نشریه علمی مهندسی فناوری اطلاعات مکانی

سال نهم، شماره چهارم، زمستان ۱۴۰۰

Vol.9, No.4, Winter 2022

۱۰۹ -۱۲۵

مقاله پژوهشی

DOR: *20.1001.1.20089635.1400.9.4.2.5*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]

[DOR: 20.1001.1.20089635.1400.9.4.2.5]

[DOI: 10.52547/jgit.9.4.109]

بهبود دقت طبقه‌بندی تصاویر ماهواره‌ای ابرطیفی با استفاده از توسعه یک شبکه عصبی کانولوشن و یادگیری عمیق

سمیه محمودی[۱]، نجمه نیسانی سامانی[۲*]، آرا تومانیان[۲]

۱- دانشجوی دکتری سنجش از دور و سیستم‌های اطلاعات جغرافیایی، دانشکده جغرافیا، دانشگاه تهران

۲- دانشیار گروه سنجش از دور و سیستم‌های اطلاعات جغرافیایی، دانشکده جغرافیا، دانشگاه تهران

تاریخ دریافت مقاله: ۱۳۹۹/۱۰/۲۵ تاریخ پذیرش مقاله: ۱۴۰۰/۱۱/۱۵

چکیده

امروزه با پیشرفت تکنولوژی، سنجنده‌های متعدد، تصاویر ابرطیفی با رزولوشن طیفی و مکانی بالا تهیه می‌کنند. تا کنون روش‌های متعددی برای طبقه‌بندی تصاویر ابرطیفی ارائه شده‌اند که هرکدام سعی دارند بر تعدادی از چالش‌های محاسباتی و پردازشی داده‌های ابرطیفی فائق آیند. با افزایش تعداد پارامترها که در نتیجه افزایش تعداد لایه‌ها در شبکه‌های عصبی مصنوعی رخ می‌دهد، کارآیی آنها را در حل مسائل پیچیده مانند طبقه‌بندی تصاویر ابرطیفی کاهش می‌دهد. در سال‌های اخیر، مفهوم یادگیری عمیق و به ویژه شبکه‌های عصبی کانولوشن، به دلیل تولید خودکار ویژگی‌ها و کاهش پارامترها نسبت به شبکه‌های عصبی چندلایه پرسپترون با به اشتراک گذاشتن پارامترها در هر لایه، بسیار مورد توجه محققان در حوزه شناسایی الگو قرار گرفته است. هدف کلی تحقیق پیش رو، توسعه یک شبکه عصبی کانولوشن برای طبقه-بندی تصاویر ابرطیفی است. نوآوری اصلی این مقاله ارائه چارچوبی در راستای بکارگیری یادگیری عمیق شامل چهار مرحله می‌باشد. مرحله اول) کاهش بعد با استفاده از روش زیرفضا، مرحله دوم) آماده‌سازی ورودی‌های شبکه عصبی کانولوشن، مرحله سوم) غنی‌سازی داده‌های آموزشی، مرحله چهارم) طراحی معماری شبکه عصبی کانولوشن. پیاده‌سازی چارچوب پیشنهادی بر روی داده‌های بنچ مارک دانشگاه پاویا، علی‌رغم بکارگیری تعداد محدودی داده آموزشی، موجب حصول صحت طبقه‌بندی ۹۸/۳ درصد شد.

کلید واژه‌ها : تصاویر ابر طیفی، یادگیری عمیق، شبکه عصبی کانولوشن، طبقه‌بندی.

* نویسنده مکاتبه کننده: تهران – خیابان انقلاب - خیابان وصال شیرازی – نبش کوچه آذین – دانشکده جغرافیا – گروه سنجش از دور و سیستم اطلاعات جغرافیایی.

تلفن: ۰۲۱۶۱۱۱۳۵۴۰

Email: nneysani@ut.ac.ir

۱- مقدمه

فناوری سنجش‌ازدور ابرطیفی در دو دهه گذشته پیشرفت چشم‌گیری داشته است. با توجه به پیشرفت صنایع الکترونیک و الکترواپتیک، امکان ساخت قطعات الکترونیکی با ابعاد بسیار کوچک و سیستم‌های پراکنش طیفی دقیق و نیز توسعه برنامه‌های سازمان‌های فضایی، طراحی و ساخت سنجنده‌های با توان تفکیک مکانی و طیفی مناسب به صورت توأم امکان‌پذیر شده است. همچنین با توسعه روزافزون سیستم‌های بدون سرنشین و تمایل کمپانی‌ها و کاربران به استفاده از پهپادها، دسترسی به داده‌های ابرطیفی با دقت مکانی و طیفی مناسب، آسان شده است. لذا چالش‌های محاسباتی اخیر، بر روی مهم‌ترین روش استخراج اطلاعات از تصاویر ابرطیفی، یعنی طبقه‌بندی، متمرکز شده است. وجود مسائل حل نشده و بسیاری از چالش-های باقی مانده از قبل و همچنین ظهور کاربردهای جدید سبب شده که همواره دستیابی به الگوریتم‌های جدید یک نیاز اساسی به شمار آید[۱ و ۲]. در رابطه با موضوع طبقه‌بندی تصاویر ابرطیفی چهار جنبه مهم مطرح است: ۱) ابعاد بالای باندهای طیفی۲) تغییرات منحنی بازتاب طیفی پیکسل‌های همجوار یا پیکسل های متعلق به یک کلاس مشخص ۳) هزینه بالای جمع‌آوری داده‌های آموزشی ۴) کیفیت داده‌ها[۳]. ابعاد بالای باندهای طیفی و کمبود تعداد نمونه‌های آموزشی سبب بروز پدیده هیوز[1] در طبقه‌بندی تصاویر ابرطیفی و در نتیجه بروز مشکل بیش‌برازش[2] می-شود[۴و ۵]. وابستگی بین باندهای طیفی از مشکلات اساسی است که باعث می‌شود نسبت تعداد نمونه‌های آموزشی به تعداد باندهای طیفی کم شود و این در حالی است‌که برخی از باندها به یکدیگر وابسته‌اند و اطلاعات متمایزکننده‌ای را در اختیار سیستم شناسایی الگو قرار نمی‌دهند. آنالیز تصاویر ابرطیفی، علاوه بر چالش‌های محاسباتی مذکور، چالش‌های پردازشی ناشی از محدودیت امکانات سخت‌افزاری را هم پیش روی محققان قرار داده است[۵ و ۷]. تا کنون روش‌های متعددی برای طبقه‌بندی تصاویر ابرطیفی ارائه شده‌اند که هرکدام سعی دارند بر تعدادی از مشکلات طبقه-بندی تصاویر ابر طیفی غلبه نمایند. ظهور یادگیری عمیق[3] در زمینه‌های تحقیقاتی مختلف به دلیل در دسترس قرارگرفتن داده‌های بزرگ و افزایش قدرت محاسباتی امکان‌پذیر شده است، به عنوان نمونه می-توان به پروژه دانشگاه برلین تحت عنوان شبکه زمین بزرگ اشاره کرد که در آن حدود ۵۹۰ هزار نمونه‌های آموزشی در ۱۹ کلاس مختلف از داده‌های سنتینل ۲ تهیه شده و برای اهداف طبقه‌بندی از آن استفاده می-شود *[۳۰]*. حجم داده‌های این پروژه حدود ۶۰ گیگابایت و به رایگان قابل استفاده برای محققان است. در پروژه دیگری محققان برای آنالیز داده‌ها از منابع مختلف (ابرطیفی، لایدار و چندطیفی) مجموعه داده مبنایی برای تلفیق و بارزسازی تصاویر ارائه دادند که این مجموعه داده نیز رایگان است *[۳۱]*. پیشرفت‌های اخیر و سیاست دسترسی رایگان به این داده ها پتانسیل مطالعات یادگیری عمیق را برای نظارت بر سطح زمین از طریق تولید نقشه‌های طبقه‌بندی و کاربری اراضی فراهم می‌کند*[۲]*. در سال‌های اخیر مفهوم یادگیری عمیق به دلیل تولید خودکار ویژگی‌ها به کمک نمونه‌های آموزشی محدود، بسیار مورد توجه محققان در حوزه استخراج اطلاعات به کمک روش‌های طبقه‌بندی قرار گرفته است. از دیگر مزایای ویژگی‌های تولید شده با تئوری یادگیری عمیق، غیرخطی بودن آنهاست که در طبقه‌بندی کلاس‌هایی که مرز غیرخطی دارند مفید است[۸]. در یک سیستم شناسایی الگو، مهمترین مرحله در تولید خروجی بهینه، تولید ویژگی مناسب و با کیفیت(تفکیک‌پذیر) است[۹]. روش‌های تولید ویژگی تا قبل از ظهور یادگیری عمیق به شکل

[1] *Hughes*
[2] *Overfitting*
[3] *Deep Learning*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

دستی و توسط کارشناس خبره انجام می‌شد، اما سیستم‌های مبتنی بر یادگیری عمیق توانایی تولید ویژگی‌های جداکننده به صورت خودکار را دارا هستند. شبکه عصبی کانولوشن[1] (*CNN*) یکی از موفق‌ترین روش‌های یادگیری عمیق است. از *CNN* های دو بعدی به طور گسترده‌ای برای استخراج ویژگی‌های مکانی و برای تشخیص اشیاء و تقسیم‌بندی تصاویر با وضوح بالا استفاده شده است[۷]. یکی از کاربردهای مهم *CNN* ها، طبقه‌بندی تصاویر ابرطیفی است. از *CNN* برای استخراج ویژگی‌های طیفی-مکانی در طبقه‌بندی این تصاویر استفاده می‌شود. کانولوشن یک بعدی برای استخراج ویژگی‌های طیفی، کانولوشن دو بعدی برای استخراج ویژگی‌های مکانی وکانولوشن سه‌بعدی برای استخراج ویژگی‌های طیفی و مکانی بطور همزمان[۱۴]. یکی از مهمترین مراحل در اجرای *CNN*ها، توجه به ورودی آنهاست. از دیدگاه طبقه‌بندی طیفی-مکانی در برخی تحقیقات مکعب ورودی (تصویر خام ابرطیفی)، بدون هیچگونه کاهش بعد یا پیش‌پردازشی به عنوان ورودی به شبکه داده می‌شود که علاوه بر اینکه اجرای آنها به دلیل افزایش حجم محاسبات (حتی با استفاده از روش‌های محاسبات موازی و پیاده‌سازی بر بستر کارت‌های گرافیک) چالش‌های فراوانی را ایجاد می‌کند، موجب کاهش کارایی این شبکه‌ها نیز می‌گردد[۲۹]. از طرفی تخمین پارامترهای مجهول شبکه در این رویکرد، نیازمند نمونه‌های آموزشی فراوانی است که از نقطه نظر عملی غیرممکن است. مشکلات در نظرگرفتن تصویر خام به عنوان ورودی از دو منظر قابل بررسی است. اول مشکل تعداد بالا و افزونگی باندهای طیفی که موجب افزایش حجم محاسبات می‌شود و دوم کاهش کیفیت نمونه‌های آموزشی (به معنای تفکیک‌پذیر بودن کلاس‌های مختلف در یک نمونه آموزشی با تعداد مشخصی از باندهای طیفی) که موجب کاهش صحت طبقه‌بندی می‌گردد. در برخی تحقیقات پیشین از روش‌های متداولی مانند آنالیز مؤلفه اصلی[2] (*PCA*)، الگوریتم‌های کاهش بعد مبتنی بر خرد جمعی و روش‌های مبتنی بر انتخاب ویژگی با الگوریتم ژنتیک برای پیش پردازش استفاده شده است[۷،۱۱ و۱۴]. هرچند که این روش‌ها تا حدی مشکل مربوط به افزونگی تعداد باندهای ورودی را برطرف کردند، اما مسأله مهمتر یعنی افزایش کیفیت نمونه‌های آموزشی کمتر مورد توجه قرار گرفته است. هرکلاس، نمونه‌های آموزشی مختص به خود را دارد و باید عوارض مربوط به آن کلاس در نمونه‌های آموزشی آن بارزسازی و از سایر کلاس‌ها تا حد ممکن تفکیک شود. اگر نمونه‌های آموزشی که ورودی شبکه‌های عصبی هستند، دارای این ویژگی باشند، سبب می‌شود که ویژگی‌هایی که به صورت خودکار با شبکه‌های عصبی کانولوشن برای هر کلاس تولید می‌شوند، کیفیت بالاتری داشته باشند و صحت طبقه‌بندی افزایش پیدا کند. در این تحقیق به کمک روش کاهش بعد مبتنی بر زیر فضا و افزایش اثر درجات خاکستری مربوط به هرکلاس، ابعاد تصویر خام ورودی به نحوی کاهش می‌یابد که در آن عوارض مربوط به هرکلاس در نمونه‌های آموزشی مربوطه بارز گردد. به کمک این روش قدرت تفکیک‌پذیری ویژگی-هایی که از تصویر خام ورودی استخراج می‌شود افزایش پیدا می‌کند. در این روش برای هر کلاس یک ویژگی استخراج می‌شود که موجب برقراری ارتباط نظیر به نظیر با تعداد کلاس‌های تصویر می‌گردد. برای تصویر خام با استفاده از نمونه‌های آموزشی به تعداد کلاس-های تصویر، ویژگی استخراج می‌شود و مشکل افزونگی و کیفیت نمونه‌های آموزشی همزمان حل می‌گردد. بر این اساس هدف کلی تحقیق پیش رو ارائه چارچوبی مبتنی بر یادگیری عمیق و *CNN* برای طبقه‌بندی تصاویر ابرطیفی است. نوآوری اصلی این مقاله، ارائه چارچوبی در راستای بکارگیری یادگیری عمیق در چهار مرحله می‌باشد، به نحوی که با افزایش کیفیت

[4] *Convolutional Neural Network*

[9] *Principle Component Analysis*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

داده‌های ورودی، صحت طبقه‌بندی افزایش یابد. در روش پیشنهادی، ابتدا طراحی معماری *CNN* انجام شده و پس از رسیدن به یک معماری مطلوب به عنوان شبکه پایه، تصویر خام ورودی به کمک روش کاهش بعد زیرفضا تبدیل شده و به عنوان ورودی به شبکه پایه در نظر گرفته شده است. در واقع برای ارزیابی از دو حالت استفاده شده است. در حالت اول، شبکه پایه بدون کاهش بعد اجرا شده و در حالت دوم، شبکه پایه با ورودی‌های کاهش بعد یافته اجرا شده است. سپس نتیجه این دو حالت مقایسه گردیده است. روش پیشنهادی شامل چهار مرحله است. مرحله اول) کاهش بعد، مرحله دوم) آماده‌سازی ورودی‌های *CNN*، مرحله سوم) غنی‌سازی داده[1]های آموزشی، مرحله چهارم) طراحی معماری *CNN*. داده‌های مورد استفاده، داده-های بنچمارک دانشگاه پاویا [2] (*PU*) است. دقت طبقه-بندی تصویر مذکور با چارچوب پیشنهادی ۹۸/۳ درصد می‌باشد. بخش‌های بعدی این مطالعه شامل مروری بر تحقیقات پیشین، مبانی نظری یادگیری عمیق و شبکه عصبی کانولوشن، داده‌ها، روش تحقیق، پیاده‌سازی و نتایج و در نهایت نتیجه‌گیری می‌باشد.

۲- مروری بر تحقیقات پیشین

برای پیاده‌سازی الگوریتم *CNN* بر روی تصاویر ابرطیفی، چالش‌های متعددی وجود دارد. تعداد زیاد باندهای طیفی (و در نتیجه هزینه بالای محاسباتی)، پیچیدگی انتخاب روش مناسب کاهش بعد، کمبود نمونه‌های آموزشی، تولید مدل‌های پیش آموزش، تنظیم پارامترهای مربوط به کرنل‌ها و استفاده همزمان از اطلاعات مکانی و طیفی در طبقه‌بندی از جمله این چالش‌ها می‌باشد[۳، ۱۰، ۱۱، ۱۲و۱۳] . از میان این چالش‌ها، انتخاب روش مناسب کاهش بعد به دلیل اینکه با بحث کمبود نمونه‌های آموزشی نیز در ارتباط است، از اهمیت بسزایی برخوردار است. به این معنی که در صورت موفقیت‌آمیز بودن روش کاهش بعد، عملاً مشکل کمبود نمونه‌های آموزشی نیز به صورت غیرمستقیم تعدیل خواهد شد. لذا در این بخش روش-هایی که مرتبط با این موضوع است، بررسی می‌گردد.

چن[3] و همکاران(۲۰۱۴) از روش خود رمز نگار[4] برای طبقه‌بندی تصاویر ابرطیفی استفاده کرده‌اند. خود رمز نگارها در واقع یک فضای فشرده از تصویر خام ورودی ارائه می‌دهند که در آن بتوان با داشتن پارامترهای مدل خود رمز نگار تصویر اصلی را بازسازی کرد. طبقه-بندی در سه حالت فقط با باندهای طیفی، فقط با ویژگی‌های مکانی و استفاده‌ی همزمان از اطلاعات طیفی و مکانی انجام شده است که صحت طبقه‌بندی در حالت استفاده از اطلاعات طیفی و مکانی بیشتر شده است[۱۴].

چن[5] و همکاران در از روش رایج *PCA* برای کاهش بعد استفاده کردند و کارایی شبکه *CNN* پیشنهادی خود را با ویژگی‌های استخراج شده در سه سناریوی استفاده از اطلاعات طیفی، اطلاعات مکانی و اطلاعات طیفی - مکانی مورد ارزیابی قرار دادند[۱۷]. همچنین در [۱۸] از روش *PCA* برای کاهش بعد تصاویر ابرطیفی به عنوان یک راهکار پیش‌پردازشی جهت پیاده‌سازی *CNN* استفاده شده است. روش نظارت‌نشده‌ی *PCA*، تضمینی جهت حفظ جدایی بین کلاس‌های تصویر نمی‌دهد و فقط سعی در حفظ محتوای اطلاعاتی تصویر دارد، به همین دلیل، استفاده از *PCA* کارآیی لازم را ندارد و تضمین رسیدن به دقت‌های بالا بخصوص در تصاویر با ابعاد بزرگ را نمی‌دهد. ژائو[6] و همکاران(۲۰۱۶) روش *PCA* را با اطلاعات حاشیه‌ای محلی کلاس‌ها برای افزایش جدایی بین کلاس‌ها ترکیب کردند[۱۹]. یوسونگ [7] و همکاران(۲۰۱۷)، یک

[6] *Data augmentation*
[7] *Pavia University*

[8] *Chen*
[9] *Auto encoders Based Methods*
[10] *Chen*
[11] *Zhao*
[12] *Yousong*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

روش انتخاب ویژگی را بر اساس مقایسه باندهای طیفی با مؤلفه اول استخراج شده از *PCA* برای بارزسازی تصاویر ابرطیفی ارائه دادند[۲۰]. در این مقاله تنها از ویژگی‌های استخراج شده از *CNN* استفاده شد و طبقه-بندی با روش یادگیری مفرط[۱] انجام شد. رومرو[۲] و همکاران(۲۰۱۵) از نسخه کرنلایز شده *PCA* برای کاهش بعد در چارچوب یک روش نظارت‌نشده استفاده کردند[۲۱]. دنگ[۳] و همکاران(۲۰۱۸) از ساختار چند مقیاسی و ترکیب اطلاعات طیفی - مکانی با استراتژی رأی‌گیری برای طبقه‌بندی تصاویر ابرطیفی با *CNN* استفاده کردند[۲۲]. نویسندگان، ویژگی‌های استخراج شده را در مقیاس‌های مختلف به عنوان ورودی به *CNN* دادند. دقت روش ارائه شده بر روی تصاویر چندطیفی با رزولوشن بالا مناسب بود، اما کارایی آن بر روی تصاویر ابرطیفی قابل توجه نبود. روش مبتنی بر خرد جمعی در [۲۳] در چارچوب یک روش خودسازمان‌ده، مورد ارزیابی قرار گرفت و دقت طبقه بندی قابل توجهی ارائه نداد. در تحقیقات اشاره شده، تنها روش کاهش بعد از منظر حجم داده‌ها در اولویت قرار گرفته است، در صورتی که کاهش حجم داده‌ها بایستی با حفظ قدرت تفکیک‌پذیری همراه باشد و حداکثر محتوای اطلاعاتی باندهای اصلی نیز حفظ شود. روش‌های کاهش بعدی مانند روش زیرفضا به دلیل ماهیت نظارت شده آنها، امکان این تبدیل را فراهم می کنند و موجب کاهش بعد همراه با حفظ و حتی افزایش قدرت تفکیک‌پذیری ویژگی‌های استخراج شده می‌گردند. دسته‌ای دیگر از روش‌ها، *CNN* را بر روی تصاویر ابرطیفی بدون کاهش بعد پیاده‌سازی کرده‌اند. مشکل اساسی آنها این است که به حجم بسیار زیادی از نمونه‌های آموزشی نیازمند هستند. بطوری‌که در روش‌های نیازمند کاهش بعد، تعداد نمونه‌های آموزشی بین ۲ تا ۱۰ درصد از کل پیکسل‌های تصویر است، اما در روش‌های بدون کاهش بعد، تعداد نمونه-های آموزشی بین ۵۰ تا ۸۰ درصد کل پیکسل‌های تصویر است که این فرض، کارایی این روش‌ها را برای انجام کاربردهای واقعی از بین می‌برد.

۳- مبانی نظری یادگیری عمیق و *CNN*

به طورکلی روش‌های یادگیری عمیق به چهار دسته تقسیم می‌شوند: ۱) روش‌های بر پایه *CNN*، ۲) روش-های برپایه ماشین بولتزمن محدود[۴]، ۳)روش‌های بر پایه اتوانکدر ۴) روش‌های برپایه بازنمایی تنک[۵]. در میان چهار دسته فوق، روش‌های مبتنی بر *CNN* در کاربردهای بینایی ماشین[۶] محبوبیت گسترده‌ای پیدا کرده است[۱۰]. از مزایای روش *CNN* می‌توان به کاهش تعداد پارامترهای وزن به کمک مفهوم به اشتراک‌گذاری[۷] وزن‌ها اشاره کرد. همچنین در تولید ویژگی‌ها در این روش، همسایگی‌های یک پیکسل در نظرگرفته می‌شود که با فرض یکسان بودن خواص آماری پیکسل‌های یک بخش در تصاویر سازگار است[۱۰]. در شبکه‌های عصبی چندلایه‌ی پرسپترون[۸] (*MLP*) نورون‌ها بصورت کامل به یکدیگر متصل هستند و با افزایش ابعاد تصاویر، تعداد اتصالات و پارامترهای آزاد بسیار زیاد شده و در نتیجه داده‌های آموزشی بسیار حجیمی لازم خواهد بود. بنابراین ممکن است در زمان آموزش به دلیل متناسب نبودن تعداد نمونه‌های آموزشی، مشکل بیش‌برازش رخ دهد[۲۴]. همچنین شبکه‌های عصبی *MLP* در برابر نویز مقاومتی از خود نشان نمی‌دهند. یکی از مهم‌ترین نقاط ضعف شبکه‌های عصبی *MLP* عدم ارائه مکانیزمی برای تولید خودکار ویژگی است. با توجه به مشکلات *MLP* روش-هایی در قالب *CNN* ارائه شده است[۲۵]. ایده این

[13] *Extreme learning method*
[14] *Romero*
[15] *Deng*
[16] *Restricted Boltzmann Machine Methods*
[17] *Sparse Coding Based Methods*
[18] *Machine vision*
[19] *Weight Sharing*
[20] *Multi-Layer Perceptron*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

روش‌ها انجام یک سری عملیات مشابه در ناحیه‌های منظمی از یک تصویر است که این ناحیه‌ها، نواحی دریافت محلی[1] نامیده می‌شوند. این عملیات منجر به تولید خودکار ویژگی خواهد شد. در این مدل با ارائه و بکارگیری مفهومی به نام اشتراک وزن‌ها، مشکل بالا بودن تعداد پارامترهای آزاد حل می‌شود و همچنین ویژگی‌های معنی‌داری از تصویر مرجع تولید می‌شود[۹]. این ویژگی سبب می‌شود که این مدل‌ها در مقایسه با *MLP*، ظرفیت تعمیم‌پذیری و کلی‌نگری بیشتری داشته باشند. *CNN* از سه لایه کانولوشن[2]، پولینگ[3]و اتصال کامل[4] تشکیل شده است که در ادامه به اختصار توضیح داده خواهند شد. لایه کانولوشن از کانولوشن تصویر ورودی به کمک ناحیه‌های دریافت محلی بوجود می‌آید و در این لایه مفهوم سلسله مراتبی بودن و به شی نزدیک شدن تحقق پیدا می‌کند. حرکت فیلتر یا فیلترها بر روی باندهای ورودی به شکل تاگل مومنت[5] است[۲۶]. تعیین تعداد فیلتربانک‌ها و لایه‌های کانولوشن و تنظیم پارامترهای آن از چالش‌های موجود است. در لایه پولینگ، کاهش سایز تصویر انجام می‌شود. علت اصلی اعمال این اپراتور، کاهش حجم محاسبات و تضمین ناورداء[6]بودن ویژگی‌ها نسبت به سایز تصویر است[۲۷]. علاوه بر مزیت فوق، غیرخطی بودن ویژگی‌های تولید شده در مرحله کانولوشن به کمک تابع فعالسازی[7]در لایه پولینگ صورت می‌گیرد. عملیات پولینگ، کاهش سایز را در امتداد ابعاد مکانی (عرض و ارتفاع) انجام می‌دهد که نتیجه این کار، یک توده با اندازه کوچکتر خواهد بود. در اصل از طریق این عملیات است که در انتهای شبکه کانولوشن یک بردار امتیاز ایجاد می‌شود. به دلیل اینکه این عملیات بر روی تمامی برش‌ها اعمال می‌گردد، عمق خروجی همان عمق ورودی به لایه پولینگ می‌باشد و تغییری نمی‌کند. روش‌های مختلفی برای اعمال پولینگ وجود دارد، اما روش‌های میانگین‌گیری و انتخاب ماکزیمم با کاهش نسبت دو و تابع فعالسازی رلو[8]از رایج‌ترین روش‌ها هستند. بعد از مرحله‌ی پولینگ، ویژگی‌های دوبعدی ایجاد شده در مرحله اتصال کامل به یک بردار تک بعدی تبدیل می‌شود. این لایه همانند شبکه عصبی *MLP* امکان اجرای آموزش پیش‌رو[9] را برای *CNN* فراهم می‌کند[۹]. توابع فعال‌سازی انواع متفاوتی دارند. چند نمونه از معروف‌ترین توابع فعال‌سازی، تابع سیگموید، تابع تانژانت هایپربولیک، تابع یکسوساز خطی رلو ،تابع خطی نمایی و سافتمکس[10] است. تمامی نورون‌های واقع در لایه مخفی اول ویژگی مشابهی را در نواحی مختلف تصویر شناسایی می‌نمایند. در نهایت خروجی لایه ورودی یا نورون‌های لایه مخفی به نام نقشه ویژگی شناخته می‌شوند.

۴- داده‌ها

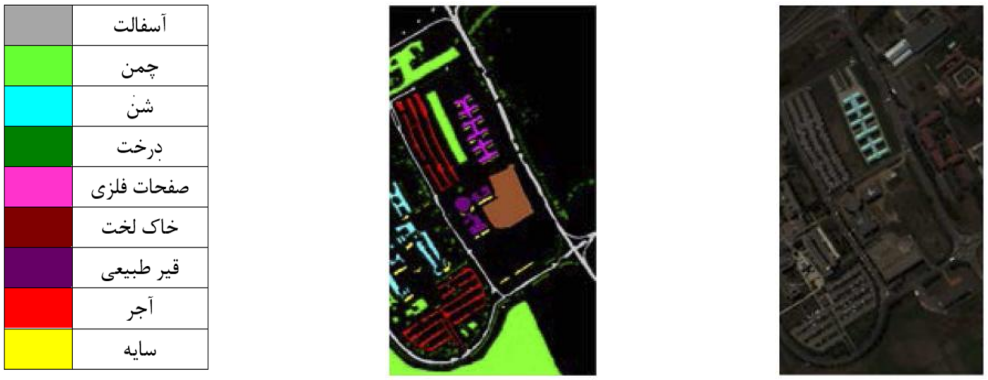
در تحقیق حاضر از داده‌های بنچمارک دانشگاه پاویا با مشخصاتی که در ادامه بیان خواهد شد، استفاده شده است. تصویر *PU* توسط سنجنده هوابرد روسیس[11] طی یک پرواز بر فراز دانشگاه پاویا در شمال ایتالیا و در سال ۲۰۰۱ اخذ شده است. داده‌های تصحیح نشده شامل ۱۱۵ باند طیفی در محدوده طول موج‌های ۰/۴۳ تا ۰/۸۶ میکرومتر و دارای توان تفکیک مکانی ۱/۳ متر هستند. داده‌های تصحیح شده با اندازه ۶۱۰ در ۳۴۰ پیکسل شامل ۱۰۳ باند طیفی پس از حذف باندهای نویزی می‌باشند. تصویر مذکور دارای ۹ کلاس است و تعداد نمونه‌های دارای برچسب ۴۲۷۷۶ می‌باشد. در شکل (۱) تصویر و داده‌های دارای برچسب قابل مشاهده می‌باشند.

---

21 Local Receptive Field
22 Convolution
23 Pooling
24 Fully connected
25 Toggle Movement
26 Invariant
27 Activation Function

28 Rectifies Linear Unit (ReLU)
29 Feed forward
30 Softmax
31 Reflective Optics System Imaging Spectrometer (ROSIS)

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]



(الف) (ب) (ج)

شکل۱: داده‌ها: الف) ترکیب رنگی کاذب تصویر *PU* ، ب) داده‌های دارای برچسب، ج) راهنمای کلاس‌های تصویر

۵- روش تحقیق

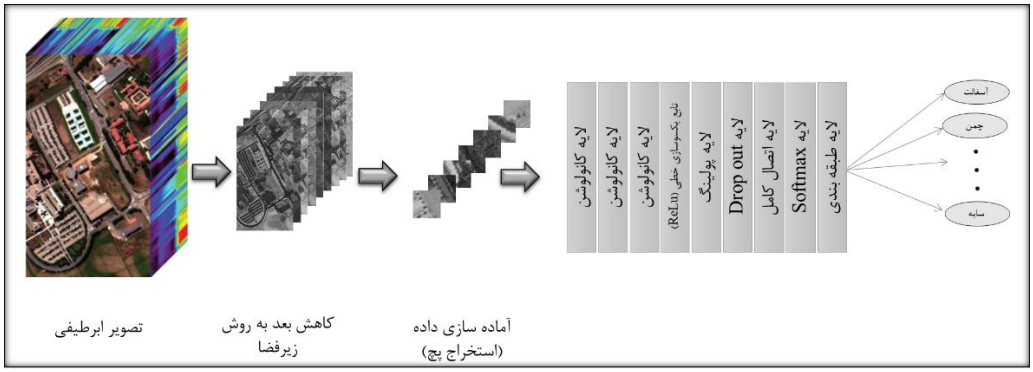
روش تحقیق، شامل چهار مرحله است. شکل (۲) این مراحل را نشان می‌دهد. مرحله اول) در این مرحله برای مواجهه با مشکل ابعاد بالای تصاویر ابرطیفی که منجر به هزینه محاسباتی چندین برابر خواهد شد و همچنین اجتناب از بکارگیری باندهایی که دارای وابستگی هستند و سبب افت صحت طبقه‌بندی می‌شوند، کاهش بعد صورت می‌گیرد. با توجه به اینکه فرآیند کاهش بعد و طبقه‌بندی *CNN* بسیار به یکدیگر وابسته هستند، در این چارچوب، روش کاهش بعد زیرفضا به نحوی بکارگیری می‌شود که کیفیت داده‌های آموزشی در هر کلاس افزایش یابد و منجر به بهبود طبقه‌بندی *CNN* شود. دلیل استفاده از روش مذکور برای کاهش بعد، توانایی آن در بارزسازی هریک از کلاس‌های تصویر در هر یک از ویژگی‌های بدست آمده از این روش می‌باشد. اصول کلی حاکم بر روش کلاس‌مبنا، بارزسازی هر کلاس مورد نظر در هر ویژگی با استفاده از داده‌های آموزشی می‌باشد.

مرحله دوم) پس از اینکه در مرحله اول ویژگی‌های متناظر با کلاس‌های تصویر *PU* استخراج می‌شوند، آماده‌سازی ورودی‌های *CNN* انجام می‌شود، که مستلزم انجام آنالیز حساسیت برای یافتن مقدار مناسب اندازه پنجره‌ای است که پیرامون هر داده آموزشی بایستی اتخاذ شود. نحوه انجام آنالیز حساسیت و نتایج آن در بخش۲-۵ بیان شده است.

مرحله سوم) مرحله غنی‌سازی داده‌های آموزشی، در این مرحله، با استفاده از اپراتورهای چرخش و قرینه-سازی شامل دو چرخش با زوایای ۹۰ و ۱۸۰ درجه و یک قرینه، تعداد داده‌های آموزشی به ۳ برابر افزایش می‌یابد. چرخش و قرینه‌سازی به منظور تولید داده‌های افزونه برای بهبود فرآیند یادگیری عمیق انجام می‌شود.

مرحله‌ی چهارم) طراحی معماری *CNN*، که با توجه به تحقیقات پیشین و انجام آنالیز حساسیت، یک معماری *CNN* با سه لایه کانولوشن و یک لایه اتصال کامل طراحی شد. آنالیز حساسیت مذکور در زیر بخش۳-۵ توضیح داده شده است. مشخصات معماری *CNN* طراحی شده نیز در جدول (۱) آمده است.

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]



شکل۲: روش تحقیق

جدول ۱: مشخصات معماری *CNN* طراحی شده

نام لایه	توضیحات
تصویر ورودی	تصاویر ۲۹ در ۲۹ با نرمالیزاسیون به روش میانگین صفر
کانولوشن	۳ لایه کانولوشن با ۶۴ فیلتر کانولوشن ۱*۲۲*۲۲ با گام*[1,1]* و پدینگ *[0,0]*
ReLU	تابع یکسوسازی خطی
Max Pooling	با اندازه ۲ در ۲ پولینگ
Dropout	۵۰ درصد
اتصال کامل	یک لایه اتصال کامل با ۹ نورون
Softmax	*Softmax*
خروجی طبقه‌بندی	*Cross entropy* با "۱"، "۲" تا "۹"

۵-۱- روش کاهش بعد زیرفضا

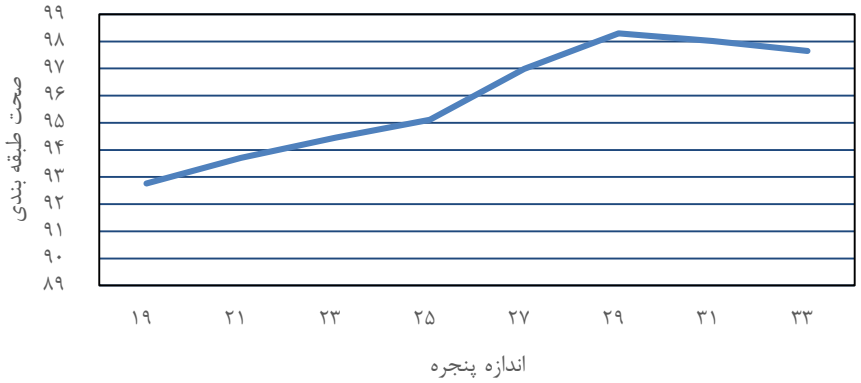
حسگرهای ابرطیفی تصاویر را در صدها باند طیفی اخذ می‌کنند. این منجر به مشکل مسأله با ابعاد بالا می‌شود که یکی از موانع اصلی برای طبقه‌بندی خودکار تصاویر است. هدف اصلی ما تولید نمونه‌های آموزشی با کیفیت همزمان با کاهش ابعاد است. روش کاهش بعد زیر فضا این امکان را فراهم می‌آورد که هم ابعاد طیفی تصویر کاهش یابد و هم در هر پچ ورودی بخش مربوط به کلاس مورد نظر وضوح بیشتری داشته باشد. استفاده از روش زیرفضا بر این فرض استوار است که توزیع نمونه-های آموزشی هر کلاس تقریباً بر روی یک فضای فرعی با ابعاد کمتر از فضای ویژگی اصلی قرار دارد. مرسوم-ترین روش برای تعریف یک زیرفضای l بعدی، بکارگیری یک مجموعه مستقل از بردارهای پایه $\{u_1,u_2,...,u_l\}$ است که در آن یک ماتریس با ابعاد $d\times l$ و رتبه l بازنمایی می‌شود (d: ابعاد فضای اصلی). $u_{(c)}$را مجموعه‌ای از بردارهای اورتونرمال پایه $r^{(c)}$– بعدی برای هر زیرفضای متعلق به هر کلاس در نظر بگیرید ($c=1,2,...,l$). مولفه‌های $u_{(c)}$ یعنی $u_{(c)}=\{e_1^{(c)},...,e_{r(c)}^{(c)}\}$ از $E_{(c)}=\{e_1^{(c)}, e_2^{(c)},...\}$ که ماتریس بردارهای ویژه ماتریس همبستگی ماتریس $R_{(c)}=E_{(c)}\Lambda E_{(c)}^{'}$ است. در اینجا Λ ماتریس مقادیر ویژه و e بردار ویژه متناظر با مقادیر ویژه ماتریس $R_{(c)}$ است که به شکل نزولی مرتب شده‌اند. برای تعیین سایز $u_{(c)}$ حدود ۹۹ درصد محتوای اطلاعاتی را مبنا قرارداده‌ایم. بدین ترتیب فضای ویژگی نهایی به دست

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

آمده (داده‌های تبدیل یافته) از کنار هم قرار دادن نرم هر یک از زیر فضاها (که حاصل تصویر فضای ویژگی اصلی است) به همراه نرم بردار ویژگی اصلی تشکیل می‌گردد. این فضای تبدیل یافته به عنوان ورودی در نظر گرفته می‌شود.

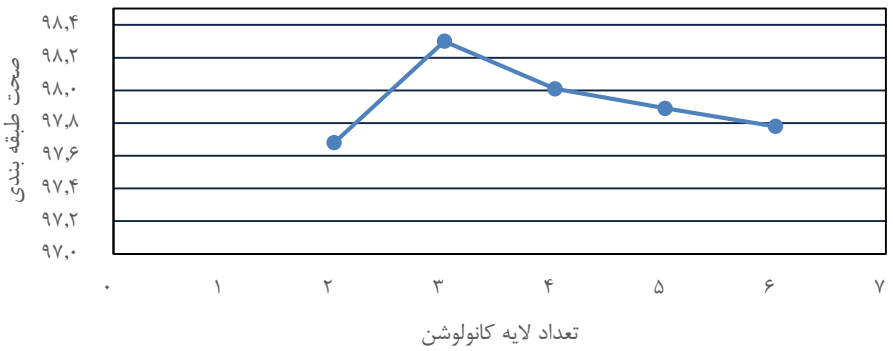
۵-۲- آنالیز حساسیت

عموماً یکی از چالش‌های مهم در طراحی معماری یک *CNN*، تعیین ابعاد پنجره ورودی به *CNN*، تعداد لایه‌های کانولوشن، سایز پولینگ و تعداد لایه‌های اتصال کامل است که به مجموعه پارامترهای یادشده، فراپارمترهای یک شبکه *CNN* می‌گویند. در این مطالعه، جهت تعیین ابعاد پنجره ورودی به *CNN* و تعداد لایه‌های کانولوشن از آنالیز حساسیت استفاده شده است. بدین ترتیب که ابتدا با بررسی مطالعات پیشین، به این نتیجه رسیده شد که ابعاد ورودی مناسب بین ۱۹ تا ۳۳ است[۲۸]، سپس برای دستیابی به ابعاد پنجره ورودی مناسب با ثابت نگه داشتن سایز پارامترها، میزان حساسیت صحت طبقه‌بندی نسبت به ابعاد پنجره ورودی بررسی شد که در شکل (۳) نتایج حاصل نشان داده شده است. تعداد لایه‌های کانولوشن یکی از فراپارمترهای اساسی می‌باشد. علی‌رغم اینکه *CNN* تعداد پارامترهای بسیار کمتری نسبت به *MLP* دارد، اما همچنان تعداد نمونه‌های آموزشی برای دستیابی به شبکه‌های بسیار عمیق کافی نیست. لذا تعداد لایه‌های کانولوشن در مطالعات پیشین بین ۲ تا ۶ لایه است که برای تعیین تعداد مناسب لایه‌های کانولوشن، با ثابت نگه داشتن سایر پارامترها (از جمله ابعاد پنجره ورودی که در آنالیز حساسیت قبلی ۲۹ * ۲۹ تعیین شد)، حساسیت صحت طبقه‌بندی نسبت به تغییر تعداد لایه‌ها ارزیابی شد. همان‌طور که در شکل(۴) نشان داده شده است تعداد ۳ لایه انتخاب شد.



شکل۳: نتایج آنالیز حساسیت صحت طبقه‌بندی نسبت به اندازه پنجره ورودی به *CNN* با ثابت نگه داشتن سایر پارامترها

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]



شکل۴: نتایج آنالیز حساسیت صحت طبقه‌بندی نسبت به تعداد لایه‌های کانولوشن با ثابت نگه داشتن سایر پارامترها

۶- پیاده‌سازی و نتایج

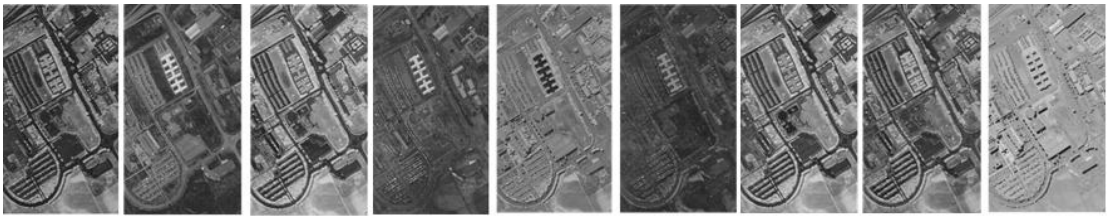
برای پیاده‌سازی چارچوب پیشنهادی از جعبه ابزار مت-کانو نت[1] نرم‌افزار متلب و از کارت گرافیک انویدیا جی فورس جی تی ایکس ۱۰۸۰[2] جهت پردازش‌های مبتنی بر جی‌پی‌یو[3] استفاده شد. شکل (۵) ویژگی‌های استخراج شده از مرحله اول چارچوب پیشنهادی یعنی مرحله کاهش بعد به روش زیرفضا را نمایش می‌دهد. در این شکل مشخص است که به ازای هر کلاس از کلاس‌های ۹ گانه تصویر، یک ویژگی متناظر با آن کلاس استخراج شده است.

جدول(۲) تعداد معلومات و مجهولات مسئله را نشان می‌دهد. همان‌طور که از این جدول مشخص است ۳۹۳۰ پیکسل دارای برچسب داریم که با توجه به اینکه ۹ ویژگی و ۳ اپراتور غنی‌سازی داده(دو تا چرخش با زوایای ۹۰ و ۱۸۰ درجه و یک قرینه‌سازی) داریم تعداد کل معلومات می‌شود ۳۹۳۰*۹*۳. با توجه به اینکه در معماری *CNN* پیشنهادی، ۶۴ فیلتر ۲۲در۲۲ داریم، بنابراین تعداد مجهولات مربوط به وزن برابر خواهد بود با ۲۲*۲۲*۶۴ و همچنین ۶۴ پارامتر مجهول بایاس خواهیم داشت. مجهولاتمان در لایه اتصال کامل هم برابر خواهد بود با ۹*۶۴ و بدین ترتیب تعداد کل مجهولاتمان ۳۱۶۱۶ خواهد بود. پارامترهای *CNN* با استفاده از روش گرادیان کاهشی[4] به دست آمدند. ورودی‌ها به بازه (۱و۱-) نرمالایز شده و میزان نرخ یادگیری عدد ۰٫۰۱ و تعداد اپک ها ۴۰۰ در نظر گرفته شد. لازم به ذکر است که حدود اعداد ذکر شده از تحقیقات پیشین استخراج و با تغییر آن‌ها در یک محدوده مشخص اعداد یاد شده به دست آمدند.

در شکل(۶) برخی از فیلترهای کانولوشن بدست آمده از *CNN* به صورت تصادفی انتخاب و ارائه شده است. شکل (۷) هم خروجی حاصل از اعمال برخی از فیلترهای کانولوشن بر روی داده‌های آموزشی تبدیل شده به قطعات ۲۹ در ۲۹ را نشان می‌دهد. شکل (۸) ، روند افزایش صحت مدل را در تکرارهای مختلف نشان می‌دهد. کاملاً مشخص است که هرچه تعداد تکرارها افزایش می‌یابد صحت مدل، روند افزایشی یافته و سپس ثابت می‌شود. جدول (۳) بیانگر ماتریس خطای طبقه‌بندی می‌باشد. در سطر و ستون آخر این جدول، دقت‌های کاربر و مولد برای هر کلاس از کلاس‌های ۹گانه تصویر ارائه شده‌اند. همچنین دقت کلی طبقه-بندی ۹۸/۳ درصد محاسبه شد. شکل (۹) هم نمایانگر نقشه خروجی طبقه‌بندی می‌باشد.

[32] *MATCONVNET*
[33] *NVIDIA GEFORCE GTX-1080*
[34] *GPU*
[35] *Gradient descend strategy*

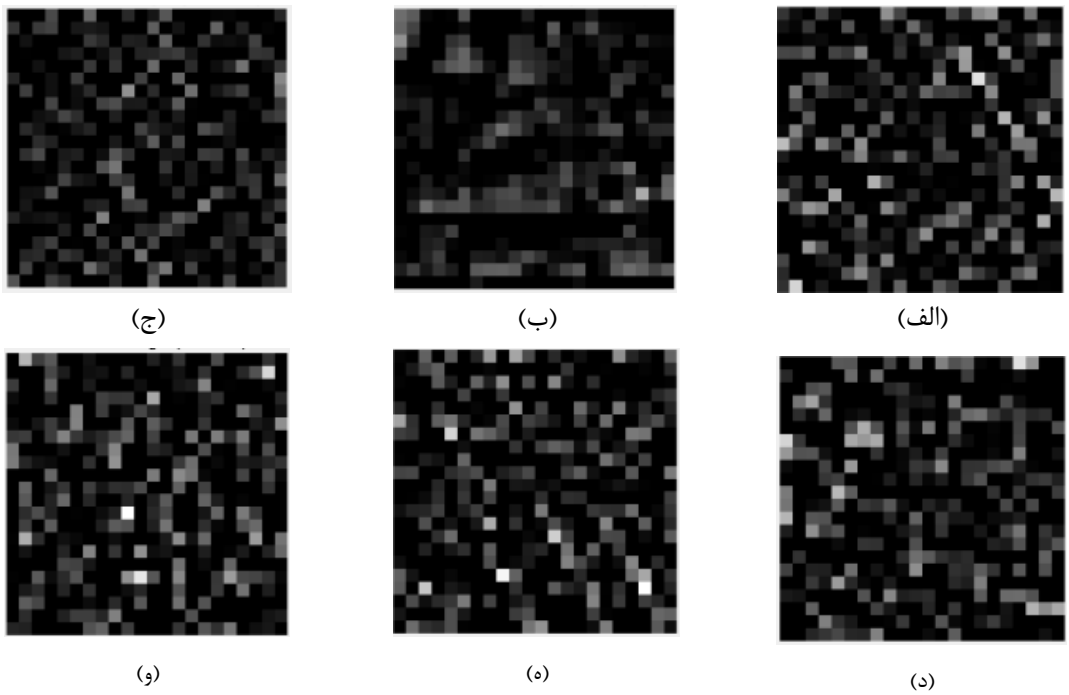
[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]



شکل۵: ویژگی‌های استخراج شده از مرحله کاهش بعد

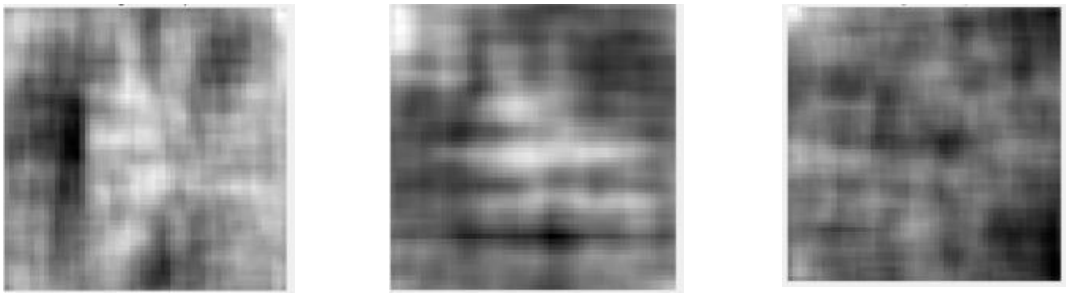
جدول ۲: معلومات و مجهولات مسأله

تعداد معلومات		تعداد مجهولات		
پیکسل‌های دارای برچسب	۳۹۳۰	لایه کانولوشن	وزن	۶۴*۲۲*۲۲
			بایاس	۶۴
تعداد ویژگی‌ها	۹	لایه اتصال کامل		۶۴*۹
تعداد کل داده‌های آموزشی	۳۹۳۰*۹*۳	تعداد کل پارامترها		۳۱۶۱۶



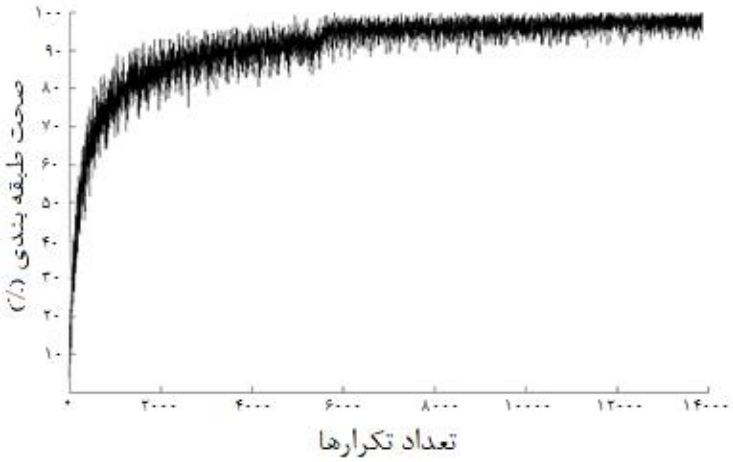
شکل۶: فیلترهای کانولوشن بدست آمده از شبکه‌ی *CNN* (الف) کرنل اول، ب) کرنل دهم، ج) کرنل بیستم، د) کرنل چهلم، ه) کرنل پنجاهم، و) کرنل شصت و چهارم)

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]



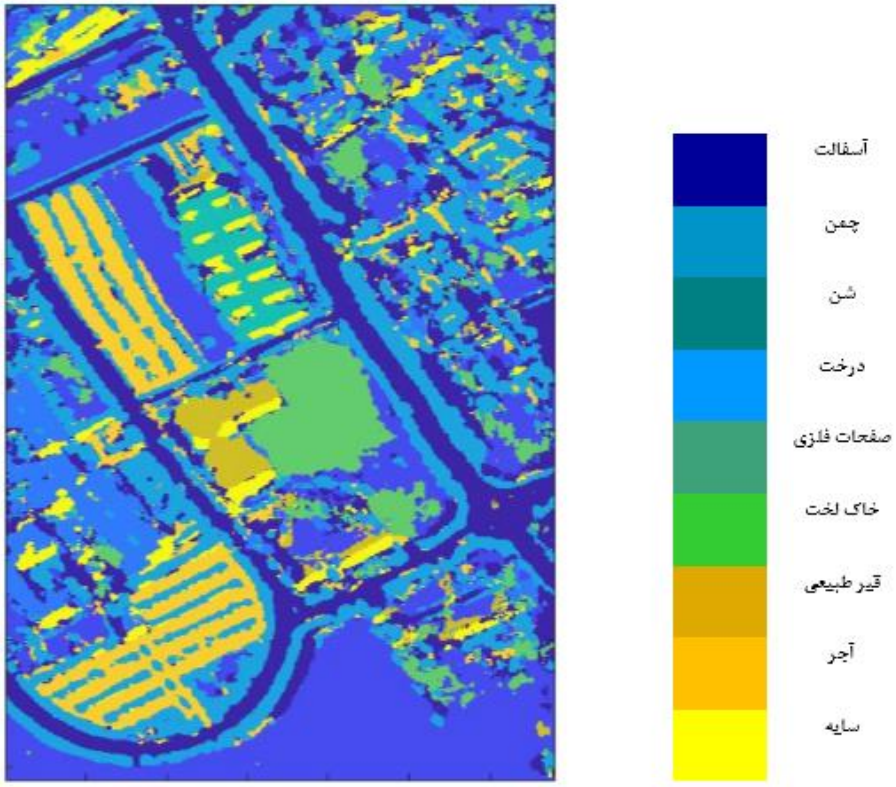
(الف) (ب) (ج)

شکل۷: خروجی حاصل از اعمال برخی از فیلترهای کانولوشن بر روی داده‌های آموزشی تبدیل شده به قطعات ۲۹ در ۲۹ (الف) نقشه ویژگی اول، ب) نقشه ویژگی دهم، ج) نقشه ویژگی بیستم)



شکل۸: روند افزایش صحت مدل در تکرارهای مختلف (محور افقی تعداد تکرارها و محور قائم صحت مدل را نشان می‌دهد.)

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]



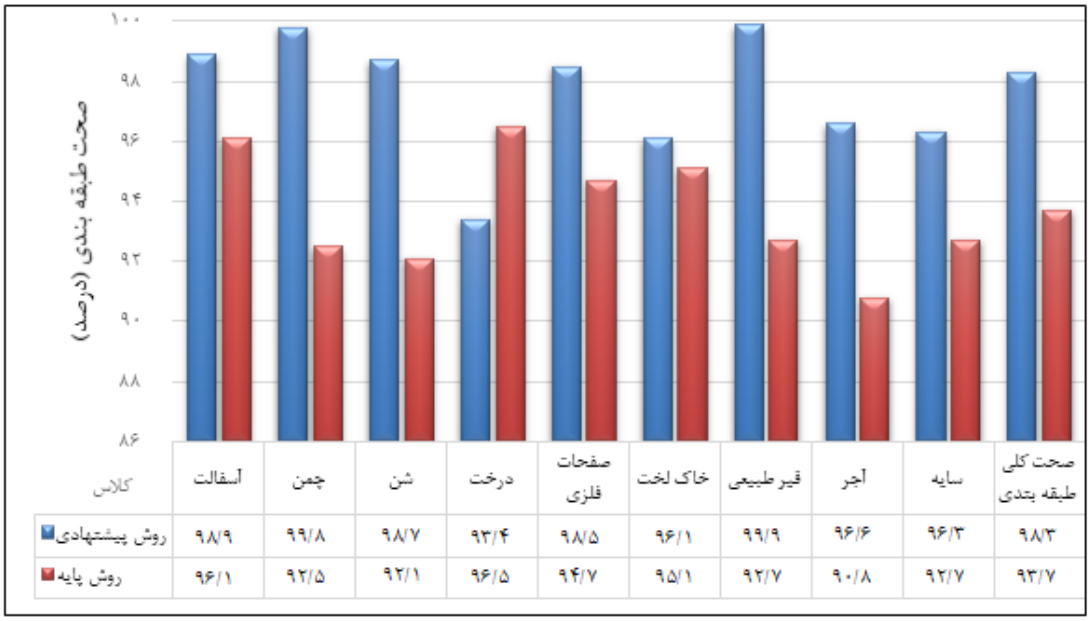
شکل۹: خروجی طبقه‌بندی

جدول۳: ماتریس خطای طبقه‌بندی

کلاس‌ها (برچسب پیش بینی شده)	۱	۶۰۲۸	۶۵	۰	۱	۰	۰	۰	۱	۲	٪ ۹۸٫۹
	۲	۱	۱۷۶۰۲	۰	۱	۰	۳۷	۰	۰	۰	٪ ۹۹٫۸
	۳	۳	۱۰	۱۷۰۷	۰	۰	۰	۳	۳	۴	٪ ۹۸٫۷
	۴	۱۳	۱۵۷	۰	۲۵۰۰	۰	۰	۰	۳	۵	٪ ۹۳٫۴
	۵	۱	۱۳	۰	۰	۱۰۸۹	۰	۰	۱	۲	٪ ۹۸٫۵
	۶	۶	۱۶۴	۰	۰	۰	۴۴۶۰	۰	۰	۰	٪ ۹۶٫۱
	۷	۱	۰	۰	۰	۰	۰	۹۵۲	۰	۰	٪ ۹۹٫۹
	۸	۲۹	۷۲	۰	۱۱	۰	۰	۰	۳۱۶۰	۰	٪ ۹۶٫۶
	۹	۱	۲۶	۰	۰	۰	۰	۰	۰	۷۰۳	٪ ۹۶٫۳
		٪ ۹۹٫۱	٪ ۹۷٫۲	٪ ۱۰۰	۹۹/۱ ٪	٪ ۱۰۰	۹۹٫۲ ٪	۹۹٫۷ ٪	٪ ۹۹٫۷	۹۸٫۲ ٪	٪ ۹۸٫۳
		۱	۲	۳	۴	۵	۶	۷	۸	۹	
داده‌های آموزشی (برچسب شناخته شده)											

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

جدول(۱) مشخصات *CNN* پایه را نشان می‌دهد. این شبکه در دو حالت مورد استفاده قرار گرفته است. در حالت اول تصویر ابرطیفی ورودی به کمک روش زیرفضا کاهش بعد یافته و به عنوان ورودی به شبکه پایه داده شده است. نتایج این حالت در جدول(۳) نشان داده شده است. برای ارزیابی دقیق‌تر و بررسی کمی میزان تأثیرگذاری روش پیشنهادی، در حالت دوم تصویر ابرطیفی بدون کاهش بعد به شبکه پایه داده شد. شکل(۱۰) نتایج حاصل از مقایسه کلاس به کلاس و همچنین صحت کلی طبقه‌بندی را نشان می‌دهد. همان‌طور که ملاحظه می‌شود صحت کلی طبقه‌بندی روش پیشنهادی در مقایسه با شبکه پایه ۴/۶ درصد افزایش یافته است. همچنین در اکثر کلاس‌ها (بجزدرخت)، روش پیشنهادی صحت بالاتری را ارائه داده است. در کلاس چمن و قیر طبیعی بیشترین میزان افزایش یعنی ۷/۳ و ۷/۲ درصد را شاهد هستیم که ناشی از قابلیت بالای روش کاهش بعد زیرفضا در بارزسازی کلاس‌هایی با بافت متفاوت است.



کلاس	آسفالت	چمن	شن	درخت	صفحات فلزی	خاک لخت	قیر طبیعی	آجر	سایه	صحت کلی طبقه بندی
روش پیشنهادی	۹۸/۹	۹۹/۸	۹۸/۷	۹۳/۴	۹۸/۵	۹۶/۱	۹۹/۹	۹۶/۶	۹۶/۳	۹۸/۳
روش پایه	۹۶/۱	۹۲/۵	۹۲/۱	۹۶/۵	۹۴/۷	۹۵/۱	۹۲/۷	۹۰/۸	۹۲/۷	۹۳/۷

شکل۱۰: مقایسه صحت طبقه‌بندی روش پیشنهادی و روش پایه

۷- نتیجه‌گیری

طبقه‌بندی تصاویر ماهواره‌ای ابر طیفی همواره با چالش‌های زیادی روبرو بوده است. از طرفی با پیشرفت تکنولوژی و افزایش قدرت تفکیک مکانی تصاویر، با وجود افزایش کلی دقت، چالش‌های مربوط به طبقه-بندی این تصاویر نیز تشدید گردیده است. تفاوت در نمونه‌های مختلف متعلق به یک کلاس و شباهت‌های موجود در اشیاء متعلق به کلاس‌های متفاوت، عملکرد نسبی الگوریتم‌های طبقه‌بندی را تحت تأثیر قرار داده است. در تصاویر در نظر گرفته شده، تفکیک کلاس‌های مشابه نظیر درخت و فضای سبز، تعیین کلاس درست در مناطقی که دو کلاس روی هم قرار گرفته‌اند و همچنین سایه‌های موجود در تصویر چالش‌های عمده پیشرو بوده است.

در این تحقیق یک چارچوب مبتنی بر مفهوم یادگیری عمیق و شبکه عصبی کانولوشن برای طبقه‌بندی تصاویر ابر طیفی ارائه شد. با توجه به اینکه فرآیند کاهش بعد و طبقه‌بندی *CNN* بسیار به یکدیگر وابسته هستند، در

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

این مطالعه، روش کاهش بعد زیرفضا به نحوی بکارگیری شد که کیفیت داده‌های آموزشی در هر کلاس افزایش یابد و منجر به بهبود طبقه‌بندی *CNN* شود. همچنین افزایش تعداد نمونه‌های آموزشی به کمک روش‌های غنی‌سازی داده جهت غلبه بر مشکل بیش‌برازش صورت گرفت. بطور کلی چارچوب پیشنهادی، شامل چهار مرحله اعم از: ۱) کاهش بعد، ۲) آماده‌سازی ورودی‌های *CNN*، ۳) غنی‌سازی داده، ۴) طراحی معماری *CNN* می‌باشد. چارچوب پیشنهادی بر روی تصاویر بنچ مارک دانشگاه پاویا پیاده‌سازی و صحت کلی ۹۸٫۳ درصد حاصل شد.

مراجع

[1] *W. Zhao and S. Du, " Learning multiscale and deep representations for classifying remotely sensed imagery", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 113, pp. 155–165, Mar. 2016, doi: 10.1016/j.isprsjprs.2016.01.004.*

[2] *M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, " Deep learning classifiers for hyperspectral imaging: A review", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 158, pp. 279–317, 2019.*

[3] *G. Camps-Valls and L. Bruzzone, " Kernel-based methods for hyperspectral image classification", IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 6, pp. 1351–1362, Jun. 2005, doi: 10.1109/TGRS.2005.846154.*

[4] *G. Hughes, " On the mean accuracy of statistical pattern recognizers", IEEE Transactions on Information Theory, vol. 14, no. 1, pp. 55–63, Jan. 1968, doi: 10.1109/TIT.1968.1054102.*

[5] *L. O. Jimenez and D. A. Landgrebe, " Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data", IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 28, no. 1, pp. 39–54, Feb. 1998, doi: 10.1109/5326.661089.*

[6] *A. Plaza, J. Plaza, A. Paz, and S. Sanchez, " Parallel Hyperspectral Image and Signal Processing [Applications Corner]", IEEE Signal Processing Magazine, vol. 28, no. 3, pp. 119–126, May 2011, doi: 10.1109/MSP.2011.940409.*

[7] *Y. Chen, K. Zhu, L. Zhu, X. He, P. Ghamisi, and J. A. Benediktsson, 'Automatic Design of Convolutional Neural Network for Hyperspectral Image Classification', IEEE Transactions on Geoscience and Remote Sensing, vol. 57, 2019.*

[8] *F. Hu, G.-S. Xia, Z. Wang, X. Huang, L. Zhang, and H. Sun, " Unsupervised Feature Learning Via Spectral Clustering of Multidimensional Patches for Remotely Sensed Scene Classification", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, no. 5, pp. 2015–2030, May 2015, doi: 10.1109/JSTARS.2015.2444405.*

[9] *Y. Bengio, A. Courville, and P. Vincent, " Representation Learning: A Review and New Perspectives", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798–1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.*

[10] *Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, " Deep learning for visual understanding: A review", Neurocomputing, vol. 187, pp. 27–48, Apr. 2016, doi: 10.1016/j.neucom.2015.09.116.*

[11] *W. Hu, Y. Huang, and L. Wei, " Deep Convolutional Neural Networks for Hyperspectral Image Classification". [Online]. Available: https://www.hindawi.com/journals/js/2015/258619/. [Accessed: 16-Oct-2019].*

[12] *K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, " Deep supervised learning for hyperspectral data classification through convolutional neural networks", in 2015 IEEE International Geoscience and Remote Sensing*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4544–4554, Aug. 2016, doi: 10.1109/TGRS.2016.2543748.

[20]Y. Li, W. Xie, and H. Li, " Hyperspectral image reconstruction by deep convolutional neural network for classification", Pattern Recognition, vol. 63, pp. 371–383, Mar. 2017, doi: 10.1016/j.patcog.2016.10.019.

[21]A. Romero, C. Gatta, and G. Camps-Valls, " Unsupervised Deep Feature Extraction for Remote Sensing Image Classification", p. 21, 2015.

[22]Z. Deng, H. Sun, S. Zhou, J. Zhao, L. Lei, and H. Zou, " Multi-scale object detection in remote sensing imagery with convolutional neural networks", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 145, pp. 3–22, Nov. 2018, doi: 10.1016/j.isprsjprs.2018.04.003.

[23]P. Ghamisi, Y. Chen, and X. X. Zhu, " A Self-Improving Convolution Neural Network for the Classification of Hyperspectral Data", IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 10, pp. 1537–1541, Oct. 2016, doi: 10.1109/LGRS.2016.2595108.

[24]Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, " Gradient-based learning applied to document recognition", Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.

[25]M. D. Zeiler and R. Fergus, " Visualizing and Understanding Convolutional Networks", in Computer Vision – ECCV 2014, vol. 8689, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 818–833.

[26]J. Gu et al., " Recent advances in convolutional neural networks", Pattern Recognition, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.

[27]H.-J. Yoo, " Deep Convolution Neural Networks in Computer Vision: a Review", Symposium (IGARSS), 2015, pp. 4959–4962, doi: 10.1109/IGARSS.2015.7326945.

[13]A. Vallet and H. Sakamoto, " A Multi-Label Convolutional Neural Network for Automatic Image Annotation", Journal of Information Processing, vol. 23, pp. 767–775, Nov. 2015, doi: 10.2197/ipsjjip.23.767.

[14]Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, " Deep Learning-Based Classification of Hyperspectral Data", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 6, pp. 2094–2107, Jun. 2014, doi: 10.1109/JSTARS.2014.2329330.

[15]F. Hu, G.-S. Xia, J. Hu, and L. Zhang, " Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery", Remote Sensing, vol. 7, no. 11, pp. 14680–14707, Nov. 2015, doi: 10.3390/rs71114680.

[16]T. Li, J. Zhang, and Y. Zhang, " Classification of hyperspectral image based on deep belief networks", 2014 IEEE International Conference on Image Processing (ICIP), pp. 5132–5136, 2014, doi: 10.1109/icip.2014.7026039.

[17]Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, " Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks", IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6232–6251, Oct. 2016, doi: 10.1109/TGRS.2016.2584107.

[18]M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, " A new deep convolutional neural network for fast hyperspectral image classification", ISPRS Journal of Photogrammetry and Remote Sensing, vol. 145, pp. 120–147, Nov. 2018, doi: 10.1016/j.isprsjprs.2017.11.021.

[19]W. Zhao and S. Du, " Spectral–Spatial Feature Extraction for Hyperspectral Image Classification: A Dimension Reduction and Deep Learning Approach",

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

IEIE Transactions on Smart Processing and Computing, vol. 4, pp. 35–43, Feb. 2015, doi: 10.5573/IEIESPC.2015.4.1.035.

[28] M. E. Paoletti et al., " Capsule Networks for Hyperspectral Image Classification", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, vol. 57, no. 4, pp. 2145–2160, 2018.

[29] T.Alipour-Fard, H. Arefi, " Structure Aware Generative Adversarial Networks for Hyperspectral Image Classification", JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING , vol. 13, pp. 5424–5438, 2020.

[30] G. Sumbul et al., " Bigearthnet: A large-scale benchmark archive for remote sensing image understanding", IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, pp. 5901-5904, 2019.

[31] J. Li, K. Zheng, J. Yao, L. Gao and D. Hong, " Deep Unsupervised Blind Hyperspectral and Multispectral Data Fusion", IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, 2022.

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.9, No.4, Winter 2022

Research Paper

The Accuracy Improvement of the Hyperspectral Satellite Image Classification by Using the Development of a Convolutional Neural Network and Deep Learning

Somayeh Mahmoudi[1], Najmeh Neisani Samani[2*], Ara Toomanian[2]

1- PhD Student in Remote Sensing and GIS, Faculty of Geography, University of Tehran
2- Associate Professor, Department of Remote Sensing and GIS, Faculty of Geography, University of Tehran

Abstract

Nowadays, with the advancement of technology, numerous sensors provide high spatial and spectral resolution images. So far, several methods have been proposed for hyperspectral image classification, each of which seeks to overcome a number of computational and processing challenges of hyperspectral data. The efficiency of multi-layer perceptron neural networks is greatly reduced due to the increase in the number of parameters along with the increase of the layers, which is essential in complex topics such as hyperspectral image classification. In recent years, the concept of deep learning, especially convolutional neural networks (CNN), has attracted the attention of pattern recognition researchers due to the automatic generation of features and the reduction of parameters compared to the multi-layer perceptron neural networks by sharing the parameters in each layer. The goal of the present study is to develop a convolutional neural networks (CNN in order to classify hyperspectral images. The innovation of this study is to provide a framework to use deep learning. The proposed framework includes four steps. The first step is to reduce dimension by using the sub-space method, the second step is to prepare the CNN inputs, the third step is to augment the teaching data, and the fourth step is to design the CNN architecture. Implementation of the proposed framework on the benchmark data of the University of Pavia, despite the use of a limited number of educational data, led to the classification accuracy of 98.3%.

Key words : *Hyperspectral Images, Classification, Deep Learning, Convolutional Neural Network.*

Correspondence Address*: Tehran, Enghelab Ave, Ghods Ave, Azin Alley, Faculty of Geography, Remote Sensing and GIS Department, University of Tehran.*
Tel*: : +98 2161113540*
Email : *nneysani@ut.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2025-09-07]
[DOR: 20.1001.1.20089635.1400.9.4.2.5]
[DOI: 10.52547/jgit.9.4.109]