



نسخ از دور

GIS ایران



سنجش از دور و GIS ایران
Iranian Remote Sensing & GIS

سال پانزدهم، شماره سوم، پاییز ۱۴۰۲

Vol. ۱۵, No. ۳, Autumn ۲۰۲۳

GIS

۶۰-۴۱

مقاله پژوهشی

بهبود طبقه‌بندی تصاویر ابرطیفی با استفاده از مدل ترکیبی شبکه‌های کپسول و درخت تصمیم تقویتی

پویا احمدی^۱، طیبه مناقبی^۲، حمید عبادی^۳، بهنام اصغری بیرامی^۴

۱. دانشجوی دکتری سنجش از دور، گروه مهندسی نقشه‌برداری و اطلاعات مکانی، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران
۲. استادیار دانشکده مهندسی نقشه برداری و اطلاعات مکانی، دانشگاه تهران، تهران، ایران
۳. استاد گروه مهندسی فتوگرامتری و سنجش از دور، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران
۴. دکتری سنجش از دور، گروه مهندسی فتوگرامتری و سنجش از دور، دانشکده مهندسی نقشه برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

تاریخ پذیرش مقاله: ۱۴۰۱/۰۱/۲۹

تاریخ دریافت مقاله: ۱۴۰۰/۰۸/۱۰

چکیده

با گسترش دانش سنجش از دور، استفاده از تصاویر هایپراسپکترال روزبه‌روز افزایش و عمومیت می‌یابد. طبقه‌بندی یکی از محبوب‌ترین موضوعات در سنجش از دور ابرطیفی است. طی دو دهه گذشته، روش‌های بسیاری برای مقابله با مشکل طبقه‌بندی داده‌های هایپراسپکترال پیشنهاد شده است. در پژوهش حاضر، ساختاری مبتنی بر یادگیری شبکه‌های کپسول برای طبقه‌بندی تصاویر ابرطیفی به کار رفته است؛ به‌گونه‌ای که ساختار شبکه بتواند، با استفاده از یک لایه کانولوشنی و یک لایه کپسول، بهترین حالت تولید ویژگی‌ها را داشته باشد و در عین حال از بیش‌برازش شبکه روی نمونه‌های آموزشی جلوگیری کند. نتایج به‌دست‌آمده نشان از کیفیت بالای ویژگی‌های تولیدی در ساختار پیشنهادی دارد. در راستای بهبود دقت طبقه‌بندی، رویکرد استخراج ویژگی از طریق شبکه طراحی شده و طبقه‌بندی با استفاده از الگوریتم درخت تقویتی XGBoost، با روش طبقه‌بندی از طریق شبکه عمیق سراسری مقایسه شد تا، علاوه بر بررسی و کیفیت‌سنجی ویژگی‌های عمیق برداری تولیدی به‌روشنی پیشنهادی در طبقه‌بندی‌کننده‌های گوناگون، میزان توانایی شبکه‌های عمیق سراسری نیز، در کاربرد طبقه‌بندی، بررسی شود. رویکرد کپسول پیشنهادی شامل سه لایه اصلی است: (۱) Prime با کپسول‌هایی به‌اندازه ۸ و ۳۲ فیلتر ۹×۹ و گام حرکتی ۲؛ (۲) Digitcaps دارای ده کپسول شانزده‌بعدی؛ (۳) لایه تماماً متصل. نتایج بررسی دو رویکرد برای شبکه عمیق و نیز ترکیب شبکه‌های کپسول با الگوریتم درخت تقویتی XGBoost مقایسه شد. رویکردهایی همچون SVM، RF-۲۰۰، LSTM، GRU، و GRU-Pretanh برای مقایسه رویکرد پیشنهادی براساس پیکربندی‌هایی در نظر گرفته شدند که در تحقیقات به آنها اشاره شده بود. برای ارزیابی مدل پیشنهادی، مجموعه داده Indian Pines نیز، شامل شانزده کلاس متفاوت، به کار رفت. با استفاده از روش پیشنهادی ترکیبی، طبقه‌بندی تصاویر با دقت ۹۹٪ روی داده‌های آموزش و دقت ۹۷/۵٪ روی داده‌های تست انجام می‌شود.

کلید واژه‌ها: طبقه‌بندی، تصاویر ابرطیفی، شبکه‌های کپسول، درخت تصمیم تقویتی، مدل ترکیبی.

* نویسنده مکاتبه‌کننده: تهران، خیابان ولیعصر، دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی. تلفن: ۰۹۱۸۵۳۳۶۰۴

Email: ahmadi.pouya@ut.ac.ir

https://orcid.org/۰۰۰۹۰۰۰۱۰۷۳۰۸۰۱۰۵۴

https://dx.doi.org/۱۰.۴۸۳۰۸/gisj.۲۰۲۳.۱۰۲۳۴۷



Copyright: © ۲۰۲۳ by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

۱- مقدمه

سنجنده‌های ابرطیفی^۱ نوع خاصی از سنجنده‌های طیف‌سنجی‌اند که پهنای باند مورد نظر را به صدها باند باریک مجاور هم تقسیم می‌کنند. دردسترس قرارگرفتن سنجنده‌هایی، با توان تفکیک طیفی^۲ بالا و تعداد باند زیاد، دستیابی به جزئیات بیشتر درمورد کلاس‌ها را امکان‌پذیر کرده و باعث افزایش دقت و اعتبار طبقه‌بندی شده است. با توجه به تعداد بسیار باند‌ها، تصاویر ابرطیفی کاربردهای فراوانی می‌یابند (Ghassemian et al., 2003). استفاده از تصاویر ابرطیفی در کاربردهایی مانند کشاورزی و برآورد دقیق محصولات، جنگل‌داری و مطالعه پوشش‌های گیاهی، مطالعه منابع آب و شناسایی آلودگی‌ها، اکتشاف منابع طبیعی و معدنی، زمین‌شناسی و نفت، کاربردهای صلح‌آمیز همچون شناسایی و پاک‌سازی مناطق جنگی آلوده، کاربردهای نظامی و شناسایی ادوات استتار شده نتایج مطلوبی به دست آورده است.

مهم‌ترین چالش، در پردازش تصاویر ابرطیفی، اندک‌بودن تعداد نقاط آموزشی به نسبت تعداد باند‌های تصویر است. درواقع، مسئله این است که اگر تعداد نقاط آموزشی به نسبت تعداد باند‌های داده، کمتر از یکصد نمونه به‌ازای هر کلاس باشد، با افزایش تعداد باند‌ها، خطای تخمین پارامترهای طبقه‌بندی‌کننده بیشتر می‌شود (Keshavarz & Ghasemian, 2005). با توجه به این مسئله، یک راه‌حل برای بهبود عملکرد طبقه‌بندی‌کننده‌های آماری، کاهش فضای ویژگی پیش‌از طبقه‌بندی این تصاویر است که به افزایش نسبت تعداد نقاط آموزشی به تعداد ویژگی منجر می‌شود. برای این منظور، دو دسته الگوریتم به‌کار می‌رود: الگوریتم‌های انتخاب باند^۳ و الگوریتم‌های کاهش ابعاد فضای ویژگی^۴. در سال‌های اخیر، الگوریتم‌های متعددی برای استخراج ویژگی مطرح و استفاده شده‌اند. استخراج ویژگی با مرز تصمیم‌گیری^۵ (Lee & Landgrebe, 1993)، نگاشت جست‌وجوگر (Jimenez & Landgrebe, 1999)، تحلیل ممیز و تبدیل موجک^۶ (Kaewpajit et al., 2003)، تحلیل مؤلفه‌های

اصلی^۷، تحلیل مؤلفه‌های مستقل^۸ (Du et al., 2008) از جمله الگوریتم‌هایی‌اند که به‌منظور کاهش ابعاد فضای ویژگی و بهبود طبقه‌بندی تصاویر ابرطیفی به‌کار رفته‌اند. بین سال‌های ۱۹۹۸ تا ۲۰۰۰، گوالتیری و چتری^۹ (۲۰۰۰) کارآیی ماشین‌های بردار پشتیبان^{۱۰} در طبقه‌بندی تصاویر ابرطیفی را بررسی کردند و نشان دادند این روش امکان آنالیز تصاویر ابرطیفی را در فضای ویژگی با ابعاد زیاد و بدون نیاز به فرایند کاهش ویژگی، فراهم می‌کند. ملگانی و بروزونه^{۱۱} (۲۰۰۴) نیز کارآیی ماشین‌های بردار پشتیبان در طبقه‌بندی تصاویر ابرطیفی را مطالعه کردند. لیانگ^{۱۲} و همکاران (۲۰۰۵) ماشین بردار رگرسیون^{۱۳} را برای طبقه‌بندی تصاویر ابرطیفی پیشنهاد دادند. از دید آنها، پیش‌بینی‌های احتمالی، تخمین اتوماتیک پارامترها و امکان انتخاب هسته‌های اختیاری از مزایای RVM بر SVM است. فوکوناگا و اسبر^{۱۴} (۲۰۰۸) مشکل اندک‌بودن تعداد نمونه‌های آموزشی در طبقه‌بندی تصاویر ابرطیفی را بررسی و نشان دادند تعداد نقاط آموزشی مورد نیاز برای طبقه‌بندی‌کننده‌های خطی، متناسب با افزایش تعداد باند‌ها، به‌صورت خطی افزایش می‌یابد. بروزونه و همکاران (۲۰۰۶) روش TSV^{۱۵} را برای طبقه‌بندی داده‌هایی مطرح کردند که نمونه‌های آموزشی آنها کاملاً مورد اعتماد نیستند. فاول^{۱۶} و همکاران (۲۰۰۷) استفاده

1. Hyperspectral
2. Spectral Resolution
3. Band Selection
4. Dimension reduction
5. Feature Extraction Based on Decision Boundaries
6. Automatic Wavelet Spectral Analysis
7. Principal Component Analysis
8. Independent Component Analysis
9. Gualtieri & Chetri
10. Support Vector Machine
11. Melgani & Bruzzone
12. Liyang
13. Regression Vector Machine
14. Fukunaga & Sber
15. Transductive SVM
16. Fauvel

هو^۹ و همکاران (۲۰۱۵) شبکه کانونولوشنی عصبی پنج‌لایه‌ای برای داده‌های ابرطیفی ارائه دادند. در این روش، از اطلاعات طیفی بدون در نظر گرفتن اطلاعات مکانی استفاده شد. یو^{۱۰} و همکاران (۲۰۱۹) شبکه کانونولوشنی کپسول را برای تشخیص وسایل نقلیه، با استفاده از تصاویر سنجش از دور پیشنهاد دادند. در رویکرد پیشنهادی، تصویر سنجش از دور با استفاده از استراتژی قطعه‌بندی سوپریکسل^{۱۱} به چندین تکه^{۱۲} تقسیم شد و به صورت ورودی برای تشخیص وجود وسایل نقلیه به کار رفت. گائو^{۱۳} و همکاران (۲۰۲۰)، رویکرد شبکه عصبی کپسول چندمقیاسه را برای تشخیص تغییرات از تصاویر رادار با دهانه مصنوعی^{۱۴} طراحی کردند. در این مطالعه، مدل کپسول صرفاً برای مشخص کردن رابطه بین ویژگی‌ها استفاده شد و یک انطباقی کانونولوشن، برای افزایش قدرت شناسایی رابطه بین ویژگی‌ها به کار رفت. در ادامه، در مورد شبکه‌های عصبی کانونولوشن سه‌بعدی (CNN)، به منظور طبقه‌بندی تصاویر ابرطیفی با استفاده هم‌زمان از ویژگی‌های طیفی و مکانی، بحث و بررسی شد و در نهایت مدل شبکه کپسول سه‌بعدی، با استفاده از کپسول‌های طیفی-مکانی، جایگزین شبکه‌های کانونولوشنی شدند (Paoletti et al., 2018; 2019). شو^{۱۵} و همکاران (۲۰۲۰) یک شبکه عصبی کپسول با شبکه کانونولوشنی اکتاو^{۱۶} را برای طبقه‌بندی

هم‌زمان از ویژگی‌های طیفی و مکانی در طبقه‌بندی تصاویر ابرطیفی به روش SVM را بررسی کردند. نتایج این بررسی روی تصویر ROSIS مؤید افزایش کم‌دقت طبقه‌بندی، بدون کاهش ابعاد فضای ویژگی و بهبود چشمگیر دقت، با استفاده از الگوریتم کاهش ابعاد فضای ویژگی بود. اکبری^۱ و همکاران (۲۰۱۶) با استخراج اطلاعات از پیکسل‌های همسایگی در فرایند انتخاب نمونه‌ها، کوشیدند دقت الگوریتم‌های طبقه‌بندی طیفی-مکانی در تصاویر ابرطیفی را بهبود ببخشند. غفاری^۲ و همکاران (۲۰۱۶) انتخاب باندهای بهینه به منظور بهبود جداسازی طیفی تصاویر ابرطیفی را بررسی کردند. آنها روش جدیدی برای انتخاب باند متشکل از اولویت‌بندی‌ها، در راستای کاهش اثر تغییرپذیری طیفی و کاهش همبستگی بین باندها مطرح کردند و نشان دادند می‌توان به حذف مؤثر بسیاری از باندهای غیر ضروری اقدام کرد.

در چند سال اخیر، شبکه‌های عصبی کانونولوشن (CNN)^۳ به‌منزله روش منتخب بر مبنای یادگیری عمیق^۴ برای تولید ویژگی از تصاویر، در نظر گرفته شده است زیرا ساختاری مناسب برای کار با تصاویر دارد (Ghamisi et al., 2017). نکته مهم در این زمینه تعیین ساختار شبکه متناسب با داده‌های ورودی به شبکه است. لوئو^۵ (۲۰۱۸) تلاش داشت از شبکه‌های کپسول^۶ استفاده کند؛ وی رویکردی مبتنی بر CNN را با مدل مبتنی بر CapsNet مقایسه کرد. نتایج این مطالعه مؤید آن بود که روش مبتنی بر CapsNet مزایای مورد انتظار را فراهم نمی‌کند و هیچ بینش عمیقی نشان داده نشده است. وانگ^۷ و همکاران (۲۰۱۸) یک ساختار CapsNet بر اساس شبکه‌های عصبی بازگشتی (RNN)^۸ را به منظور تجزیه و تحلیل احساسات و عواطف مطرح کردند. در پژوهشی دیگر، شبکه‌های عصبی کپسول این‌گونه تعریف شد که نوع کاملاً جدیدی از معماری یادگیری عمیق را نشان می‌دهد که قصد دارد بر محدودیت‌ها و اشکالات CNN، همچون از دست رفتن اطلاعات در مرحله Max Pooling غلبه کند (Sabour et al., 2017).

1. Akbari
2. Ghaffari
3. Convolution Neural Networks
4. Deep Learning
5. Luo, Zou
6. CapsNet
7. Wang
8. Recurrent Neural Network
9. Hu
10. Yu
11. Superpixel Segmentation Strategy
12. Patch
13. Gao
14. Synthetic Aperture Radar
15. Xu
16. Octave Convolution

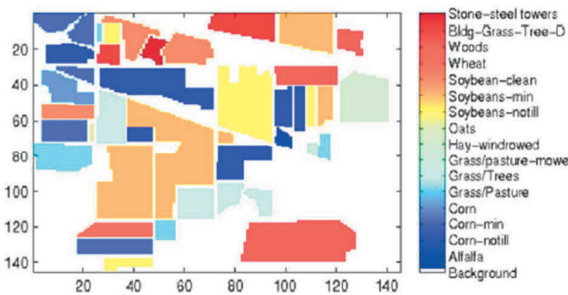
۲- داده‌ها و معرفی منطقه مورد مطالعه

تصاویر ابرطیفی در جامعه سنجش از دور، به‌منظور استفاده از ترکیب صدها کانال طیفی در صحنه‌ای واحد، کاربرد گسترده‌ای دارند. طبقه‌بندی این تصاویر، به‌دلیل ماهیت پیچیده صحنه تصویر (مثلاً داده‌های بسیار، پیکسل‌های مختلط و نمونه‌های آموزشی محدود)، مسئله‌ای پرچالش به‌شمار می‌آید و به‌همین دلیل تلاش‌های بسیاری برای پرداختن به آن انجام شده است. مجموعه داده‌های مورد استفاده، در این تحقیق، داده‌های جامع از زمین‌های پائین در امریکا است که در ژوئن سال ۱۹۹۲ جمع‌آوری شده‌اند و شانزده کلاس را شامل می‌شوند که اغلب مربوط به پوشش زمینی‌اند (Ghassemian et al., 2003). در شکل ۱، این تصویر به‌همراه نقشه کاربری نشان داده شده است.

این مجموعه داده، به‌منزله داده‌های پایه، برای ارائه مدل‌های طبقه‌بندی در بیشتر تحقیقات بررسی و استفاده شده‌اند. داده‌های این مجموعه این امکان را فراهم می‌کند که مقایسات جامعی از رویکردها، در تحقیق پیش رو، انجام شود. ابعاد این تصاویر ابرطیفی ۱۴۵×۱۴۵×۲۰۰ است و اطلاعات هر یک از دسته‌ها در جدول ۱ و پوشش زمینی در شکل ۷ آمده است.

تصاویر ابرطیفی، طراحی کردند. در رویکرد طراحی‌شده، شبکه کانولوشنی معماری‌های موازی را سازگار و ویژگی‌های تصاویر را استخراج کرد.

در این پژوهش، از ساختاری مبتنی بر یادگیری عمیق برای طبقه‌بندی تصاویر ابرطیفی استفاده شده است؛ به‌گونه‌ای که بتوان از ویژگی‌های معماری شبکه کپسول برای رفع معایب شبکه کانولوشنی، به‌منظور طبقه‌بندی تصاویر ابرطیفی و مقایسه نتایج این دو شبکه، استفاده کرد. در این مطالعه، با توجه به اینکه خروجی‌های شبکه کپسول به‌صورت بردار است، می‌توان ارتباط بین ویژگی‌های استخراجی را از طریق اندازه و جهت بردار، استخراج کرد؛ هر یک از المان‌های بردار خروجی نشان‌دهنده توصیف خاصی از موجودیت تصویر است. رویکرد استخراج ویژگی، از طریق شبکه طراحی شده و طبقه‌بندی با استفاده از الگوریتم درخت تقویتی XGBoost^۱ با روش طبقه‌بندی از طریق شبکه عمیق سراسری مقایسه شده است تا، علاوه بر بررسی و کیفیت‌سنجی ویژگی‌های عمیق تولیدی به‌روش پیشنهادی در طبقه‌بندی‌کننده‌های متفاوت، میزان توانایی شبکه‌های عمیق سراسری نیز در کاربرد طبقه‌بندی، بررسی شود.



(ب)



(الف)

شکل ۱. ترکیب رنگی تصویر Indian Pines (الف)؛ نقشه کاربری تصویر Indian Pines (ب)

1. Xtreme Gradient Boosting

تکنیک‌های یادگیری ماشین^۱ سنتی داشته است (Xu et al., 2018; Li, 2018). تفاوت ساختار شبکه‌های عصبی کانولوشنی که یکی از تکنیک‌های یادگیری عمیق محسوب می‌شوند، با شبکه‌های عصبی مصنوعی، محدودیت‌های اعمال شده در ارتباطات بین نورون‌ها در شبکه‌های کانولوشنی است. اگرچه این محدودیت‌ها عمومیت شبکه‌های کانولوشن را در مواجهه با داده‌های متنوع کاهش می‌دهد، ساختار این شبکه‌ها با توجه به قیود خاص آن، کاملاً منطبق و مناسب با ساختار تصاویر است. همچنین در شبکه‌های عصبی کپسول، با توجه به اینکه خروجی شبکه بردار است، می‌توان ارتباط بین ویژگی‌های استخراجی از شبکه‌های کانولوشنی را در نظر گرفت (Maggiori et al., 2016). براین اساس، پژوهش حاضر بر فرایندهای تولید ویژگی‌های عمیق و طبقه‌بندی بر مبنای آن متمرکز شده است؛ در ادامه، جزئیات روش‌های به کار رفته در هر مرحله را بیان می‌کنیم.

۳-۲- شبکه‌های عصبی کانولوشنی

شبکه‌های عصبی کانولوشنی نوعی از شبکه‌های رو به جلو هستند که اولین بار، برای تشخیص تصویر به کار رفتند. این شبکه‌ها از مهم‌ترین و پرکاربردترین

به‌منظور جلوگیری از ایجاد مشکل در تحلیل‌های عددی، مقادیر باندها نرمال‌سازی شد. برای نرمال‌سازی داده‌ها، از دو روش بهره گرفته شد؛ در روش تمامی باندها، مقدار درجهٔ خاکستری هر پیکسل در هر باند بر جمع مقادیر درجات خاکستری آن پیکسل در تمامی باندها تقسیم می‌شود و در روش تک‌باند نیز، مقدار درجهٔ خاکستری هر پیکسل در هر باند بر جمع مقادیر درجات خاکستری تمامی پیکسل‌های آن باند تقسیم می‌شود. تأثیر روش نرمالیزه کردن در نتایج طبقه‌بندی با استفاده از هستهٔ گاوسین بررسی و نشان داده شد، با هر دو روش، نتایج مشابهی به دست می‌آید؛ از این رو در این پژوهش، روش تمامی باندها استفاده شد.

۳- مبانی نظری

۳-۱- رویکردهای مبتنی بر یادگیری عمیق

در چند دههٔ اخیر، با افزایش قدرت پردازنده‌های گرافیکی و در دسترس بودن مجموعه‌های حجیم داده‌ها، استفاده از تکنیک‌های یادگیری عمیق، در بسیاری از شاخه‌های مهندسی و کاربردی، مطرح شده است. به موازات آن حوزه‌های پردازش تصویر نیز تحت تأثیر قرار گرفته‌اند؛ به گونه‌ای که در بسیاری از کاربردهای پردازش تصویر، یادگیری عمیق کارایی بیشتری از

جدول ۱. پوشش زمینی کلاس‌های تصویر

پوشش زمینی	تعداد پیکسل‌های کلاس در تصویر	شمارهٔ کلاس	پوشش زمینی	تعداد پیکسل‌های کلاس در تصویر	شمارهٔ کلاس
جو دوسر	۲۰	۹	یونجه	۵۴	۱
سویای بدون شخم	۹۶۸	۱۰	ذرت بدون شخم	۱۴۳۴	۲
سویای کم‌شخم	۲۴۶۸	۱۱	ذرت کم‌شخم	۸۳۴	۳
سویای شخم‌زده	۶۱۴	۱۲	ذرت	۲۳۴	۴
گندم	۲۱۲	۱۳	سبزه/چمن‌زار	۴۹۷	۵
بیشه	۱۲۹۴	۱۴	سبزه/درختان	۷۴۷	۶
ساختمان/سبزه	۳۸۰	۱۵	سبزه/چمن‌زار کوتاه	۲۶	۷
سنگ/برج	۹۵	۱۶	کاه و خاشاک	۴۸۹	۸

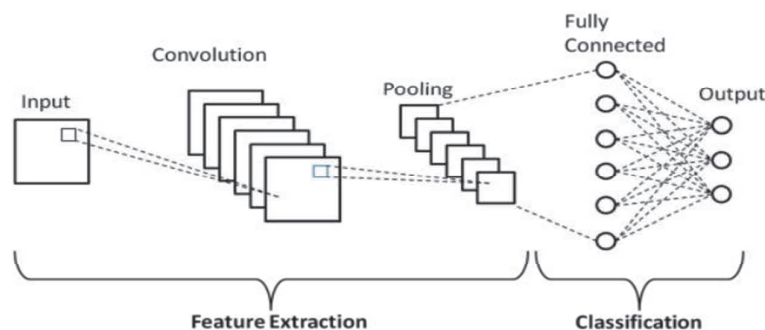
(بیشترین داده‌ها متعلق به کلاس‌های ۱۱، ۲ و ۱۴ است)

1. Machine Learning

- واحد اصلاح‌شده خطی (ReLU):^۲ مزیت اصلی استفاده از این تابع بر سایر توابع این است که هم‌زمان تمامی نورون‌ها را فعال نمی‌کند؛ بدین معنی که در یک زمان، فقط نورون‌های اندکی فعال می‌شوند و این باعث می‌شود شبکه به صورت تنک دربیاید و ویژگی‌های مهمی را فرابگیرد؛
- لایه Dropout: اصطلاح «Dropout» به واحدهای حذف‌شده (پنهان و آشکار) در شبکه عصبی اشاره دارد. این تکنیک را اولین بار جفری هینتون^۳ به کار برد (Bengio & Ian, 2015). Dropout در واقع از تکنیک‌های Regularization تصادفی است که طی هر انتشار رو به جلوی واحدها در لایه‌های پنهان یا آشکار، به صورت تصادفی یا موقت، مطابق یک بردار از متغیرهای تصادفی برنولی بیرون انداخته می‌شوند.
- لایه تماماً متصل: پس از آخرین لایه تجمیع، لایه تماماً متصل^۴ وجود دارد که ویژگی‌های میانی دوبعدی را به بردارهای یک‌بعدی، به منظور ادامه فرایند یادگیری، نگاشت می‌کند. از سویی، می‌توان گفت این لایه وظیفه محاسبه امتیاز دسته‌های طبقه‌ها را برعهده دارد. لایه تماماً متصل به ما این امکان را می‌دهد که خروجی شبکه را در قالب برداری با اندازه ثابت ارائه دهیم.

روش‌های یادگیری عمیق، به‌ویژه در زمینه استخراج ویژگی و طبقه‌بندی تصاویر سنجش از دور، محسوب می‌شوند (LeCun et al., 2015). شبکه‌های عصبی کانولوشنی لایه‌های متعددی دارند که هر یک از لایه‌ها یک سطح از ویژگی را نمایندگی می‌کنند. این سطوح متفاوت نمایندگی با کنار هم قرار گرفتن ساختارهای ساده و غیرخطی به دست می‌آیند و هر یک از ساختارها ویژگی‌ها را به سطح انتزاع بالاتری ارتقا خواهد داد. شبکه کانولوشنی چهار مشخصه اصلی دارند: ارتباطات محلی، اشتراک وزن، لایه ادغام و مشخصه عمق که این مشخصات آن را از سایر روش‌ها متمایز می‌سازد. ساختار شبکه‌های کانولوشنی بخش‌های متوالی زیر را شامل می‌شود (Bengio & Ian, 2015):

- لایه کانولوشنی: این لایه، به منظور استخراج ویژگی‌های ضمنی از ماتریس ورودی، از فیلترهای گوناگونی استفاده می‌کند؛
- لایه تجمیع: این‌گونه معمول است که به صورت دوره‌ای، پس از هر لایه کانولوشن، برای کاهش اندازه ویژگی‌های میانی از یک لایه تجمیع استفاده شود. کاهش حجم پارامترها سبب کاهش محاسبات شبکه می‌شود و از این رو می‌توان بیش‌برازش^۱ را کنترل کرد؛



شکل ۲. معماری شبکه‌های کانولوشنی

منبع: LeCun et al., 2015

1. Overfitting
2. Rectified Linear Unit
3. Geoffrey Everest Hinton
4. Fully Connected Layer

اینکه شبکه دچار اشتباه نشود، باید شبکه کانولوشنی با تصاویری از چرخش‌های متفاوت آموزش ببینند؛ این کار نیازمند داده‌های آموزشی بسیار زیادی است. با توجه به معایب بیان‌شده، استفاده از شبکه کپسول برای رفع معایب شبکه‌های کانولوشنی پیشنهاد می‌شود.

۳-۳- شبکه کپسول

شبکه‌های عصبی کپسول روش کاملاً جدیدی از معماری یادگیری عمیق است که سعی دارد بر محدودیت‌ها و اشکالات شبکه کانولوشنی غلبه کند (Xue et al., 2015). خروجی‌های شبکه عصبی کپسول در مجموعه‌ای از کپسول‌ها، به‌صورت بردار ذخیره



شکل ۴. تصویر چهره‌ای که نحوه قرارگیری و ابعاد اجزای آن تغییر یافته است

می‌شوند و کپسول‌ها محاسبات پیچیده‌ای را روی این ویژگی‌ها انجام می‌دهند؛ سپس نتایج این محاسبات را به بردار کوچکی که حاوی اطلاعات بسیار مفیدی است، نگاشت می‌کنند. از این‌رو می‌توان ارتباط بین ویژگی‌های استخراجی شبکه را از طریق اندازه و جهت بردار، حفظ کرد و هر یک از المان‌های بردار خروجی نشان‌دهنده توصیفی خاص از تصویر است (Candemir et al., 2014). یکی از مهم‌ترین تفاوت‌های شبکه‌های کپسول با شبکه‌های کانولوشنی در شکل ویژگی‌های استخراج شده است که به‌صورت برداری با ویژگی اندازه و جهت است؛ در حالی که در شبکه‌های کانولوشنی و

۳-۲-۱- مزایای شبکه کانولوشنی

مهم‌ترین مزیت استفاده از شبکه عصبی کانولوشن توانایی آن در استخراج خودکار ویژگی‌های تصاویر، با استفاده از مفهوم یادگیری عمیق است. همچنین استفاده از لایه pooling در شبکه‌های کانولوشنی باعث کاهش حجم محاسبات می‌شود و تضمین می‌کند، حتی زمانی که ویژگی‌های استخراجی دچار تغییرات و چرخش‌های کوچکی شوند، خروجی شبکه همچنان یکسان بماند. این ویژگی در طبقه‌بندی اهمیت بسیاری دارد. استفاده از max pooling نیز باعث می‌شود، با تغییرات اندک و جابه‌جایی بسیار کم در تصویر، شبکه همچنان قادر به شناسایی شیء مورد نظر باشد (شکل ۳).



شکل ۳. مزیت استفاده از لایه pooling در شبکه‌های کانولوشنی
منبع: Bengio & Ian, 2015

۳-۲-۲- معایب شبکه کانولوشنی

یکی از کاربردهای شبکه‌های کانولوشنی استخراج ویژگی‌ها برای طبقه‌بندی تصاویر است اما، زمانی که نمونه معرفی شده به مدل آموزش دیده CNN تغییراتی داشته باشد، شبکه در تشخیص این تغییرات دچار مشکل می‌شود. دلیل این مشکل استفاده از لایه Max-pooling است که فضای ارتباطی بین ویژگی‌ها را از بین می‌برد. در شکل ۴، تصویری از چهره درهم‌ریخته و غیرواقعی ارائه شده است که شبکه‌های CNN به‌اشتباه، آن را چهره تشخیص می‌دهند زیرا ارتباط بین ویژگی‌های استخراجی و ابعاد ویژگی‌ها را در نظر نمی‌گیرند. برای

• لایه PrimeCaps

مطابق نسخه اولیه، این لایه دارای کپسول‌هایی به‌اندازه ۸، ۳۲ فیلتر ۹×۹ و گام حرکتی ۲ است.

• DigitCaps

نسخه اولیه این لایه ده کپسول شانزده‌بعدی است که با توجه به مسئله تحقیق حاضر، این مقدار به دو کپسول شانزده‌بعدی تغییر یافت. در این لایه، مقدار الگوریتم مسیریابی نیز، همانند نسخه اولیه، با مقدار ۳ تنظیم شد.

• لایه‌های تماماً متصل

برخلاف پژوهش هینتون که در آن لایه‌های تماماً متصل در نظر گرفته نشده است، در این مطالعه، لایه‌های مورد اشاره منظور شده‌اند و در آن، می‌توان از چندین لایه استفاده کرد. در واقع، این لایه سعی در فراگرفتن ویژگی‌های استخراجی از لایه DigitCaps دارد. در این تحقیق، آخرین لایه تماماً متصل به‌اندازه طبقه‌های مد نظر این مطالعه (شانزده‌تا) است.

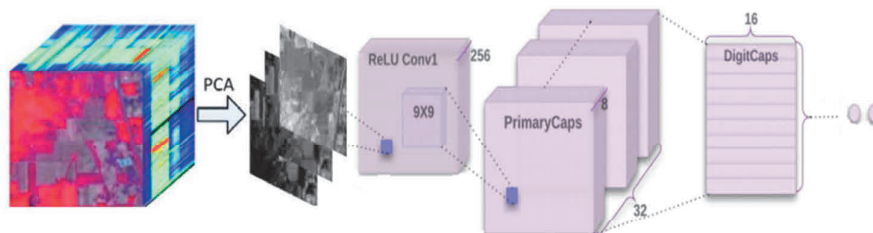
شکل ۵ ساختار کلی رویکرد شبکه‌های کپسول را نشان می‌دهد.

برای یک کپسول، مطابق شکل ۶، ورودی شبکه بردار u_i و خروجی آن بردار v_j است و پارامترها از یک لایه به لایه دیگر منتقل می‌شوند.

برای ارسال کپسول از لایه‌ای به لایه بعدی، به فرایند خاصی نیاز است که در ادامه، به آن می‌پردازیم.

برای این کار، یک ماتریس تبدیل w_i روی خروجی کپسول u_i لایه پیشین اعمال می‌شود و سپس وزن جمعی s_j براساس وزن‌های c_{ij} ، مطابق رابطه (۱)، محاسبه می‌شود (Candemir et al., 2014).

$$\begin{aligned} \widehat{U}_{jl} &= w_{ij}U_j \\ s_j &= \sum c_{ij} \cdot \widehat{U}_{jl} \end{aligned} \quad \text{رابطه (۱)}$$



شکل ۵. شمای کلی از رویکرد شبکه‌های کپسول

شبکه‌های تماماً متصل، خروجی‌ها به‌صورت عددی (اسکالر) هستند. شبکه‌های کپسول، مطابق شکل ۴، دارای لایه‌های Conv1d، PrimaryCaps و DigitCaps هستند (Xue et al., 2015; Candemir et al., 2014). با توجه به اینکه ویژگی‌ها به‌صورت بردار ذخیره شده‌اند، با تغییر آنها در جای‌جای تصویر، اندازه بردار حفظ می‌شود اما جهت آن تغییر می‌کند و در نتیجه، برای شناسایی شیء با چرخش‌های متفاوت، به نمونه‌های آموزشی بیشتری نیاز نخواهد بود.

۱-۳-۳- اجزای شبکه کپسول پیشنهادی و محاسبه خروجی

شبکه کپسول طراحی شده، در این پژوهش، بسیار شبیه به شبکه کپسولی است که هینتون آن را توصیف کرده (Candemir et al., 2014)؛ با این تفاوت که تغییراتی روی آن اعمال شده است. لایه‌های شبکه طراحی شده بدین‌قرار است:

• لایه کانولوشن

این لایه دارای ۲۵۶ فیلتر کانولوشنی به‌اندازه ۹×۹ و گام حرکتی ۲ است که روی آن، تابع فعال‌ساز «ReLU» اعمال شده است. یکی از تغییرات اعمال شده در این مرحله، در قیاس با شبکه هینتون، تغییر مقدار گام حرکتی از ۱ به ۲ است زیرا مدل اصلی روی مجموعه داده‌های mnist استفاده شده بود و تصاویر mnist به‌صورت ۲۸×۲۸ است. با استفاده از گام حرکتی ۲، مقدار پارامترهایی که مدل در لایه بعد دریافت می‌کند، بسیار کاهش می‌یابد.

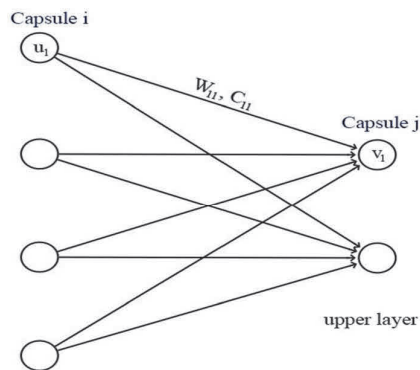
بهبود طبقه‌بندی تصاویر ابرطیفی ...

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \cdot \frac{s_j}{\|s_j\|} \quad \text{رابطه (۲)}$$

$$v_j \approx \|s_j\| s_j \quad \text{for small } s_j$$

$$v_j \approx \frac{s_j}{\|s_j\|} \quad \text{for larg } s_j$$

لایه آخر تماماً متصل، در مقایسه با رویکردهای سنتی یادگیری ماشین مانند SVM و XGBoost، دقت کمتر و میزان محاسبات بیشتری دارد؛ از این رو در رویکرد پیشنهادی، مدل XGBoost به‌منزله جایگزین این لایه‌ها به‌کار رفت.



شکل ۶. ارسال پارامترها از یک لایه به کپسول‌های لایه بعدی

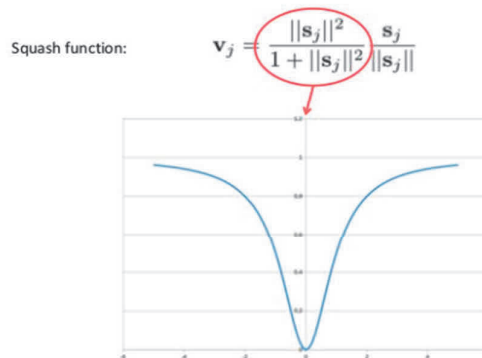
۳-۴- مدل درخت تقویتی

مدل درخت تقویتی روشی بسیار مؤثر و پرکاربرد در یادگیری ماشین است که چن و گواسترین^۲ (۲۰۱۶) آن را پیشنهاد کرده‌اند. این شبکه نوعی تکنیک اجرایی کارآمد برای افزایش صحت و عملکرد ماشین افزایش گرادیان (GBM)^۳ و به‌ویژه، افزایش دقت طبقه‌بندی درخت‌های رگرسیون است. همچنین یکی از محبوب‌ترین و کارآمدترین عملی‌سازی و اجراهای الگوریتم Gradient Boosted Trees و روش یادگیری نظارت‌شده‌ای مبتنی بر تقریب عملکرد با بهینه‌سازی توابع خاص است (Paoletti et al., 2019). Boosting در ML از تکنیک‌های یادگیری متوالی برای تبدیل فرضیه‌ای ضعیف به فرضیه‌ای قوی، به‌منظور افزایش دقت مدل به‌شمار می‌رود (Chen & Guestrin, 2016). ساختار مدل پیشنهادی، پس‌از حذف لایه تماماً متصل و جایگزینی آن با XGBoost، در شکل ۸ نشان داده شده است.

۱. در شبکه‌های کپسول، برای محاسبه خروجی کپسول، از مسیریابی پویا با محاسبه ضریب اتصالی به صورت تکراری استفاده می‌کنند. این فرایند مشابه الگوریتم پس‌انتشار در شبکه‌های عصبی مصنوعی است؛ با این تفاوت که عمل مسیریابی به‌تعداد محدود انجام می‌شود (به‌صورت پیش‌فرض، سه بار).

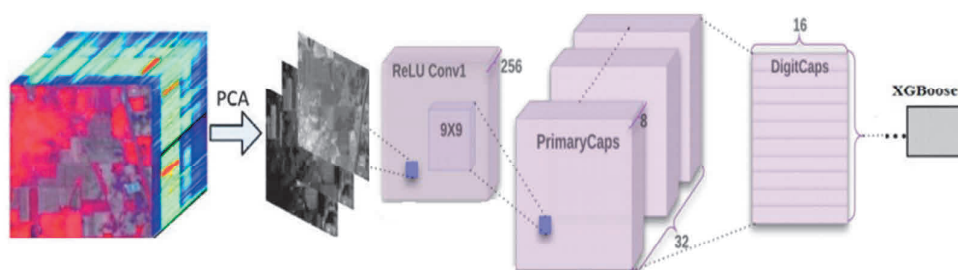
2. Chen & Guestrin
3. Gradient Boosting Machine

Squash



شکل ۷. شکل تابع Squash
منبع: Candemir et al., 2014

در این رابطه، c_{ij} ها ضرایب اتصالی‌اند که با استفاده از مسیریابی پویا محاسبه می‌شوند و $\sum c_{ij}$ برای جمع آنها در نظر گرفته شده است. به‌طور مفهومی، c_{ij} معیاری برای اندازه‌گیری میزان احتمال فعال‌سازی کپسول j با کپسول i است. در شبکه‌های کپسول، به‌جای اعمال ReLU، تابع squash را روی خروجی s_j اعمال می‌کنند. این تابع بردارهای کوچک را به ۰ و بردارهای بزرگ را به بردارهای واحدی کاهش می‌دهد و به‌همین دلیل خروجی یک کپسول v_j ، مطابق رابطه (۲)، طولی بین [0-1] دارد. در شکل ۷، تابع squash ارائه شده و فرایند اعمال squash روی بردار ورودی s_j نیز در ادامه، آورده شده است (Candemir et al., 2014).



شکل ۸. شمای کلی از رویکرد پیشنهادی تلفیق شبکه‌های کپسول و درخت تصمیم تقویتی

• شبکه‌های عصبی بازگشتی RNN-LSTM

LSTM^۱ نوعی مدل یا ساختار برای داده‌های ترتیبی است که در سال ۱۹۹۵ برای توسعه شبکه‌های عصبی بازگشتی ظهور یافت (Keshavarz & Ghasemiyani, 2005). عبارت Long Term Memory به وزن‌های یادگرفته‌شده و Short Memory به حالت‌های درونی سلول‌ها اطلاق می‌شود. تغییر مهم و اصلی این ساختار، در قیاس با RNN، جایگزین کردن لایه میانی RNN با یک بلوک است (Keshavarz & Ghasemiyani, 2005). مهم‌ترین ویژگی LSTM امکان یادگیری وابستگی بلندمدت است که از طریق شبکه‌های عصبی بازگشتی، امکان‌پذیر نبود. برای پیش‌بینی گام زمانی بعدی، لازم است مقادیر وزن‌ها در شبکه به‌روزرسانی شوند که این کار مستلزم حفظ اطلاعات گام‌های زمانی ابتدایی است. در تحقیق مورد بررسی، شبکه‌های LSTM به‌منظور طبقه‌بندی تصویر ابرطیفی، اجرا شد. نتایج پژوهش در جدول ۳ آمده است (Keshavarz & Ghasemiyani, 2005).

• شبکه‌های عصبی بازگشتی RNN-GRU

معماری GRU^۲ (Keshavarz & Ghasemiyani, 2005) به‌منظور حل کاستی‌های شبکه‌های عصبی بازگشتی قدیمی مانند مشکل محوشدگی گرادیان و نیز معایب معماری LSTM مطرح شده است. GRU اغلب نسخه‌ای تغییر یافته از LSTM در نظر گرفته می‌شود زیرا هر دوی این معماری‌ها طراحی مشابهی دارند و در برخی موارد،

در این پژوهش، به‌منظور بررسی صحت روش پیشنهادی، نتایج با سایر روش‌های منابع پیشین نیز بررسی شد؛ در ادامه، توضیح مختصر روش‌ها خواهد آمد.

• رویکرد ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان نوعی از طبقه‌بندی‌کننده‌های آماری است که عملکرد کلی آن بیشتر بر الگوهای طبقه‌بندی و تخمین رگرسیون غیرخطی تمرکز دارد (Landgrebe, 2002). ماشین‌های بردار پشتیبان، از طریق انتقال فضای ورودی به فضایی با بعد بالاتر، می‌توانند مسائل غیرخطی را حل کنند؛ این کار از طریق توابع کرنل امکان‌پذیر است. کرنل‌های متفاوت بسیاری همچون کرنل گوسی، کرنل خطی، کرنل چندجمله‌ای و کرنل تانژانت هایدربولیک وجود دارند. در پژوهش انجام‌شده، کرنل گوسی به‌منظور طبقه‌بندی تصاویر ابرطیفی به کار رفته است.

• رویکرد RF-200

الگوریتم جنگل تصادفی (Landgrebe, 2002) مبتنی بر دسته‌ای از درخت‌های تصمیم‌گیری است که برچسب نهایی هر پیکسل را براساس برآوردی از تصمیمات درخت‌ها و اغلب بر مبنای رأی اکثریت، تعیین می‌کند و برای بسیاری از مجموعه داده‌ها، طبقه‌بندی با صحت بالایی را تضمین می‌کند.

1. Long Short Term Memory
2. Gated Recurrent Unit

در شکل ۹، چهار عبارت FN^9 و TN^8 ، FP^7 ، TP^6 بیانگر تعداد نمونه‌های صحیح است که به‌درستی پیش‌بینی شده‌اند. عبارت FP گویای تعداد نمونه‌های نادرستی است که به‌درستی پیش‌بینی شده‌اند. TN بیانگر تعداد نمونه‌هایی نادرستی است که نادرست دسته‌بندی شده‌اند. عبارت FN نیز به تعداد نمونه‌های درستی اشاره دارد که نادرست دسته‌بندی شده‌اند.

در طبقه‌بندی دودویی، معیار Precision برای اشاره به نسبت بین نمونه‌های پیش‌بینی‌شده مثبت که دقیقاً مثبت‌اند، در قیاس با کل پیش‌بینی‌های مثبت انجام‌شده به کار می‌رود. این مسئله با رابطه (۳) نشان داده می‌شود (Ghamisi et al., 2017).

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{رابطه (۳)}$$

به‌طور کلی عبارت Precision به معنی توانایی بالا برای پیش‌بینی درست است. معیار Recall بیانگر تعداد اسناد پیش‌بینی‌شده مثبت، به نسبت کل اسناد موجود، شامل پیش‌بینی مثبت و منفی است که طبق رابطه (۴) نشان داده می‌شود.

$$\text{Recall} = \frac{TP}{TP + FN} \quad \text{رابطه (۴)}$$

عبارت Recall به این معنی است که طبقه‌بندی مورد نظر توانسته بسیاری از اسناد را به‌درستی طبقه‌بندی کند.

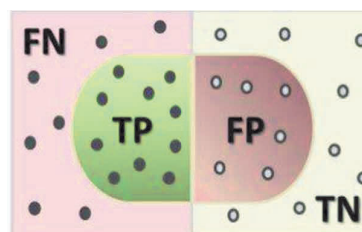
به‌صورت یکسان، نتایج عالی در بردارند. همان‌گونه که بیان شد، یکی از راه‌ها برای حل مشکل محوشدگی گرادپان در شبکه عصبی سنتی، استفاده از GRU است. این نوع معماری از مفهومی به‌نام دروازه به‌روزرسانی^۱ و دروازه بازنشانی^۲ استفاده می‌کند. این دو دروازه، در اصل، دو بردارند که با استفاده از آنها تصمیم گرفته می‌شود چه اطلاعاتی به خروجی منتقل شود و چه اطلاعاتی انتقال نیابد. نکته خاص درباره دروازه‌های مورد نظر این است که این دروازه‌ها را می‌توان طوری آموزش داد که اطلاعات مربوط به گام‌های زمانی بسیار قبل را حفظ کند؛ بدون آنکه حین گذر زمان (طی گام‌های زمانی گوناگون) دستخوش تغییر شوند.

• رویکرد RNN-GRU-PRetanh

در این رویکرد، شبکه بازگشتی با بلوک‌های GRU و نیز تابع فعال‌ساز تانژانت هایدربولیک پارامتریک^۳، برای طبقه‌بندی تصویر ابرطیفی استفاده شده است (Keshavarz & Ghasemiyani, 2005). نتایج این رویکرد در جدول ۳ آمده است.

۵-۳- معیارهای ارزیابی

پیش از طرح و بررسی نتایج، معیارهای ارزیابی مورد استفاده را بررسی می‌کنیم. طبقه‌بندی دوکلاسه^۴ معیارهای ارزیابی گوناگونی دارند که قبل از بررسی آنها، پارامترهای به‌کاررفته در این معیارها توضیح داده می‌شود. اگر طبقه‌بندی دوکلاسه روی مجموعه‌ای داده تست اعمال شود، نتایج زیر (شکل ۹) به دست می‌آید (Domingos, 2012).



شکل ۹. اعمال طبقه‌بندی دودویی روی یک مجموعه داده تست^۵

1. Update gate
2. Reset gate
3. Parametric Rectified Tanh (PRetanh)
4. Binary Classification
5. Apply Binary Classifications to a Test Data Set
6. True Positive
7. False Positive
8. True Negative
9. False Negative

جست‌وجوی پارامترهای بهینه در مدل‌های یادگیری عمیق است. ۸۰٪ داده‌های این تحقیق، به‌منزله داده‌های آموزشی، برای آموزش مدل و ۲۰٪ به‌صورت داده‌های آزمون ارزیابی مدل، استفاده خواهند شد.

۴- نتایج

در این بخش، نتایج حاصل از شبکه کپسول پیشنهادی و شبکه CNN با استفاده از معیارهای ارزیابی Precision، Recall و F1-score (Ghamisi et al., 2017) مقایسه می‌شود. رویکرد پیشنهادی نیز با پنج رویکرد مرسوم طبقه‌بندی مقایسه شده است که در ادامه، نتایج هر یک از این مدل‌ها بیان می‌شود.

۴-۱- نتایج رویکرد مدل شبکه کانولوشنی

در این بخش، نتایج اجرای شبکه CNN روی داده پاینس هند بیان شده است. نتایج در دو حالت، به‌ازای هر کلاس مجزا، به‌ازای تمامی حالت‌ها و برای مقایسه با سایر رویکردها گزارش داده شده است. برای تحلیل‌های بیشتر، نمودار خطا و متریک‌های اندازه‌گیری در شکل ۱۰ آورده شده است. جدول ۲ نیز نتیجه اجرای شبکه CNN روی مجموعه داده‌های تست، بعد از پنجاه بار تکرار مدل است که به‌ازای هر کلاس ارائه شده است. ستون Count تعداد نمونه به‌ازای هر کلاس را نشان می‌دهد.

در این نمودارها، محور X نشان‌دهنده تعداد دفعات اجرای الگوریتم روی داده‌ها و محور Y گویای معیار به‌دست‌آمده، به‌ازای هر بار اجرای الگوریتم است. در شکل ۱۱ نیز، ماتریس خطای مرتبط با رویکرد اول ارائه شده است. جدول ۳ دقت نتایج به‌دست‌آمده با استفاده از ساختار اول و مقایسه آن با نتایج دیگر پژوهش‌ها را درمورد این مجموعه از داده‌ها، بازگو می‌کند.

1. Patch Size

2. <https://optuna.org/>

معیار F-Measure که میانگین هارمونیک نیز نامیده شده، ترکیب دو معیار Precision و Recall است. این عبارت معمولاً F-Score متعادل نیز نامیده و به‌صورت رابطه (۵) نشان داده می‌شود (Ghamisi et al., 2017).

$$F_{score} = 2 \times \frac{Precision * Recall}{Precision + Recall} \quad \text{رابطه (۵)}$$

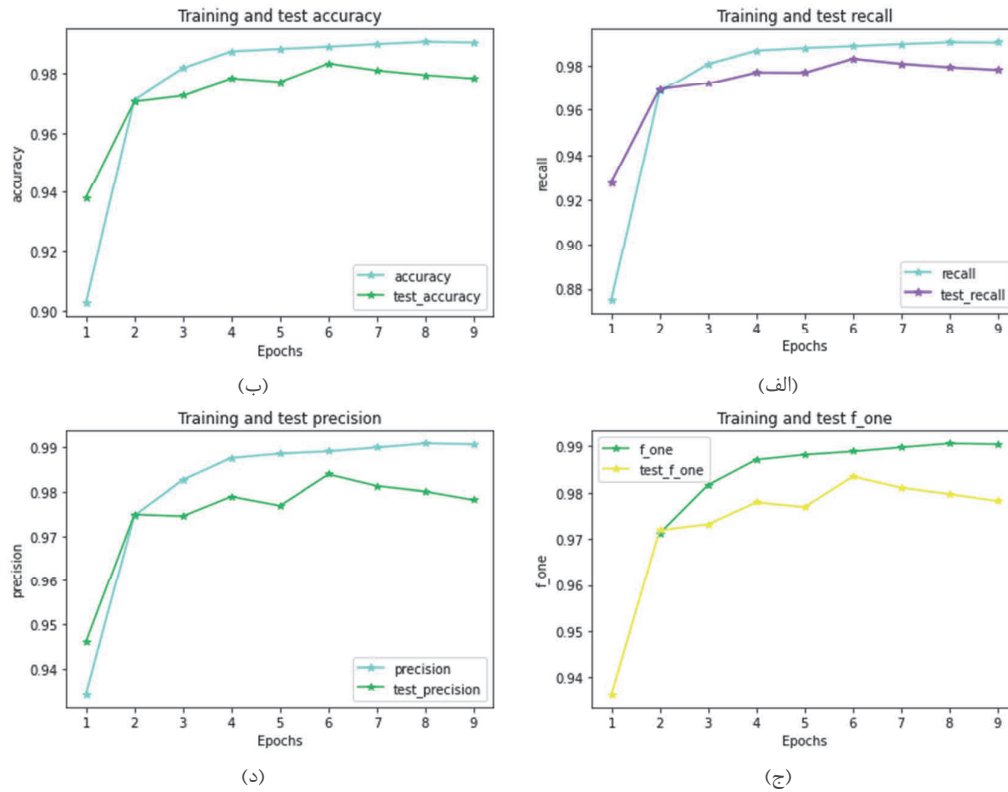
Accuracy یکی دیگر از معیارهایی است که در طبقه‌بندی دودویی کاربرد دارد. برای محاسبه این معیار، نسبت TP و TN به تعداد کل موارد پیش‌بینی شده در نظر گرفته می‌شود؛ بنابراین رابطه (۶) برای آن به‌دست می‌آید.

$$Accuracy = \frac{TP + FN}{TP + FP + FN + TN} \quad \text{رابطه (۶)}$$

۶-۳- پارامترهای شبکه

در پژوهش پیش رو، اندازه نمونه‌های آموزشی $5 \times 5 \times 30$ است؛ 5×5 به اندازه دسته^۱ اشاره دارد و ۳۰ نتیجه اعمال الگوریتم کاهش بعد PCA روی باندهای تصویر ابرطیفی است. در بررسی‌های انجام‌شده ما، اندازه ۳۰ بهینه‌ترین حالت برای کاهش بعد PCA است؛ در این حالت تغییرات در معیارها بسیار ناچیز است و به‌طور کلی در افزایش بعد، بهبود حاصل نمی‌شود. در مجموع، اندازه فیلترهای کانولوشنی در شبکه‌های CNN یا کپسول ممکن است ۳، ۵ و یا بیشتر باشد. در این تحقیق، از فیلتر 9×9 استفاده شده است. لایه کانولوشنی، در ساختار رویکرد پیشنهادی، دارای ۲۵۶ فیلتر کانولوشنی به‌اندازه 9×9 و گام حرکتی ۲ است که روی آن تابع فعال‌ساز ReLU اعمال شده است. در این ساختار رویکرد همچنین، کپسول‌هایی به‌اندازه ۸ با ۳۲ فیلتر 9×9 و گام حرکتی ۲ به‌کار رفت و میزان Dropout، در تمامی لایه‌های شبکه، 0.5 در نظر گرفته شد. تمامی پارامترهای مدل با استفاده از Optuna^۲ به‌دست آمده‌اند. Optuna کتابخانه‌ای متن‌باز برای

بهبود طبقه‌بندی تصاویر ابرطیفی ...

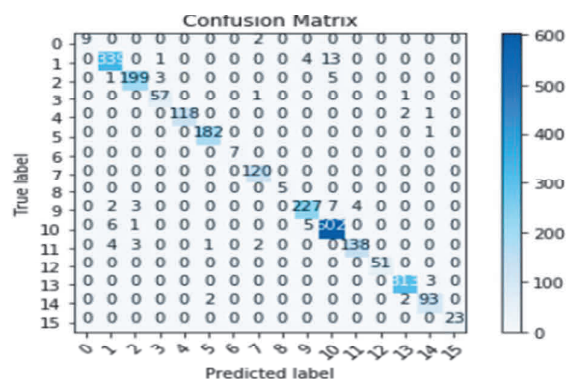


شکل ۱۰. نمودار Recall (الف)؛ نمودار Accuracy (ب)؛ نمودار Precision (ج)؛ نمودار F1 روی داده‌های آموزش و تست با رویکرد پیشنهادی اول (د)

جدول ۲. نتایج حاصل از ساختار پیشنهادی اول روی مجموعه داده‌های تست

Count	معیارهای ارزیابی			شماره کلاس
	F1-score	Recall	Precision	
۱۱	۹۰	۸۲	۱۰۰	۱
۳۵۷	۹۶	۹۵	۹۶	۲
۲۰۸	۹۶	۹۶	۹۷	۳
۵۹	۹۵	۹۷	۹۳	۴
۱۲۱	۹۹	۹۸	۱۰۰	۵
۱۸۳	۹۹	۱۰۰	۹۹	۶
۷	۱۰۰	۱۰۰	۱۰۰	۷
۱۲۰	۹۸	۱۰۰	۹۶	۸
۵	۱۰۰	۱۰۰	۱۰۰	۹
۲۴۳	۹۵	۹۴	۹۶	۱۰
۶۱۴	۹۷	۹۸	۹۶	۱۱
۱۴۸	۹۵	۹۳	۹۷	۱۲
۵۱	۱۰۰	۱۰۰	۹۸	۱۳
۳۱۶	۹۹	۹۹	۹۸	۱۴
۹۷	۹۵	۹۶	۹۵	۱۵
۲۳	۱۰۰	۱۰۰	۱۰۰	۱۶
۲۵۶۳	۹۶/۸۶	۹۶/۸۷	۹۶/۸۸	میانگین کل

پویا احمدی و همکاران



شکل ۱۱. ماتریس خطای^۱ رویکرد اول

جدول ۳. مقایسه نتایج حاصل از ساختار پیشنهادی اول با نتایج موجود (مقادیر برحسب درصد است)

RF-200	SVM-RBF	RNN-LSTM	RNN-GRU- than	RNN-GRU- Pretanh	رویکرد اول	شماره کلاس
۵۸/۸۴	۶۰/۷۷	۴۶/۰۳	۶۸/۹۳	۷۰/۵۹	۹۷/۵۶	۱
۵۸/۴۲	۷۷/۶۸	۶۱/۳۷	۴۰/۹۴	۷۰/۲۸	۹۷/۵۲	۲
۸۲/۶۱	۷۹/۳۵	۸۶/۹۶	۷۸/۸۰	۸۱/۵۲	۸۵/۳۳	۳
۸۵/۹۱	۹۱/۰۵	۸۷/۰۲	۸۷/۹۲	۹۰/۱۶	۸۷/۹۲	۴
۸۰/۴۹	۸۴/۳۶	۸۶/۶۶	۸۷/۵۲	۹۱/۹۷	۸۵/۲۲	۵
۹۴/۷۶	۹۲/۰۳	۷۹/۴۹	۹۷/۲۷	۹۶/۱۳	۹۷/۴۹	۶
۷۷/۳۴	۶۹/۶۱	۵۹/۶۹	۸۲/۷۹	۸۴/۷۵	۷۴/۶۲	۷
۵۹/۴۳	۵۹/۳۱	۶۴/۸۹	۵۰/۵۸	۵۹/۶۴	۶۷/۹۹	۸
۶۳/۴۸	۷۹/۶۱	۶۰/۴۶	۷۹/۴۳	۸۶/۱۷	۸۷/۵۸	۹
۵۹/۰۶	۹۷/۵۳	۹۸/۷۷	۹۸/۷۷	۹۹/۳۸	۹۸/۷۷	۱۰
۸۸/۲۶	۸۵/۲۱	۷۵/۳۲	۸۴/۷۳	۸۴/۹۷	۸۷/۶۲	۱۱
۵۴/۸۵	۶۳/۶۴	۷۱/۸۲	۶۱/۲۱	۷۷/۵۸	۷۲/۴۲	۱۲
۹۷/۷۸	۱۰۰	۹۱/۱۱	۸۸/۸۹	۹۵/۵۶	۹۳/۳۳	۱۳
۵۸/۹۷	۸۷/۱۸	۷۹/۴۹	۷۹/۴۹	۸۴/۶۲	۷۱/۷۹	۱۴
۸۱/۸۲	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۱۵
۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۶

است با استفاده از شبکه‌های کپسول، تاحد بسیاری از این مشکلات کاسته شود. در هر دو رویکرد پیشنهادی، برای کنترل Overfitting از Early Stopping استفاده شده است. این سیاست، طی دو دور اجرایی، بررسی می‌کند که اگر مقدار Loss تغییری نکرده باشد، یادگیری را متوقف کند. برای یادگیری پارامترهای هر دو مدل نیز، Optuna به کار رفت و بهینه‌ترین پارامترها برای طبقه‌بندی انتخاب شد.

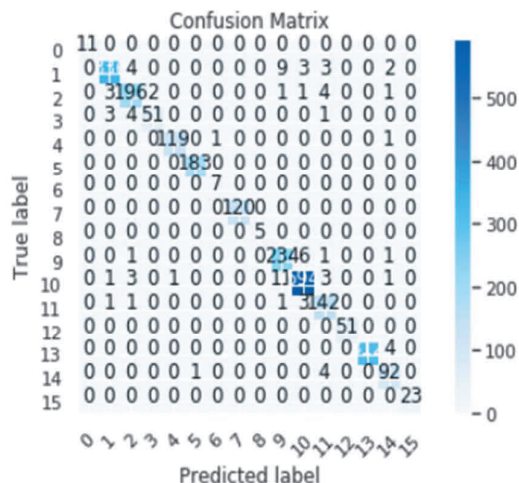
مقادیر ساختار اول، در مقایسه با رویکردهای قبلی، دقت شایان توجهی به دست آورده است. دقت این مدل، در بیشتر دامنه‌ها، کمتر از رویکرد RNN-GRU-Pretanh است و همچنین فقط توانسته روی داده‌های کلاس Buildings-grass-trees دقت برای کسب کند. این درحالی است که در سایر طبقه‌ها، میزان دقت نسبتاً کمتر بوده است. این مدل به دلیل داشتن عمق و تعداد فیلتر کمتر، در طبقه‌بندی تصاویر، دقت کمتری داشته از این‌رو در رویکرد پیشنهادی دوم، سعی شده

نتایج خروجی از الگوریتم روی مجموعه داده‌های تست، پس از سی بار یادگیری شبکه را بیان می‌کند که به‌ازای هر کلاس، گزارش داده شده است. در شکل ۱۲، ماتریس خطای مرتبط با رویکرد دوم ارائه شده است. نمودار خطا و سایر متریک‌های اندازه‌گیری در شکل ۱۳ آورده شده است.

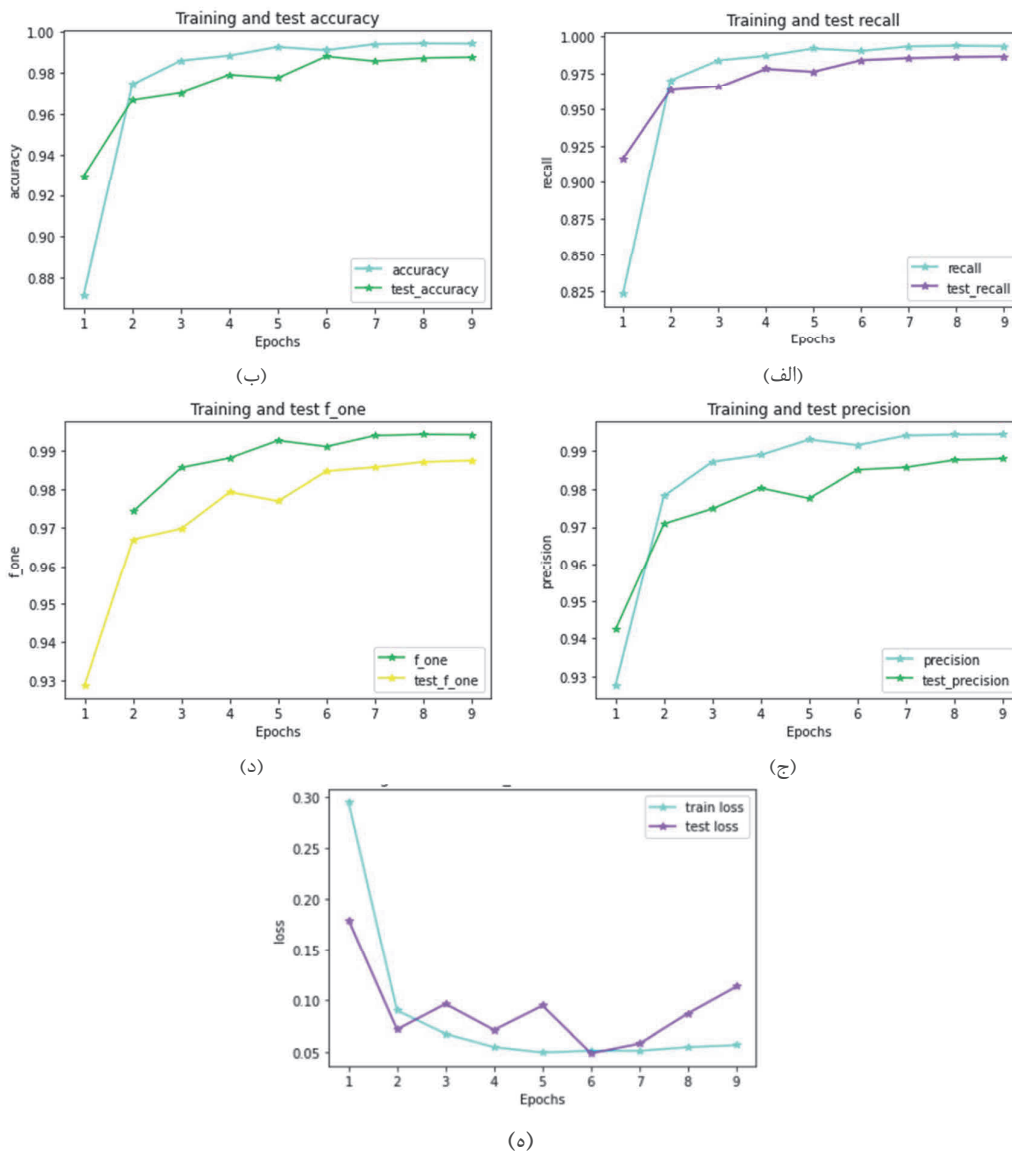
۴-۲- نتایج رویکرد پیشنهادی دوم شبکه‌های کپسول در رویکرد پیشنهادی دوم نیز، از ۸۰٪ داده‌ها برای آموزش و از ۲۰٪ آنها برای تست استفاده شده است. مشابه رویکرد پیشنهادی اول، نتایج در دو حالت (به‌ازای هر کلاس مجزا، به‌ازای تمامی حالت‌ها) برای مقایسه با سایر رویکردها گزارش شده است. جدول ۴

جدول ۴. نتیجه حاصل از رویکرد پیشنهادی دوم روی مجموعه داده‌های تست

Count	معیارهای ارزیابی			شماره کلاس
	F1-score	Recall	Precision	
۱۱	۱۰۰	۱۰۰	۱۰۰	۱
۳۵۷	۹۶	۹۴	۹۸	۲
۲۰۸	۹۴	۹۴	۹۴	۳
۵۹	۹۱	۸۷	۹۶	۴
۱۲۱	۹۹	۹۸	۹۹	۵
۱۸۳	۱۰۰	۱۰۰	۹۹	۶
۷	۹۳	۱۰۰	۸۸	۷
۱۲۰	۱۰۰	۱۰۰	۱۰۰	۸
۵	۱۰۰	۱۰۰	۱۰۰	۹
۲۴۳	۹۴	۹۶	۹۲	۱۰
۶۱۴	۹۷	۹۷	۹۸	۱۱
۱۴۸	۹۳	۹۶	۹۰	۱۲
۵۱	۱۰۰	۱۰۰	۱۰۰	۱۳
۳۱۶	۹۹	۹۹	۱۰۰	۱۴
۹۷	۹۲	۹۵	۹۰	۱۵
۲۳	۱۰۰	۱۰۰	۱۰۰	۱۶
۲۵۶۳	۹۶/۶۲	۹۶/۶۰	۹۶/۶۸	میانگین کل



شکل ۱۲. ماتریس خطای رویکرد دوم



شکل ۱۳. نمودار Recall (الف)؛ نمودار Accuracy (ب)؛ نمودار Precision (ج)؛ F1 (د)؛ نمودار loss روی داده‌های آموزش و تست، با رویکرد پیشنهادی دوم (ه)

تأثیری در هیچ‌یک از معیارهای ارزیابی ندارد و هر چهار معیار، حدوداً پس از هفت مرحله یادگیری شبکه، بدون تغییر باقی می‌مانند. در جدول ۵، مقایسه دقت‌های به‌دست‌آمده از طریق رویکرد پیشنهادی دوم، با نتایج حاصل از دیگر پژوهش‌ها درباره این مجموعه از داده‌ها، آمده است.

منحنی‌های حاصل، با رویکرد پیشنهادی دوم، کاملاً متفاوت از منحنی‌های حاصل از رویکرد پیشنهادی اول است. در این رویکرد، اختلاف بین دو منحنی آموزش و تست در تمامی حالات، کمتر شده است و این منحنی‌ها به مجانبی افقی میل کرده‌اند. این نکته نشان می‌دهد افزایش تعداد یادگیری شبکه، از مقداری معین،

جدول ۵. مقایسه نتایج حاصل از ساختار پیشنهادی دوم، با نتایج سایر پژوهشگران

شماره کلاس	رویکرد اول	رویکرد دوم	RNN-GRU-Pretanh	RNN-GRU-than	RNN-LSTM	SVM-RBF	RF-2000
۱	۹۷/۵۶	۹۵/۷۱	۷۰/۵۹	۶۸/۹۳	۶۴/۰۳	۶۰/۷۷	۵۴/۸۴
۲	۹۷/۵۲	۸۲/۷۷	۷۰/۲۸	۴۰/۹۴	۶۱/۳۷	۷۷/۶۸	۵۸/۴۲
۳	۸۵/۳۳	۸۶/۵۲	۸۱/۵۲	۷۸/۸۰	۸۶/۹۶	۷۹/۳۵	۸۲/۶۱
۴	۸۷/۹۲	۹۱/۹۳	۹۰/۱۶	۸۷/۹۲	۸۷/۰۲	۹۱/۰۵	۸۵/۹۱
۵	۸۵/۲۲	۹۲/۹۷	۹۱/۹۷	۸۷/۵۲	۸۶/۶۶	۸۴/۳۶	۸۰/۴۹
۶	۹۷/۴۹	۹۸/۱۲	۹۶/۱۳	۹۷/۲۷	۷۹/۴۹	۹۲/۰۳	۹۴/۷۶
۷	۷۴/۶۲	۸۴/۷۰	۸۴/۷۵	۸۲/۷۹	۵۹/۶۹	۶۹/۶۱	۷۷/۳۴
۸	۶۷/۹۹	۶۷/۶۴	۵۹/۶۴	۵۰/۵۸	۶۴/۸۹	۵۹/۳۱	۵۹/۴۳
۹	۸۷/۵۸	۷۶/۹۰	۸۶/۱۷	۷۹/۴۳	۶۰/۴۶	۷۹/۶۱	۶۳/۴۸
۱۰	۹۸/۷۷	۹۹/۶۰	۹۹/۳۸	۹۸/۷۷	۹۸/۷۷	۹۷/۵۳	۵۹/۰۶
۱۱	۸۷/۶۲	۸۵/۹۰	۸۴/۹۷	۸۴/۷۳	۷۵/۳۲	۸۵/۲۱	۸۸/۲۶
۱۲	۷۲/۴۲	۷۸/۵۹	۷۷/۵۸	۶۱/۲۱	۷۱/۸۲	۶۳/۶۴	۵۴/۸۵
۱۳	۹۳/۳۳	۹۶/۵۰	۹۵/۵۶	۸۸/۸۹	۹۱/۱۱	۱۰۰	۹۷/۷۸
۱۴	۷۱/۷۹	۷۹/۶۲	۸۴/۶۲	۷۹/۴۹	۷۹/۴۹	۷۸/۱۸	۵۸/۹۷
۱۵	۹۰/۹۱	۹۱/۹۱	۹۱/۹۱	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۸۱/۸۲
۱۶	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰

الگوریتم‌های طبقه‌بندی و افزایش قدرت تفکیک طیفی تصاویر، با وجود افزایش کلی دقت، چالش‌های طبقه‌بندی این تصاویر نیز بیشتر شده است.

یکی از مهم‌ترین مشکلات طبقه‌بندی این تصاویر، در مدل‌های یادگیری ماشین، انتخاب ویژگی‌هاست زیرا این مدل‌ها نیاز به انتخاب ویژگی به صورت دستی دارند و هرچه ویژگی‌های مناسب‌تری استخراج شود، این مدل‌ها می‌توانند طبقه‌بندی بهتری انجام دهند. از این رو با یادگیری عمیق، به‌منزله مدل‌هایی که می‌توانند ویژگی‌ها را به صورت خودکار انتخاب کنند، جایگزین مناسبی برای طبقه‌بندی این تصاویر شمرده می‌شوند. در این نوع یادگیری، ویژگی‌ها براساس فیلترهایی استخراج می‌شود که کاربر هیچ دخالتی در آن ندارد. برای افزایش دقت طبقه‌بندی این تصاویر، تکنیک‌های یادگیری عمیق شبکه‌های کانولوشنی روی این نوع تصاویر اعمال می‌شود. در پژوهش حاضر، از روش یادگیری عمیق شبکه‌های عصبی کانولوشنی و نیز شبکه‌های کپسول، برای افزایش دقت طبقه‌بندی تصاویر ابرطیفی استفاده شد. ابعاد تصویر ورودی به

نتایج رویکرد پیشنهادی دوم، در بیشتر کلاس‌ها، بهتر از رویکرد اول بوده و این ساختار توانسته است دقت بیشتر از RNN-GRU-Pretanh به دست آورد؛ به این دلیل که رویکرد دوم شبکه‌های کپسول، با استخراج مطلوب ویژگی‌ها، توانسته است مهندسی ویژگی خودکار مناسبی برای مدل استاندارد XGBoost ایجاد کند. همچنین با استفاده از رویکرد دوم در دامنه‌های Alfalfa، Corn-min، Grass-pasture، Soybean-clean، Soybean-notill، Grass-trees و Wheat Buildings-grass-trees دقت بیشتری حاصل شده است. ساختار دوم در سایر کلاس‌ها نیز دقت مورد قبولی داشته و البته علت اندک بودن دقت این ساختار، روی این طبقه‌ها، کم بودن تعداد نمونه‌ها در آنهاست که در نتیجه آن، مدل به خوبی نتوانسته است این داده را آموزش ببیند و یادگیری را کامل کند.

۵- بحث و نتیجه‌گیری

طبقه‌بندی تصاویر ماهواره‌ای همواره با چالش‌های بسیاری روبه‌رو بوده است. با پیشرفت و گسترش

- Information to Select Symbols**, Scientific - Research Quarterly of Geographical Data (SEPEHR), 25(98), PP. 5-14.
- Bengio, Y. & Ian, J., 2015, **Goodfellow, and Aaron Courville. "Deep Learning"**, Nature, 521.7553, PP. 436-444.
- Bruzzone, L., Chi, M. & Marconconi, M., 2006, **A Novel Transductive SVM for Semisupervised Classification of Remote-Sensing Images**, IEEE Transactions on Geoscience and Remote Sensing, 44(11), PP. 3363-3373.
- Candemir, S., Jaeger, S., Palaniappan, K., Musco, J.P., Singh, R.K., Xue, Z. & McDonald, C.J., 2014, **Lung Segmentation in Chest Radiographs Using Anatomical Atlases with Nonrigid Registration**, IEEE Transactions on Medical Imaging, 33(2), PP. 577-590.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Vila-Frances, J. & Calpe-Maravilla, J., 2006, **Composite Kernels for Hyperspectral Image Classification**, IEEE Geoscience and Remote Sensing Letters, 3(1), PP. 93-97.
- Camps-Valls, G., Shervashidze, N. & Borgwardt, K.M., 2010, **Spatio-Spectral Remote Sensing Image Classification with Graph Kernels**, IEEE Geoscience and Remote Sensing Letters, 7(4), PP. 741-745.
- Chen, T. & Guestrin, C., 2016, **Xgboost: A Scalable Tree Boosting System**, In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining.
- Dehghani, H., 2006, **Classification of Distance Measurement Images with Large Dimensions and Limited Educational Examples**, Ph.D. Thesis, Tarbiyat Modarres University, Department of Electronic Engineering, Tehran, Iran.
- Domingos, P., 2012, **A Few Useful Things to Know about Machine Learning**, Communications of the ACM.
- Du, P., Tan, K., Zhang, W. & Yan, Zh., 2008, **ANN Classification of OMIS Hyperspectral Remotely Sensed Imagery: Experiments and Analysis**, Congress on Image and Signal Processing, IEEE.

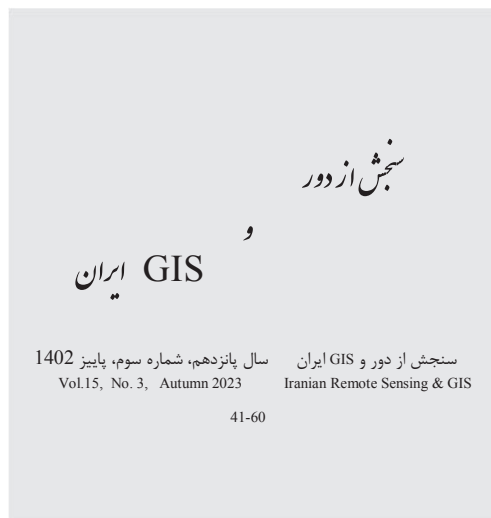
شبکه از عوامل تأثیرگذار در عملکرد استخراج ویژگی با استفاده از شبکه‌های کپسول است. در پژوهش حاضر، ابعاد $5 \times 5 \times 30$ در نظر گرفته شد. سه معیار Precision، Recall و F1-Score برای ارزیابی نتایج به کار رفت و نتایج، به صورت میانگین و انحراف معیار برای بیست بار اجرای شبکه، تهیه شد. از مجموع داده‌های در نظر گرفته شده، ۸۰٪ برای آموزش طبقه‌بندی و ۲۰٪ برای ارزیابی طبقه‌بندی استفاده شد. به منظور بهبود دقت طبقه‌بندی تصاویر نیز، دو رویکرد متفاوت در زمینه استفاده از شبکه‌های عصبی کانولوشنی و شبکه کپسول، در طبقه‌بندی تصاویر هوایی ابرطیفی به کار رفت. یکی از مهم‌ترین تفاوت‌های شبکه‌های کپسول با شبکه‌های کانولوشنی در ذخیره داده‌ها، در قالب یک بردار است. ویژگی بردار بودن ویژگی اندازه و جهت را نیز به همراه خواهد داشت و می‌توان ارتباط بین ویژگی‌ها را نیز استخراج کرد و به دقت بالاتری رسید. در رویکرد اول، از این شبکه‌ها در حالت یکپارچه برای استخراج ویژگی و در ادامه، طبقه‌بندی استفاده شد و در رویکرد دوم، ویژگی‌های استخراج شده از طریق شبکه‌های کپسول، به منزله ورودی، به مدل درخت تقویتی XGBoost برای طبقه‌بندی معرفی شد. در رویکرد اول، از دو لایه کانولوشنی و دو لایه تماماً متصل استفاده شد که حصول دقت ۹۸٪ روی داده‌های آموزش و دقت ۹۶/۳۰٪ روی داده‌های تست را در پی داشت. در رویکرد دوم، از یک لایه کانولوشنی، یک لایه کپسول و XGBoost استفاده شد که به دقت ۹۹٪ روی داده‌های آموزش و نیز دقت ۹۷/۵٪ روی داده‌های تست انجامید. رویکرد پیشنهادی اول فقط در یکی از طبقه‌ها توانسته است دقت برابر با رویکرد برتر در سایر پژوهش‌ها به دست آورد؛ در صورتی که رویکرد دوم، در بیشتر موارد، توانسته است به دقت بالاتر دست یابد.

۶- منابع

- Akbari, D., Safari, A.R. & Homayouni, S., 2016, **Improving Spectral-Spatial Classification of Supercritical Images by Using Spatial**

- Fauvel, M., Chanussot, J., Benediktsson, J.A. & Sveinsson, J., 2007, **Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles**, IEEE Transactions on Geoscience and Remote Sensing, 46(11), PP. 2012-2020.
- Fukunaga, A. & Sber, G., 2008, **Providing an Optimal Method Based on Deep Learning for Spectral Classification Images with High Resolution Spatial Resolution in Semi-Urban Areas**, Journal of Geomatics Science and Technology, 9(2), PP. 151-170.
- Gao, Y., Gao, F., Dong, J. & Li, H.C., 2020, **SAR Image Change Detection Based on Multiscale Capsule Network**, IEEE Geosci. Remote Sens. Lett., 18(3).
- Ghaffari, O., Voldan-Zoj, M.J. & Mokhtarzadeh, M., 2016, **Selecting the Band in Order to Optimize the Spectral Separation of Supercritical Images**, The 1st National Conference on Geospatial Information Technology, K.N. Toosi University of Technology, Tehran, Iran.
- Ghamisi, P., Plaza, J., Chen, Y., Li, J. & Plaza, A.J., 2017, **Advanced Spectral Classifiers for Hyperspectral Images: A Review**, IEEE Geoscience and Remote Sensing Magazine, 5(1), PP. 8-32.
- Ghassemian, H., Keshavarz, A. & Landgrebe, D., 2003, **Hyper-Spectral Image Processing and Analyses**, Space Magazine, 1(3), PP. 32-41.
- Gualtieri, J.A. & Chetri, S.R., 2000, **Support Vector Machines for Classification of Hyperspectral Data**, Proc. IGARSS, Honolulu, HI, PP. 813-815.
- Gualtieri, J.A. & Cromp, R.F., 1999, **Support Vector Machines for Hyperspectral Remote Sensing Classification**, Proceedings Vol. 3584, 27th AIPR Workshop: Advances in Computer-Assisted Recognition.
- Gualtieri, J.A., Chetri, S.R., Cromp, R.F. & Johnson, L.F., 1999, **Support Vector Machines Classifiers as Applied to AVIRIS Data**, In Summaries 8th JPL Airborne Earth Science Workshop, JPL Pub., 99-17, PP. 217-227.
- Hu, W., Huang, Y., Wei, L., Zhang, F. & Li, H., 2015, **Deep Convolutional Neural Networks for Hyperspectral Image Classification**, J. Sens, 2015, P. 258619.
- Jimenez, L. & Landgrebe, D.A., 1999, **Hyperspectral Data Analysis and Feature Reduction via Projection Pursuit**, IEEE Trans. On Geoscience and Remote Sensing, 37(6), PP. 2653-2667.
- Kaewpajit, S., Moigne, J.L., Ghazawi, T.E., 2003, **Automatic Wavelet Spectral Analysis for Reduction of Hyperspectral Imagery**, IEEE Trans. on Geoscience and Remote Sensing, 41(4), PP. 863-871.
- Keshavarz, A. & Ghasemiyah, H., 2005, **A High-Speed Vector Machine-Based Algorithm for Classifying Hyperspectral Images Using Spatial Correlation**, Iranian Journal of Electrical and Computer Engineering, 3(1), PP. 37-44.
- Landgrebe, D.A., 2002, **Hyperspectral Image Data Analysis**, IEEE Signal Processing Magazine, 19(1), PP. 17-28.
- LeCun, Y., Bengio, Y. & Hinton, G., 2015, **Deep Learning**, Nature, 521, PP. 436-444.
- Lee, C. & Landgrebe, D.A., 1993, **Feature Extraction Based on Decision Boundaries**, IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4), PP. 388-400.
- Li, H., 2018, **Deep Learning for Natural Language Processing: Advantages and Challenges**, National Science Review., 5(1), PP. 24-26.
- Liyang, W., Yongyi, Y., Nishikawa, R.M., Wernick, M.N. & Edwards, A., 2005, **Relevance Vector Machine for Automatic Detection of Clustered Micro Calcifications**, IEEE Trans. Med. Imag., 24(10), PP. 1278-1285.
- Luo, Y., Zou, J., Yao, C., Li, T. & Bai, G., 2018, **HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image.**, arXiv 2018, arXiv:1802.10478.
- Maggiori, E., Tarabalka, Y., Charpiat, G. & Alliez, P., 2016, **High-Resolution Semantic Labeling with Convolutional Neural Networks**, ArXiv: 1611.01962.

- Melgani, F. & Bruzzone, L., 2004, **Classification of Hyperspectral Remote Sensing Images with Support Vector Machines**, IEEE Transactions on Geoscience and Remote Sensing, 42(8), PP.1778-1790.
- Paoletti, M.E., Plaza, J. & Plaza, A., 2018, **A New Deep Convolutional Neural Network for Fast Hyperspectral Image Classification**, ISPRS Journal of Photogrammetry and Remote Sensing, 145, PP. 120-147. doi:10.1016/j.isprsjprs.2017.11.021.
- Paoletti, M.E., Haut, J.M., Fernandez-Beltran, R., Plaza, J., Plaza, A., Jun, L. & Pla, F., 2019, **Capsul Networks for Hyperspectral Image Classification**, IEEE Transactions on Geoscience and Remote Sensing, 57(4), PP. 2145-2160. doi:10.1109/TGRS.2018.2871782.
- Sabour, S., Frosst, N. & Hinton, G.E., 2017, **Dynamic Routing between Capsules**, In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4-9 December 2017, PP. 3859-3869.
- Shahhoseeini, R., 2009, **Evaluation of Support Vector Machines in the Classification of Hyperspectral Remote Sensing Data**, Ms.c Thesis, University of Tehran, Tehran, Iran.
- Wang, Y., Sun, A., Han, J., Liu, Y. & Zhu, X., 2018, **Sentiment Analysis by Capsules**, In Proceedings of the 2018 World Wide Web Conference on World Wide Web, Lyon, France, 23-27 April 2018; PP. 1165-1174.
- Xu, J.L., Esquerre, C. & Sun, D.W., 2018. **Methods for Performing Dimensionality Reduction in Hyperspectral Image Classification**, Journal of Near Infrared Spectroscopy, 26(1), PP. 61-75.
- Xu, Q., Wang, D.Y. & Luo, B., 2020, **Faster Multiscale Capsule Network with Octave Convolution for Hyperspectral Image Classification**, IEEE Geosci. Remote Sens. Lett, 18(2).
- Xue, Z., You, D., Candemir, S., Jaeger, S., Antani, S., Long, L.R. & Thoma, G.R., 2015, **Chest X-Ray Image View Classification**, In Proceedings of the 28th International Symposium on Computer-Based Medical Systems (CBMS) (PP. 66-71), Brazil: Ribeirão Preto.
- Yu, Y., Gu, T., Guan, H., Li, D. & Jin, S., 2019, **Vehicle Detection from High-Resolution Remote Sensing Imagery Using Convolutional Capsule Networks**, IEEE Geosci. Remote Sens. Lett., 16, PP. 1894-1898.



Improving the Classification of Hyperspectral Images Using the Combined Model of CapsNet and the Extreme Gradient Boosting

Ahmadi P.^{1*}, Managhebi T.², Ebadi H.³, Asghari Beirami B.⁴

1. Ph.D. Student, Faculty of Geomatics Engineering, K.N. Toosi University of Technology, Tehran, Iran
2. Assistant Professor, Faculty of Mapping and Spatial Information Engineering, University of Tehran, Tehran, Iran
3. Prof. Photogrammetry & Remote Sensing Dept, K.N. Toosi University of Technology, Tehran, Iran
4. Ph. D. in Remote sensing engineering, Faculty of Geomatics Engineering, K.N. Toosi University of Technology, Tehran, Iran

Abstract

With the development of remote sensing science, the use of hyperspectral images is becoming more widespread. Classification is one of the most popular topics in hyperspectral remote sensing. In the last two decades, a number of methods have been proposed to address the problem of hyperspectral data classification. In the present study, a structure based on learning capsule networks has been used to classify hyperspectral images, so that the network structure can have the most optimal generation of features by using a convolution layer and a capsule layer, and at the same time Avoid overfitting the on training data. The obtained results show the high quality of production features in the proposed structure.

In order to improve the classification accuracy, the feature extraction approach through the designed network and the classification by the Extreme Gradient Boosting was compared with the classification method by the global deep network. The proposed capsule approach consists of 3 basic layers: 1) Prime caps, which are capsules of size 8 and 32 with 9×9 filters and movement step 2, 2) Digitcaps with 10 16-dimensional capsules, and 3) fully connected layer. The results of examining two approaches for deep networking as well as combining capsule networks with XGBoost reinforcement tree algorithm were compared. Approaches such as SVM, RF-200, LSTM, GRU and GRU-Pretanh were considered to compare the proposed approach based on the configurations mentioned in their research.

Up in addition to the study and quality measurement of production vector deep features by the proposed method in different classifiers, the ability of deep global networks in the application of classification should also be examined. The results of examining two approaches for deep network and also combining CapsNet with XGBoost show that by using the proposed combined method, images are classified with 99% accuracy on training data and 97.5% accuracy on test data.

Keywords: Classification, Hyperspectral images, CapsNet, XGBoost, Fusion model.

* Correspondence Address: No. 1346, ValiAsr Street, Mirdamad Cross, Faculty of Geomatics Engineering, K.N. Toosi University of Technology, Tehran. Tel: 09185363604
 Email: ahmadi.pouya@ut.ac.ir
<https://orcid.org/0009-0001-7308-1054>
<https://dx.doi.org/10.48308/gisj.2023.102347>



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).