه بازنمایی مغز

*عهده‌دار مکاتبات:

زهرا حیدران داروقه امنیه

نشانی: گروه مهندسی برق، واحد دولت آباد، دانشگاه آزاد اسلامی، اصفهان، ایران

پست الکترونیکی: z.heidaran@iauda.ac.ir

در یادگیری عمیق، شبکه عصبی پیچشی^۱ یک دسته از شبکه عصبی عمیق^۲ است که بیشتر برای تجزیه و تحلیل تصویر بصری استفاده می‌شود [۱]. در سال‌های اخیر طبقه‌بندی تصاویر در کاربردهای مختلفی مانند سیستم‌های امنیتی، ترویجی و آموزشی مورد استفاده قرار گرفته است. طراحی سیستم‌های خودکار برای استخراج ویژگی‌های اساسی از تصاویر یک چالش مهم به حساب می‌آید [۲-۴]. CNN نسبت به سایر مدل‌های مشابه محبوب‌تر شده است و تحقیقات گسترده‌ای توسط شرکت‌های بزرگی مانند گوگل، فیس‌بوک، آمازون، مایکروسافت و غیره انجام شده است. تشخیص شیء خروجی امر مهمی در یادگیری عمیق و الگوریتم‌های یادگیری ماشین است که به مجموعه‌ای از وظایف مرتبط برای شناسایی اشیاء در تصاویر یا فیلم‌ها اشاره دارد و درک بهتری از عملکرد مغز جهت تشخیص یک تصویر، منجر به پیشرفت‌های بیشتری در ساختار ماشین‌های هوشمند می‌گردد.

اتفاقی که در مسیر بینایی می‌افتد این است که سیگنال‌ها (اسپایک‌ها) بعد عبور از شبکه چندین مرحله را می‌گذرانند که در این مراحل آن گیرنده‌هایی که نورون‌های بینایی دارند یک سری الگوهایی را از دل این نمایش اسپایک‌ها استخراج می‌کنند که این الگوها در قشر بینایی اولیه ناحیه ۱۷ برودمن (ناحیه‌ای تقریباً مرکزی از مغز است) که معمولاً $V1$ نامیده می‌شود ممکن است که یک سری الگوهای ساده مثل وجود خطوط افقی عمودی مورب باشد ولی هرچه که از $V1$ به سمت $V2$ ، $V3$ و $V4$ می‌رود این الگوها پیچیده‌تر می‌شود.

در [۵] آنالیز تشابه بازنمایی^۳، یک قالب مشترک و ساده برای خلاصه کردن و مقایسه مستقیم مجموعه‌های داده از انواع مختلف روش‌ها برای ارزیابی شباهت‌های مرتبه دوم آنها ارائه می‌کنند و یک RDM^4 معمولاً به‌عنوان یک ماتریس $C \times C$ از تمام تخمین‌های زوجی از عدم تشابه بین شرایط C در بین ویژگی‌ها بر اساس برخی متریک ساخته می‌شود.

در [۶] برای ساختن یک RDM ، همه محرک‌ها با یکدیگر مقایسه می‌شوند و در نتیجه ماتریسی به‌دست می‌آید که در امتداد قطر آن متقارن است و محرک‌ها را می‌توان با محاسبه شباهت یا عدم تشابه بین محرک‌ها مقایسه کرد و استفاده از معیار تشابه (مثلاً r پیرسون بین ۱ (همبستگی مثبت کامل) و -۱ (همبستگی منفی کامل)) که به صورت رابطه (۱) قابل محاسبه است): در

^۱ Convolutional Neural Network (CNN)

^۲ Deep Neural Network

^۳ Representational Similarity Analysis (RSA)

^۴ Representational Dissimilarity Matrix

مقابل اندازه گیری عدم تشابه (1-) هنگام ساخت RDM ها تأثیر آماری بر نتایج ندارد، اما به نظر می رسد اندازه گیری عدم تشابه مطلوب است زیرا معمولاً در تکنیک های دیگر استفاده می شود و مزایایی برای درک مفهومی نتایج دارد و عدم تشابه با سازماندهی شهودی محرک ها در فضا همراه است که در آن محرک های غیر مشابه بیشتر از هم فاصله دارند و بنابراین می توان در تجسم به سبک شبکه ای ترسیم کرد.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (1)$$

که در آن r ضریب همبستگی، x_i مقادیر متغیر X در یک نمونه، y_i مقادیر متغیر Y در یک نمونه \bar{x} مقادیر متغیر X ، \bar{y} مقادیر متغیر Y است.

در [7] RDM یک ماتریس متقارن $n \times n$ است، که در آن n تعداد شرایط آزمایشی است ($n=45$ در این مطالعه) و مقادیر قطری نشان دهنده عدم تشابه (یا فاصله) برای هر جفت شرایط در یک جنبه ی خاص است.

در [8] تجزیه و تحلیل شباهت بازنمایی یک رویکرد مدل رمزگذاری برجسته فزاینده برای MVPA⁵ است. هدف مدل های رمزگذاری پیش بینی الگوهای پاسخ فعالیت عصبی از توصیف شرایط تجربی است. RSA چارچوبی برای مقایسه هندسه های انتزاعی فضاها بازنمایی مختلف فراهم می کند. این با محاسبه معیارهای شباهت زوجی بین الگوهای مختلف فعالیت و مونتاژ این معیارها در یک ماتریس دو بعدی به نام ماتریس عدم تشابه بازنمایی به دست می آید. بنابراین یک RDM روابط شباهت بین الگوهای فعالیت را در پاسخ به تغییرات در شرایط تجربی خلاصه می کند. سپس دانشمندان علوم اعصاب می توانند از روش های آماری برای آزمایش شباهت RDM های مختلف تولید شده از فضاها بازنمایی متمایز استفاده کنند. این یک حس کمی به دست می دهد که تا چه حد دو فضای بازنمایی متفاوت شبیه هم هستند.

در [9] یک روش رایج برای تجزیه و تحلیل هندسه های بازنمایی، تحلیل شباهت بازنمایی است. یک فرآیند دو مرحله ای است: در مرحله ی اول، RSA هندسه بازنمایی را با تخمین فاصله بازنمایی برای هر جفت شرایط تجربی (مثلاً محرک های مختلف) مشخص می کند و آنها را در یک ماتریس عدم تشابه بازنمایی جمع می کند. در اینجا از اصطلاح کلی تر "عدم تشابه" استفاده کرده تا معیارهای عدم تشابه را در برگیرد که فاصله یا متریک به معنای ریاضی نیست. یک RDM برای جمعیت عصبی در ناحیه مورد نظر مغز و برای هر بازنمایی مدل محاسبه می شود. در مرحله ی دوم، هر مدل با دقت پیش بینی آن از RDM داده

⁵ Multi Variate Pattern Analysis

ارزیابی می‌شود. **RSA** به طور گسترده‌ای مورد استفاده قرار می‌گیرد و از زمانی که مدل‌های بازنمایی قابل محاسبه تصویر مانند شبکه‌های عصبی عمیق رایج‌تر شده‌اند، محبوبیت بیشتری به دست آورده است. اخیراً پیشرفت‌های مهمی با تخمین فواصل بازنمایی و اندازه‌گیری‌های بهتر دقت پیش‌بینی **RDM** صورت گرفته است. برای تخمین، برآوردگرهای فاصله با جانبداری و بی‌طرفانه با قابلیت اطمینان بهبود یافته پیشنهاد شده‌اند. برای کمی‌سازی بهتر دقت پیش‌بینی **RDM**، توزیع نمونه‌برداری از تخمین‌گرهای فاصله به دست آمده است و معیارهای دقت پیش‌بینی **RDM** که وابستگی‌های بین تخمین‌های عدم تشابه را در نظر می‌گیرد، پیشنهاد شده است.

در [۱۰] داده‌ها به یک فضای هندسی - ماتریس عدم تشابه بازنمایی - که تفاوت‌های نسبی بین پاسخ‌های چندگانه را برجسته می‌کند، در **RSA**، پیش‌بینی می‌شود. از آنجایی که این **RDM** ها نسبت به نوع داده‌ای که از آن مشتق شده‌اند بی‌تفاوت هستند، برای مثال، **RDM** از یک نقطه زمانی خاص در **EEG**، می‌تواند با **RDM** های یک خوشه وکسل خاص در **fMRI** مقایسه شود، بنابراین دقت زمانی خوب **EEG** با دقت فضایی خوب **fMRI** ترکیب می‌شود.

در [۱۱] تحلیل شباهت بازنمایی مدل‌های محاسبات مغز را با بررسی اینکه چگونه الگوهای فعالیت عصبی شرایط تجربی را منعکس می‌کنند، آزمایش می‌کند. به جای پیش‌بینی مستقیم الگوهای فعالیت، مدل‌ها هندسه بازنمایی را پیش‌بینی می‌کنند، همانطور که توسط ماتریس عدم تشابه بازنمایی تعریف می‌شود که نشان می‌دهد چگونه الگوهای فعالیت‌های مختلف مشابه یا غیرمشابه با شرایط تجربی مختلف مرتبط هستند.

در [۱۲] از طریق تصویربرداری تشدید مغناطیسی عملکردی^۶، فعالیت مغز را می‌توان برای درک ساختار بازنمایی اطلاعات در یک منطقه تجزیه و تحلیل کرد، به ویژه با تجزیه و تحلیل‌های چند متغیره، که می‌تواند محرک‌ها و شرایط را در ابعاد مختلف بررسی کند. یکی از این تکنیک‌ها، تحلیل شباهت بازنمایی، میزان (عدم شباهت) بین الگوهای فعالیت برای محرک‌های مختلف را اندازه‌گیری می‌کند تا فضای بازنمایی در یک منطقه را تعیین کند. این را می‌توان به عنوان یک ماتریس عدم تشابه بازنمایی نشان داد که در آن عدم تشابه الگوهای فعالیت بین محرک‌ها یا شرایط منعکس می‌شود. یک روش متداول برای مقایسه فضاهای بازنمایی مناطق مختلف، اندازه‌گیری شباهت (همبستگی **RDM**) های مناطق مختلف، از طریق یک **RSA** "مرتبه دوم" است. هنگام مقایسه **RDM** ها در مناطق مختلف، یک **RSA** مرتبه دوم **RSA** های کلی (در طول زمان فرو ریخته) مناطق مختلف

^۶ Functional Magnetic Resonance Imaging (fMRI)

را مقایسه می‌کند. با این حال، ترکیب تمام نقاط زمانی به این روش، اطلاعاتی را که ممکن است در نوسانات قدرت تشابه بازنمایی در طول زمان وجود داشته باشد، حذف می‌کند.

در [۱۳] RSA مبتنی بر مفهوم ماتریس‌های عدم تشابه بازنمایی است که بازنمایی‌های محرک درونی یک سیستم را از نظر تفاوت‌های پاسخ زوجی مشخص می‌کند. با هم، مجموعه تمام مقایسه‌های زوجی ممکن تخمینی از آرایش هندسی محرک‌ها در فضای فعال‌سازی با ابعاد بالا ارائه می‌کند. بازنمایی دو سیستم اگر بر تمایزات یکسانی در بین محرک‌ها تأکید داشته باشند، مشابه در نظر گرفته می‌شوند، یعنی به میزانی که RDM های آنها موافق هستند. مقایسه در سطح RDM ها، که می‌تواند در فضاهای منبع با مبدأ و ابعاد مختلف محاسبه شود، در نتیجه مشکل تعریف یک نگاشت مطابقت بین واحدهای مشاهده را کنار می‌گذارد.

در [۱۴] می‌توان ماتریسی به نام ماتریس عدم تشابه بازنمایی ساخت که درجه تمایز هر جفت محرک یا شرایط را نشان می‌دهد. RDM یک ماتریس متقارن مربعی است که حاوی یک سلول برای هر جفت محرک یا شرایط تجربی است و به عنوان امضای بازنمایی‌ها عمل می‌کند. RSA یک همبستگی مرتبه دوم بین RDM های مدل و RDM های الگوی فعال‌سازی را محاسبه می‌کند. مدل‌های RDM شباهت بین محرک‌ها را که توسط یک مدل محاسباتی یا فرضیه در مورد ساختار فضای محرک پیش‌بینی شده است، نشان می‌دهند. RDM های الگوی فعال‌سازی برای مجموعه ای از وکسل‌ها با استفاده از تابع عدم تشابه محاسبه می‌شوند.

بنابراین ماتریس عدم تشابه بازنمایی نوعی از نمایش داده است که در آن، نمونه‌ها به صورت برداری از فاصله با تصاویر مرجع نمایش داده می‌شوند و یک تصویر به جای اینکه در فضای پیکسل نمایش داده شود، به صورت برداری از فاصله‌ها با مجموعه‌ای از تصاویر مرجع نمایش داده می‌شود. ساختار مقاله به این شرح است: در قسمت دوم مدل پیشنهادی بیان می‌شود. در قسمت سوم به فاز ارزیابی و مقایسه با دو مدل مرجع پرداخته خواهد شد و در پایان، نتیجه گیری این پژوهش در قسمت چهارم بررسی خواهد شد.

۲- مدل پیش‌پردازش پیشنهادی

اکنون بسیاری از کمپانی‌های عرصه‌ی هوش مصنوعی ساختاری برای شبکه‌های عمیق ارائه داده‌اند [۱۵] که تقریباً همه‌ی آنها از ساختار اولیه شبکه عمیق پیچشی پیروی می‌کنند. با توجه به بررسی‌های محدود مدل‌سازی شبکه‌ی در شبکه‌های عمیق در این

قسمت پیش پردازشی قبل از اعمال تصویر به شبکه پیچشی همراه با فرآیند تقسیم‌بندی و درون‌یابی ارائه می‌شود که چنین عملی باعث کاهش انرژی و استخراج لبه‌های بین رنگی شده و باعث می‌شود نورون‌های کمتری در شبکه عمیق فعالیت داشته باشند که از فعالیت نورانی زائد می‌کاهد و منجر به تنگی می‌گردد. شناخت رفتار شبکیه و چشم انسان بسیار اهمیت دارد تا مدل‌سازی صحیحی صورت گیرد [۱۶-۱۸].

همچنین مدلی از شبکه عمیق طراحی می‌شود که شبکه طراحی شده مبتنی بر شبکه پیچشی است که انطباق بیشتری به مسیر بینایی مغز دارد. شبکه‌ی طراحی شده از استانداردهای پایه و اولیه شبکه پیچشی پیروی دارد. شبیه‌سازی بر روی دسته‌های مختلف تصاویر از مجموعه داده کلنک آزمایش می‌شود. یکی از معایب مجموعه داده‌ی کلنک وجود عکس‌های مشابه از یک نمونه در هر گروه است که منجر به تضعیف نتیجه در پاسخ به نمونه‌های تنها (تک) می‌شود و نتایج با مدل پایه شبکه عمیق پیچشی مقایسه می‌شود. در این کار از شبکه پیچشی عمیق اولیه استفاده شده است و تلاش می‌شود تأثیر پیش پردازشی با فرآیند تقسیم‌بندی و درون‌یابی بر روی شبکه عمیق ارزیابی گردد.

ماهیت RDM [۱۹-۲۰] به تجزیه و تحلیل انجام شده بر روی سیگنال‌های EEG^۱، BEH^۲ و fMRI^۳ که توسط شبکه عصبی تولید می‌شود بستگی دارد (EEG، BEH و fMRI) به عنوان تصویربرداری رفتاری، الکتروانسفالوگرافی و تصویربرداری تشدید مغناطیسی کارکردی تعریف شده‌اند).

در الگوریتم ۱ و شکل ۱ مدل پیش‌پردازش پیشنهادی را می‌توان بدین صورت تحلیل نمود که ابتدا تصویر ورودی که به یک دسته‌ی مرتبی از تصاویر تقسیم شده شامل ۱۰ نمونه در ۲۰ دسته جمعا ۲۰۰ نمونه دریافت و به شبکیه اعمال می‌گردد. همان‌طور که در جدول ۱ نشان داده شده است، وقتی در مباحث علوم اعصاب از RDM استفاده و BEH اعمال شود قطعا $n_trials=1$ ، $n_cons=20$ ، $n_subs=10$ (۱۰ تا دسته مختلف و در هر دسته ۲۰ نمونه با تعداد آزمایش‌ها ۱) بنابراین دیتاها n_cons ، n_subs & n_trials و خروجی $[n_subs, n_cons, n_cons]$ خواهد بود که با یک تابع view تغییر ابعاد انجام خواهد شد.

جدول ۱- مقادیر پارامترها

^۲ Behavioral

^۱ Electroencephalography

^۳ Facebook AI Research lab (FAIR)

تعداد آزمایش‌ها	۱
تعداد نمونه در هر دسته	۲۰
تعداد دسته	۱۰

یک سری از توابع گوسی بر روی پاسخ گیرنده نور مشخصی اعمال می‌شود. از این ویژگی چندین لبه‌یابی در مسیر بینایی رخ می‌دهد و از فیلتر شبکه عبور و ماتریس RDM شبکه به تصویر تنک شده اعمال می‌گردد. در صورتیکه جستجوی عکس‌های مشابه از یک نمونه در هر گروه خاتمه یافته خروجی را خواهیم داشت. در غیر اینصورت مسیر بینایی مغز با پیشروی در شبکه پیش رو قرار گرفته و با توجه به وجود سلولهای ساده و پیچیده در مسیر بینایی مغز طول موجهایی با فرکانسهای مختلف ایجاد می‌گردد. در ناحیه‌ی V2 و V4 (فرآیند بینایی در مغز به صورت سلسله مراتبی از ناحیه‌ی V1 شروع می‌شود و با گذر از نواحی V2 و V4 به ناحیه‌ی گیجگاهی تحتانی مغز می‌رسد) ترکیب‌های مختلفی از سلول‌های ساده و پیچیده برای ساختن سلول‌هایی با قابلیت پاسخ به تحریک‌های پیچیده و پیچیده‌تر به وجود می‌آید. نقطه‌ی مشترک این ترکیب‌ها آن است که در هر سلول پیچیده که از ترکیب سلول‌های ساده به وجود می‌آید، پاسخ به تحریک نسبت به اندازه و محل محرک در حوزه‌ی دید سلول نایسته می‌شود و در تشکیل پاسخ سلول‌های ساده از سلول‌های پیچیده‌ی مرحله‌ی قبل پیچیدگی تحریک‌ها افزایش می‌یابد. این ترکیبات تا آنجا ادامه می‌یابد که در قشر مغز فراگیجگاهی سلول‌هایی یافت می‌شوند که به تحریک‌های فوق‌العاده پیچیده‌ای همچون چهره‌ی انسان در یک زاویه‌ی خاص پاسخگو هستند [۱۶-۱۸] و [۲۱، ۲۲] که نیاز به مسیر پیشروی بینایی مغز بخوبی جهت تشخیص تحریک‌های پیچیده چهره انسان محسوس می‌باشد. مجدداً عمل لبه‌یابی و اعمال ماتریس RDM بر روی فرکانسها با طول موجهای مختلف صورت گرفته تا هنگامی که عکس‌های مشابه از یک نمونه در هر گروه جستجو خاتمه یابد. در غیر اینصورت مرحله‌ی بعد بخش‌بندی تصویر پردازشی است که نواحی جالب (لبه‌ها) در تصویر را از سایر نواحی جدا می‌کند.

الگوریتم ۱- الگوریتم پیش پردازش

Input: images (number of channels, length, width) → Retina

Output: images and Update weights

1: for k = 1 do

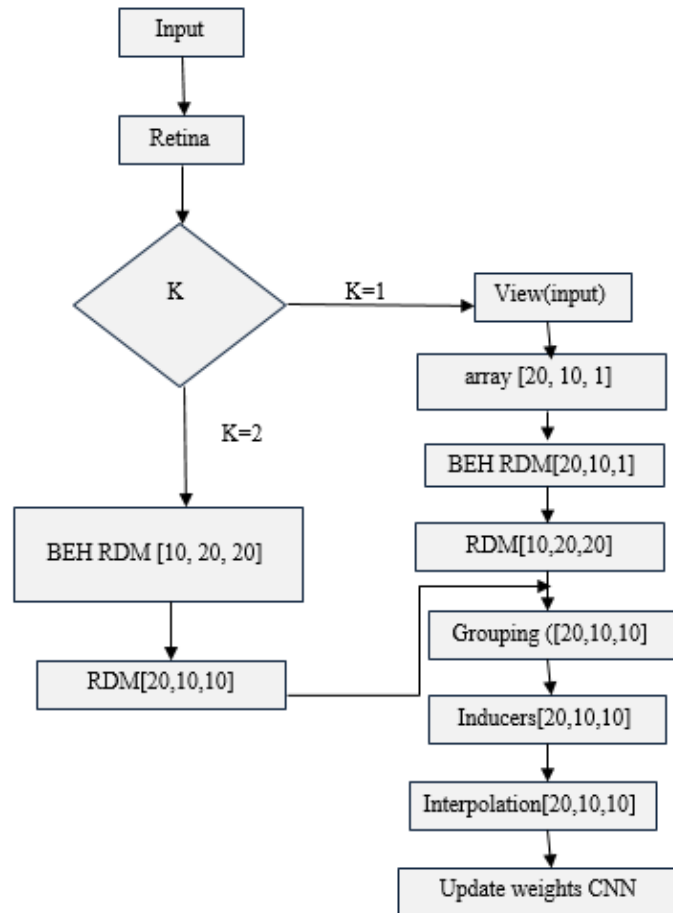
2: View (images (3, 32, 32)) → array [20, 10, 1]

3: BEH RDM [20, 10, 1] → Dissimilarity matrix [10,20,20]

4: end for

5: for k = 2 do

- 6: BEH RDM [10, 20, 20] → Dissimilarity matrix [20,10,10]
- 7: end for
- 8: Group process (Dissimilarity matrix [20,10,10])
- 9: Inductive process(Dissimilarity matrix [20,10,10])
- 10: Interpolation process (Dissimilarity matrix [20,10,10])
- 11: Update weights CNN



شکل ۱- فلوچارت الگوریتم پیش پردازش

اکنون بایستی تابعی برای RDM تعریف نموده تا بهترین خروجی RDM یعنی صفر نمودن مقادیر قطر اصلی ماتریس تعریف شده را در نظر گرفته و در نتیجه مقادیر دیگر ماتریس نشان هندهی بالاترین درجه تمایز خواهند بود که به اینصورت بتوان یک تنکی با افزایش تعداد صفرها در ماتریس عدم تشابه بازنمایی ایجاد کرد. بدین منظور متغیر $L(e_{j,i})$ که در معادله‌ی با الهام از رابطه ارائه شده در [۱۸] و با هدف صفر نمودن مقادیر قطر اصلی ماتریس تعریف شده را در نظر گرفته و در نتیجه مقادیر دیگر ماتریس نشان هندهی بالاترین درجه تمایز خواهند بود.

$$L(e_{j,i}) = S_{j,i,j} - S_{j,i,j} - \min \log \left(1 + \frac{e^{S_{ij.1}} + e^{S_{ij.2}} + e^{S_{ij.3}} + \dots + e^{S_{ij.N}}}{e^{S_{ij.N}}} \right) \quad (2)$$

که در آن N تعداد دسته‌ها، $L(e_{ji})$ تابع هزینه RDM بوده که در مسائل آموزش، هدف کاهش تابع هزینه خواهد بود که به صورت معادله‌ی (۱) تعریف خواهد شد و S_{jiz} شباهت تصویر i از دسته j به مرکز دسته z است. در رابطه‌ی (۲)، مقدار S_{jiz} می‌بایست برای کاهش $L(e_{ji})$ به حداقل برسد. یعنی دسته‌ای که مشابه است دارای تمایز صفر می‌باشد و در قطر اصلی مقدار صفر را خواهیم داشت و کمترین مقدار خود را می‌پذیرد یعنی مقادیر قطر اصلی را به حداقل می‌رساند. مقدار جمله دوم باید تا حد امکان برای همه دسته‌هایی که به دسته خود تعلق ندارند به حداکثر برسد و به این ترتیب میزان عدم تجانس سایر عناصر غیر قطر اصلی نزدیک به یک می‌باشد که مینیمایز کل عبارت در نظر گرفته خواهد شد. به عنوان مثال، اگر $0/2$ ، $0/9$ و $0/8$ عناصر با قطر غیر اصلی باشند، عنصری که کمترین مقدار را دارد، یعنی $0/8$ انتخاب می‌شود و مقدار آن تا ۱ افزایش می‌یابد.

۲-۱- مجموعه داده

در این مقاله از مجموعه داده **Cifar10** هم برای انجام شبیه سازی‌ها استفاده نموده که این مجموعه داده اغلب برای ارزیابی راه حل‌های CNN استفاده شده است. مجموعه داده **Cifar10** (مؤسسه تحقیقات پیشرفته کانادا، ۱۰ دسته) زیرمجموعه‌ای از مجموعه داده‌های **Tiny Images** است و از ۶۰۰۰۰ تصویر رنگی 32×32 تشکیل شده است. همچنین جهت ارزیابی مدلها از نمونه داده کلکتک ۲۵۶ هم استفاده گردیده که مجموعه داده کلکتک ۲۵۶ شامل تصاویری از ۳۰۶۰۷ تصویر دنیای واقعی شامل ۲۵۷ دسته می‌باشد که هر دسته حداقل با ۸۰ تصویر نمایش داده می‌شود. برای ارزیابی عملکرد رویکرد پیشنهادی، از **Google Colab** و نرم افزار **matlab** استفاده شده است. **Google Colab** یک سند اجرایی، تعاملی و مبتنی بر ژوپیتر با سخت‌افزار رایگان بوده، که به **Google Drive** متصل است و می‌توان با دیگران برای کدنویسی همزمان شریک شد. نوت‌بوک‌های **Colab** کدها را روی سرورهای ابری گوگل اجرا می‌کنند. بنابراین، مجموعه داده‌ها را می‌توان بر روی سرورهای ابری وارد کرده و به راحتی می‌توان یک طبقه‌بندی کننده تصویر را آموزش داد.

چندین چارچوب برای توسعه برنامه‌های شبکه عصبی در **Google Colab** وجود دارد، از جمله **Tensorflow** (عرضه شده توسط **Google** در سال ۲۰۱۵)، **PyTorch** (عرضه شده توسط **Facebook** در سال ۲۰۱۶) و **Keras** که نسخه جدید **Tensorflow** می‌باشد. در این مدل از چارچوب **PyTorch** استفاده شده زیرا کار با آن در مقایسه با سایر چارچوب‌ها آسان‌تر است. **PyTorch** یک کتابخانه یادگیری ماشین منبع باز مبتنی بر کتابخانه **Torch** است که برای بینایی کامپیوتر و برنامه‌های

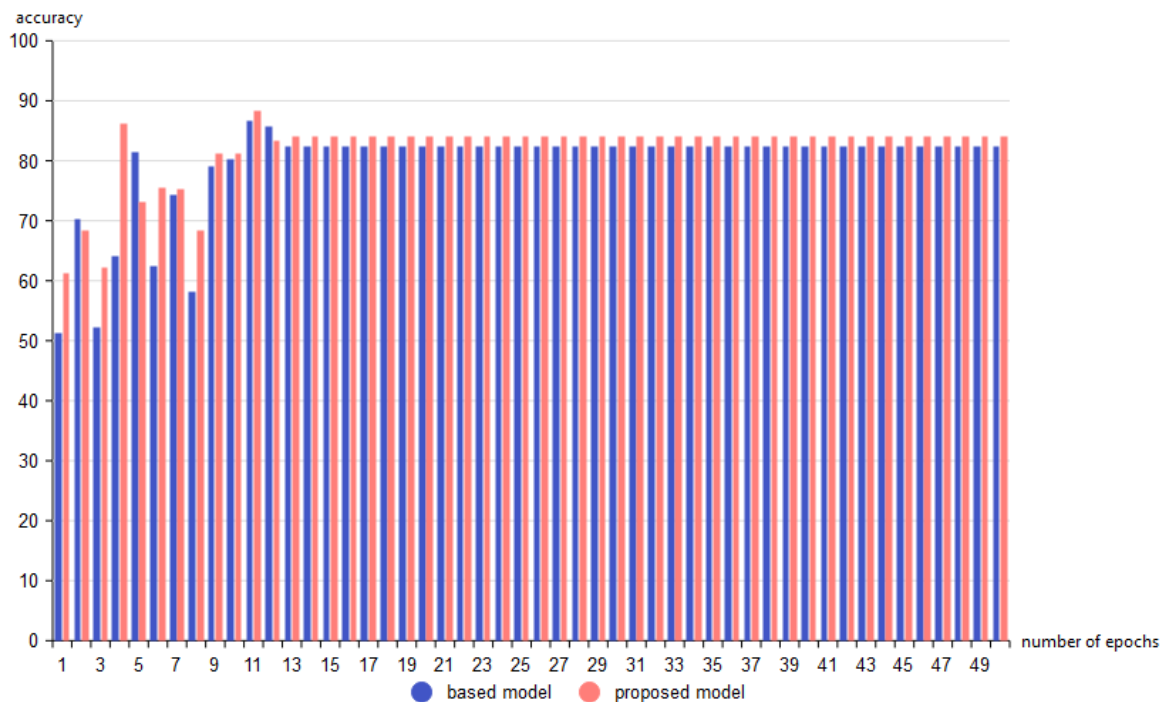
پردازش طبیعی استفاده می‌شود و توسط آزمایشگاه تحقیقاتی هوش مصنوعی فیس بوک^۹ در سال ۲۰۱۶ توسعه یافته است [۱۹]، [۲۳].

۲-۲- شبیه سازی مدل پیشنهادی

مدل پیشنهادی دودسته و سه دسته است زیرا زمان محاسبات و پیچیدگی ارزیابی کاهش یابد. مدل از دو منظر رنگ و شکل ارزیابی می‌گردد. یعنی دو و سه دسته تصاویری که انتخاب نموده یکبار در دو دسته مشخصه رنگ متفاوت باشند و در مدل سه دسته مشخصه شکل متفاوت باشند. یکی از مهم‌ترین پارامترها در آموزش شبکه‌های پیچشی طول اندازه-دسته است. منظور از اندازه-دسته تعداد نمونه‌های مورد نیاز جهت اصلاح وزن‌ها در هر تکرار است که طول آن ۵ انتخاب می‌شود. این عدد متداول بین نمونه کارهای مشابه است.

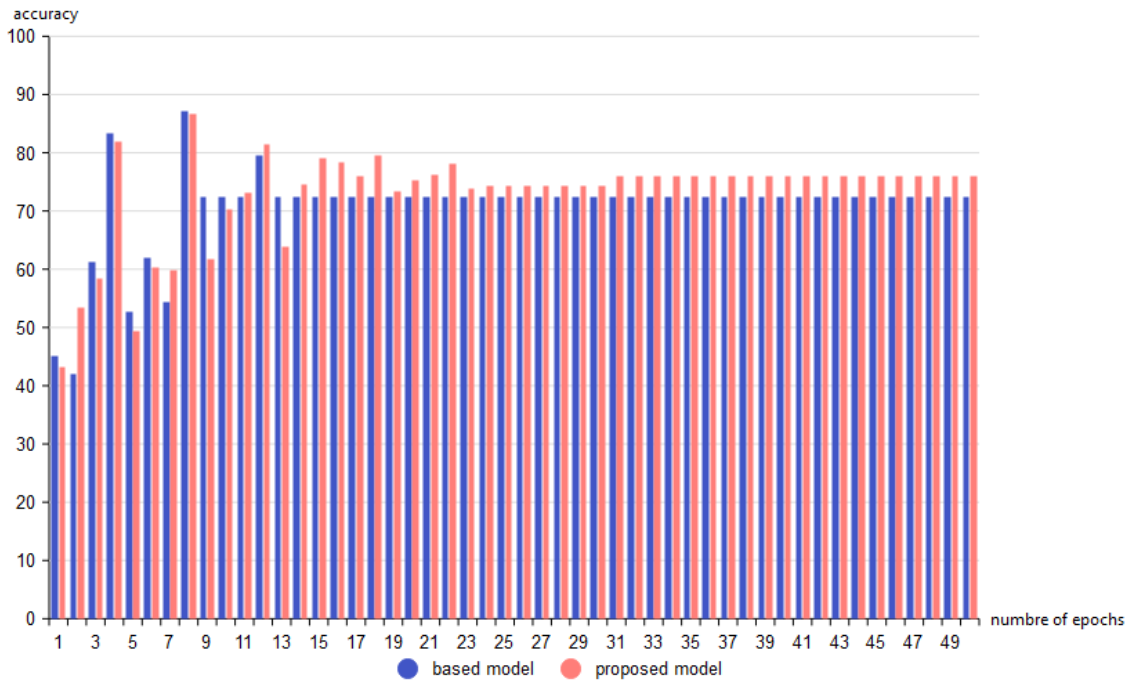
در آزمایشها ۲۵ درصد تصاویر برای آزمایش و بقیه برای آموزش شبکه استفاده می‌شوند. البته این آزمایشها به تعداد ۵ مرتبه تکرار شده‌اند و نتایج میانگین هستند.

در آزمایش اول از تصاویر دو دسته دارای شکل یکسان و رنگ متفاوت جهت آموزش مدل استفاده شد. نرخ بازشناسی به ازای ۵۰ دوره در شکل ۲ نشان داده شده است. با توجه به نمودار در صورت استفاده از مدل پیشنهادی نتیجه زودتر از حالت مدل پایه پایدار می‌شود و همچنین نرخ بازشناسی بالاتری خواهد داشت.



شکل ۲- نرخ بازشناسی مدل پیش پردازش پیشنهادی دو دسته

در ادامه آزمایش با تغییر جزئی مدل آن را برای سه دسته دارای رنگ یکسان و شکل متفاوت انجام گردید. در شکل ۳ نتیجه بازشناسی به ازای ۵۰ دوره نشان داده شده است. مشاهده می شود که مدل ارائه شده نرخ بازشناسی کمی (حدود ۴ درصد) بهتر ثبت کرده است.



شکل ۳- نرخ بازشناسی مدل پیش پردازش پیشنهادی سه دسته

با توجه به نتایج در دو آزمایش میتوان این نتیجه گیری را کرد که رمزنگاری رنگی زمانی تأثیر مثبت دارد که مشخصه‌ی غالب تفکیک پذیری دسته‌ها رنگ باشد. در آزمایش اول دو دسته بسیار به هم از نظر شکل شبیه هستند ولی غالباً رنگ متفاوتی دارند و مدل پیشنهادی توانسته کارایی بهتری ثبت کند ولی در حالت دوم یعنی بین سه دسته وسیله نقلیه که مشخصه شکل متفاوت دارند مدل نرخ بازشناسی بهتری ثبت نکرده البته در حالت پایدار کمی بهتر بوده است. بدلیل اینکه RDM با ایجاد تنگی (کاهش فعالیت نورون‌ها) کمک کرده تا وزنه‌های CNN سریع‌تر همگرا شود و این پارامتر جدیدی که اضافه شده شتاب دهنده یا کاتالیزور فرایند آموزش CNN با افزایش تعداد صفرهای ماتریس عدم تشابه بازنمایی می باشد.

۳- فاز ارزیابی مدل پیش پردازش پیشنهادی

معماری شبکه عصبی پیچشی با کارایی بالا، یکی از مهم ترین الزامات برای دستگاه‌های با منابع محدود (مانند دستگاه‌های تعبیه شده و موبایل) است. قدرت محاسباتی و اندازه حافظه دو محدودیت مهم این دستگاهها هستند. اخیراً معماری‌هایی برای غلبه بر این محدودیت‌ها با در نظر گرفتن تجهیزات سخت افزاری-نرم افزاری خاص پیشنهاد شده است.

اولین مدلی که برای مقایسه انتخاب شد [۲۴]، بلوک‌های سبک وزن باقیمانده با اتصال متراکم برای تضمین نظارت عمیق، جریان گرادیان کارآمد، و قابلیت‌های استفاده مجدد از شبکه عصبی پیچشی پیشنهاد شده‌اند. روش پیشنهادی، هزینه آموزش و فرآیندهای استنباط را بدون استفاده از تجهیزات سخت‌افزاری-نرم‌افزاری خاص با کاهش تعداد پارامترها و عملیات محاسباتی در حین دستیابی به دقت عملی کاهش می‌دهد.

نتایج تجربی گسترده نشان می‌دهد که معماری پیشنهادی از نظر اندازه مدل، پارامترهای مورد نیاز و حتی دقت کارآمدتر از AlexNet و VGGNet است.

مدل پیشنهادی در CIFAR10, SVHN, Fashion MNIST, MNIST, ImageNet و CIFAR100 قابل ارزیابی است که در این مقاله با CIFAR10 ارزیابی انجام شده است. در مجموعه داده‌های Fashion MNIST به نتایج پیشرفته و در سایر موارد به نتایج معقول دست می‌یابد. نتایج به دست آمده نشان‌دهنده برتری روش پیشنهادی نسبت به مدل‌های کارآمد مانند SqueezeNet است. همچنین با مدل‌های کارآمد پیشرفته مانند CondenseNet و ShuffleNet قابل مقایسه است. یکی از اهداف اصلی این کار طراحی یک مدل فوق‌العاده کوچک برای تحلیل میزان پیچیدگی معماری‌های CNN جدید بود که از نظر خطای طبقه‌بندی به عملکردی پیشرفته دست یافته‌اند.

برای این هدف، یک مدل کوچک CNN بر اساس دو ایده اصلی اتصالات باقیمانده و لایه‌های پیچشی به هم پیوسته متراکم پیشنهاد شد. این معماری RdenseCNN پیشنهادی دارای تعدادی پارامتر و هزینه محاسباتی کم با یک خطای طبقه‌بندی ممکن بود. این شامل لایه‌های متراکم متصل از نظر برخی از بلوک‌های لانه با اتصالات پرش برای حفظ جریان اطلاعات در یک مدل عمیق CNN بود.

مدل پیشنهادی ۲۶ برابر پارامترهای کمتر با ۸ درصد خطای طبقه‌بندی بهتر از اولین مدل CNN گسترده (AlexNet) داشت. شایان ذکر است که کوچکترین مدل پیشنهادی در ابزارها و برنامه‌های چندرسانه‌ای با AlexNet در اندازه مدل $\times 60$ دقت یکسانی (۱٪ دقت بهتر) داشت.

نتایج نشان داد که برخی از سطوح پیچیدگی در مدل های CNN به طور جدی در مدل های اولیه CNN، توجهات غیرقابل قبولی دارند. در نتیجه، این مدل های پیچیده غیر قابل اجرا را می توان با یک مدل بسیار کوچکتر و کارآمدتر، به ویژه در کاربردهای منابع محدود، جایگزین کرد. به این ترتیب، روش پیشنهادی برای کاربردهای حافظه و توان محدود که دقت معقول در تعداد بسیار کم پارامترها و هزینه محاسباتی قابل قبول است، مناسب تر است.

هزینه هنگفت استنتاج شبکه های عصبی عمیق را می توان با فشرده سازی شبکه کاهش داد.

هرس اتصالات یکی از رویکردهای غالب مورد استفاده برای فشرده سازی شبکه است.

با این حال، تکنیک های هرس موجود از یک یا چند مورد از محدودیت های زیر رنج می برند:

(۱) آنها زمان و انرژی مصرف شده در مرحله آموزش سنگین محاسباتی را به دلیل اضافه شدن مراحل هرس و تنظیم دقیق افزایش می دهند.

(۲) آنها بر اساس اطلاعات موضعی در مورد آمار یک لایه خاص، با نادیده گرفتن تأثیر انتشار خطا در شبکه، لایه ها را هرس می کنند.

(۳) آنها فاقد ابزار کارآمد برای تعیین اهمیت جهانی کانال ها هستند.

(۴) به دلیل استفاده از هرس بدون ساختار، ممکن است در هنگام پیاده سازی بر روی پلتفرم های اصلی (GPUها و TPUها) به هیچ مزیت انرژی منجر نشوند و برای بهره مندی از مزایای آن به سخت افزار تخصصی نیاز دارند.

مدل دوم انتخاب شده برای مقایسه [۲۵]، برای پرداختن به مسائل فوق، یک روش ساده و در عین حال مؤثر برای هرس تدریجی کانال در حین آموزش با استفاده از یک معیار مبتنی بر داده به نام امتیاز مربوط به ویژگی ارائه کرده است.

تکنیک پیشنهادی نیاز به بازآموزی اضافی را با هرس کم اهمیت ترین کانال ها به شیوه ای ساختار یافته در فواصل زمانی ثابت در طول مرحله تمرین منظم حذف می کند.

هرس توسط امتیازهای مربوط به ویژگی هدایت می شود، که به ارزیابی مؤثر سهم هر کانال در قدرت تمایز شبکه کمک می کند.

اثربخشی روش پیشنهادی را بر روی معماری‌هایی مانند VGG و ResNet با استفاده از مجموعه داده‌هایی مانند CIFAR10، CIFAR100 و ImageNet نشان داده و با موفقیت به فشردگی مدل قابل توجهی دست یافته در حالی که دقت کمتر از ۱ درصد معامله می‌گردد.

برخلاف تکنیک‌های هرس موجود، تکنیک پیشنهادی نیازی به آموزش مدل تا زمان همگرایی قبل از هرس ندارد. در عوض، این مدل در سطح جهانی به شکلی ساختار یافته در طول مرحله آموزش واقعی هرس می‌شود. این امر باعث کاهش پیچیدگی محاسباتی و زمانی آموزش همراه با استنتاج می‌شود.

از این رو، روش هرس تدریجی کانال پیشنهادی امکان آموزش بر روی دستگاه‌های لبه (با منابع محدود) را فراهم می‌کند که ممکن است با تکنیک‌های هرس موجود امکان‌پذیر نباشد.

روش هرس از یک متریک مبتنی بر داده، امتیازات مربوط به ویژگی‌ها برای تعیین کانال‌های اضافی یا کم اهمیت استفاده می‌کند. جایگزینی امتیاز مربوط به ویژگی با معیارهای بسیار ساده‌تر مانند استاندارد L1 یا L2 می‌تواند به پیچیدگی محاسباتی بهتر منجر شود. با این حال، این معیارها تنها آمار لایه‌های جداگانه را در نظر می‌گیرند و تأثیر انتشار خطا در شبکه را نادیده می‌گیرند و منجر به کاهش دقت بالاتر پس از هرس می‌شوند. امتیازات مربوط به ویژگی با استفاده از انتشار ارتباط لایه‌ای محاسبه می‌شود که یک تکنیک قابل توضیح پیشرفته است.

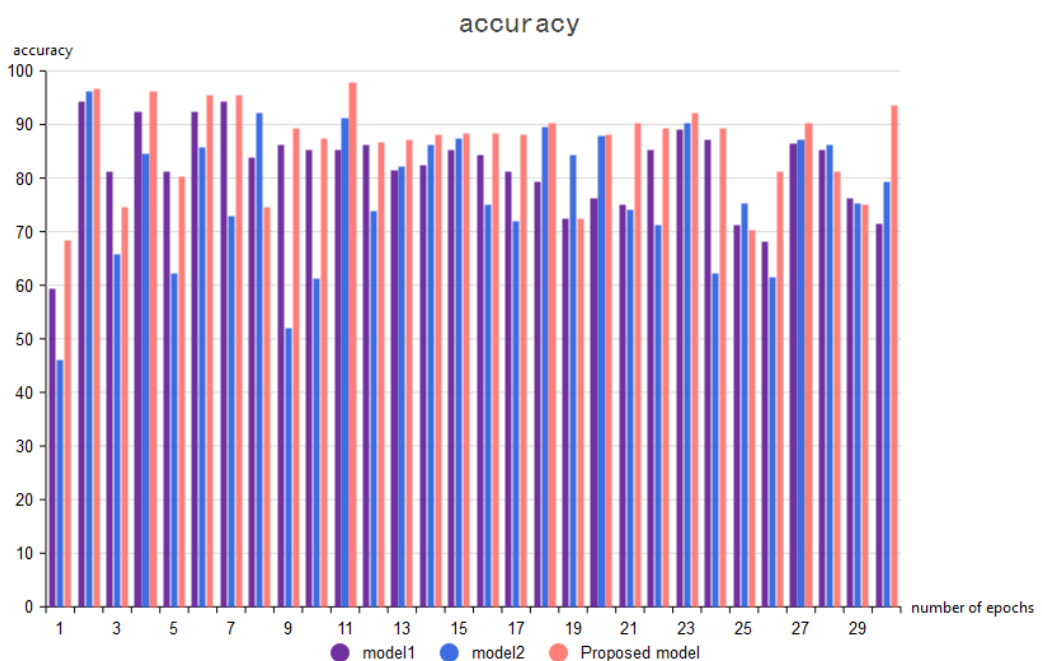
از داده‌های آموزشی برای تعیین تأثیر متوسط کانال‌ها بر خروجی پیش‌بینی شده استفاده می‌کند.

به طور خاص، امتیازات مربوط به ویژگی شامل آمار کل شبکه با در نظر گرفتن فعال‌سازی هر کانال و مسیر انتشار آن به لایه نهایی است. از این رو، تکنیک پیشنهادی از یک متریک بهتر برای تعیین کمیت سهم یک کانال نسبت به قدرت تمایز شبکه بدون مرحله بهینه‌سازی استفاده می‌کند.

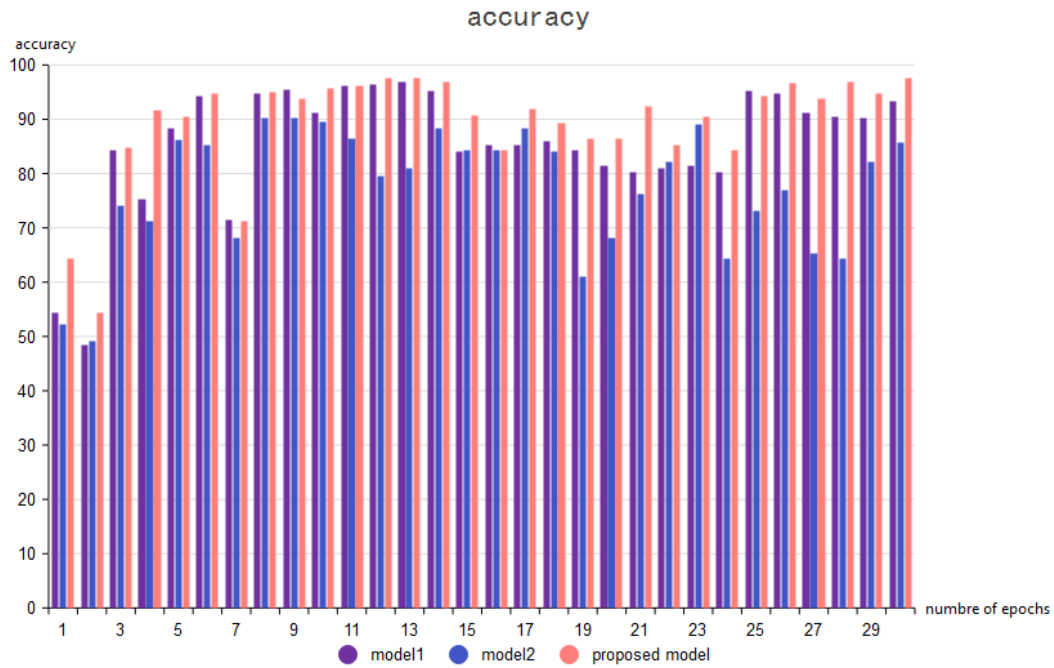
شبکه‌های عصبی پیچشی برای بسیاری از وظایف بینایی کامپیوتری حیاتی هستند و برای تنظیمات کم‌منبع نیاز به اجرای کارآمد انرژی دارند. در این مدل، یک تکنیک هرس تدریجی کانال را در حین آموزش فشردگی و شتاب CNN بر اساس امتیازهای مربوط به ویژگی ارائه گردیده است.

اهمیت کانال با استفاده از امتیازهای مربوط به ویژگی هر چند دوره در طول آموزش ارزیابی می شود و کانال های کم اهمیت هرس می شوند. روش هرس پیشنهادی عاری از بازآموزی تکراری است که پیچیدگی محاسباتی و زمانی هرس شبکه عصبی عمیق را کاهش می دهد. اثربخشی تکنیک هرس با استفاده از مجموعه داده ها و معماری های معیار نشان داده شده است. مشاهده می گردد که تکنیک پیشنهادی می تواند فشرده سازی و شتاب قابل توجهی را با کاهش دقت کمتر از ۱ درصد به دست آورد.

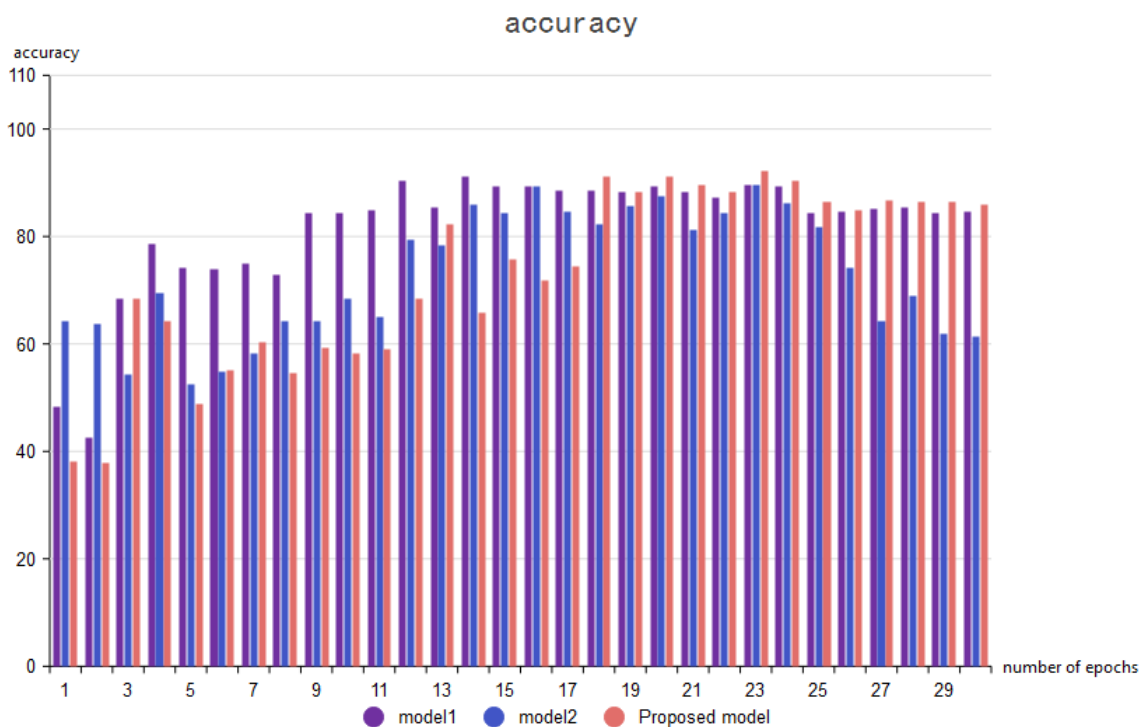
در شکل های ۴، ۵ و ۶ از تصاویر دو دسته و سه دسته و نرخ تشخیص حیوان در تصاویر جهت آموزش مدل استفاده شد که نرخ بازشناسی به ازای ۳۰ دوره نشان داده شده است.



شکل ۴- نتایج ارزیابی نرخ بازشناسی دو دسته مدل پیش پردازش پیشنهادی و مدل های مرجع در مجموعه داده ی Cifar10



شکل ۵- نتایج ارزیابی نرخ بازشناسی سه دسته مدل پیش پردازش پیشنهادی و مدل‌های مرجع در مجموعه داده Cifar10



شکل ۶- نتایج ارزیابی نرخ بازشناسی حیوان مدل پیش پردازش پیشنهادی و مدل‌های مرجع در مجموعه داده

Cifar10

طبق ارزیابی انجام شده در جدول ۲ و بررسی پارامترهای ارزیابی برای مدل پیش‌پردازش پیشنهادی و مدل‌های مرجع بصورت کمی و همچنین در جدول ۳ مقایسه پارامترهای ارزیابی برای مدل پیش‌پردازش پیشنهادی و مدل‌های مرجع، بر روی مجموعه

داده‌ی Cifar10 بررسی و میزان بهبود پارامترها با مدل پیش‌پردازش پیشنهادی مقایسه شده است و مشاهده می‌شود که این مدل نرخ بازشناسی بهتری ثبت کرده است.

جدول ۲- پارامترهای ارزیابی برای مدل پیش‌پردازش پیشنهادی و مدل‌های مرجع

نرخ تشخیص حیوان	نرخ بازشناسی سه گروه	نرخ بازشناسی دو گروه	پارامترهای ارزیابی
۸۱/۶۹	۸۵/۵۵	۸۱/۹۸	مدل ۱
۷۳/۰۱	۷۷/۳۷	۷۷/۰۲	مدل ۲
۷۳/۰۳	۸۹/۲۹	۸۶/۱۹	مدل پیش‌پردازش پیشنهادی

جدول ۳- مقایسه‌ی پارامترهای ارزیابی برای مدل پیش‌پردازش پیشنهادی و مدل‌های مرجع

مقایسه مدل‌ها و مدل پیشنهادی			
نرخ تشخیص حیوان	نرخ بازشناسی سه گروه	نرخ بازشناسی دو گروه	پارامترها ارزیابی
-۱۱	۴/۸۷	۵/۵۱	مقایسه مدل ۱ با مدل پیش‌پردازش پیشنهادی
۰/۵۲	۱۶/۶۲	۱۴/۵۷	مقایسه مدل ۲ با مدل پیش‌پردازش پیشنهادی

۴- نتیجه‌گیری

در این مقاله یک الگوریتم پیش‌پردازش با الهام از مسیر بینایی و اعمال تابع BEH RDM ارائه شد که با تعریف تابع RDM از طریق افزایش تعداد صفرهای ماتریس عدم تشابه بازنمایی با صفر نمودن درایه‌های قطر اصلی و یک نمودن سایر درایه‌ها باعث ایجاد تنگی گردید تا وزنه‌های CNN سریع‌تر همگرا شده و این مدل جدید نقش شتاب دهنده یا کاتالیزور فرایند آموزش CNN را دارا می‌باشد. نتایج ارزیابی آن در زیر مجموعه‌هایی از مجموعه داده کلتک و داده‌های دو گروهه و سه گروهه، دارای نرخ بازشناسی بیشتر بوده است.

طبق پیش‌بینی اعمال کدینگ رنگی ویژگی رنگ را در شبکه با ارزش‌تر قرار می‌دهد و بنابراین در شبکه‌هایی که رمزنگاری رنگی اعمال شده باشد مشخصه رنگ جهت تفکیک دسته‌ها ارجح‌تر از دیگر مشخصه‌ها مانند ریخت و شکل شده بود. از این رو در مسائلی که رنگ عامل مهمی در بازشناسی بود، همان‌طور که در شکل‌های (۲) و (۳) مشاهده می‌گردد، مدل پیشنهادی توانست نرخ بازشناسی بهتری را ثبت کند. همچنین همان‌طور که در شکل‌های (۴)، (۵) و (۶) و جدول‌های (۲) و (۳) قابل مشاهده است مقایسه مدل پیش‌پردازش پیشنهادی با مدل‌های مرجع در حیطه‌ی پارامترهای ارزیابی بررسی گردید که برتری نرخ بازشناسی مدل پیشنهادی را ثبت کند.

۴-۱- محدودیت‌ها و پیشنهادهایی برای کارهای آینده

یکی از محدودیت‌های مدل پیشنهادی زمان شبیه‌سازی بالای آن می‌باشد. بنابراین، تشخیص سریع تصویر را می‌توان با تشخیص متمرکز تعمیم داد تا زمان شبیه‌سازی کاهش یابد. انعطاف‌پذیری مدل پیشنهادی، مدل‌سازی پدیده‌های مختلف را ساده می‌کند که در نتیجه محاسبات را به میزان قابل توجهی گران‌تر می‌کند.

پیشنهادهایی جهت تحقیقات آینده به شرح ذیل می‌باشد: ۱- انتخاب ویژگی‌های برتر تصویر با استفاده از روش‌های بهینه‌سازی مانند الگوریتم ژنتیک و حذف ویژگی‌های با اطلاعات متفاوت کمتر، ۲- ارزیابی پارامترهای قابل آموزش، سرعت آموزش و زمان آموزش مدل، مدل نمودن دیگر نواحی مسیر بینایی و اضافه نمودن آن‌ها به مدل، ۳- گنجاندن زمان در مدل که در این صورت می‌توان بازشناسی تصویر سریع را به بازشناسی با تمرکز و توجه تعمیم داد، ۴- امکان جای کار زیاد مدل پیشنهادی در حوزه ارتقای زیستی که شاید با مطالعات بیشتر بتوان ایده‌های زیستی بسیاری را در ارتقای مدل بکار برد که این امر مستلزم مطالعه زیاد در حوزه یافته‌های علوم اعصاب بینایی است، ۵- استفاده از مجموعه داده‌های دیگر، از جمله Mnist جهت ارزیابی و مقایسه‌ی نتایج ارزیابی با مدل مبتنی بر CNN.

- [1] R. Khallaf, & M. Khallaf, "Classification and analysis of deep learning applications in construction: A systematic literature review", *Automation in Construction*, 129, 103760, 2021.
- [2] M. Momeny, M.A. Sarram, A.M. Latif, R. Sheikhpour, & Y. D. Zhang, "A Noise Robust Convolutional Neural Network for Image classification", *Results in Engineering*, 100225, 2021.
- [3] M. Shariati, M. S. Mafipour, J. H. Haido, S. T. Yousif, A. Toghroli, N. T. Trung, , & A. Shariati, "Identification of the most inuencing parameters on the properties of corroded concrete beams using an Adaptive Neuro-Fuzzy Inference System (ANFIS)", *Steel Compos Struct*, 34(1),155, 2020
- [4] L. Liu, Z. Wang, T. Qiu, Q. Chen, Y. Lu, & C. Y. Suen. "Document image classification: Progress over two decades", *Neurocomputing*, 453, 223-240, 2021.
- [5] J. B. Ritchie, H. L. Masson, S. Bracci & H. P. O. de Beeck. "The unreliable influence of multivariate noise normalization on the reliability of neural dissimilarity". *NeuroImage*, 245, 118686, 2021.
- [6] H. Popal, Y. Wang & I. R. Olson, "A guide to representational similarity analysis for social neuroscience". *Social Cognitive and Affective Neuroscience*, 14(11), 1243-1253, 2019, 2019.
- [7] X. Wang, Y. Xu, Y. Wang, Y. Zeng, J. Zhang, Z. Ling & Y. Bi, " Representational similarity analysis reveals task-dependent semantic influence of the visual word form area". *Scientific reports*, 8(1), 1-10, 2018.
- [8] P. H. Kieval, "Mapping representational mechanisms with deep neural networks". *Synthese*, 200(3), 1-25, 2022.
- [9] H. H. Schütt, A. D. Kipnis, J. Diedrichsen & N. Kriegeskorte, "WStatistical inference on representational geometries". *arXiv preprint arXiv:2112.09200*, 2021.
- [10] I. Muukkonen, K. Ölander, J. Numminen & V. R. Salmela, "Spatio-temporal dynamics of face perception", *NeuroImage*, 209, 116531, 2022.
- [11] J. Diedrichsen, E. Berlot, M. Mur, H. H. Schütt, M. Shahbazi & N. Kriegeskorte, "Comparing representational geometries using whitened unbiased-distance-matrix similarity", *arXiv preprint arXiv:2007.02789*, 2020.
- [12] M. N. Coutanche, E. Akpan & R. R. Buckser. "Representational connectivity analysis: identifying networks of shared changes in representational strength through jackknife resampling". *BioRxiv*, 2020.
- [13] J. Mehrer, C. J. Sporer, N. Kriegeskorte & T. C. Kietzmann. "Individual differences among deep neural network models". *Nature communications*, 11(1), 1-12, 2020.

[14] J. Jung, M. Visser, R. J. Binney & M. A. Lambon Ralph, “Establishing the cognitive signature of human brain networks derived from structural and functional connectivity”, *Brain Structure and Function*, 223(9), 4023-4038, 2018.

[15] S. A. Aketi, S. Roy, A. Raghunathan & K. Roy, K, “Gradual channel pruning while training using feature relevance scores for convolutional neural networks”. *IEEE Access*, 8, 171924-171932, 2020.

[16] E. Kandel, “Principles of Neural Science”, Fifth Edition. McGraw-Hill Education, 2013.

[17] Hubel, D. Eye, Brain, and Vision. [Online]. Available: <http://hubel.med.harvard.edu/index.html>.

[18] J. P. Frisby, and J. V. Stone, “Seeing: The computational approach to biological vision”, The MIT Press, 2010.

[19] Z.H.D. Amnyieh, S.M.J.R. Fatemi, M. Rastgarpour, G.A. Ghazvini, "CNN-RDM: A new image processing model for improving the structure of deep learning based on representational dissimilarity matrix", *Journal of Supercomputing*, vol. 9, pp. 1-25, Sept. 2022 (doi: 10.1007/s11227-022-04661-7).

[۲۰] حیدران داروفه امنیه. زهرا، رستگار فاطمی، سید محمد جلال، رستگارپور. مریم، آقای قزوینی. گلناز (۲۰۲۲). افزایش دقت شبکه‌های عصبی پیچشی مبتنی بر مدل چهار-جریان با فیلترهای پردازش تصویر و نگاشت خطی‌ساز فضای عدم تشابه. *روش‌های هوشمند در صنعت برق*. ۱-۲۸.

[21] A. Krizhevsky, I. Sutskever and G. E. Hinton, “Imagenet classification with deep convolutional neural networks”. in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[22] A. Khan et al, “Optimal Skipping Rates: Training Agents with Fine-Grained Control Using Deep Reinforcement Learning”, *Journal of Robotics*, 2019: p. 10, 2019.

[23] D. Rao, & B. McMahan, “Natural language processing with PyTorch: build intelligent language applications using deep learning”, “O'Reilly Media, Inc.”2019.

[24] S. A. Aketi, S. Roy, A. Raghunathan & K. Roy, “Gradual channel pruning while training using feature relevance scores for convolutional neural networks”, *IEEE Access*, 8, 171924-171932, 2020.

[25] F. Fooladgar, & S. Kasaei, (2020), “Lightweight residual densely connected convolutional neural network. *Multimedia Tools and Applications*”, 79(35), 25571-25588, 2020.