

Increasing the Accuracy of Convolutional Neural Networks Based on 4-Stream Model with Image Processing Filters and Linear Mapping of Dissimilarity Space**Zahra Heydaran-Darogheh-Amnyieh¹, Seyed Mohammad Jalal Rastegar-Fatemi¹, Maryam Rastgarpour², Golnaz Aghaee-Ghazvini³**¹Department of Electrical Engineering- Saveh Branch, Islamic Azad University, Saveh, Iran²Department of Computer Engineering- Saveh Branch, Islamic Azad University, Saveh, Iran³Department of Computer Engineering- Dolatabad Branch, Islamic Azad University, Isfahan, Iran
z.heidaran@iauda.ac.ir, jalal.pe77@iau.ac.ir, m.rastgarpour@iau-saveh.ac.ir, g.ghaee@iauda.ac.ir**Abstract**

With the expansion and success of convolutional networks, the topic of deep learning has attracted increasing attention in recent years; Since convolutional networks include many layers, optimal learning of network layers is of great importance. In this paper, a new model, called the 4-stream model, is presented with the aim of helping to linearize the data space using representational dissimilarity transformation, and the effects of this transformation on standard classifications for artificial data and Cifar10 images are investigated. Then, two models based on data preprocessing with dissimilarity transform representation and Sobel and Edge Detector filters are analyzed. The 4-stream model increased the accuracy by 3.2% due to the increase in the number of model parameters, and hence the capacity of the network. Besides, adding the dissimilarity representation wherever the classifier cannot perform a high-resolution classification by merely using the main features, can help to increase the discriminability of classes by adding linear features.

Keywords: convolutional neural network, deep learning, dissimilarity vector space, prototype, representational dissimilarity matrix

Received: 26 August 2022

Revised: 25 November 2022

Accepted: 14 December 2022

Corresponding Author: Dr. Seyed Mohammad Jalal Rastegar-Fatemi

Citation: Z. Heydaran-Darogheh-Amnyieh, S.M.J. Rastegar-Fatemi, M. Rastgarpour, G. Aghaee-Ghazvini, "Increasing the accuracy of convolutional neural networks based on 4-stream model with image processing filters and linear mapping of dissimilarity space", Journal of Intelligent Procedures in Electrical Technology, vol. 16, no. 61, pp. 1-28, June 2025 (in Persian).

<https://dorl.net/dor/20.1001>.....

مقاله پژوهشی

افزایش دقت شبکه‌های عصبی کانولوشنی مبتنی بر مدل چهار-جریان با فیلترهای پردازش تصویر و نگاشت خطی ساز فضای عدم تشابه

زهرا حیدران داروقه‌امنیه^۱، سید محمدجلال رستگارفاطمی^۱، مریم رستگارپور^۲، گلناز آفایی‌قزوینی^۳

۱- گروه برق- واحد ساوه، دانشگاه آزاد اسلامی، ساوه، ایران

۲- گروه کامپیوتر- واحد ساوه، دانشگاه آزاد اسلامی، ساوه، ایران

۳- گروه کامپیوتر- واحد دولت آباد، دانشگاه آزاد اسلامی، اصفهان، ایران

z.heidaran@iauda.ac.ir, jalal.pe77@iau.ac.ir, m.rastgarpour@iau-saveh.ac.ir, g.aghaee@iauda.ac.ir

چکیده: در سال‌های اخیر با گسترش و موفقیت شبکه‌های کانولوشنی، موضوع یادگیری عمیق بیش از پیش مورد توجه قرار گرفته است. از آنجا که شبکه‌های کانولوشنی شامل لایه‌های زیادی هستند، یادگیری بهینه لایه‌های شبکه از اهمیت بالایی برخوردار است. در این مقاله، مدل جدیدی به نام چهار-جریان، با هدف کمک به خطی کردن فضای داده از طریق تبدیل عدم تشابه بازنمایی ارائه و تأثیر این تبدیل روی طبقه‌بندهای استاندارد برای داده‌های مصنوعی و تصاویر سیفار-۱۰ بررسی و دو مدل مبتنی بر پیش‌پردازش داده با تبدیل عدم تشابه بازنمایی و فیلترهای سوبل و آشکارساز لبه تحلیل شده است. مدل چهار-جریان به دلیل بالا رفتن تعداد پارامترهای مدل و به تبع آن ظرفیت شبکه میزان $3/2$ درصد افزایش دقت داشته است و اضافه نمودن بازنمایی عدم تشابه در جایی که طبقه‌بند نتواند با ویژگی‌های اصلی، تفکیک‌پذیری بالایی انجام دهد، می‌تواند تا حدودی با افزودن ویژگی‌های خطی به تفکیک‌پذیری کلاس‌ها کمک کند.

کلمات کلیدی: سیستم کانولوشنی، فضای برداری عدم تشابه، ماتریس عدم تشابه بازنمایی، مرجع، یادگیری عمیق

تاریخ ارسال مقاله: ۱۴۰۱/۶/۴

تاریخ بازنگری مقاله: ۱۴۰۱/۹/۴

تاریخ پذیرش مقاله: ۱۴۰۱/۹/۲۳

نام نویسنده‌ی مسئول: دکتر سید محمدجلال رستگارفاطمی

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

۱- مقدمه

در یادگیری عمیق، شبکه عصبی کانولوشنی^۱ (CNN) یک کلاس 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]. شبکه عصبی کانولوشنی نسبت به سایر مدل‌های مشابه محبوب‌تر شده 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. عدم تشابه بازنمایی^۳ یک روش جایگزین به جای شیوه سنتی نمایش داده در فضای ویژگی است. این نحوه نمایش در طبقه‌بندی، شناسایی الگو^۴، مسأله‌های با تعداد کلاس زیاد و نیز مسأله‌ای که تعداد کمی مثال برای آموزش موجود است کاربرد دارد [۵]. همچنین عدم تشابه نقش مهمی در ادراک، شناخت و شیوه دسته‌بندی اطلاعات در انسان ایفا می‌کند [۶]. در دهه ۲۰۰۰ تا ۲۰۱۰ میلادی و قبل از شروع انقلاب یادگیری عمیق، روش عدم تشابه به‌طور گسترده در یادگیری ماشین مورد استفاده بوده است، اما در سال‌های اخیر با موفقیت‌های چشم‌گیر یادگیری عمیق که با آموزش انتها به انتها در شبکه‌های عصبی نیاز به مهندسی ویژگی را به حداقل می‌رساند، میزان توجه جامعه علمی به عدم تشابه بازنمایی تا حدودی کاهش یافت. هدف این مقاله، تحلیل و بررسی اثر مثبتی است که عدم تشابه بازنمایی در بهبود یادگیری عمیق به‌طور بالقوه می‌تواند ایفا کند. به‌طور کلی نحوه به‌کارگیری عدم تشابه بازنمایی مطابق جدول (۱) در سه دسته قرار می‌گیرد. در ادامه، مطالعه‌های انجام شده در هر یک از این روش‌ها شرح داده می‌شوند:

۱-۱- فضای عدم تشابه

در این شیوه از محاسبه عدم تشابه، فاصله بین دو بردار به‌صورت یک عدد اسکالر محاسبه می‌شود. برای محاسبه فاصله از سنج‌های مختلفی از جمله فاصله اقلیدسی^۵، فاصله همبستگی^۶ یا هر کمیت دیگری می‌توان استفاده کرد. در مرجع [۶]، برای ساخت ماتریس عدم تشابه، ابتدا تعداد K مرجع^۷ انتخاب شده، فاصله هر کدام از نمونه‌های آموزشی با تمام مرجع‌ها سنجیده می‌شود. به این ترتیب هر مثال در فضای جدید با K عدد اسکالر نمایش داده می‌شود. سپس دو طبقه‌بند، یکی K -نزدیکترین همسایگی^۸ (KNN) و دیگری طبقه‌بندی‌کننده مبتنی بر چگالی معمولی خطی/کوآدراتیک منظم^۹ (RLNC/RQNC) را روی داده‌های جدید آزموده و توانسته با طبقه‌بندی‌کننده مبتنی بر چگالی معمولی خطی/کوآدراتیک منظم دقت بالاتری نسبت به K -نزدیکترین همسایگی به‌دست آورد. در مرجع [۷]، از دو راه مختلف برای ساختن طبقه‌بند با عدم تشابه استفاده شده است. در روش اول، عدم تشابه روی فضای اقلیدسی اجرا شده و در روش دوم، از فاصله هسته^{۱۰} استفاده شده است. در نهایت نشان داده شده که روش فاصله هسته توانسته برای K -نزدیکترین همسایگی نتیجه بهتری نسبت به فاصله اقلیدسی بگیرد و همچنین روش‌های کاهش ابعاد فضا نظیر تجزیه و تحلیل اجزای اصلی^{۱۱} (PCA) اثر منفی نويز را روی روش اقلیدسی کاهش داده است. در مرجع [۸]، چنانچه تعداد اندکی مرجع (بین ۳ تا ۱۲ درصد از کل داده‌های آموزش) به شیوه هوشمندانه انتخاب شود، دقت نسبت به زمانی که از طبقه‌بند K -نزدیکترین همسایگی روی کل داده‌های آموزشی استفاده شود، بهتر می‌گردد. همچنین با انتخاب تعداد کمتر مرجع، پیچیدگی زمانی به مراتب پایین‌تر خواهد بود. در مرجع [۹]، بازده طبقه‌بند K -نزدیکترین همسایگی و ماشین بردار پشتیبانی^{۱۲} (SVM) مبتنی بر ویژگی‌های اصلی داده با طبقه‌بند K -نزدیکترین همسایگی مقایسه شده که در نتیجه طبقه‌بند K -نزدیکترین همسایگی بر عدم تشابه یا بازده بیشتری دارند یا حداقل در همان محدوده بازده قرار می‌گیرند. در مرجع [۱۰]، از عدم تشابه برای شناسایی الگو استفاده گردیده است. در این مقاله، چنین مطرح شده که مقایسه میان اشیاء، نه فقط در فضای ویژگی و به‌صورت برداری، بلکه به صورت‌های دیگر نظیر بازنمایی ساختاری^{۱۳}، بازنمایی هسته^{۱۴} و مدل کلاس^{۱۵} می‌تواند انجام پذیرد و در این حیث جا برای توسعه مفهوم عدم تشابه وجود دارد. در مرجع [۱۱]، با استفاده از عدم تشابه، یک طبقه‌بند برای تشخیص حالت بدن انسان ساخته شده است.

Table (1): How to use dissimilarity of representation

جدول (۱): نحوه بکارگیری عدم تشابه بازنمایی

ردیف	نوع روش	شیوه محاسبه	کاربرد
۱	فضای عدم تشابه	محاسبه فاصله دوبدو میان مثال‌ها و مرجع‌ها به صورت اسکالر	تبدیل داده‌هایی مانند گراف که نمایش برداری ندارند به بردار، کاهش ابعاد فضا، ایجاد ویژگی‌های خطی و تفکیک‌پذیر
۲	بردار عدم تشابه ^{۱۶}	محاسبه بردار تفاضل میان مثال‌ها و مرجع‌ها	تبدیل مسأله با تعداد زیاد کلاس به مسأله باینری، حل مسائلی که تعداد کمی مثال برای هر کلاس دارند
۳	ماتریس عدم تشابه	محاسبه فاصله دوبدو میان تمام مثال‌ها با یکدیگر به صورت ماتریس	مقایسه سیگنال‌ها با یکدیگر، شناسایی الگوها، کاربردهای پزشکی

هر کنش و حرکت در انسان توالی از حالت‌های بدن در طول زمان است که توسط زاویه اعضای بدن قابل شناسایی است. آن‌ها برای شناسایی و طبقه‌بندی حرکات انسان، یک سری مرجع برای حالت‌های پایه‌ای بدن می‌سازند و در هر لحظه حالت فعلی بدن با این مرجع‌ها مقایسه شده و عدم تشابه محاسبه می‌شود. در واقع این مرجع‌ها یک دستگاه مختصات جدید ارائه می‌دهند که هر حالت انسان در این دستگاه مختصات، به صورت پراکنده^{۱۷} و توسط تعداد اندکی از محورها قابل بیان است. سپس طبقه‌بندی در این مختصات جدید اجرا شده و توانسته به بازده بالاتر نسبت به روش‌های زمان خود دست یابد. در مرجع [۱۲]، مسأله جاسازی نمودار^{۱۸} را با عدم تشابه انجام داده که در آن گراف به فضای برداری نگاشت می‌شود. هدف از نگاشت این است که بتوان داده‌های متنوعی که در حوزه کامپیوتر به صورت گراف نمایش می‌یابد را ابتدا به فضای برداری تبدیل کرد تا بتوان طبقه‌بندی‌های موجود را روی آن اجرا نمود. محققان ادعا کردند که این شیوه پتانسیل بالایی دارد و می‌تواند از روش‌هایی که مستقیماً روی گراف عمل می‌کنند، بهتر باشد. در مرجع [۱۳]، از عدم تشابه برای طبقه‌بندی متون استفاده شده است. هدف از این کار طبقه‌بندی اسناد متنی به دسته‌های از پیش تعریف شده است. به طور سنتی طبقه‌بندی اسناد با استفاده از روش کوله کلمات^{۱۹} انجام می‌گرفت که دارای معایبی از جمله تعداد زیاد ویژگی و بالا بودن ابعاد فضا بود. محققان، از عدم تشابه برای کاهش ابعاد فضا استفاده نموده و طبقه‌بندی‌های مختلف را روی فضای کاهش یافته اعمال کردند. مطابق بررسی انجام شده، با ساخت فضاهای مختلف از عدم تشابه با معیارهای گوناگون فاصله و با مجموعه‌های مختلفی از مرجع‌ها و آموزش طبقه‌بند مجزا روی هر یک از این فضاها و در نهایت ترکیب خروجی این‌ها با استفاده از اکثریت^{۲۰} توانسته به دقت بالاتری نسبت به روش‌های موجود برسد.

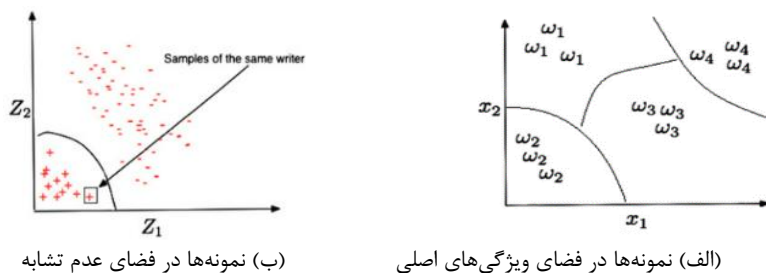
۱-۲- عدم تشابه بردار

در این شیوه از محاسبه عدم تشابه، بردارهای مثال از بردارهای مرجع تفریق شده و به صورت یک بردار جدید استفاده می‌شود. در مرجع [۱۴]، از عدم تشابه برای تبدیل مسأله بزرگ چندکلاسه^{۲۱} به مسأله دوکلاسه^{۲۲} (باینری) استفاده شده است. در مسائل بزرگ چندکلاسه تعداد زیادی دسته‌بندی وجود دارد که برای هر دسته تنها تعداد کمی مثال موجود است، به عنوان نمونه، شناسایی گونه پرندگان، درستی سنجی امضاء و شناسایی نویسنده متن که در آن هر شخص یا هر گونه یک کلاس مجزا در نظر گرفته می‌شود. در مرجع [۱۵]، برای تبدیل مسأله چند کلاسه به مسأله دودویی، ابتدا از مثال‌های نمونه تعداد K مرجع انتخاب شده است و سپس با تفاضل گرفتن بردار ویژگی‌ها بین مثال‌های نمونه و مرجع به صورت دوبدو، هر جفت بردار تبدیل به یک بردار می‌شود و در نهایت به این صورت برچسب^{۲۳} زده می‌شود: اگر هر دو بردار اولیه از یک کلاس باشند برچسب مثبت (یک) و اگر از دو کلاس مجزا باشند برچسب منفی (صفر). به این صورت مسأله طبقه‌بندی اولیه با M دسته و N مثال به یک مسأله با 2 دسته و $N \times K$ مثال تبدیل شده و یک طبقه‌بند استاندارد روی مسأله جدید آموزش داده می‌شود. آنگاه در زمان تست مثال، تست از تمام K مرجع تفاضل گرفته شده و K بار طبقه‌بند اجرا می‌شود که K عدد باینری صفر یا ۱ در خروجی وجود خواهد داشت و صفر بیانگر متفاوت بودن دو بردار تست و مرجع و ۱ بیانگر هم‌کلاس بودن دو بردار تست و مرجع است. در نهایت برچسب کلاسی از مرجع به عنوان پاسخ برگردانده می‌شود که بیشترین ۱ را از K عدد دارا باشد. در مرجع [۱۶]، دو فیلتر الگوی باینری محلی^{۲۴} (LBP) و کوانتیزاسیون فاز محلی^{۲۵} (LPQ) برای استخراج ویژگی بافت^{۲۶} از متن استفاده شده و در کنار آن از بردار عدم تشابه برای تبدیل مسأله به مسأله باینری بهره برده است. همچنین تأثیر افزایش یا کاهش تعداد مرجع‌ها در دقت نهایی مدل

ارزیابی شده است. شکل (۱) تبدیل مسأله ۴ کلاسه به مسأله باینری را نشان می‌دهد که در آن بردارهای جدید از تفاضل دو بردار قدیمی بدست می‌آید و برچسب مثبت نشان‌دهنده هم‌کلاس بودن دو بردار تشکیل‌دهنده آن و برچسب منفی نشان‌دهنده غیرهم‌کلاس بودن دو بردار تشکیل‌دهنده آن است. در مرجع [۱۷] با استفاده از تبدیل چندکلاسه به دوکلاسه، مسأله تأیید اصالت امضاء بررسی شده است. مسأله تأیید امضاء به دلیل اینکه تعداد کمی مثال (نمونه امضاء) برای هر کلاس وجود دارد و تعداد کلاس‌ها (افراد) بسیار زیاد است، چالش برانگیز می‌شود. محققان، با استفاده از عدم تشابه و تبدیل مسأله به باینری و استفاده از حداکثر ادغام^{۲۷} به جای اکثریت در زمان استحصال برچسب نهایی توانستند دقت مدل را بالا ببرند. در مرجع [۱۸]، تشخیص گونه‌های جنگل از روی تصاویر ماهواره با استفاده از عدم تشابه انجام می‌شود که از یک دیتاست شامل ۲۲۴۰ تصویر از ۱۱۲ گونه جنگل استفاده شده است. هر تصویر از فیلترهای مختلفی مانند الگوی باینری محلی، فیلتر گابور^{۲۸}، کوانتیزاسیون فاز محلی، تغییر شکل ویژگی ثابت مقیاس^{۲۹} (SIFT) و ویژگی قوی تسریع^{۳۰} (SURF) عبور داده شده و به تناظر آن فضاهای متعددی از عدم تشابه ایجاد گردیده است که هر کدام به یک طبقه‌بند مجزا داده می‌شود و در نهایت از خروجی همه اینها برآیند^{۳۱} گرفته شده که با این روش دقت به میزان قابل ملاحظه‌ای بهبود یافته است. در مرجع [۱۹]، مسأله شناسایی نوع پرند از روی سیگنال صوتی بررسی شده که به دلیل تنوع گونه‌های پرندگان، تعداد کلاس‌ها بسیار زیاد است. ابتدا از حوزه زمان-فرکانس، ویژگی استخراج می‌شود، سپس از عدم تشابه برای تبدیل مسأله به باینری استفاده شده است و به این ترتیب برای دیتاست شامل ۹۱۵ کلاس دقت ۷۱ درصد به دست می‌آید. محققان، عنوان می‌کنند که در مسائل با تعداد بسیار زیاد کلاس استفاده از بردار عدم تشابه و تبدیل مسأله به دو کلاسه گزینه بهتری نسبت به روش‌هایی که مستقیماً مسأله را به صورت چندکلاسه حل می‌کنند، خواهد بود. شکل (۲) ایده ارائه شده را نشان می‌دهد.

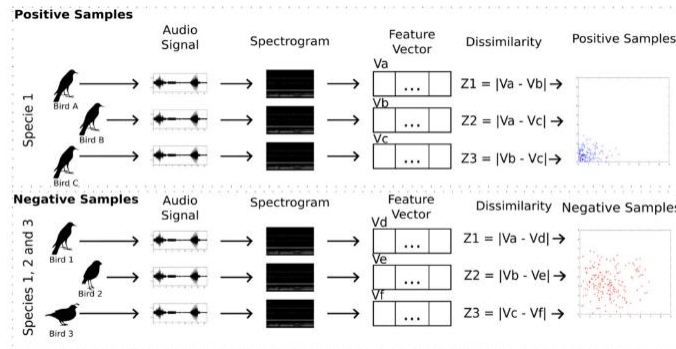
۱-۳- ماتریس عدم تشابه بازنمایی^{۳۲}

تحلیل شباهت بازنمایی^{۳۳} (RSA) و ماتریس عدم تشابه بازنمایی (RDM) کاربردهای متنوعی در علوم پزشکی دارد. بررسی الگوهای فعالیت مغزی از روی سیگنال الکتروانسفالوگرافی^{۳۴} (EEG)، تحلیل تصاویر تصویربرداری رزونانس مغناطیسی کاربردی^{۳۵} (fMRI) برای درک بهتر وابستگی بخش‌های مختلف مغز، ارائه مدل‌های ریاضی و محاسباتی و تطبیق آن با تئوری‌های پزشکی و روانشناسی از کاربردهای تحلیل شباهت بازنمایی و ماتریس عدم تشابه بازنمایی است. برای مثال، در مرجع [۲۰]، فعالیت مغزی فرد در هنگام مشاهده تصاویر روی مانیتور، به صورت سیگنال‌های الکتروانسفالوگرافی، در یک آزمایش، ذخیره می‌گردد. سپس ماتریس عدم تشابه بازنمایی، از عدم تشابه میان سیگنال‌های الکتروانسفالوگرافی، محاسبه می‌شود. از روی این ماتریس و همچنین تعبیه کردن^{۳۶} آن توسط روش مقیاس بندی چند بعدی^{۳۷} (MDS) می‌توان به ساختار ذهنی فرد دست یافت. شکل (۳) این روش را نشان می‌دهد. اینکه کدام تصاویر از نظر فرد به یکدیگر شبیه‌اند و کدام مفاهیم با همدیگر ارتباط دارند، به روانشناس کمک می‌کند تا شناخت بهتری نسبت به بیماری فرد داشته باشد. همچنین با مقایسه ماتریس عدم تشابه بازنمایی برای فرد بیمار و سالم امکان تشخیص بهتر بیماری، طبقه‌بندی کردن^{۳۸} بیماری‌ها و نیز رصد روند بهبود بیمار تحت درمان میسر می‌شود. شکل (۴) تحلیل عملکرد مغز با این روش را نشان می‌دهد. در مرجع [۲۱]، به کمک ماتریس عدم تشابه بازنمایی و تحلیل شباهت بازنمایی می‌توان در علوم اعصاب به پیش‌بینی رفتار افراد پرداخت، تفاوت فعالیت نورون‌ها را در گروه‌های مختلف افراد تحلیل کرد و تئوری‌های روانشناسی را با داده‌های تجربی مورد آزمون قرار داد.

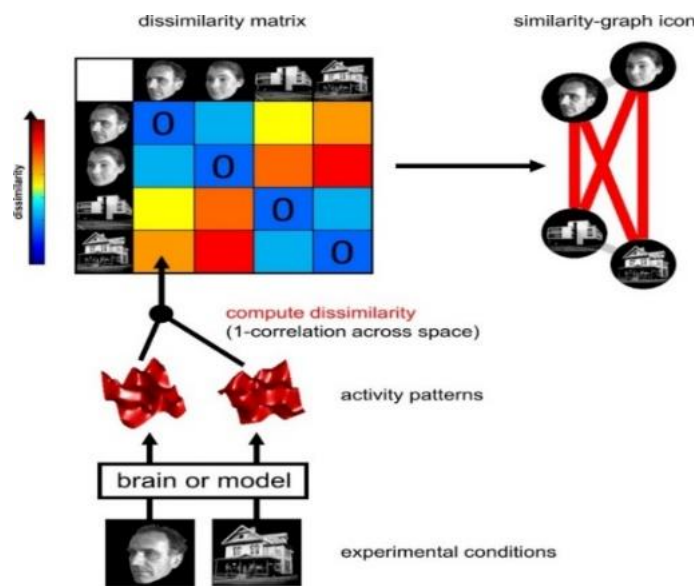


شکل (۱): تبدیل مسأله ۴ کلاسه به باینری [۱۶]

Figure (1): Converting the 4-class problem to binary, a) Samples in the space of main features, b) Samples in the space of dissimilarity [16]



شکل (۲): تبدیل مسأله چندکلاسه به باینری در تشخیص گونه پرندگان [۱۹]
 Figure (2): Converting multi-class problems to binary in bird species recognition [19]

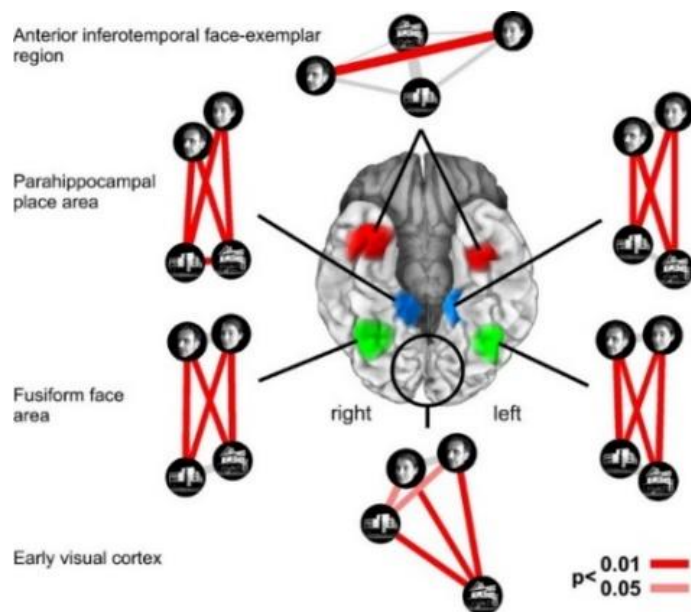


شکل (۳): نحوه ساخت ماتریس عدم تشابه بازنمایی و تعبیه کردن آن [۲۰]
 Figure (3): How to make the RDM matrix and embed it [20]

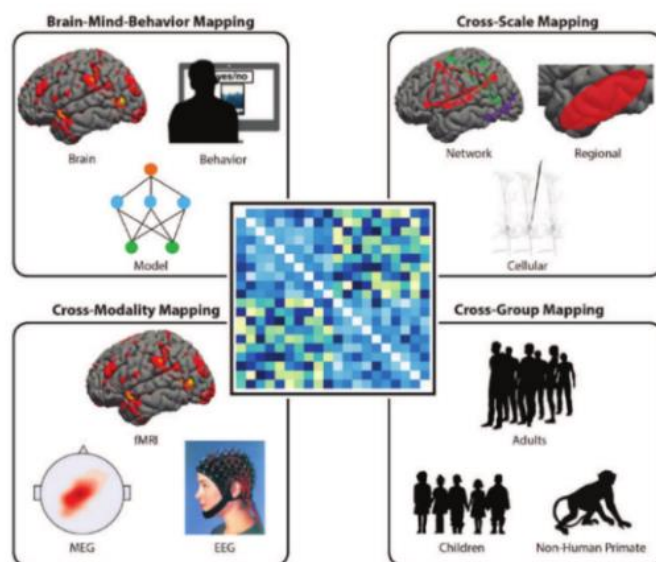
محققان، یک راهنمای جامع از نحوه بکارگیری ماتریس عدم تشابه بازنمایی و تحلیل شباهت بازنمایی در مسائل پزشکی و علوم اعصاب ارائه داده‌اند که شکل (۵) این مفهوم را نشان می‌دهد. شکل (۶) نمونه‌ای از تحلیل ماتریس عدم تشابه بازنمایی است که در آن یک فرد در معرض نمایش تعدادی تصاویر قرار می‌گیرد.

در مرجع [۲۱]، از یک طرف مطابق شکل (۶-الف)، پاسخ نورون‌های مغز ذخیره و به ازای هر تصویر دوبدو مقایسه می‌شود و ماتریس عدم تشابه بازنمایی آن به دست می‌آید و از طرف دیگر مطابق شکل (۶-ب)، پاسخ رفتاری فرد، ذخیره و ماتریس عدم تشابه بازنمایی آن محاسبه می‌شود و از جهت دیگر مطابق شکل (۶-ج)، از شباهت مفهومی تصاویر، یک ماتریس عدم تشابه بازنمایی دیگر به دست می‌آید و در فاز نهایی مانند شکل (۶-د)، این ماتریس‌ها با یکدیگر مقایسه می‌گردد تا میزان انطباق فعالیت مغزی و پاسخ رفتاری و مدل مفهومی با یکدیگر سنجیده شود و سلامت جسمی، ذهنی و روانی فرد مورد بررسی قرار گیرد. در عین حال هر یک ماتریس‌های عدم تشابه بازنمایی فوق، به تنهایی هم قابل تحلیل بوده و اطلاعات ارزشمندی در اختیار پزشک قرار می‌دهد.

در مرجع [۲۲]، آزمایشی بر روی افراد دارای اعتیاد به سیگار انجام شده است. به این صورت که این افراد در معرض پیام‌های ترک دخانیات قرار می‌گیرند و هم‌زمان تصاویر تصویربرداری روزنانس مغناطیسی کاربردی از قشر جلوی مغز^{۳۹} این افراد گرفته و سپس میزان استعمال سیگار آنها یک ماه بعد پرسیده می‌شود. نتیجه نشان می‌دهد که با استفاده از تحلیل ماتریس عدم تشابه بازنمایی و تحلیل شباهت بازنمایی شدت نمایش پیام در قشر جلوی مغز فرد با تغییر رفتار وی در آینده همبستگی دارد.



شکل (۴): تحلیل عملکرد قسمت‌های مختلف مغز با عدم تشابه بازنمایی [۲۰]
Figure (4): Analysis of the function of different parts of the brain with RDM [20]



شکل (۵): کاربرد تحلیل عدم تشابه بازنمایی در مقایسه نواحی مختلف مغز، گروه‌ها و مدل‌ها [۲۱]
Figure (5): Application of RDM analysis in comparing different brain areas, groups and models [21]

در مرجع [۲۳]، افراد شرکت‌کننده در معرض تصاویری مانند خانه، چهره و غیره قرار داده می‌شوند و همزمان با برخی تصاویر، شوک الکتریکی به فرد اعمال می‌گردد. بر اساس نتیجه آزمایش، پاسخ نوروها برای تصاویر همراه با شوک، از دسته‌های متفاوت، می‌تواند پاسخ انقباض مردمک چشم را که در آزمایش جداگانه دیگری به دست می‌آید، پیش‌بینی کند. در حالی که این پیش‌بینی بدون تحلیل شباهت بازنمایی و فقط با مقایسه میانگین فعالیت نوروها امکان پذیر نبوده است.

۱-۴- استفاده از عدم تشابه در یادگیری عمیق

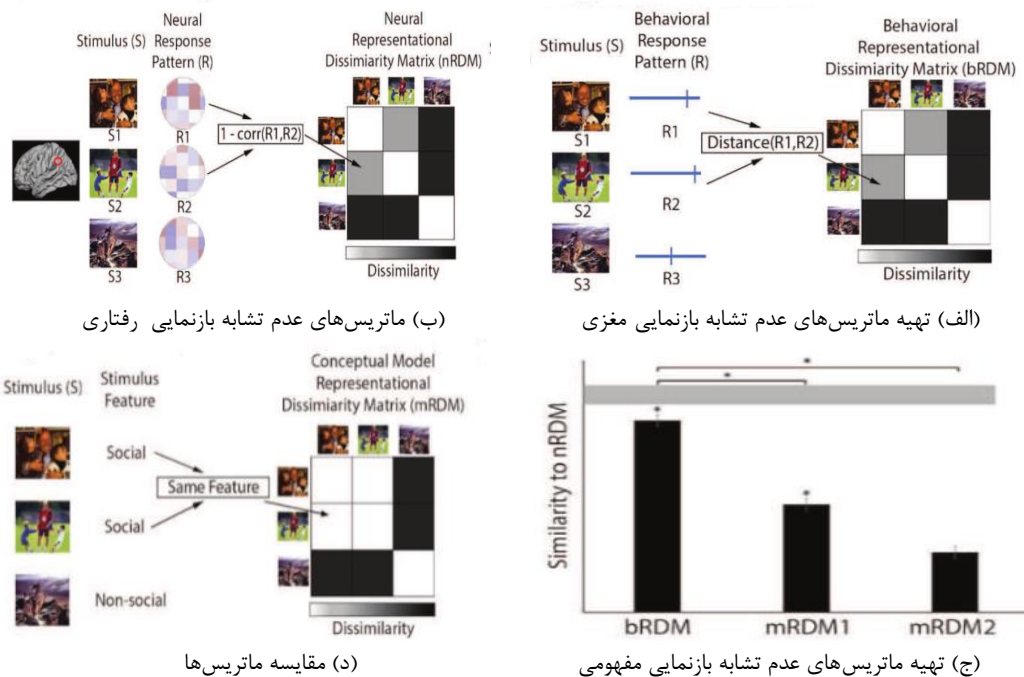
در مرجع [۵]، برای مسائل طبقه‌بندی با تعداد زیاد کلاس و تعداد کم مثال آموزشی، ویژگی‌های استخراج شده از شبکه‌های پیش از آموزش^{۴۰} می‌تواند برای ساخت فضای عدم تشابه استفاده شود. به‌طور مشخص، به این صورت که مسأله چندکلاسه با تفاضل گرفتن بردارهای تولید شده توسط شبکه عصبی به یک مسأله باینری تبدیل گردد. باید اشاره کرد که موفقیت این

استراتژی خیلی به انتخاب دقیق مرجع‌ها بستگی دارد. این انتخاب باید به گونه‌ای باشد که توزیع داده در تمامی کلاس‌ها را به طور کامل پوشش دهد. همچنین تعداد مرجع‌ها اهمیت دارد، به طوری که هرچه تعداد بیشتری مرجع وجود داشته باشد، پوشش بیشتری از توزیع داده حاصل می‌شود و دقت مدل بالاتر می‌رود. از طرف دیگر یک مدل یادگیری عمیق می‌تواند مستقیماً جهت تولید بردارهای مناسب برای عدم تشابه آموزش ببیند. این ایده در شکل (۷) نشان داده شده است.

در شکل (۷) دو شبکه عصبی کانولوشنی می‌توانند هم‌زمان به نحوی آموزش دیده شوند تا بردارهایی به عنوان ویژگی ساخته شود که بین دو ورودی از یک کلاس، مقدار عدم تشابه به صفر نزدیک باشد و بین دو ورودی از دو کلاس متفاوت عددی غیر صفر به دست آید.

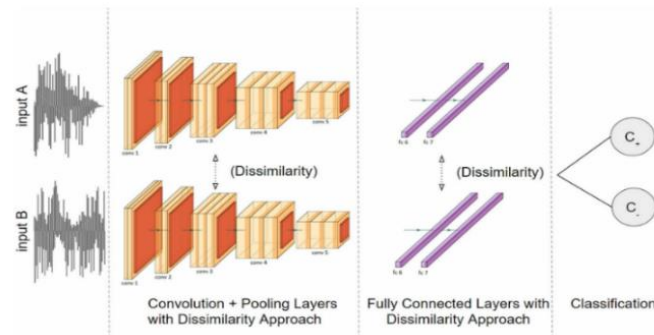
در مرجع [۲۴]، با دادن مجموعه‌ای از تصاویر به شبکه عصبی کانولوشنی و استخراج ویژگی از لایه آخر خروجی شبکه، ماتریس عدم تشابه بازنمایی محاسبه شده و با تغییر تابع زیان، شبکه عصبی کانولوشنی را به نحوی آموزش می‌بیند که درایه‌های ماتریس روی کلاس‌های مشابه مقادیر کمینه و روی کلاس‌های متفاوت مقادیر بیشینه را داشته باشند. به این ترتیب با استفاده از عدم تشابه و ماتریس عدم تشابه بازنمایی و تابع زیان جدید، مدل ارائه شده می‌تواند دقت را روی داده‌ها زمانی که تنها ۵ نمونه به ازای هر کلاس وجود دارد از ۵۱ درصد به ۶۰ درصد افزایش دهد. در مرجع [۲۵]، ماتریس‌های عدم تشابه بازنمایی، 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 ماتریس عدم تشابه بازنمایی معمولاً به عنوان 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 مرجع [۲۶]، داده‌ها به یک فضای هندسی "ماتریس عدم تشابه بازنمایی" که تفاوت‌های نسبی بین پاسخ‌های چندگانه را برجسته می‌کند، نگاشت می‌شود. این ماتریس‌ها، نسبت به نوع داده‌ای که از آن مشتق شده‌اند، مجزا هستند. برای مثال، ماتریس عدم تشابه بازنمایی از یک نقطه زمانی خاص در الکتروانسفالوگرافی، می‌تواند با ماتریس‌های عدم تشابه بازنمایی یک خوشه وکسل خاص در تصویربرداری رزونانس مغناطیسی کاربردی مقایسه شود، بنابراین دقت زمانی خوب الکتروانسفالوگرافی با دقت فضایی خوب تصویربرداری رزونانس مغناطیسی کاربردی ترکیب می‌شود.



شکل (۶): تهیه ماتریس‌های عدم تشابه بازنمایی مغزی، رفتاری و مفهومی و انطباق آن‌ها [۲۱]

Figure (6): Preparation of cerebral, behavioral and conceptual RDM matrices and their adaptation, a) Preparation of brain RDM matrices, b) Behavioral BDM matrices, c) Preparation of conceptual RDM matrices, d) Comparison of matrices [21]



شکل (۷): آموزش شبکه عصبی کانولوشنی برای تولید بهتر بردارهای عدم تشابه [۵]

Figure (7): Convolutional neural network training for better generation of dissimilarity vectors [5]

در مرجع [۲۷]، تحلیل شباهت بازنمایی مدل‌های محاسبات مغز، با بررسی این که چگونه الگوهای فعالیت عصبی شرایط تجربی را منعکس می‌نمایند، آزمایش می‌شود. به جای پیش‌بینی مستقیم الگوهای فعالیت، مدل‌ها هندسه بازنمایی را پیش‌بینی می‌کنند. همان‌طور که توسط ماتریس عدم تشابه بازنمایی تعریف می‌شود که چگونه الگوهای فعالیت‌های مختلف مشابه یا غیرمشابه با شرایط تجربی مختلف مرتبط هستند.

از آنجایی که در پرتو پیشرفت‌های چشمگیر یادگیری عمیق، توجه جامعه علمی به موضوع عدم تشابه کاهش یافته است، این مقاله قصد دارد اثر روش‌های مبتنی بر عدم تشابه را که قبلاً در روش‌های کلاسیک مورد توجه بودند، این بار در ترکیب با یادگیری عمیق مورد تحلیل قرار دهد. به عبارت دیگر، هدف، تحلیل و بررسی اثر مثبتی است که عدم تشابه بازنمایی در بهبود یادگیری عمیق و یا در کنار آن به‌طور بالقوه می‌تواند ایفا کند. از مرور مقالات مرتبط و پژوهش‌های انجام شده در مبحث عدم تشابه می‌توان به این جمع‌بندی رسید که در چند سال گذشته، بحث عدم تشابه و کمکی که می‌تواند در کنار یادگیری عمیق به پردازش تصویر کند، از کانون توجه دور مانده است. تعداد مقالات چاپ شده در این مبحث نسبت به سایر مقالات بسیار کمتر بوده و مقالاتی که به‌طور تخصصی به این موضوع پرداخته باشند، کمیاب و عمدتاً مربوط به سال‌های قبل ۲۰۱۰ میلادی هستند. از این‌رو انگیزه مقاله، بررسی دقیق و همه‌جانبه تبدیل فضای عدم تشابه در خطی‌سازی فضا و نحوه تأثیر آن بر بهبود طبقه‌بندی‌های مختلف از جمله رگرسیون لجستیک^{۴۱}، ماشین بردار پشتیبان، K -نزدیکترین همسایگی، بی‌ساده لوح^{۴۲}، درخت تصمیم^{۴۳}، جنگل تصادفی^{۴۴} و تقویت گرادیان^{۴۵} است.

در مقالات پیشین، تنها طبقه‌بندی‌های ماشین بردار پشتیبان و K -نزدیکترین همسایگی استفاده شده و هیچکدام اثر نگاشت داده‌ها خصوصاً داده‌های از جنس تصویر را به فضای عدم تشابه به‌طور گسترده مورد آزمون قرار نداده بودند. هدف دیگر استفاده از نگاشت فضای عدم تشابه به عنوان پیش‌پردازش برای کمک به یادگیری عمیق است. در اکثر پژوهش‌های قبل، به کلی از یادگیری عمیق در فضای عدم تشابه استفاده نشده و یا در موارد اندکی که از شبکه‌های عصبی عمیق استفاده شده است، نگاشت فضای عدم تشابه بعد از شبکه عصبی قرار گرفته می‌شود، به نحوی که ابتدا از لایه‌ها ویژگی استخراج می‌گردد و سپس این ویژگی‌ها توسط ماتریس عدم تشابه بازنمایی نگاشت شده‌اند. نگاشت فضای عدم تشابه به عنوان پیش‌پردازش قبل از شبکه عصبی عمیق می‌تواند با انجام خطی‌سازی فضای داده برای شبکه‌های عصبی عمیق کمک کننده باشد و این نگاشت به بهبود شبکه‌های جدید و پیشرفته پردازش تصویر نظیر شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی^{۴۶} (CVT) و شبکه عصبی کانولوشنی متراکم باقیمانده^{۴۷} (RDenseCNN) کمک خواهد کرد. ساختار مقاله در ادامه 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، نتیجه‌گیری این پژوهش در قسمت چهارم بررسی خواهد شد.

۲- مدل پیشنهادی چهار-جریان^{۴۸}

همان‌طور که اشاره شد، ماتریس عدم تشابه بازنمایی نوعی از نمایش داده است که در آن، نمونه‌ها به‌جای اینکه در فضای ویژگی‌های اصلی خود نمایش یابند، به‌صورت برداری از فاصله با تصاویر مرجع نمایش پیدا می‌کنند.

به عنوان مثال، یک تصویر به جای اینکه در فضای پیکسل نمایش داده شود، به صورت برداری از فاصله‌ها با مجموعه‌ای از تصاویر مرجع نمایش می‌یابد. معیار فاصله می‌تواند با روش‌های مختلف اندازه‌گیری شود. در مدل ارائه شده، عمدتاً از فاصله اقلیدسی و در مواردی از فاصله همبستگی استفاده شده است.

لیستی از معروف‌ترین معیارهای فاصله، در جدول (۲) آمده که x_s و x_t یک بردار n بعدی، $D(x_s, x_t)$ فاصله دو بردار، علامت - و میانگین و ترانهاده بردار است. شیوه نمایش عدم تشابه، کاربردهای متنوعی دارد که از جمله آن‌ها می‌توان به خطی‌سازی فضا برای تسهیل طبقه‌بندی در مرجع [۲۸]، تبدیل مسأله چندکلاسه به دو کلاسه در مرجع [۲۹]، ساخت ماتریس عدم تشابه بازنمایی از روی سیگنال‌های مغز و مقایسه آن‌ها برای کاربردهای پزشکی و روانشناسی در مرجع [۷] اشاره کرد. شکل (۸) نحوه انتقال تصویر به فضای برداری عدم تشابه^{۴۹} را نشان می‌دهد. دیتاست سیفار-۱۰^{۵۰} (مؤسسه تحقیقات پیشرفته کانادا، ۱۰ کلاس) شامل ۶۰۰۰ تصویر با سایز $3 \times 32 \times 32$ است. در حالت کلی، برای هر آزمایش یک‌بار تصاویر به صورت خام بعد از نرمال‌سازی به فرم یک بردار 1×3072 به طبقه‌بند داده می‌شود و بار دیگر با محاسبه فاصله با ۱۰۰ مرجع به فضای برداری عدم تشابه برده و به فرم یک بردار 1×100 به طبقه‌بند، اعمال می‌شود.

مدل ارائه شده در چند فاز اجرا شده است. تصاویر اصلی تغییر شکل داده شده^{۵۱} به صورت بردار 1×3072 ، تصاویر عبور داده شده از فیلتر آشکارساز لبه^{۵۲}، تصاویر عبور داده شده از فیلتر سوبل^{۵۳}، تصاویر سیاه-سفید^{۵۴} شده، تصاویر سیاه-سفید عبور داده شده از فیلتر آشکارساز لبه، تصاویر سیاه-سفید عبور داده شده از فیلتر سوبل. شکل (۹) نحوه اجرای فرآیند تعدادی از فازها را نشان می‌دهد. در مدل چهار-جریان، ابتدا هر تصویر از دیتاست سیفار-۱۰ به شبکه از پیش آموزش دیده شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی ۱۳ در مرجع [۲۸] و شبکه عصبی کانولوشنی متراکم باقیمانده در مرجع [۲۹] داده شده است.

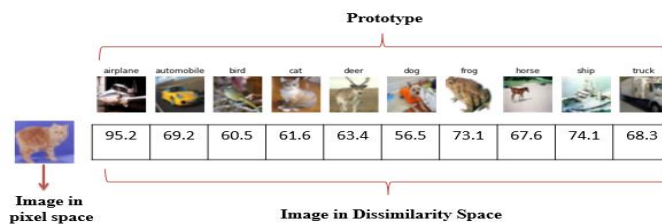
شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی یک طبقه‌بند غیرخطی، دارای ۴۶۷ لایه و با پیچیدگی بالا است و شبکه عصبی کانولوشنی متراکم باقیمانده از توالی ۳ یا ۴ بلوک ساخته شده که هر بلوک شامل ۱۶، ۲۰ یا ۲۴ لایه کانولوشنی است. هر لایه کانولوشنی از دو نرمال‌سازی دسته‌ای^{۵۸} به همراه دو کانولوشن ۲ بعدی تشکیل شده است. به منظور افزایش قدرت یادگیری و جلوگیری از پدیده محو‌گردان، اتصالات باقیمانده^{۵۹} بین تمام لایه‌ها قرار گرفته است. شبکه عصبی کانولوشنی متراکم باقیمانده به نسبت سایر شبکه‌های پردازش تصویر مانند Vgg16 و شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی، به دلیل تعداد پارامتر کمتر، دقت پایین‌تری دارد، اما پیچیدگی زمانی کمتر، باعث می‌شود در سخت‌افزارهای سبک‌تر مانند تلفن‌های هوشمند قابل استفاده باشد.

در شکل‌های (۱۰) و (۱۱) مشخصات لایه طبقه‌بند پرسپترون چند لایه^{۶۰} شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و مدل پایه استفاده شده شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و در شکل‌های (۱۲) و (۱۳) ساختار کلی شبکه عصبی کانولوشنی متراکم باقیمانده و مدل پایه استفاده شده شبکه عصبی کانولوشنی متراکم باقیمانده در ادامه آمده است.

Table (2): The most famous distance criteria

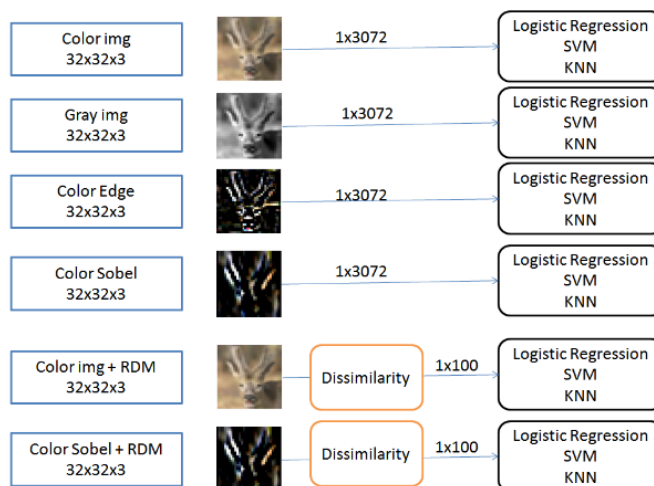
جدول (۲): معروف‌ترین معیارهای فاصله

فاصله	رابطه
فاصله اقلیدسی	$D = \sqrt{\sum_{j=1}^n x_{sj} - x_{tj}}^2$
فاصله بلوک شهری ^{۵۵}	$D = \sum_{j=1}^n x_{sj} - x_{tj} $
فاصله چیشف ^{۵۶}	$D = \max_j \{x_{sj} - x_{tj}\}$
فاصله کسینوسی ^{۵۷}	$D = 1 - \frac{\sum_{j=1}^n x_{sj} x_{tj}}{\sqrt{\sum_{j=1}^n x_{sj}^2} \sqrt{\sum_{j=1}^n x_{tj}^2}}$
فاصله همبستگی	$D = 1 - \frac{(x_s - \bar{x}_s)(x_t - \bar{x}_t)'}{\sqrt{(x_s - \bar{x}_s)(x_s - \bar{x}_s)'}} \sqrt{(x_t - \bar{x}_t)(x_t - \bar{x}_t)'}}$



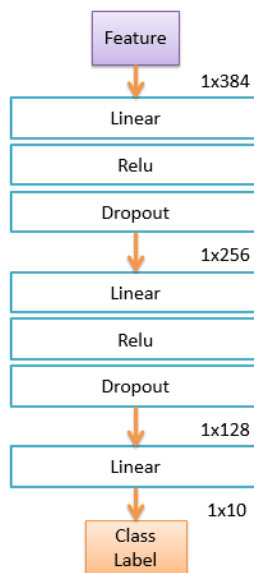
شکل (۸): فاصله تصاویر مرجع و تصاویر اصلی

Figure (8): The distance between prototype images and original images



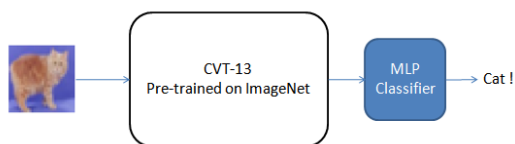
شکل (۹): نحوه اجرای فرآیند تعدادی از فازها

Figure (9): How to implement the process of a number of phases



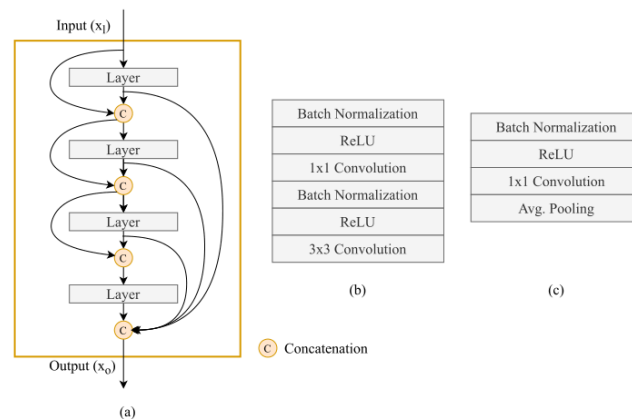
شکل (۱۰): ساختار پرسپترون چند لایه برای طبقه بندی [۲۸]

Figure (10): Multi-layer perceptron structure for classification [28]



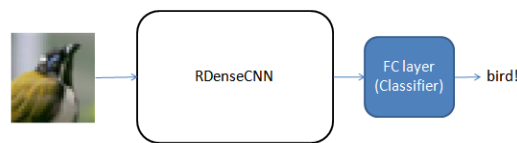
شکل (۱۱): مدل پایه استفاده شده شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی [۲۸]

Figure (11): The base model of CVT used [28]



شکل (۱۲): ساختار کلی شبکه عصبی کانولوشنی متراکم باقیمانده [۲۹]

Figure (12): General structure of the RDenseCNN network [29]

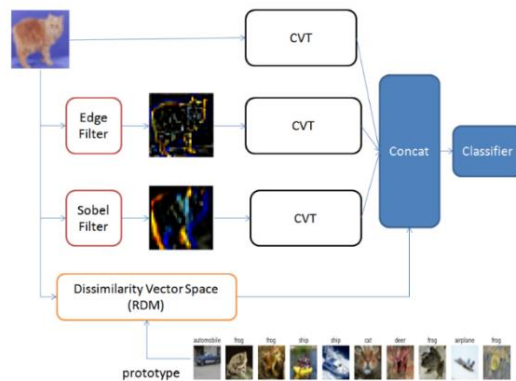


شکل (۱۳): مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده [۲۹]

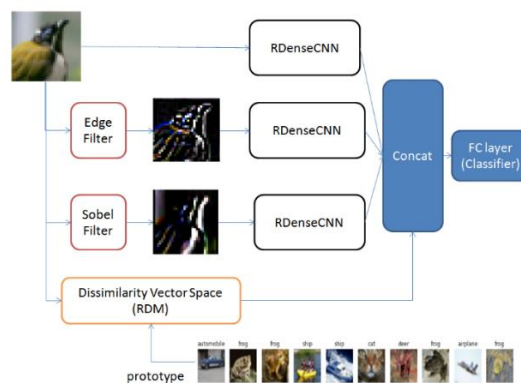
Figure (13): RDenseCNN base model [29]

در مدل چهار-جریان، ابتدا نسخه اصلی تصویر به شبکه پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و شبکه عصبی کانولوشنی متراکم باقیمانده اعمال می‌شود که از نظر دقت مشابه ویژگی‌های دو شبکه پایه کار می‌کند. سپس تصویر از فیلتر آشکارساز لبه، عبور می‌نماید و تصویر جدیدی حاصل می‌گردد، یعنی تصویر به فضای دیگر (دو-جریان) منتقل می‌شود که این فضای جدید ممکن است بتواند ویژگی‌های مناسب‌تری را به شبکه تزریق و به تفکیک کلاس‌ها کمک کند. در ادامه، تصویر از فیلتر سوبل عبور داده می‌شود و به فضای جدید (سه-جریان) انتقال می‌یابد و ویژگی‌های مطلوب‌تری برای تفکیک کلاس‌ها استخراج می‌گردد. سپس تصویر با استفاده از تصاویر مرجع، به فضای برداری عدم تشابه داده می‌شود که ۱۰۰ عدد در خروجی تولید می‌نماید و طبقه‌بند قادر است اعداد جدید اضافه شده را انتخاب کند. قبلاً فقط ویژگی‌های شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و شبکه عصبی کانولوشنی متراکم باقیمانده بوده، ولی اکنون ماتریس عدم تشابه بازنمایی به سه مدل شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و شبکه عصبی کانولوشنی متراکم باقیمانده اضافه شده است. به عبارت دیگر، تصویر به فضاهای مختلف منتقل می‌گردد تا در صورتی که شبکه از خود تصویر نتوانست اطلاعات مناسب را استخراج کند، از فیلترهای سوبل و آشکارساز لبه و ماتریس عدم تشابه بازنمایی داده‌های مناسب‌تری به دست آورد. در نهایت، این ویژگی‌ها همه به هم چسبانده^{۶۱} می‌شود، به صورت یک بردار واحد درمی‌آید و طبقه‌بند آموزش می‌بیند که از این بردار واحد، یاد بگیرد تصاویر را طبقه‌بندی کند. به عبارت دیگر، از لایه ماقبل آخر، خروجی گرفته شده و به عنوان ویژگی برای آن تصویر ذخیره می‌گردد و یک شبکه سه لایه هم به عنوان طبقه‌بند، برای پیش‌بینی کلاس‌ها روی ویژگی‌های استخراج شده، آموزش داده می‌شود. مدل چهار-جریان، مبتنی بر پیش پردازش عدم تشابه و استفاده از دو فیلتر سوبل و آشکارساز لبه به همراه تصویر اصلی به نحوی پیاده‌سازی می‌شود که در آن، لایه طبقه‌بند مشابه مدل قبل آموزش دیده و لایه الحاق برای چسباندن تمام چهار-جریان داده استفاده شده است که امکان انتخاب بردار ویژگی بیشتری را برای طبقه‌بند فراهم می‌سازد.

شکل‌های (۱۴) و (۱۵) نمای کلی مدل چهار-جریان را برای دو مدل شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و شبکه عصبی کانولوشنی متراکم باقیمانده نمایش می‌دهد. در تحلیل علت افزایش دقت، می‌توان به بالا رفتن تعداد پارامترهای مدل و به تبع آن ظرفیت شبکه در مدل چهار-جریان اشاره کرد و اینکه اضافه کردن عدم تشابه بازنمایی در جایی که طبقه‌بند نتواند با ویژگی‌های اصلی، تفکیک‌پذیری بالایی انجام دهد، می‌تواند تا حدودی با اضافه کردن ویژگی‌های خطی به تفکیک‌پذیری کلاس‌ها کمک کند.



شکل (۱۴): بلوک دیاگرام مدل چهار-جریان برای شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی
Figure (14): Block diagram of 4-stream model for CVT



شکل (۱۵): بلوک دیاگرام مدل چهار-جریان برای شبکه عصبی کانولوشنی متراکم باقیمانده
Figure (15): Block diagram of 4-stream model for RDenseCNN

۱-۲- تحلیل تولید داده‌های خطی تفکیک‌پذیر توسط نگاشت عدم تشابه

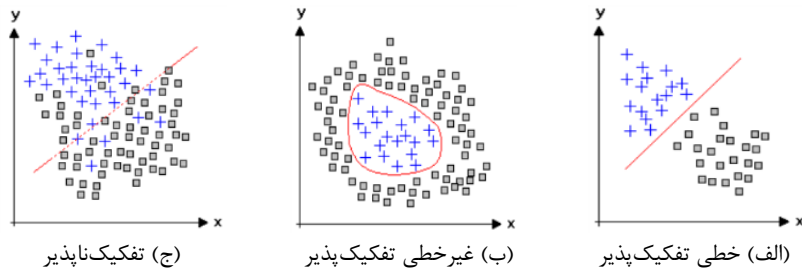
در اکثر داده‌های با منشأ طبیعی، پراکندگی نقاط به نحوی است که نقاط مجاور با احتمال بیشتری از یک کلاس یکسان هستند، یعنی کلاس‌ها در خوشه‌های محلی توزیع شده‌اند و احتمال اینکه دو نقطه همسایه از یک کلاس باشند، بسیار بیشتر از احتمال هم‌کلاس بودن دو نقطه با فاصله زیاد است. نگاشت عدم تشابه از این خاصیت بهره می‌برد. چنانچه نقطه مرجع داخل یک خوشه قرار گیرد، به‌طور میانگین فاصله D برای نقاط هم‌کلاس با مرجع، کمتر از نقاط غیر هم‌کلاس با مرجع خواهد بود. لذا استفاده از D به عنوان ویژگی، باعث ایجاد مقداری تفکیک‌پذیری خواهد شد. از آنجا که با انتخاب تعداد دلخواه K مرجع، می‌توان K ویژگی مستقل تولید کرد، در حالت حدی با افزایش K می‌توان به تفکیک‌پذیری ۱۰۰ درصد رسید. البته افزایش بیش از حد K از نظر کاربردی مطلوب نیست، به این دلیل که با نگاشت به فضای ابعاد بالاتر، علاوه بر افزایش سربار محاسباتی، افزایش تعداد پارامترهای آزاد مدل، باعث سرریز شدن و عدم تعمیم روی داده‌های تست می‌شود. بنابراین، انتخاب کمترین تعداد ویژگی با بیشترین قابلیت تفکیک‌پذیری موضوع مهمی خواهد بود. در ادامه تحلیل ریاضی قابلیت خطی‌سازی نگاشت عدم تشابه بررسی خواهد شد. در رابطه (۱) بردار نمونه‌ها نشان داده شده است:

$$\vec{X}^{(j)} = [x_1^{(j)}, x_2^{(j)}, \dots, x_N^{(j)}] \quad (1)$$

که در آن $X^{(i)}$ یک بردار N بعدی در فضای ویژگی‌های اصلی (N ویژگی)، $X_i^{(i)}$ المان i ام از $X^{(i)}$ (یک عدد اسکالر) و j از 1 الی M اندیس نمونه‌ها (نمونه‌های داده) (M مثال یا نمونه) است. در رابطه (۲) مجموعه مرجع‌ها مشخص شده است.

$$Q = \{R^{(1)}, R^{(2)}, \dots, R^{(k)}\} \quad \vec{R}^{(i)} \in R^N \quad (2)$$

که در آن Q مجموعه K مرجع، $R^{(i)}$ یک بردار N بعدی (مرجع i ام) است. در مرجع $[30]$ ، خطی تفکیک‌پذیر 62 در طبقه‌بندی باینری به‌صورت رابطه‌های (۳) و (۴) تعریف شده است.



شکل (۱۶): داده‌های خطی تفکیک‌پذیر، غیرخطی تفکیک‌پذیر و تفکیک‌ناپذیر

Figure (16): Linear separable, non-linear separable and non-separable data, a) Linear separable, b) non-linear separable, c) non-separable

$$H_1: \sum_{i=1}^N \theta_i x_i^{(j)} \geq 0 \quad \forall y_j = 1 \quad j = 1, \dots, M \quad (3)$$

$$H_2: \sum_{i=2}^N \theta_i x_i^{(j)} < 0 \quad \forall y_j = -1 \quad j = 1, \dots, M, y_j \in \{+1, -1\} \quad (4)$$

که در آن y_j برچسب نمونه j ام، $x_i^{(j)}$ المان i ام از $X^{(j)}$ و θ_i پارامترهای طبقه‌بند (پارامترهای صفحه جداکننده) است. همچنین در رابطه‌های (۳) و (۴)، هر دوی H_1 و H_2 باید برقرار باشند. به عبارت دیگر، همان‌طور که در شکل (۱۶) قابل مشاهده است، ترکیب خطی ویژگی‌ها برای مثال‌های مثبت یک طرف صفحه جداکننده و با همان ضرایب برای مثال‌های منفی در طرف دیگر قرار می‌گیرند. دو رابطه (۳) و (۴) را می‌توان مانند رابطه (۵) به فرم برداری نوشت:

$$H: y_j \theta^T X^{(j)} \geq 0 \quad \forall j = 1, \dots, M \quad (5)$$

که در آن θ^T ترانهاده بردار θ و $\theta^T X^{(j)}$ ضرب داخلی دو بردار θ و $X^{(j)}$ است. در رابطه (۶) تعریف فاصله اقلیدسی عنوان شده که در آن $D(X^{(j)}, R^{(k)})$ فاصله دو بردار $X^{(j)}$ و $R^{(k)}$ است:

$$D(\bar{X}^{(j)}, \bar{R}^{(k)}) = \sqrt{\sum_{i=1}^N (X_i^{(j)} - R_i^{(k)})^2} \quad (6)$$

در رابطه (۷) نگاهت به فضای عدم تشابه تعریف شده است که در آن تابع $D(X^{(j)}, R^{(k)})$ بردار X را با K مرجع، مقایسه و با محاسبه فاصله با هر مرجع به فضای جدید K بعدی نگاهت می‌کند.

$$D: R^N \rightarrow R^K \quad D(\bar{X}^{(j)}, Q) = [D(X^{(j)}, R^{(1)}), D(X^{(j)}, R^{(2)}), \dots, D(X^{(j)}, R^{(k)})] \quad (7)$$

۲-۲- الگوریتم ساخت فضای خطی تفکیک‌پذیر با استفاده از نگاهت عدم تشابه

۱- فرض بر این است که نقاط مجاور در داده‌ها، با احتمال بیشتری از کلاس یکسان باشند.
۲- با انتخاب مرجع R_1 ، بر طبق رابطه (۸)، یک ویژگی جدید تعریف شده که به ازای هر نقطه داده X یک کمیت اسکالر (فاصله) برمی‌گرداند.

$$\bar{D}_1^{(j)} = \text{Distance}(X^{(j)}, R^{(1)}) \quad \forall j \in [1, \dots, M] \quad (8)$$

۳- با استفاده از فرض ۱، به دلیل اینکه قسمت پایین تابع توزیع چگالی D_1 اکثراً متعلق به نقاطی از داده است که هم‌کلاس با مرجع R_1 هستند، بنابراین ویژگی D_1 ، برخی از نقاط داده را به‌طور خطی جدا می‌کند و رابطه (۹) قابل تعریف است.

$$\exists \theta_1 \quad y_j \theta_1 D_1^{(j)} \geq 0 \quad \text{for some } j \quad (9)$$

۴- با انتخاب مرجع دوم، ویژگی دیگری قابل تعریف است که توسط آن برخی دیگر از نقاط، به‌طور خطی تفکیک‌پذیر خواهند بود که بنابراین، رابطه‌های (۱۰) و (۱۱) قابل تعریف هستند.

$$\bar{D}_2^{(j)} = \text{Distance}(X^{(j)}, R^{(2)}) \quad \forall j \in [1, \dots, M] \quad (10)$$

$$\exists \theta_2 \quad y_j \theta_2 D_2^{(j)} \geq 0, \quad \text{for some } j \quad (11)$$

۵- مرحله قبل، مطابق رابطه (۱۲)، K بار تکرار می‌شود تا زمانی که تمام نقاط هم‌زمان به‌طور خطی تفکیک‌پذیر شوند.

$$y \sum_{i=1}^K \theta_i D_i^{(j)} \geq 0 \quad \forall j \in [1, \dots, M] \quad (12)$$

که به فرم برداری رابطه (۱۳) را می‌توان نوشت:

$$y_j < \theta, D^{(j)} \geq 0 \quad \forall j \quad (13)$$

$$D^{(j)} = [D_1(X^{(j)}, R^{(1)}), \dots, D_K(X^{(j)}, R^{(K)})]$$

که در آن:

$$\theta = [\theta_1, \theta_2, \dots, \theta_K]$$

و <.,.> بیانگر ضرب داخلی دو بردار است. در نتیجه نگاشت به فضای ابعاد بالاتر، در صورتی که K به حد کافی بزرگ باشد، منجر به ایجاد فضای خطی تفکیک‌پذیر خواهد شد. لازم به ذکر است این نگاشت از نظر کاربردی زمانی مطلوب خواهد بود که K حتی‌الامکان کوچک باشد و داده‌ها به فضای ابعاد پایین‌تری نگاشت شوند، زیرا طبقه‌بندی با تعداد کمتر پارامتر آزاد و تعداد ویژگی کمتر، باعث تعمیم بهتر روی داده‌های تست خواهد شد. بنابراین، نحوه انتخاب مرجع‌ها و تعداد آن‌ها، در میزان تعمیم مدل نهایی تأثیرگذار خواهد بود. از دیتاست سیفار-۱۰ برای انجام شبیه‌سازی‌ها استفاده نموده که این دیتاست اغلب برای ارزیابی راه‌حل‌های شبکه عصبی کانولوشنی استفاده شده است.

برای ارزیابی عملکرد رویکرد پیشنهادی، از گوگل کولب استفاده شده است. گوگل کولب یک سند اجرایی، تعاملی و مبتنی بر ژوپیتر با سخت‌افزار رایگان بوده که به گوگل درایو متصل است و می‌توان با دیگران برای کدنویسی همزمان شریک شد. نوت-بوک‌های کولب کدها را روی سرورهای ابری گوگل اجرا می‌کنند. بنابراین، دیتاست‌ها را می‌توان بر روی سرورهای ابری وارد کرد و به راحتی می‌توان یک طبقه‌بندی‌کننده تصویر را آموزش داد. چندین چارچوب برای توسعه برنامه‌های یادگیری عمیق وجود دارد، از جمله تنسورفلو (عرضه شده توسط گوگل در سال ۲۰۱۵)، پایتورچ (عرضه شده توسط فیس‌بوک در سال ۲۰۱۶) و کراس که نسخه جدید تنسورفلو است. در این مدل از چارچوب پایتورچ استفاده شده زیرا کار با آن در مقایسه با سایر چارچوب‌ها آسان‌تر است. پایتورچ یک کتابخانه یادگیری ماشین منبع باز مبتنی بر کتابخانه تورچ است که برای بینایی کامپیوتر و برنامه‌های پردازش طبیعی استفاده می‌شود و توسط آزمایشگاه تحقیقاتی هوش مصنوعی فیس‌بوک^{۶۵} در سال ۲۰۱۶ توسعه یافته است [۳۱].

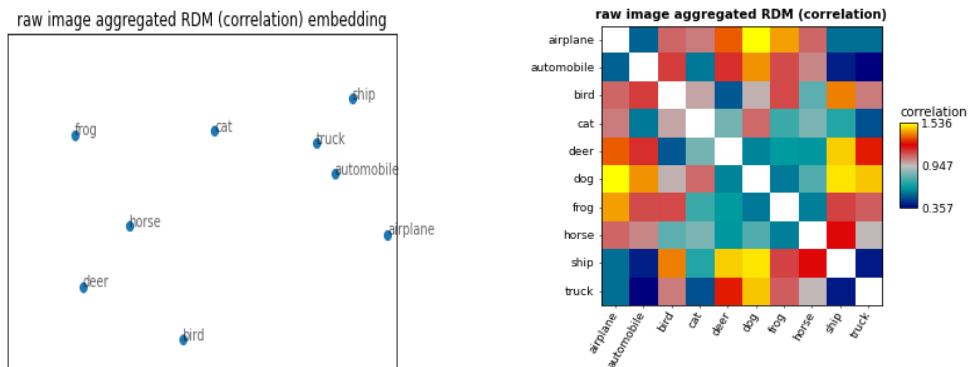
۳- نتایج شبیه‌سازی مدل چهار-جریان

شکل (۱۷) ماتریس عدم تشابه بازنمایی حاصل از تصاویر سیفار-۱۰ را نشان می‌دهد که نزدیک بودن کلاس اتومبیل به کامیون و کشتی مشهود است. از ماتریس ایجاد شده، برای تحلیل سیگنال‌های مغزی استفاده می‌شود. به این صورت که فعالیت مغزی فرد در آزمایشگاه در زمان رؤیت تصاویر مختلف ضبط شده و سپس با تهیه ماتریس عدم تشابه بازنمایی این سیگنال‌ها و تعبیه کردن آن در فضای دو بعدی می‌توان دریافت که فرد کدام تصاویر را به یکدیگر از نظر ذهنی نزدیک می‌داند. تحلیل ترس درونی فرد از یک پدیده، از طریق نزدیکی موقعیت مکانی تصویر مربوطه با سایر تصاویر خشن یا رعب‌آور قابل مشاهده است و یا رصد این که قبل و بعد از آموزش یک مفهوم مثلاً ترک دخانیات، فضای ذهنی فرد نسبت به تصویر سیگار به چه صورت تغییر کرده است، از کاربردهای دیگر تحلیل ماتریس عدم تشابه بازنمایی است.

۳-۱- بررسی تأثیر فضای برداری عدم تشابه بر خطی‌سازی فضای داده‌های مصنوعی

در شکل (۱۸)، تعداد ۱۰۰۰ نقطه در فضای دوبعدی به‌طور تصادفی تولید گردیده و به صورت غیرخطی به دو کلاس صفر و ۱ تقسیم شده‌اند به نحوی که طبقه‌بندهای خطی قادر به جداکردن نقاط نباشند. سپس نقاط به دو دسته آموزش و تست تقسیم شده‌اند و طبقه‌بندهای استاندارد، روی این داده تست شدند که نتایج در ادامه بیان می‌شوند. همان‌طور که در شکل (۱۹) نشان داده شده است، طبقه‌بندهای خطی مانند بردار پشتیبان خطی^{۶۶} و رگرسیون لجستیک قادر به تفکیک داده‌های غیرخطی فوق نیستند، زیرا میزان دقت دو طبقه‌بند خطی اشاره شده در مقایسه با سایر طبقه‌بندها کمتر است. در مرحله بعد، داده‌ها به فضای برداری عدم تشابه نگاشت شده است و سپس همین طبقه‌بندها در فضای جدید تست می‌شوند. در شکل (۲۰)، نقاط آبی

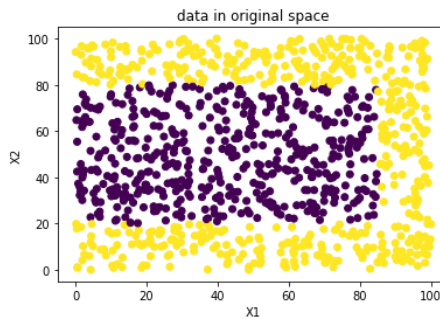
(پیکسل ۱) و قرمز (پیکسل ۲) مرجع با موقعیت x_1 و x_2 انتخاب شده و موقعیت نقاط در فضای جدید ترسیم شده است. واضح است که نقاط در فضای جدید، خوشه‌بندی مناسب‌تری برای طبقه‌بندی خطی دارند. سپس ۱۰ مرجع، به‌طور تصادفی برای هر کلاس انتخاب شده و بعد از انتقال نقاط به فضای برداری عدم تشابه، طبقه‌بندی استاندارد روی نقاط جدید اجرا شده‌اند. همان‌طور که در شکل (۲۱) مشاهده می‌شود، طبقه‌بندی خطی در فضای جدید توانسته‌اند بهتر عمل کنند. زیرا میزان دقت دو طبقه‌بندی خطی در مقایسه با مرحله قبلی، افزایش داشته است. برای مقایسه بهتر، عملکرد طبقه‌بندیها، قبل و بعد از فضای برداری عدم تشابه در شکل (۲۲) و جدول‌های (۳) و (۴) نشان داده شده است. مشخص است که روی داده‌های فوق، انتقال نقاط به فضای برداری عدم تشابه، باعث افزایش قابل توجه دقت طبقه‌بندی بردار پشتیبان خطی و رگرسیون لجستیک و بهبود مقادیر امتیاز F1، صحت و فراخوانی شده و در عین حال تأثیر چندانی بر طبقه‌بندیهای غیرخطی نداشته است.



شکل (الف) ماتریس عدم تشابه بازنمایی حاصل از تصاویر سیفار-۱۰ (ب) میزان پراکندگی مدل‌ها

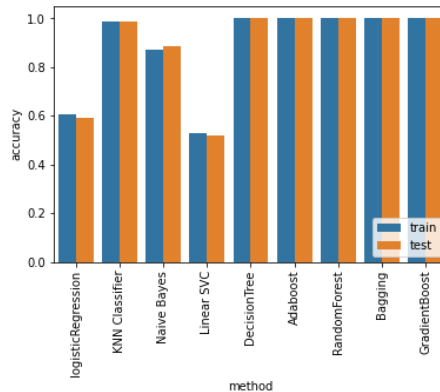
شکل (۱۷): ماتریس عدم تشابه بازنمایی حاصل از تصاویر سیفار-۱۰ و میزان پراکندگی مدل‌ها

Figure (17): RDM resulting from Cifar10 images and the amount of dispersion of the models, a) RDM resulting from Cifar10 images, b) Amount of dispersion of the models



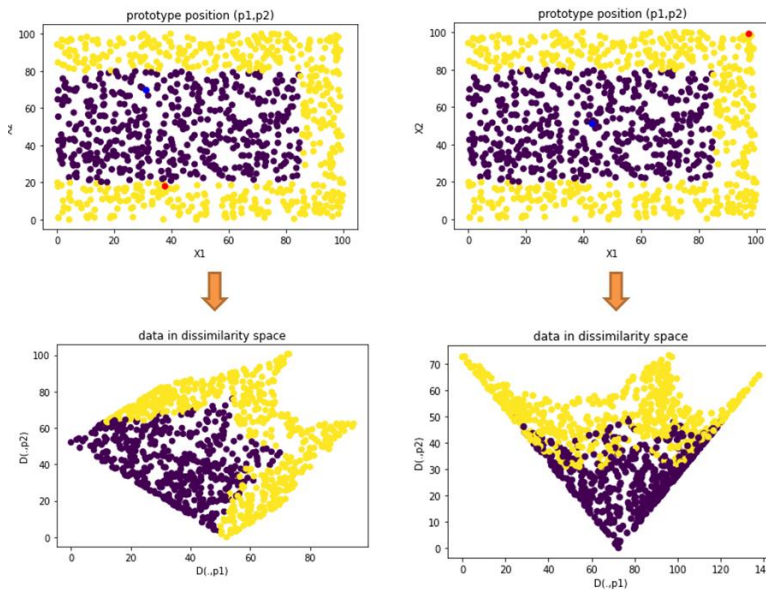
شکل (۱۸): تولید ۱۰۰۰ نقطه در فضای دوبعدی

Figure (18): Production of 1000 points in two-dimensional space

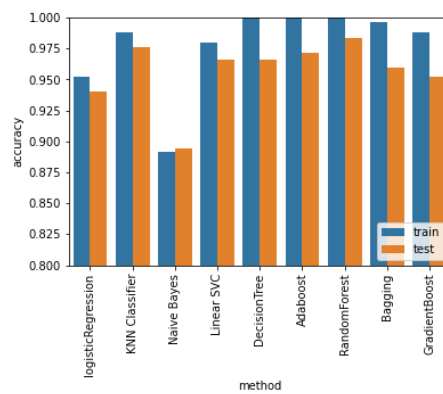


شکل (۱۹): عملکرد طبقه‌بندیها قبل از فضای برداری عدم تشابه

Figure (19): Performance of classifiers before dissimilarity vector space

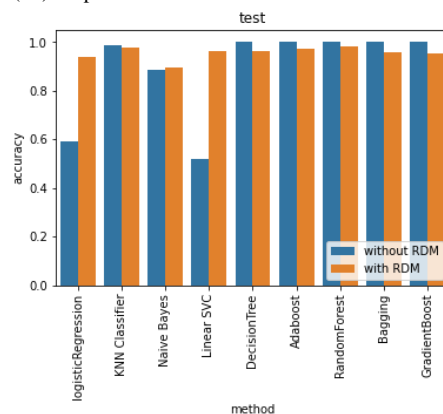


شکل (۲۰): مرجع انتخاب شده نقاط با آبی و قرمز و موقعیت نقاط در فضای جدید
 Figure (20): selected prototype of points with blue and red and the position of the points in the new space



شکل (۲۱): اجرای طبقه‌بندهای استاندارد روی نقاط جدید

Figure (21): Implementation of standard classifications on new points



شکل (۲۲): عملکرد طبقه‌بندها قبل و بعد از فضای برداری عدم تشابه
 Figure (22): Performance of classifiers before and after dissimilarity vector space

۲-۳- بررسی تأثیر فضای برداری عدم تشابه بر خطی‌سازی فضای تصاویر سیفار-۱۰
 در شکل (۲۳)، نمونه تصاویر سیفار-۱۰ و فیلترهای اعمال شده، نشان داده شده است. شکل (۲۴)، ۱۰ مرجع انتخاب شده از کلاس آهو، به تفکیک هر فیلتر نشان داده شده است.

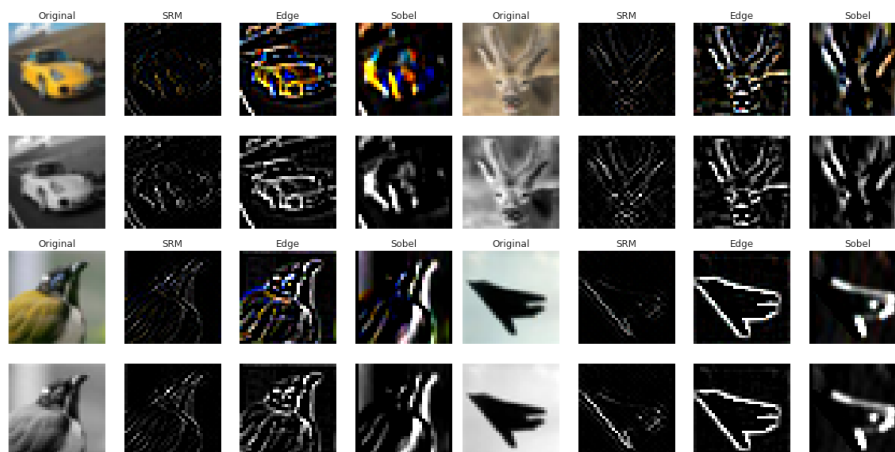
در شکل (۲۵) نتایج آزمایش نشان می‌دهد در طبقه‌بندی‌های خطی، زمانی که هر یک از تصاویر، اعم از تصویر رنگی یا تصویر سیاه-سفید یا تصویر عبور کرده از فیلتر سوبل و آشکارساز لبه، بدون اعمال ماتریس عدم تشابه به طبقه‌بندی خطی داده شوند، دقت پایین‌تر از زمانی است که همین تصویر با اعمال ماتریس عدم تشابه به فضای خطی نگاشت شود. به عبارت دیگر انتقال تصاویر به فضای برداری عدم تشابه موجب افزایش دقت در طبقه‌بندی‌های خطی می‌شود. اما برای طبقه-بندی‌های غیرخطی، نتایج متفاوت است بنحوی که گاهی باعث کاهش اندک دقت و گاهی باعث افزایش اندک دقت می‌شود.

Table (3): Accuracy of different classifications before mapping to dissimilarity vector space
جدول (۳): دقت طبقه‌بندی‌های مختلف قبل از نگاشت به فضای برداری عدم تشابه

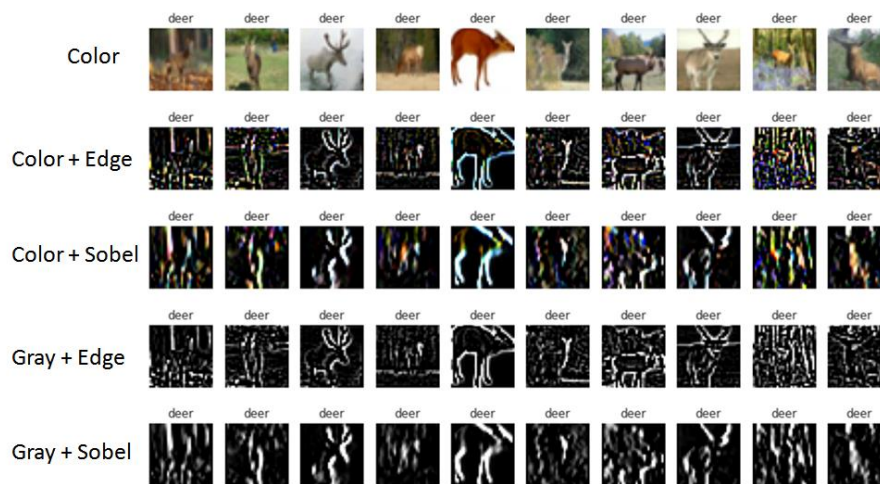
ردیف	انواع طبقه‌بند	دقت	امتیاز F1	صحت	فراخوانی
۰	رگرسیون لجستیک	۰/۵۹۲	۰/۵۹۰۷۱۶	۰/۵۹۱۱۵۹	۰/۵۹۰۷۵۰
۱	K- نزدیکترین همسایگی	۰/۹۸۶	۰/۹۸۵۹۸۴	۰/۹۸۵۹۲۲	۰/۹۸۶۰۵۴
۲	بیز ساده لوح	۰/۸۸۴	۰/۸۸۱۹۴۳	۰/۹۰۰۷۲۷	۰/۸۸۰۳۸۹
۳	بردار پشتیبان خطی	۰/۵۱۸	۰/۳۴۱۲۳۸	۰/۲۵۹۰۰۰	۰/۵۰۰۰۰۰
۴	درخت تصمیم	۱	۱	۱	۱
۵	آداپوست	۱	۱	۱	۱
۶	جنگل تصادفی	۱	۱	۱	۱
۷	بگینگ	۱	۱	۱	۱
۸	تقویت‌گرادیان	۱	۱	۱	۱

Table (4): Accuracy of different classifications before mapping to dissimilarity vector space
جدول (۴): دقت طبقه‌بندی‌های مختلف بعد از نگاشت به فضای برداری عدم تشابه

ردیف	انواع طبقه‌بند	دقت	امتیاز F1	صحت	فراخوانی
۰	رگرسیون لجستیک	۰/۹۴۰	۰/۹۳۹۹۳۸	۰/۹۳۹۸۴۱	۰/۹۴۰۰۶۶
۱	K- نزدیکترین همسایگی	۰/۹۷۶	۰/۹۷۵۹۵۳	۰/۹۷۶۳۵۲	۰/۹۷۵۶۸۰
۲	بیز ساده لوح	۰/۸۹۴	۰/۸۹۳۶۴۲	۰/۸۹۵۰۸۹	۰/۸۹۳۰۶۹
۳	بردار پشتیبان خطی	۰/۹۶۶	۰/۹۶۵۹۵۱	۰/۹۶۶۰۲۶	۰/۹۶۵۸۸۴
۴	درخت تصمیم	۰/۹۶۶	۰/۹۶۵۹۶۱	۰/۹۶۵۹۰۱	۰/۹۶۶۰۲۸
۵	آداپوست	۰/۹۷۲	۰/۹۷۱۹۷۱	۰/۹۷۱۸۶۶	۰/۹۷۲۱۰۸
۶	جنگل تصادفی	۰/۹۸۴	۰/۹۸۳۹۶۹	۰/۹۸۴۳۷۴	۰/۹۸۳۶۹۱
۷	بگینگ	۰/۹۶۰	۰/۹۵۹۹۲۲	۰/۹۶۰۳۰۹	۰/۹۵۹۶۶۰
۸	تقویت‌گرادیان	۰/۹۵۲	۰/۹۵۱۹۷۲	۰/۹۵۱۸۴۹	۰/۹۵۲۳۷۰

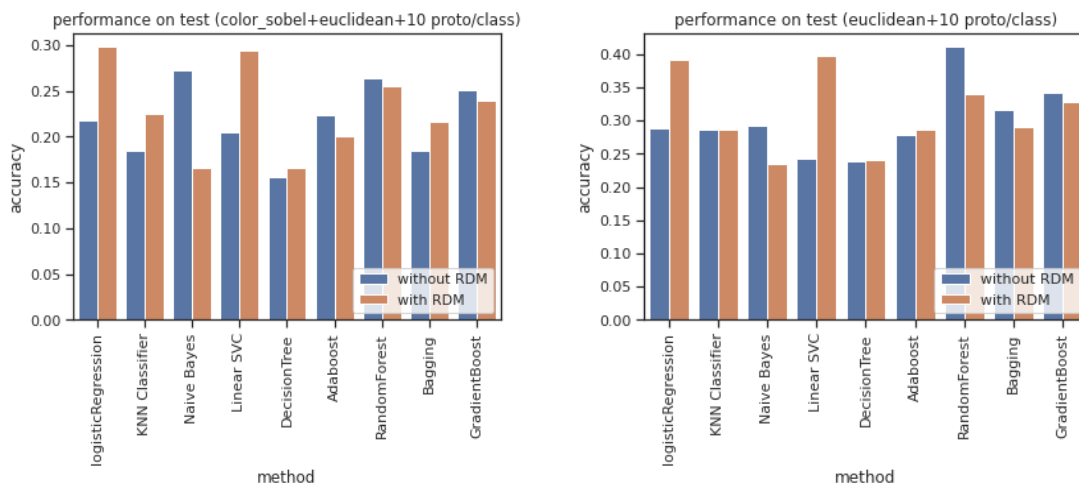


شکل (۲۳): نمونه تصاویر سیفار-۱۰ و فیلترهای اعمال شده
Figure (23): Examples of cifar10 images and applied filters



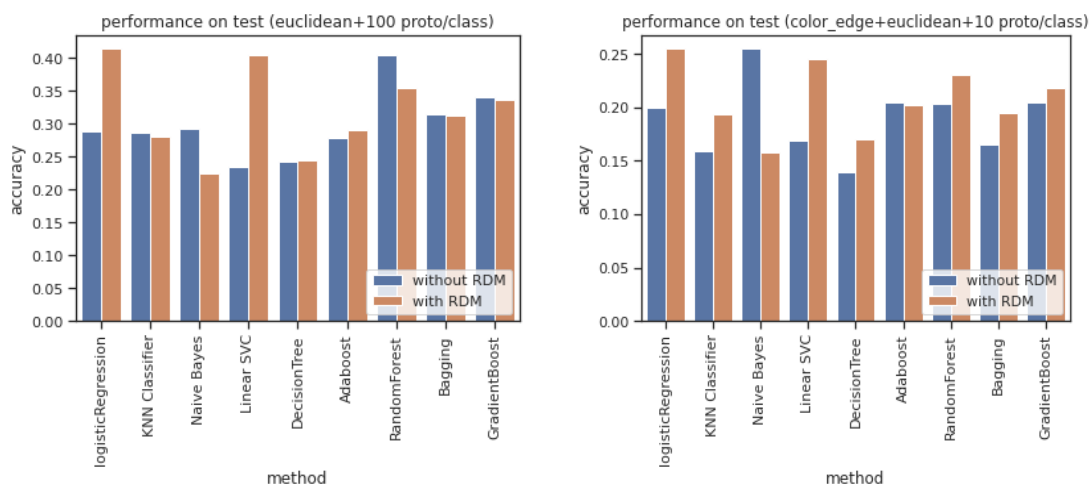
شکل (۲۴): ۱۰ مرجع انتخاب شده از کلاس آهو، به تفکیک هر فیلتر

Figure (24): 10 prototypes selected from the deer class, separated by each filter



(ب) تصویر رنگی با عبور از فیلتر سوبل و اعمال ماتریس عدم تشابه با ۱۰ مرجع

(الف) ماتریس عدم تشابه با روش اقلیدسی با ۱۰ مرجع



(ج) تصویر رنگی با عبور از آشکارساز لبه و اعمال ماتریس عدم تشابه با روش اقلیدسی با ۱۰۰ مرجع

شکل (۲۵): میزان دقت در طبقه‌بندی‌های مختلف

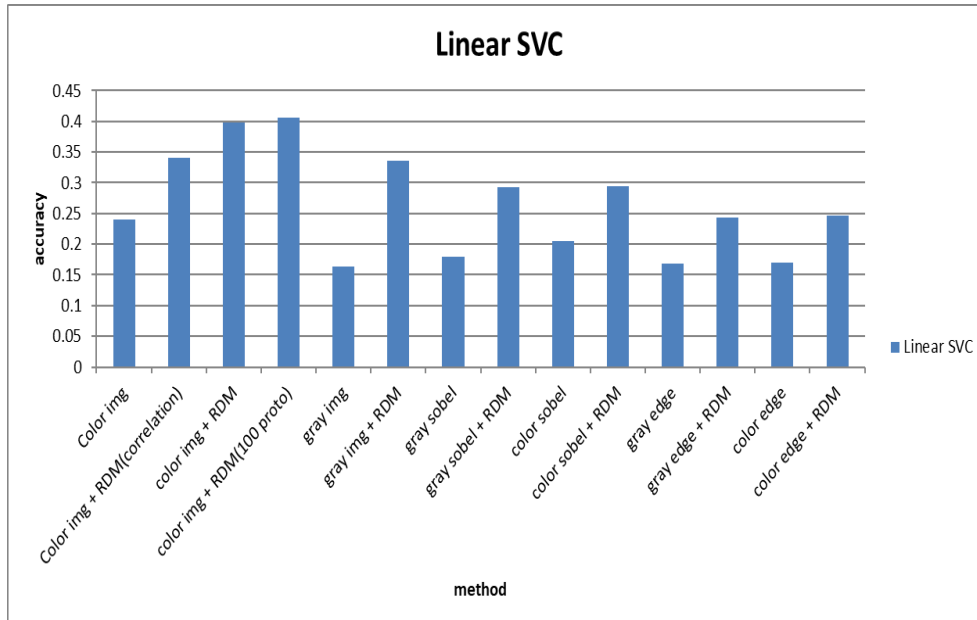
Figure (25): The level of accuracy in different classifications, a) RDM with euclidean method and 10 prototypes, b) Color image by passing the Sobel filter and applying the RDM with 10 prototypes, c) color image by passing the edge detector and applying a RDM with 10 prototypes, d) The RDM by euclidean method with 100 prototypes

دلیل این امر این است که طبقه‌بندهای غیرخطی، لزوماً با نگاشت داده به فضای خطی منتفع نمی‌شوند، بلکه ممکن است در فضای ویژگی‌های اصلی داده بهتر بتواند عمل کند و زمانی که نگاشت عدم تشابه انجام می‌گیرد، به خاطر کاهش ابعاد (در اینجا از ۳۰۷۲ به ۱۰۰) بخشی از اطلاعات حذف شوند و باعث کاهش دقت شود. لازم به ذکر است چنانچه تعداد بیشتری مرجع انتخاب گردد و نگاشت به فضای ابعاد بالاتر انجام گیرد همواره افت اطلاعات کمتری رخ داده و باعث بهبود اندک در دقت می‌شود. اما این افزایش دقت به بهای افزایش پیچیدگی محاسباتی و زمان پردازش خواهد بود.

همان‌طور که در شکل (۲۵) نشان داده شده است تحلیل طبقه‌بندهای مختلف بر روی داده‌های تست انجام شده است که نمودار (الف) ماتریس عدم تشابه با روش اقلیدسی و ۱۰ مرجع در حالت قبل و بعد از اعمال ماتریس عدم تشابه، نمودار (ب) تصویر رنگی با عبور از فیلتر سوبل و اعمال ماتریس عدم تشابه با ۱۰ مرجع، نمودار (ج) تصویر رنگی با عبور از آشکارساز لبه با اعمال ماتریس عدم تشابه با ۱۰ مرجع و نمودار (د) ماتریس عدم تشابه با روش اقلیدسی و ۱۰۰ مرجع در حالت قبل و بعد از اعمال ماتریس عدم تشابه، است. تحت این شرایط مشاهده می‌شود که دقت برای طبقه‌بندهای خطی بردار پشتیبان و رگرسیون لجستیک در حالت تبدیل، افزایش یافته است.

الف- نتایج تفکیکی برای طبقه‌بندهای خطی: در شکل‌های (۲۶) و (۲۷)، به ترتیب میزان دقت اجرای طبقه‌بندهای بردار پشتیبان خطی و رگرسیون لجستیک در داده‌های مختلف نشان داده شده که به عنوان نمونه در شکل (۲۶)، مقایسه تصاویر رنگی و ماتریس بازنمایی عدم تشابه و ۱۰۰ مرجع در ستون چهارم با تصاویر رنگی ستون سوم و ۱۰ مرجع، نشان می‌دهد که با افزایش مرجع میزان دقت افزایش یافته است، یعنی دقت با اعمال ماتریس عدم تشابه بازنمایی بیشتر از دیتای متناظر قبل خود است.

ب- نتایج تفکیکی برای طبقه‌بندهای غیرخطی: شکل‌های (۲۸) الی (۳۱) به ترتیب میزان دقت اجرای طبقه‌بندهای K-نزدیکترین همسایگی، جنگل تصادفی، تقویت گرادیان و درخت تصمیم در حالت‌های مختلف را نشان داده شده است که مشاهده می‌شود دقت طبقه‌بند با اعمال ماتریس عدم تشابه بازنمایی بیشتر یا کمتر از دیتای متناظر قبل خود است.



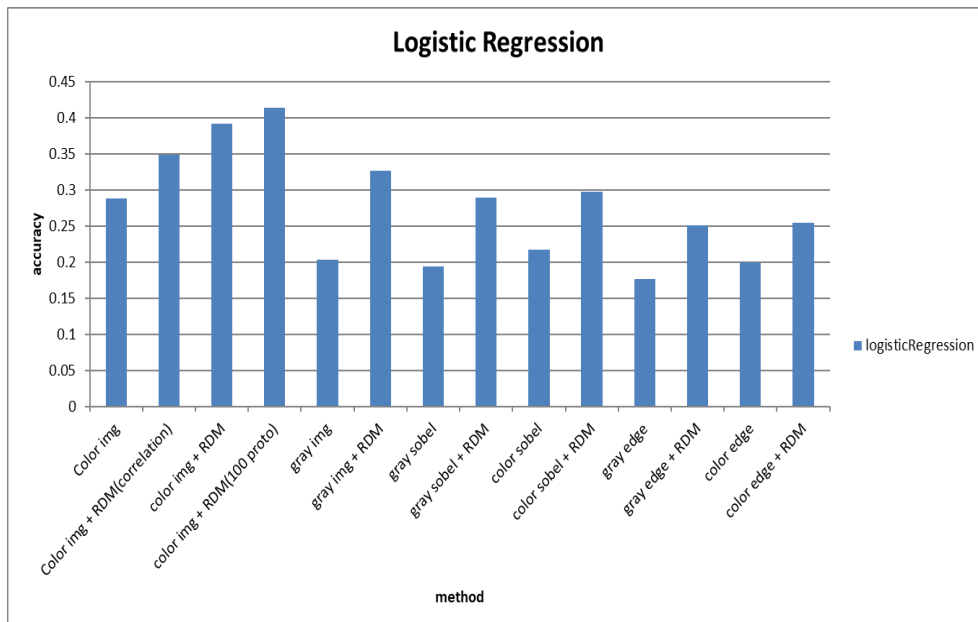
شکل (۲۶): میزان دقت اجرای طبقه‌بند بردار پشتیبان خطی در حالت‌های مختلف

Figure (26): Accuracy level of linear svc classifier in different modes

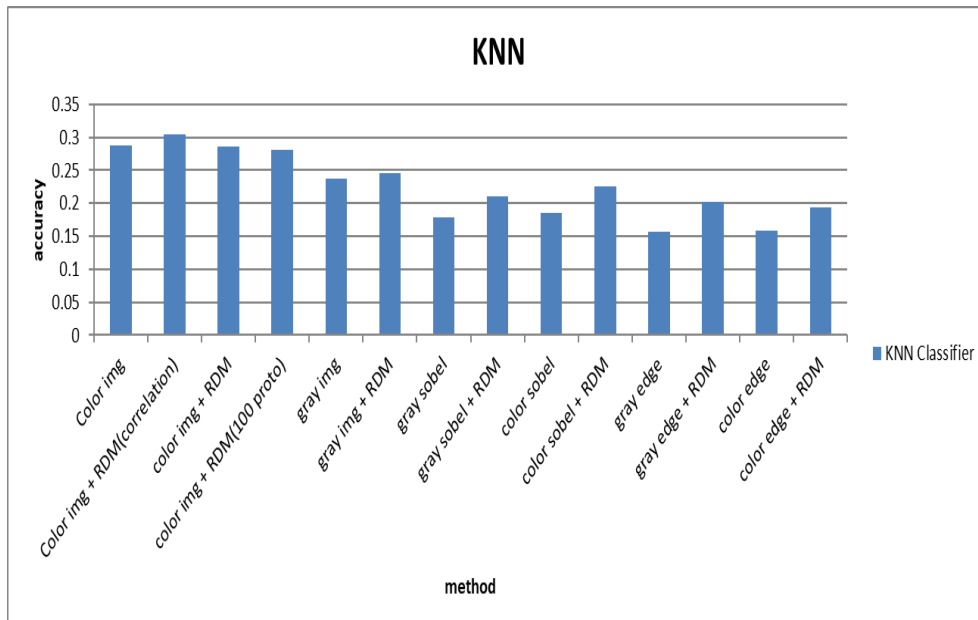
۳-۳- مدل یادگیری عمیق مبتنی بر پیش پردازش با عدم تشابه بر پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی میکروسافت

در جدول (۵) و شکل (۳۲)، به‌ازاء ۴۵۰۰۰ داده آموزش، ۱۰۰۰۰ داده تست و ۵۰۰۰ داده اعتبارسنجی، نتایج و دقت به‌دست آمده بعد از آموزش لایه پرسپترون چند لایه روی داده‌های سیفار-۱۰ قابل مشاهده است. همان‌طور که در شکل (۲۴) ملاحظه

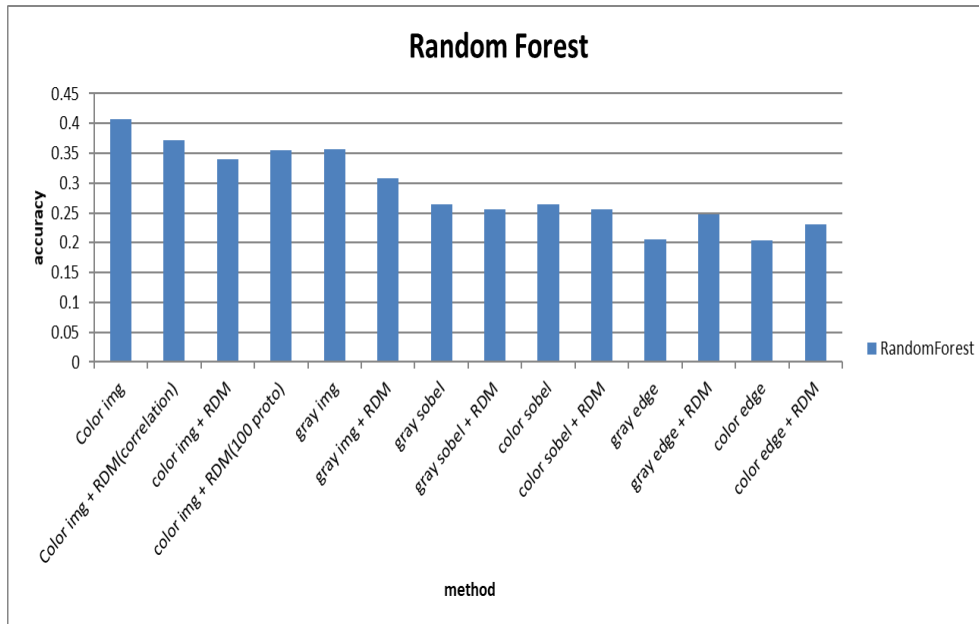
می‌شود با افزایش تعداد دوره‌ها احتمال بیش برآزش ۶۷ وجود خواهد داشت. در جدول (۶) و شکل (۳۳) نتایج و دقت به‌دست آمده بعد از آموزش لایه پرسپترون چند لایه روی داده‌های سیفار-۱۰ مدل چهار-جریان قابل مشاهده است. مقایسه دو مدل نشان می‌دهد مدل چهار-جریان نزدیک به یک درصد افزایش دقت داشته است.



شکل (۲۷): میزان دقت اجرای طبقه‌بند رگرسیون لجستیک در حالت‌های مختلف
Figure (27): The accuracy of logistic regression classifier implementation in different modes

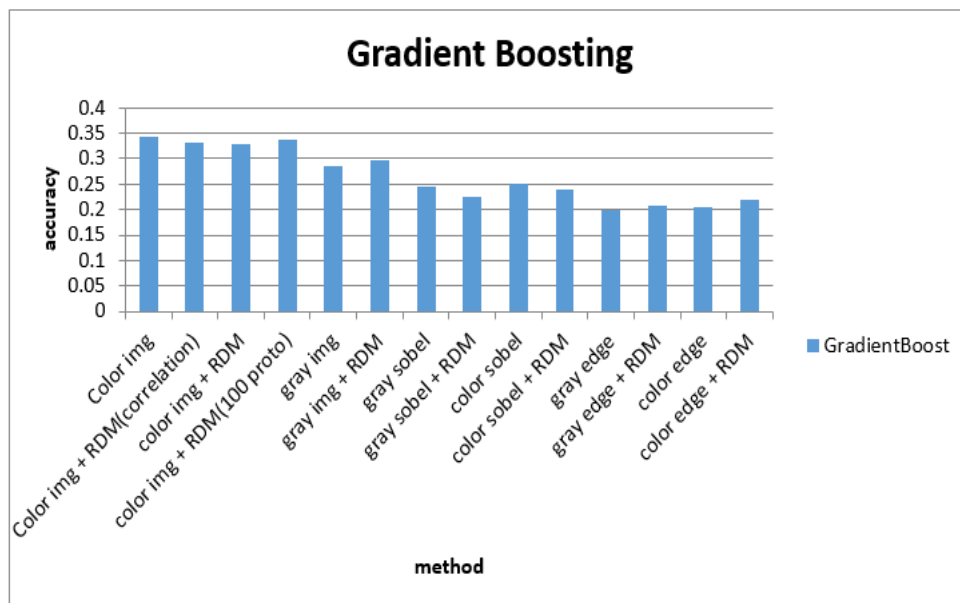


شکل (۲۸): میزان دقت اجرای طبقه‌بند K - نزدیکترین همسایگی در حالت‌های مختلف
Figure (28): The accuracy of KNN classifier implementation in different modes



شکل (۲۹): میزان دقت اجرای طبقه‌بند جنگل تصادفی در حالت‌های مختلف

Figure (29): Accuracy of Random Forest classifier in different modes



شکل (۳۰): میزان دقت اجرای طبقه‌بند تقویت گرادیان در حالت‌های مختلف

Figure (30): Accuracy of GradientBoost classifier in different modes

Table (5): The accuracy obtained after training the MLP layer of the CVT base model

جدول (۵): دقت به دست آمده بعد از آموزش لایه پرسپترون چند لایه مدل پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی

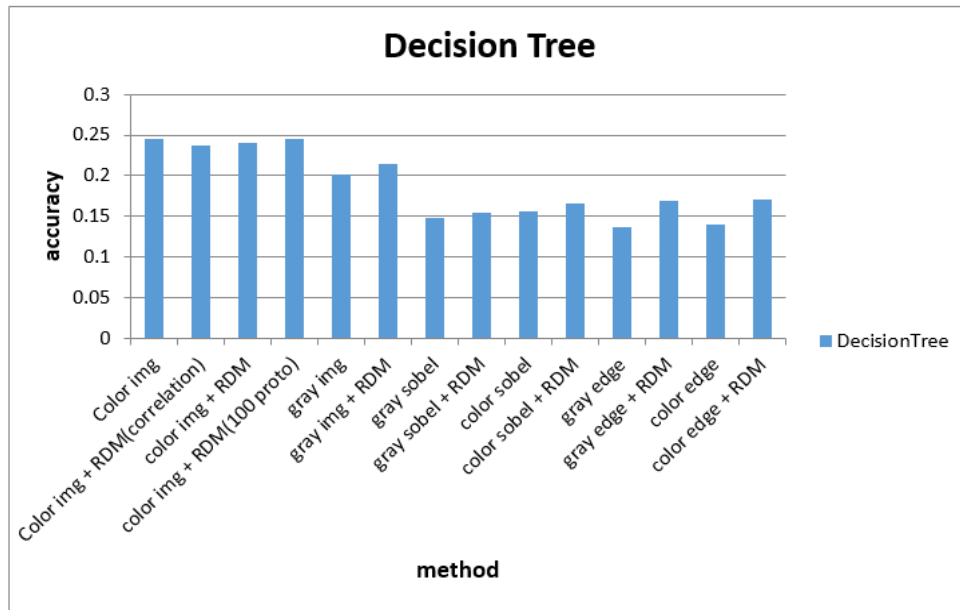
تست	اعتبارسنجی	آموزش	دقت
۰/۹۲۰۹	۰/۹۲۲۲	۰/۹۶۰۵	مدل پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی

Table (6): The accuracy obtained after training the MLP layer on Cifar10 data of the CVT-13 4Stream model

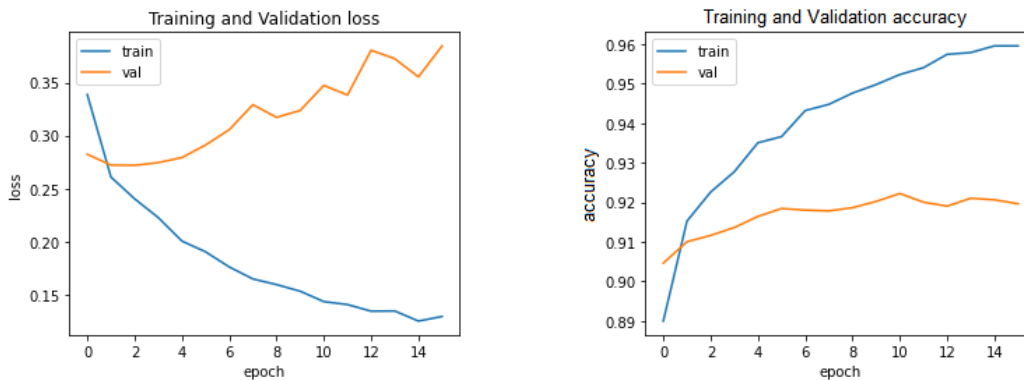
جدول (۶): دقت به دست آمده بعد از آموزش لایه پرسپترون چند لایه روی داده‌های سیفار-۱۰ مدل شبکه‌های عصبی کانولوشنی با

مبدل‌های بینایی چهار-جریان

تست	اعتبارسنجی	آموزش	دقت
۰/۹۲۹۲	۰/۹۲۹۶	۰/۹۵۷۵	مدل شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی چهار-جریان



شکل (۳۱): میزان دقت اجرای طبقه‌بند درخت تصمیم در حالت‌های مختلف
Figure (31): Accuracy of DecisionTree classifier execution in different modes

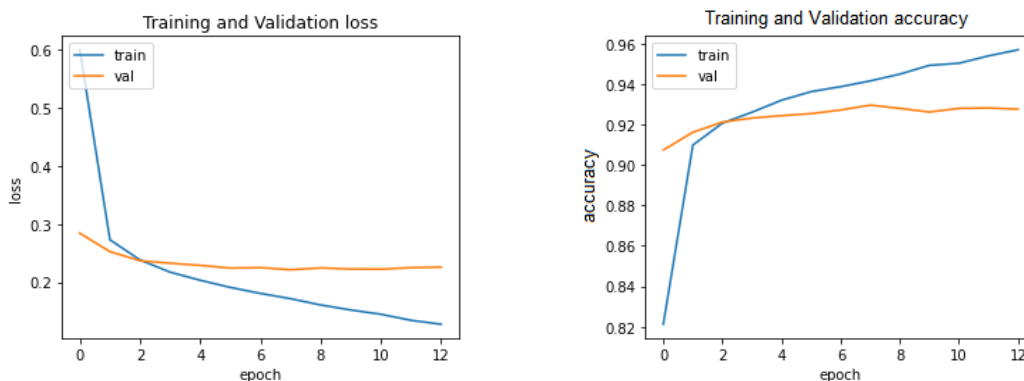


(ب) زیان آموزش و اعتبارسنجی

(الف) دقت آموزش و اعتبارسنجی

شکل (۳۲): فرایند آموزش مدل پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی

Figure (32): The process of training the basic model of CVT, a) Training and validation accuracy, b) Training and validation loss



(ب) زیان آموزش و اعتبارسنجی

(الف) دقت آموزش و اعتبارسنجی

شکل (۳۳): فرایند آموزش مدل شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی چهار-جریان

Figure (33): The process of training the model of CVT-13 4Stream, a) Training and validation accuracy, b) Training and validation loss

از آنجا که شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی یک طبقه‌بند غیرخطی، دارای ۴۶۷ لایه و با پیچیدگی بالا است عملیات استخراج ویژگی و خطی‌سازی داده‌ها به منظور تفکیک بهتر کلاس‌ها در خود شبکه وجود دارد. لذا همان‌طور که در بخش‌های قبل ملاحظه شد، اضافه کردن بازنمایی عدم تشابه تنها در مواردی منجر به بهبود قابل توجه بازده می‌شود که طبقه‌بند پایه، خطی بوده و ناتوان از جداسازی داده‌های غیرخطی باشد.

در غیر این صورت، اضافه کردن ویژگی‌های جدید تنها اندکی ممکن است منجر به بهبود شود و با افزایش تعداد دوره‌ها احتمال وقوع بیش برآزش کاهش می‌یابد. برای مقایسه بهتر دو مدل پایه و چهار-جریان، جدول (۷) و شکل (۳۴) در کنار یکدیگر رسم شده‌اند.

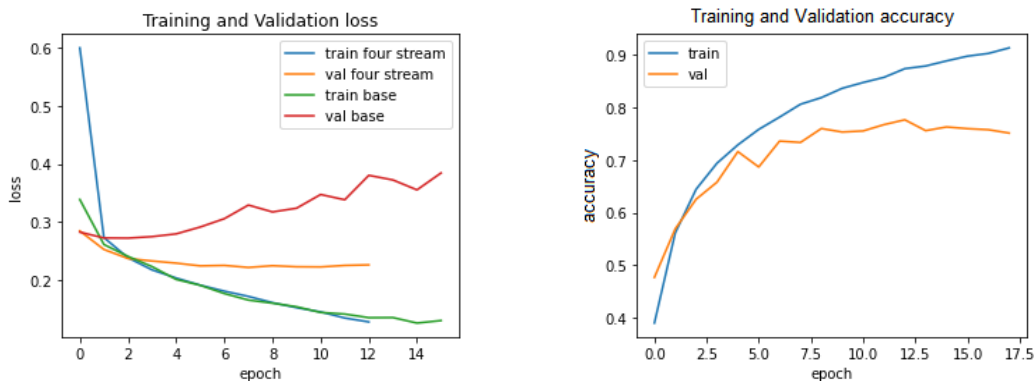
۳-۴- مدل یادگیری عمیق مبتنی بر پیش‌پردازش با عدم تشابه بر پایه شبکه عصبی کانولوشنی متراکم باقیمانده
در مدل پایه، شبکه عصبی کانولوشنی متراکم باقیمانده مطابق مرجع [۲۹]، پیاده‌سازی شده و کل شبکه با داده‌های سیفار-۱۰ از صفر آموزش دیده است. جدول (۸) و شکل (۳۵) دقت مدل را بعد از ۱۷ دوره آموزش روی سیفار-۱۰ نشان می‌دهد. کل شبکه روی سیفار-۱۰ به میزان ۱۴ دوره آموزش داده شد. نتایج به دست آمده از آموزش مدل در جدول (۹) و شکل (۳۶) به این شرح است. مقایسه دو مدل نشان می‌دهد که مدل چهار-جریان به میزان ۳/۲ درصد افزایش دقت داشته است. در تحلیل علت افزایش دقت، می‌توان به بالا رفتن تعداد پارامترهای مدل و به تبع آن ظرفیت شبکه در مدل چهار-جریان اشاره کرد و اینکه اضافه کردن بازنمایی عدم تشابه در جایی که طبقه‌بند نتواند با ویژگی‌های اصلی، تفکیک‌پذیری بالایی انجام دهد، می‌تواند تا حدودی با اضافه کردن ویژگی‌های خطی به تفکیک‌پذیری کلاس‌ها کمک کند. برای مقایسه بهتر مدل‌های پایه و چهار-جریان، جدول (۱۰) و شکل (۳۷) در کنار یکدیگر رسم شده‌اند.

Table (7): The accuracy obtained after training the MLP layer of the CVT base model and CVT-13 4Stream

جدول (۷): دقت به دست آمده بعد از آموزش پرسپترون چند لایه مدل پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و چهار-

جریان

تست	اعتبارسنجی	آموزش	دقت
۰/۹۲۰۹	۰/۹۲۲۲	۰/۹۶۰۵	مدل پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی
۰/۹۲۹۲	۰/۹۲۹۶	۰/۹۵۷۵	مدل شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی چهار-جریان



(ب) زیان آموزش و اعتبارسنجی

(الف) دقت آموزش و اعتبارسنجی

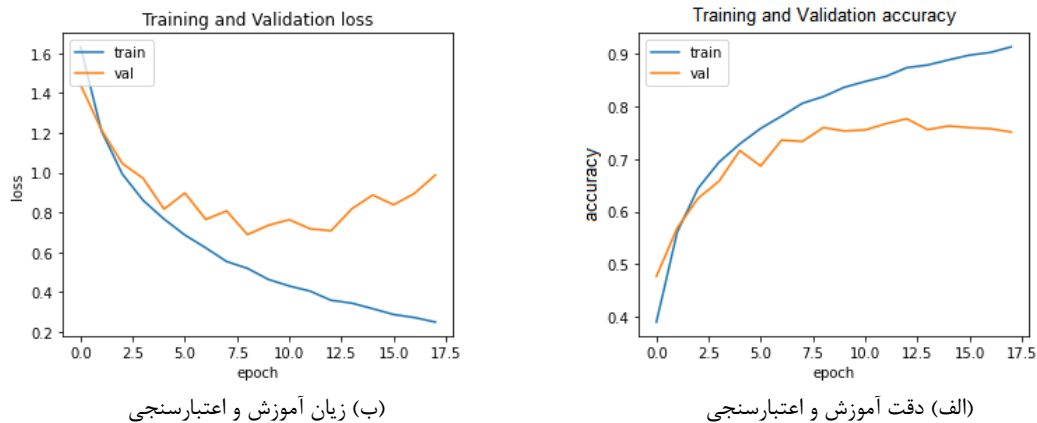
شکل (۳۴): فرایند آموزش مدل پایه شبکه‌های عصبی کانولوشنی با مبدل‌های بینایی و چهار-جریان

Figure (34): The process of teaching the basic model of CVT and CVT-13 4Stream, a) Training and validation accuracy, b) Training and validation loss

Table (8): Accuracy obtained after 17 training epochs on Cifar10 data of RDenseCNN base model

جدول (۸): دقت به دست آمده بعد از ۱۷ دوره آموزش روی داده‌های سیفار-۱۰ مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده

تست	اعتبارسنجی	آموزش	دقت
۰/۷۷۴۳	۰/۷۷۶۰	۰/۸۷۷۶	مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده



(ب) زیان آموزش و اعتبارسنجی

(الف) دقت آموزش و اعتبارسنجی

شکل (۳۵): فرایند آموزش مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده

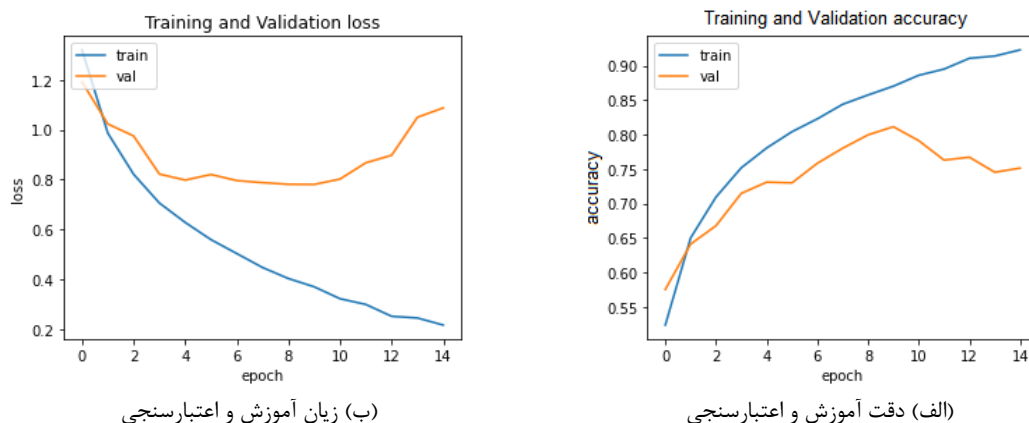
Figure (35): The process of training the basic model of RDenseCNN, a) Training and validation accuracy, b) Training and validation loss

Table (9): Accuracy obtained after 14 training courses on Cifar10 data of RDenseCNN 4-stream model

جدول (۹): دقت به دست آمده بعد از ۱۴ دوره آموزش روی داده‌های سیفار-۱۰ مدل شبکه عصبی کانولوشنی متراکم باقیمانده چهار-جریان

جریان

تست	اعتبارسنجی	آموزش	دقت
۰/۸۰۶۳	۰/۸۱۱۴	۰/۸۷۰۱	مدل شبکه عصبی کانولوشنی متراکم باقیمانده چهار-جریان



(ب) زیان آموزش و اعتبارسنجی

(الف) دقت آموزش و اعتبارسنجی

شکل (۳۶): فرایند آموزش مدل شبکه عصبی کانولوشنی متراکم باقیمانده چهار-جریان

Figure (36): The process of training the model of RDenseCNN 4-stream, a) Training and validation accuracy, b) Training and validation loss

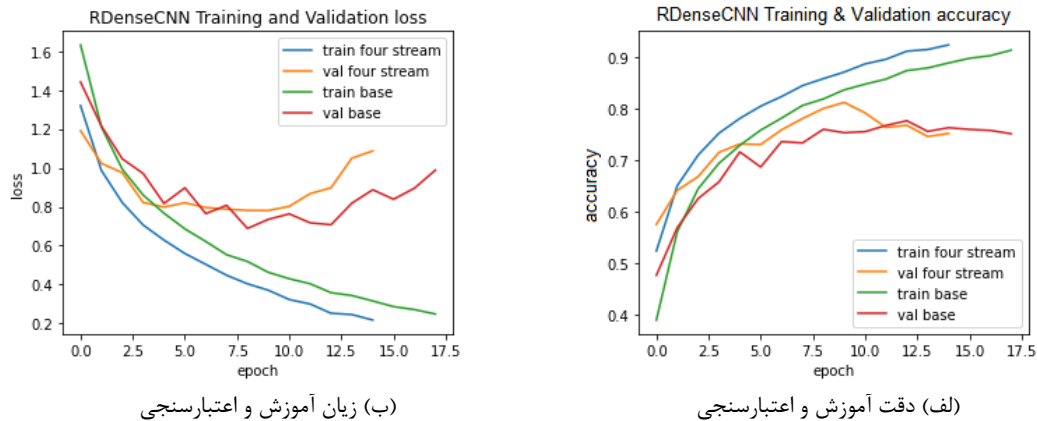
Table (10): The accuracy obtained after 17 training epochs of the basic model of RDenseCNN and RDenseCNN 4-stream

جدول (۱۰): دقت به دست آمده بعد از ۱۷ دوره آموزش مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده و چهار-جریان

تست	اعتبارسنجی	آموزش	دقت
۰/۷۷۴۳	۰/۷۷۶۰	۰/۸۷۷۶	مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده
۰/۸۰۶۳	۰/۸۱۱۴	۰/۸۷۰۱	مدل شبکه عصبی کانولوشنی متراکم باقیمانده چهار-جریان

۴- نتیجه گیری

در مدل چهار-جریان، نگاشت فضای عدم تشابه به عنوان پیش‌پردازش قبل از شبکه عصبی عمیق می‌تواند با انجام خطی‌سازی فضای داده برای شبکه‌های عصبی عمیق کمک‌کننده باشد. این تأثیر مثبت فضای برداری عدم تشابه، روی طبقه‌بندی استاندارد برای داده‌های مصنوعی و تصاویر سیفار-۱۰ بررسی و در نهایت به تحلیل دو مدل مبتنی بر پیش‌پردازش داده شبکه‌های عصبی کانولوشنی با ترانسفورماتورهای بینایی و شبکه عصبی کانولوشنی متراکم باقیمانده با عدم تشابه و فیلترهای سوبل و آشکارساز لبه پرداخته شد که بیشترین تأثیر این تبدیل زمانی خواهد بود که طبقه‌بند بعد از آن، نیاز به داده‌های خطی داشته باشد.



(ب) زیان آموزش و اعتبارسنجی

(الف) دقت آموزش و اعتبارسنجی

شکل (۳۷): فرایند آموزش مدل پایه شبکه عصبی کانولوشنی متراکم باقیمانده و چهار-جریان

Figure (37): The process of teaching the basic model of RDenseCNN and RDenseCNN 4-stream, a) Training and validation accuracy, b) Training and validation loss

از نظر دقت تفکیک، تصاویر روی مجموعه داده سیفار-۱۰، مدل مبتنی بر شبکه‌های عصبی کانولوشنی با ترانسفورماتورهای بینایی با ۹۳ درصد بازده بیشترین دقت را دارد که این به دلیل از پیش آموزش یافته بودن روی داده‌های شبکه-تصویر^{۶۸} و همچنین استفاده هم‌زمان از نسل جدید شبکه‌های عصبی ترانسفورمر به همراه کانولوشن است. بعد از آن مدل مبتنی بر شبکه عصبی کانولوشنی متراکم باقیمانده با بازده ۸۰ درصد قرار دارد که این عدد می‌تواند با پیش آموزش طولانی مدت روی دیتاست-های بزرگ بالاتر نیز برود. در نهایت طبقه‌بندهای متداول نظیر ماشین بردار پشتیبان، K- نزدیکترین همسایگی و درخت تصمیم با بازده حدود ۳۰ درصد پایین‌تر از همه قرار می‌گیرند. اگرچه تبدیل عدم تشابه می‌تواند با افزایش قابل توجهی بازده این روش‌ها را تا ۴۰ درصد بالا ببرد اما همچنان فاصله زیادی با شبکه‌های عصبی مخصوص پردازش تصویر وجود دارد که این خود نشان دهنده قدرت فوق العاده یادگیری عمیق در حل اینگونه مسائل است. یکی از محدودیت‌های مدل پیشنهادی زمان شبیه‌سازی بالای آن است. 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 این شرح است. انتخاب ویژگی‌های برتر تصویر با استفاده از روش‌های بهینه‌سازی مانند الگوریتم ژنتیک و حذف ویژگی‌های با اطلاعات متفاوت کمتر، ارزیابی پارامترهای قابل آموزش، سرعت آموزش و زمان آموزش، گنجاندن زمان در مدل با تعمیم بازشناسی تصویر سریع به بازشناسی با تمرکز و توجه و پردازش موازی، استفاده از دیتاست‌های دیگر از جمله منیست جهت ارزیابی و مقایسه نتایج ارزیابی با مدل مبتنی بر شبکه عصبی کانولوشنی.

سپاسگزاری

این مقاله مستخرج از رساله دکتری در دانشگاه آزاد اسلامی واحد ساوه است. نویسندگان بر خود لازم می‌دانند مراتب تشکر صمیمانه خود را از همکاران حوزه پژوهشی دانشگاه آزاد اسلامی و داوران محترم که ما را در انجام و ارتقای کیفی این مقاله یاری نموده‌اند، اعلام نمایند.

References

مراجع

- [1] M. Zhang, W. Li, Q. Du, "Diverse region-based CNN for hyper spectral image classification", IEEE Trans. on Image Processing, vol. 27, no. 6, pp. 2623-2634, June 2018 (doi: 10.1109/TIP.2018.2809606).
- [2] Z. Gong, P. Zhong, Y. Yu, W. Hu, S. Li, "A CNN with multi scale convolution and diversified metric for hyper spectral image classification", IEEE Trans. on Geoscience and Remote Sensing, vol. 57, no. 6, pp. 3599-3618, June 2019 (doi: 10.1109/TGRS.2018.2886022).
- [3] Y. Pei, Y. Huang, Q. Zou, X. Zhang, S. Wang, "Effects of image degradation and degradation removal to cnn-based image classification", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 43, no. 4, pp. 1239-1253, April 2021 (doi: 10.1109/TPAMI.2019.2950923).

- [4] M. Gour, S. Jain, T.S. Kumar, "Residual learning based CNN for breast cancer histopathological image classification", *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 621-635, Sept. 2020 (doi:10.1002/ima.22403).
- [5] Y.M. Costa, D. Bertolini, A.S. Britto, G.D. Cavalcanti, L.E. Oliveira, "The dissimilarity approach: A review", *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2783-2808, April 2020 (doi: 10.1007/s10462-019-09746-z).
- [6] E. Pękalska, R.P. Duin, "Dissimilarity representations allow for building good classifiers", *Pattern Recognition Letters*, vol. 23, no. 8, pp. 943-956, June 2002 (doi: 10.1016/S0167-8655(02)00024-7).
- [7] E. Pekalska, P. Paclik, R. P. Duin, "A generalized kernel approach to dissimilarity-based classification", *Journal of Machine Learning Research*, vol. 2, pp. 175-211, Dec. 2001 (doi: 10.1.1.16.3363).
- [8] E. Pekalska, R.P. Duin, "Dissimilarity-based classification for vectorial representations", *Proceeding of the IEEE/ ICPR*, pp. 137-140, Hong Kong, China, Aug. 2006 (doi: 10.1109/ICPR.2006.457).
- [9] R.P. Duin, M. Loog, E. Pękalska, D.M. Tax, "Feature-based dissimilarity space classification", *Proceeding of ICPR*, vol. 6388, pp. 46-55, Springer, Berlin, Heidelberg, Aug. 2010 (doi: 10.1007/978-3-642-17711-8_5).
- [10] R.P. Duin, E. Pękalska, "The dissimilarity space: Bridging structural and statistical pattern recognition", *Pattern Recognition Letters*, vol. 33, no. 7, pp. 826-832, May 2012 (doi: 10.1016/j.patrec.2011.04.019).
- [11] I. Theodorakopoulos, D. Kastaniotis, G. Economou, S. Fotopoulos, "Pose-based human action recognition via sparse representation in dissimilarity space", *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 12-23, Jan. 2014 (doi: 10.1016/j.jvcir.2013.03.008).
- [12] H. Bunke, K. Riesen, "Graph classification based on dissimilarity space embedding", *Proceeding of IAPR-SPR-SSPR*, vol 5342, pp. 996-1007, Springer, Berlin, Heidelberg, Dec. 2008 (doi: 10.1007/978-3-540-89689-0_103).
- [13] R.H. Pinheiro, G.D. Cavalcanti, R. Tsang, "Combining dissimilarity spaces for text categorization", *Information Sciences*, vol. 406-407, pp. 87-101, Sept. 2017 (doi: 10.1016/j.ins.2017.04.025).
- [14] C. Santos, E.J. Justino, F. Bortolozzi, R. Sabourin, "An off-line signature verification method based on the questioned document expert's approach and a neural network classifier", *Proceeding of the IEEE/IWFHR*, pp. 498-502, Kokubunji, Japan, Oct. 2004 (doi: 10.1109/IWFHR.2004.17).
- [15] S.H. Cha, S.N. Srihari, "Writer identification: statistical analysis and dichotomizer", *Proceeding of IAPR-SPR-SSPR*, vol 1876, pp. 123-132, Springer, Berlin, Heidelberg, Dec. 2000 (doi:10.1007/3-540-44522-6_13).
- [16] D. Bertolini, L. S. Oliveira, E. Justino, R. Sabourin, "Texture-based descriptors for writer identification and verification", *Expert Systems with Applications*, vol. 40, no. 6, pp. 2069-2080, May 2013 (doi: 10.1016/j.eswa.2012.10.016).
- [17] L.S. Oliveira, E. Justino, R. Sabourin, "Off-line signature verification using writer-independent approach", *Proceeding of the IEEE/IJCNN*, pp. 2539-2544, Orlando, FL, USA, Aug. 2007 (doi: 10.1109/IJCNN.2007.437-1358).
- [18] J.G. Martins, L.S. Oliveira, A.S. Britto, R Sabourin, "Forest species recognition based on dynamic classifier selection and dissimilarity feature vector representation", *Machine Vision and Applications*, vol. 26, no. 2, pp. 279-293, Apr. 2015 (doi: 10.1007/s00138-015-0659-0).
- [19] R.H. Zottesso, Y.M. Costa, D. Bertolini, L.E. Oliveira, "Bird species identification using spectrogram and dissimilarity approach", *Ecological Informatics*, vol. 48, pp. 187-197, Nov. 2018 (doi: 10.1016/j.ecoinf.2018.08.007).
- [20] N. Kriegeskorte, M. Mur, P.A. Bandettini, "Representational similarity analysis-connecting the branches of systems neuroscience", *Frontiers in Systems Neuroscience*, vol. 2, no. 4, Nov. 2008 (doi: 10.3389/neuro.06-004.2008).
- [21] H. Popal, Y. Wang, I.R. Olson, "A guide to representational similarity analysis for social neuroscience", *Social Cognitive and Affective Neuroscience*, vol. 14, no. 11, pp. 1243-1253, Nov. 2019 (doi: 10.31234/osf.io/nd8fh).
- [22] T.K. Pegors, S. Tompson, M.B. O'Donnell, E. B. Falk, "Predicting behavior change from persuasive messages using neural representational similarity and social network analyses", *NeuroImage*, vol. 157, pp. 118-128, Aug. 2017 (doi: 10.1016/j.neuroimage.2017.05.063).
- [23] R.M. Visser, H.S. Scholte, T. Beemsterboer, M. Kindt, "Neural pattern similarity predicts long-term fear memory", *Nature neuroscience*, vol. 16, no. 4, pp. 388-390, April 2013 (doi: 10.1038/nn.3345).
- [24] 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).
- [25] J.B. Ritchie, H.L. Masson, S. Bracci, H.P.O. Beeck, "The unreliable influence of multivariate noise normalization on the reliability of neural dissimilarity", *NeuroImage*, vol. 245, Article Number: 118686, Dec. 2021 (doi: 10.1016/j.neuroimage.2021.118686).

- [26] I. Muukkonen, K. Ölander, J. Numminen, V. R. Salmela, "Spatio-temporal dynamics of face perception", *NeuroImage*, vol. 209, Article Number: 116531, April 2020 (doi: 10.1101/550038).
- [27] J. Diedrichsen, E. Berlot, M. Mur, H. H. Schütt, M. Shahbazi, N. Kriegeskorte, "Comparing representational geometries using whitened unbiased-distance-matrix similarity", *Neural Data Science/Analysis*, vol. 5, no. 3, Aug. 2021 (doi: 10.51628/001c.27664).
- [28] H. Wu, B. Xiao, N. Codella, M. Liu, X. Dai, L. Yuan, L. Zhang, "Cvt: Introducing convolutions to vision transformers", *Proceeding of the IEEE/ICCV*, pp. 22-3, Montreal, QC, Canada, Oct. 2021 (doi: 10.1109/iccv-48922.2021.00009).
- [29] F. Fooladgar, S. Kasaei, "Lightweight residual densely connected convolutional neural network", *Multimedia Tools and Applications*, vol. 79, no. 35, pp. 25571-25588, Sept. 2020 (doi: 10.1007/s11042-020-09223-8).
- [30] N. Gruzling, "Linear separability of the vertices of an n-dimensional hypercube", Ph.D. Thesis, University of Northern British Columbia, 2007 (doi: 10.24124/2007/bpgub464).
- [31] R. Elshawi, A. Wahab, A. Barnawi, S. Sakr, "DLBench: a comprehensive experimental evaluation of deep learning frameworks", *Cluster Computing*, vol. 24, no. 3, pp. 2017-2038, Sept. 2021 (doi: 10.1007/s10586-021-03240-4).

زیرنویس‌ها

1. Convolutional neural network (CNN)
2. Deep neural network
3. Representational dissimilarity
4. Pattern recognition
5. Euclidean distance
6. Correlation distance
7. Prototype
8. K-nearest neighbor
9. Regularized linear/Quadratic normal density-based classifier
10. Kernel distance
11. Principal components analysis
12. Support vector machine
13. Structural representation
14. Kernel representation
15. Class model
16. Dissimilarity vector
17. Sparse
18. Graph embedding
19. Bag of word
20. Majority
21. Multi class
22. Two class
23. Label
24. Local binary pattern
25. Local phase quantization
26. Texture feature
27. Maximum fusion
28. Gabor filter
29. Scale invariant feature transform
30. Speedup robust feature
31. Ensemble
32. Representational dissimilarity matrix
33. Representational similarity analysis
34. Electroencephalography
35. Functional magnetic resonance imaging
36. Embedding
37. Multi dimensional scaling
38. Classify
39. Prefrontal cortex
40. Pre-train
41. Logistic regression

42. Naivebayes
43. Decision tree
44. Random forest
45. Gradient boosting
46. Convolutions to vision transformers
47. Residual dense convolutional neural network
48. 4-stream
49. Dissimilarity vector space
50. Canadian institute for advanced research
51. Reshape
52. Edge detector
53. Sobel
54. Gray
54. City Block distance
55. Chebishev distance
56. Cosine distance
57. Batch normalization
58. Residual
59. Multi layer perceptron
60. Concat
61. Linear separability
62. Transpose
63. Probability density function (PDF)
64. Facebook ai research lab (FAIR)
65. Linear support vector classification
66. Overfitting
67. Image net