

بهبود دقت طبقه‌بندی تصاویر ابرطیفی با استفاده از شبکه‌های عصبی پیچشی

مهسا تکیه‌نژاد^۱، عطاالله ابراهیم‌زاده^{۲*}، ملیحه احمدی^۳

۱- کارشناسی ارشد مخابرات سیستم، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل

۲- استاد گروه مخابرات سیستم، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل

۳- دکتری مخابرات سیستم، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی شیراز

تاریخ دریافت مقاله: ۱۴۰۱/۰۷/۲۳ تاریخ پذیرش مقاله: ۱۴۰۲/۰۳/۱۳

چکیده

طبقه‌بندی تصاویر ابرطیفی اهمیت زیادی در تجزیه و تحلیل تصاویر سنجش از دور دارند. روش‌های یادگیری عمیق با موفقیت برای طبقه‌بندی داده‌های سنجش از دور مورد استفاده قرار گرفته‌اند. در سال‌های اخیر، شبکه‌های عصبی پیچشی (CNNs) کاربرد قابل ملاحظه‌ای در طبقه‌بندی تصاویر ابرطیفی داشته‌اند که هر یک سعی در غلبه بر چالش‌های محاسباتی و پردازشی داده‌های ابرطیفی داشته‌اند. با افزایش تعداد پارامترها و لایه‌های شبکه‌های عصبی پیچشی کارایی آن‌ها در حل مسائل پیچیده کاهش می‌یابد. به همین دلیل در مقاله حاضر یک معماری جدید از شبکه‌های عصبی پیچشی معرفی شده است تا علاوه بر داشتن عملکرد بالای شبکه، زمان محاسباتی را نیز به صورت چشمگیری کاهش دهد. در روش پیشنهادی از اطلاعات طیفی-مکانی به عنوان ورودی شبکه و به جهت کاهش باندهای طیفی از فرآیند آنالیز مولفه‌های اصلی استفاده می‌شود. همچنین برای جلوگیری از بیش‌برازش از ترکیب نرمال‌سازی بسته‌ای و حذف تصادفی استفاده می‌گردد. در این روش شبکه عصبی پیچشی دو بعدی شامل لایه‌های پیچشی، لایه‌های ادغام و لایه‌های کاملاً متصل می‌باشد. علاوه بر استفاده از آنالیز مولفه‌های اصلی، یافتن پنجره مناسب در روش پیشنهادی مورد توجه قرار گرفته است. جهت مقایسه مدل پیشنهادی با دیگر معماری‌ها، آزمایش‌های این مقاله بر مجموعه داده‌های ایندین پاینس، دانشگاه پاولیا و سالیناس صورت گرفته است. نتایج شبیه‌سازی نشان می‌دهد که روش پیشنهادی علاوه بر داشتن دقت بالای موفقیت (۱۰۰ درصد) در طبقه‌بندی و مدت زمان کمتر، دارای پیچیدگی کمتری نسبت به مدل‌های موجود می‌باشد.

کلید واژه‌ها: طبقه‌بندی تصاویر ابرطیفی، شبکه عصبی پیچشی، آنالیز مولفه‌های اصلی، انتخاب پنجره مناسب.

* نویسنده مکاتبه کننده: مازندران - بابل - خیابان شریعتی - دانشگاه صنعتی نوشیروانی بابل - دانشکده برق و کامپیوتر - گروه مخابرات سیستم.

تلفن: ۰۱۱۳۲۳۳۹۲۱۴

۱- مقدمه

طیف‌سنجی تصاویر^۱ که به نام تصویربرداری ابرطیفی نیز شناخته می‌شود، اندازه‌گیری، تحلیل و تفسیر طیف‌هایی است که از یک شیء در طول موج‌های کوتاه، متوسط و بلند به وسیله سنجنده‌های هواپرد یا فضاپرد ماهواره‌ای گرفته می‌شود [۱]. این تکنیک نسبتاً جدید در اوایل دهه ۱۹۸۰ با تولید یک سنجنده طیف‌سنج هواپرد که بعدها "آویریس" (Aviris)^۲ نامیده شد، آغاز گردید و توانست به کمک باندهای مختلف و متنوع تصویربرداری، اطلاعات جامع و کاملی از سطح زمین در اختیار محققان قرار دهد. هدف اصلی در تصویربرداری ابرطیفی، به دست آوردن محتوای طیفی برای هر پیکسل از تصویر است. اثر طیفی مربوط به مواد مختلف منحصر به فرد است مانند اثر انگشت و در نتیجه به دست آوردن آن به شناسایی مواد کمک می‌کند. تصویربرداری ابرطیفی باندهای طیفی داده‌ها را برای هر پیکسل در تصویر یک صحنه به دست می‌آورد. حسگرهای تصویربرداری ابرطیفی اطلاعات را به عنوان مجموعه‌ای از تصاویر جمع‌آوری می‌کنند که نشان - دهنده باندهای مختلف طیف الکترومغناطیسی است. این تصاویر برای تشکیل یک مکعب داده ابرطیفی ترکیب می‌شوند که می‌توانند پردازش و تجزیه و تحلیل شوند تا داده‌های طیفی را برای طیف گسترده‌ای از کاربردها تفسیر کنند. تصاویر ابرطیفی (HSI)^۳ با قابلیت قدرت تفکیک طیفی بالا موجب پدید آمدن پیشرفت‌های چشم‌گیری در حوزه‌های مختلف شده است. تصاویر ابرطیفی توانایی زیادی برای بهره‌گیری در مسائل مختلف سنجش از دور دارند، اما با مشکلات مختلفی از جمله ابعاد زیاد داده‌ها و کمبود نمونه‌های آموزشی روبرو هستند. به همین دلیل روش‌های خاص پردازش و طبقه‌بندی^۴ برای این نوع از تصاویر طراحی و

معرفی شده‌اند. روش‌های طبقه‌بندی به سه روش کلی از جمله طبقه‌بندی تحت نظارت^۵، بدون نظارت^۶ و نیمه نظارتی^۷ تقسیم می‌شوند. روش طبقه‌بندی تحت نظارت یک روش رایج طبقه‌بندی تصاویر ابرطیفی می‌باشد. در این روش، شبکه با دریافت اطلاعات برچسب زده شده آموزش می‌بیند و سعی می‌کند الگوی بین داده‌ها و برچسب‌هایشان را به صورت یک تابع، یاد گرفته و برچسب داده‌های جدید و دیده نشده را پیش‌بینی کند. برخی روش‌های طبقه‌بندی نظارت‌شده شامل روش ماشین بردار پشتیبان^۸ (SVM)، شبکه‌های عصبی مصنوعی^۹ (ANN) و طبقه‌بند حداقل فاصله^{۱۰} (MDC) می‌باشند.

فرید ملگانی و همکاران (۲۰۰۴) از ماشین بردار پشتیبان برای طبقه‌بندی تصاویر ابرطیفی استفاده کردند که از آن زمان به‌طور گسترده به عنوان یک روش تحت نظارت مورد استفاده قرار گرفته است. SVM قادر است با استفاده از ویژگی‌های طیفی و بافتی، طبقه‌بندی را با دقت خوبی انجام دهد [۲]. یوشی چن و همکاران (۲۰۱۴) از رمزگذار خودکار پشته‌ای^{۱۱} (SAE) به عنوان طبقه‌بند در تصاویر ابرطیفی استفاده کردند. پس از آموزش تمام لایه‌ها در SAE، برای تنظیم دقیق و طبقه‌بندی، همه لایه‌ها به هم متصل شده و یک طبقه‌بند رگرسیون لجستیک^{۱۲} (LR) در شبکه قرار می‌گیرد [۳].

تانگ لی و همکاران (۲۰۱۴)، مدلی را پیشنهاد دادند که در آن شبکه باور عمیق^{۱۳} (DBN) با دو لایه پنهان ویژگی‌های طیفی-مکانی پیکسل‌های مکعب ابرطیفی را استخراج می‌کند به طوری که ابتدا ابعاد طیفی مکعب

^۵ Supervision Classification^۶ Unsupervised Classification^۷ Semi-supervised Classification^۸ Support Vector Machine^۹ Artificial Neural Networks^{۱۰} Minimum Distance Classifier^{۱۱} Stacked Autoencoder^{۱۲} Logistic Regression^{۱۳} Deep Belief Network^۱ Imaging Spectroscopy^۲ Airborne Visible Infrared Imaging Spectrometer^۳ Hyper Spectral Images^۴ Classification

این مطالعه شامل یک لایه پیچشی، یک لایه ادغام بیشینه^{۱۷}، یک لایه کاملاً متصل^{۱۸} (FC) و لایه خروجی است. اگرچه دقت طبقه‌بندی بدست آمده، توانایی CNN را در طبقه‌بندی تصاویر ابرطیفی نشان داده است، اما حداکثر دقت کلی ۹۲/۶٪ نشان می‌دهد که نتایج می‌توانند بیشتر از این‌ها بهبود یابند [۱۱]. فیلیپ سوان و همکاران (۱۹۸۱) برای کاهش ابعاد طیفی داده‌های ورودی، یک CNN برای استخراج ویژگی‌ها پیشنهاد کردند. بدین صورت که برای استخراج ویژگی‌های پیکسل‌ها، ابتدا بعد داده اصلی با استفاده از^{۱۹} (BLDE) کاهش می‌یابد و تنها چند باند اصلی ابتدایی حفظ می‌شود، سپس یک پنجره مربع در اطراف هر پیکسل آموزشی تشکیل می‌شود. این پنجره‌ها به عنوان مجموعه داده‌های آموزشی برای آموزش CNN به شیوه‌ای تحت نظارت استفاده می‌شوند. ویژگی‌ها در آخرین لایه چارچوب CNN مسطح شده‌اند و بردارهای ویژگی را تشکیل می‌دهند [۱۲]. یوشی چن و همکاران (۲۰۱۶) برای بهره‌برداری از اطلاعات طیفی و مکانی، یک^{۲۰} (3D-CNN) طراحی کرده‌اند. این معماری می‌تواند اطلاعات طیفی و مکانی را به‌طور همزمان استخراج کند. پنجره سه‌بعدی در اطراف پیکسل هدف، در مکعب ابرطیفی تشکیل می‌شود و به عنوان ورودی، این پنجره سه‌بعدی را به شبکه تغذیه می‌کند. پس از چندین لایه پیچشی و ادغام، ویژگی‌های استخراج‌شده به شکل یک بردار خواهد بود که حاوی اطلاعات عمیق طیفی و مکانی پیکسل هدف است. این شبکه می‌تواند پارامترهای آموزش پذیر کمتری داشته باشد اما هزینه محاسبات به دلیل لایه پیچشی سه‌بعدی در امتداد باندهای طیفی افزایش می‌یابد [۱۳].

کنستانتینز ماکانتاتیز و همکاران (۲۰۱۶) مشابه یوشی

ابرطیفی، با استفاده از آنالیز مولفه اصلی^{۱۴} (PCA) کاهش می‌یابد و تنها سه جزء اصلی اول حفظ می‌شوند. در مرحله بعد، یک پنجره سه‌بعدی به اندازه ۷×۷×۳ در اطراف هر نمونه آموزشی (پیکسل) تشکیل می‌شود که در ادامه تبدیل به بردار شده و به عنوان ورودی به شبکه داده می‌شود. طرح طبقه‌بندی نشان داده شده در مقاله-ی ذکر شده دارای چهار لایه ی ورودی، دو ماشین بولتزمن محدود^{۱۵} (RBM)، و یک لایه رگرسیون لجستیک می‌باشد. در ساختار هر دو روش DBN و SAE با توجه به وجود اتصالات کامل بین لایه‌های مختلف تعداد فراوانی از نمونه‌های آموزشی برای آموزش تعداد زیاد پارامترها احتیاج می‌باشد، که این امر به دلیل کمبود نمونه‌های آموزشی یک مشکل اساسی است. علاوه بر این، SAE و DBN نمی‌توانند اطلاعات مکانی را به صورت موثر استخراج کنند، زیرا ورودی این روش‌ها باید به صورت برداری باشد و باید اطلاعات مکانی، قبل از آموزش استخراج شوند و در یک بردار همراه اطلاعات طیفی وارد شبکه شوند [۴]. برای حل این مشکلات در سال‌های اخیر شبکه‌های عصبی پیچشی^{۱۶} (CNN) معرفی شده‌اند و بسیار مورد توجه قرار گرفت‌اند. CNN با در نظر گرفتن ارتباطات محلی می‌تواند اطلاعات مکانی را استخراج کند. همچنین در این شبکه‌ها استراتژی اشتراک وزن باعث کاهش چشمگیری در پارامترهای قابل آموزش شبکه می‌شود. برخی از مطالعات اخیر از شبکه عصبی پیچشی به عنوان ساختار شبکه عمیق برای استخراج اطلاعات طیفی و مکانی از شیوه‌های تحت نظارت و بدون نظارت استفاده کرده‌اند [۹، ۸، ۷، ۶، ۵، ۴، ۱۰]. یانگیو هوانگ و همکاران (۲۰۱۵)، مدلی را پیشنهاد کردند که در آن هر بردار نمونه (طیف نمونه) را به عنوان یک تصویر دو بعدی در نظر می‌گیرد، بنابراین ورودی شبکه، امضای طیفی هر پیکسل است. ساختار CNN مورد استفاده در

¹⁷ Max Pooling

¹⁸ Fully Connected

¹⁹ Balanced Local Discriminant Embedding

²⁰ 3 Dimension Convolutional Neural Network

¹⁴ Principle Component Analysis

¹⁵ Restricted Boltzmann Machine

¹⁶ Convolutional Neural Network

می باشد. ساختار تحقیق بدین صورت است که در قسمت دوم به روش شناسی پرداخته می شود که در آن شبکه های عصبی پیچشی و انواع لایه های آن و همچنین معماری شبکه پیشنهادی بررسی می گردد، در قسمت سوم منطقه های مورد مطالعه بررسی می شود، در قسمت چهارم نتایج و بحث ارائه شده است و در قسمت پایانی، نتیجه گیری و پیشنهادات ارائه خواهد شد.

۲- روش شناسی

در این بخش روش ها و ابزارهای تحقیق معرفی می شود. ابزارهای تحقیق شامل شبکه های عصبی پیچشی و لایه های مختلف آن و همچنین توصیف روش پیشنهادی می باشد.

۲-۱- شبکه های عصبی پیچشی

در میان انواع مختلف شبکه های عصبی عمیق، شبکه های عصبی پیچشی بیشتر مورد مطالعه قرار گرفته اند. *CNN* یک معماری یادگیری عمیق شناخته شده است که از مکانیسم ادراک بصری طبیعی موجودات زنده الهام گرفته است. این شبکه ها به سرعت برای حل مشکلات بینایی رایانه مانند طبقه بندی تصویر و تشخیص اشیا در سال های اخیر توسعه یافته اند. *CNN* یک شبکه عصبی است که از یک عملیات ریاضی به نام کانولوشن در لایه های خود استفاده می کند [۱۶]. عملیات کانولوشن به صورت زیر در رابطه (۱) تعریف می شود:

رابطه (۱)

$$(x * w)(t) = \int_{-\infty}^{\infty} x(a)w(t-a)da.$$

در رابطه (۱)، x ورودی، w هسته و t بعد می باشد. اگر x و w در امتداد محور t گسسته باشند، عملیات کانولوشنی به صورت زیر در رابطه (۲) بیان می شود:

رابطه (۲)

$$(x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a)$$

و چن و همکارانش در بررسی قبلی، مدلی پیشنهاد کردند که در آن، برای استخراج اطلاعات طیفی و مکانی به طور همزمان از مکعب ابرطیفی داده استفاده شده است. به عبارت دیگر، یک پنجره سه بعدی در اطراف هر پیکسل هدف تشکیل می شود که حاوی اطلاعات طیفی و مکانی است. با این حال، بر خلاف مدل یوشی چن و همکارانش، که این پنجره سه بعدی را به عنوان ورودی مستقیم به *CNN* می دهد، در این مدل $(R-PCA)^{21}$ روی هر پنجره سه بعدی برای کاهش ابعاد داده های ورودی اعمال می شود [۱۴].

سوالپا کومار روی و همکاران (۲۰۲۰)، یک مدل پیچشی که ترکیبی از شبکه عصبی پیچشی دوبعدی و سه بعدی می باشد را برای استفاده از مزایای قابلیت یادگیری خودکار ویژگی های دو بعدی و سه بعدی ارائه دادند، که این روش سعی در غلبه بر کاستی های مدل های قبلی دارد. لایه های پیچشی دوبعدی و سه بعدی برای مدل پیشنهادی به گونه ای تدوین شده اند که از ویژگی های طیفی و مکانی تا حد بسیاری برای دستیابی به حداکثر دقت ممکن استفاده کنند. این روش شامل سه لایه پیچشی سه بعدی، یک لایه پیچشی دوبعدی و سه لایه کاملاً متصل می باشد که مدل "هیبرید اس ان" (*HybridSN*)²² نامگذاری شده است. روش پیشنهادی با روش های دیگر از نظر دقت طبقه بندی مقایسه شد و مشاهده گردید که این روش، از روش های دیگر نتایج بهتری ارائه می دهد [۱۵].

بررسی ها نشان می دهد که اکثر تحقیقات ارائه شده از لایه های پیچشی سه بعدی برای طبقه بندی استفاده می کنند که این مسئله پیچیدگی کار را زیاد می کند. انتخاب پنجره مناسب می تواند دقت طبقه بندی را افزایش دهد. همچنین کاهش تعداد باندهای طیفی می تواند به کاهش پیچیدگی سیستم نهایی منجر شود. روش پیشنهادی این مقاله به دنبال تحقق این دو هدف

²¹ Random Principle Component Analysis

²² Hybrid Spectral Convolutional Neural Network

۲-۱-۱- لایه پیچشی

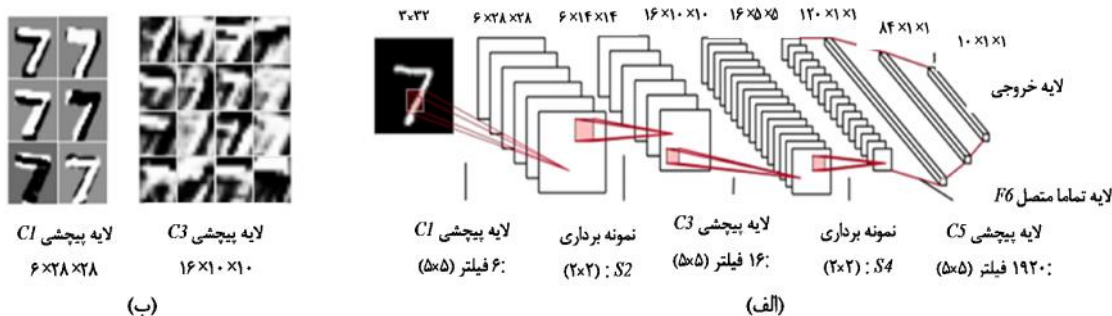
هدف لایه پیچشی یادگیری نقشه ویژگی‌های ورودی است. همانطور که در شکل (۱-الف) نشان داده شده است، لایه پیچشی از چندین هسته تشکیل می‌شود که برای محاسبه نقشه ویژگی‌های مختلف استفاده می‌شود. شکل (۱-ب) نقشه‌های ویژگی رقم هفت را نشان می‌دهد که توسط دو لایه پیچشی اول آموزش دیده شده است. همچنین، نقشه‌های ویژگی هر لایه در یک بلوک متفاوت نمایش داده می‌شود بطوری‌که، هسته‌های لایه پیچشی اول برای شناسایی ویژگی‌های سطح پایین مانند لبه‌ها و منحنی‌ها طراحی شده‌اند، درحالی‌که هسته‌های لایه‌های بالاتر یاد می‌گیرند که ویژگی‌های انتزاعی بیشتری را رمزگذاری کنند. با انباشتن چندین لایه پیچشی و ادغام، می‌توان به تدریج نمایش ویژگی‌های سطح بالاتر را استخراج نمود.

برای ورودی‌ها و هسته‌هایی که بیش از یک بعد دارند، می‌توان عملیات کانولوشنی را نیز تعریف کرد. به عنوان مثال، یک عملیات کانولوشنی بین یک تصویر دو بعدی I و یک هسته دو بعدی K در رابطه (۳) بیان می‌شود:

رابطه (۳)

$$S(i, j) = (H * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n)$$

در رابطه (۳)، خروجی S یک نقشه ویژگی (Feature Map) ناشی از عملیات کانولوشنی در نظر گرفته می‌شود. بسته به مقادیر یا وزن هسته K ، ماتریس S ویژگی‌های متفاوت استخراج شده از I را نشان می‌دهد. یک معماری CNN از چندین لایه پیچشی تشکیل شده است. شهود پشت این امر این است که هر چه شبکه عمیق‌تر باشد، ویژگی‌های استخراج شده پیچیده‌تر و در نتیجه عمل کرد بهتری به دست می‌آید [۱۷]. با این حال لازم به ذکر است که با افزایش لایه‌ها پیچیدگی شبکه نیز افزایش می‌یابد.



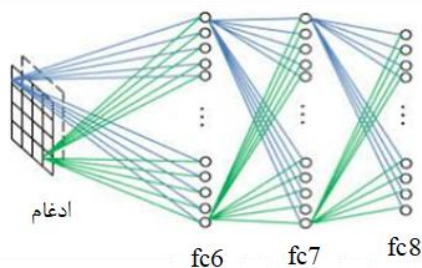
شکل ۱: (الف) معماری شبکه LeNet-5، (ب) تجسم ویژگی‌ها در شبکه LeNet-5 [۱۸]

ویژگی مد نظر، ابعاد هسته آن انتخاب می‌شود. لایه ادغام دارای پارامتر قابل آموزش نمی‌باشد و صرفاً یک نمونه برداری ساده و مؤثر انجام می‌دهد. این لایه اندازه نقشه‌های ویژگی و پارامترهای شبکه را کاهش می‌دهد. پیاده‌سازی آن به دو صورت ادغام بیشینه و ادغام میانگین (Mean Poling) است. در روش ادغام بیشینه، معمولاً مربع‌هایی با اندازه ۴x۴ انتخاب می‌شوند و روی

۲-۱-۲- لایه ادغام

این لایه برای نمونه برداری از نقشه‌های ویژگی در امتداد ابعاد مکانی شامل عرض و ارتفاع، برای کاهش مقدار پارامترها و محاسبات در شبکه و در نتیجه، برای کنترل بیش‌برازش استفاده می‌شود. لایه ادغام بر روی تمام نورون‌های لایه پیچشی عمل می‌کند و معمولاً بعد از یک لایه پیچشی قرار می‌گیرد که با توجه به کیفیت

نورون در لایه دیگر متصل می‌کند. لایه کاملاً متصل این امکان را فراهم می‌سازد تا نتیجه شبکه، در قالب یک بردار با اندازه مشخص ارائه شود. از این بردار می‌توان برای کلاس‌بندی استفاده کرد و یا اینکه از آن، در جهت ادامه پردازش‌های بعدی بهره برد. مشکل اساسی لایه‌های کاملاً متصل این است که دارای تعداد بسیار زیادی پارامتر هستند که موجب هزینه پردازشی بسیار بالا و طولانی شدن مدت زمان آموزش می‌شود. در شکل (۳) عملیات لایه کاملاً متصل نشان داده شده است [۱۹].

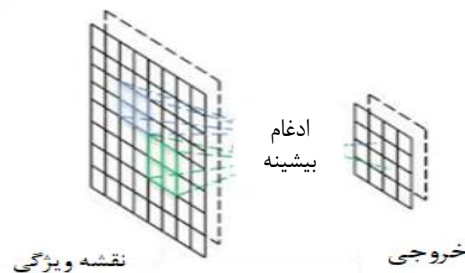


شکل ۲: عملیات لایه ادغام [۲۰]

تصویر حرکت داده می‌شوند و در هر چهار پیکسل، مقدار بیشینه آن‌ها انتخاب شده و به لایه بعدی انتقال داده می‌شود. در شکل (۲) یک لایه ادغام مشاهده می‌شود که بر روی صفحه 8×8 اعمال شده و اندازه فیلتر ادغام آن 2×2 است. در نتیجه، ابعاد صفحه به 4×4 کاهش پیدا کرده است. با مقایسه بین این دو مدل ادغام، ثابت شده است که ادغام بیشینه موجب همگرایی سریع‌تر و انتخاب ویژگی‌های متمایزتر می‌شود.

۲-۱-۳- لایه کاملاً متصل

لایه‌های کاملاً متصل هر نورون در یک لایه را، به یک



شکل ۳: عملیات لایه کاملاً متصل [۲۰]

و B تعداد باندهای طیفی است. PCA معمولاً برای کاهش ابعاد طیفی استفاده شده است. از آنجایی که تعداد باندهای طیفی B مشخص است، پس از اعمال PCA ، بعد طیفی به B_r کاهش می‌یابد. هدف آن است که با در نظر گرفتن تعداد کمی از مولفه‌های اصلی، حداکثر اطلاعات که حدود 99.9% می‌باشد، حفظ شوند. در داده‌های ابرطیفی پیکسل‌های مجاور از نظر ویژگی‌های طیفی - مکانی کاملاً مشابه هم هستند. بنابراین با این ایده، بعد مکانی یعنی (H, W) به تکه‌های مربع کوچک با اندازه پنجره‌های $P \times P$ تقسیم می‌شود، بطوریکه پیکسل مرکزی و پیکسل‌های اطراف آن برچسب یکسانی خواهند داشت. فرض کنید یک پیکسل $I_{h,w}$ در مکان (h, w) با برچسب $g_{x,y}$ با استفاده از پنجره مربعی T به اندازه $P \times P$ وجود دارد بطوریکه $I_{h,w}$ در مرکز پنجره باشد و با در نظر گرفتن ابعاد طیفی

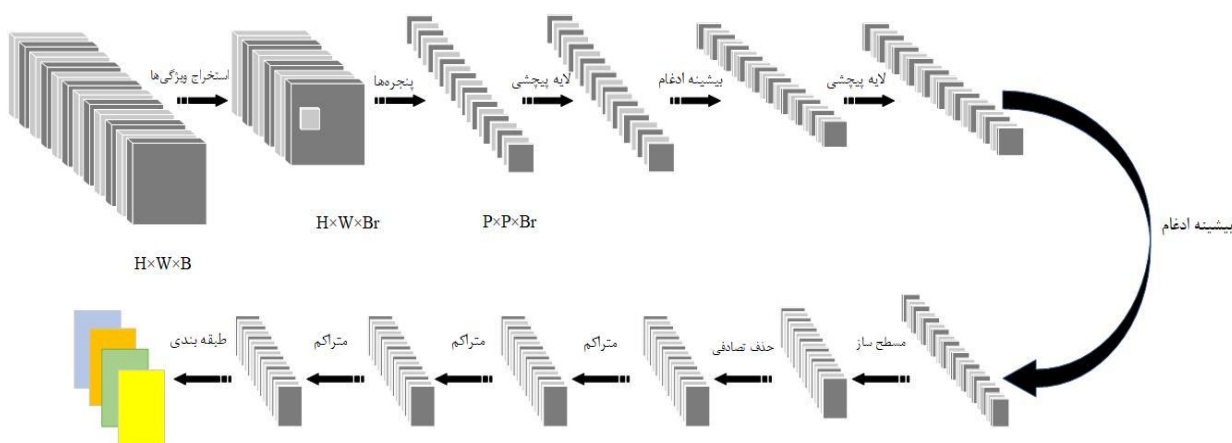
۲-۲- شرح تحقیق

رویکرد ما براساس یک چارچوب یکپارچه است که اطلاعات طیفی و مکانی را در یک مرحله واحد با یکدیگر ترکیب می‌کند و همزمان ویژگی‌های طیفی - مکانی سطح بالا را استخراج می‌کند. بطور خاص، در این پژوهش بهره‌برداری از یک شبکه عصبی پیچشی دوبعدی پیشنهاد شده است. ویژگی‌های طیفی با ویژگی‌های بازتابی در هر پیکسل برای هر باند طیفی مرتبط است، درحالی که ویژگی‌های مکانی با در نظر گرفتن همسایگان آن استخراج می‌شود. پس از اعمال PCA ، تصویر ابرطیفی گرفته شده به پنجره‌هایی تجزیه می‌شود، که هر کدام حاوی اطلاعات طیفی و مکانی برای یک پیکسل خاص می‌باشند. داده ابرطیفی با ابعاد $H \times W \times B$ می‌باشد، بطوریکه در آن H ارتفاع، W عرض

۲-۱-۲ معماری مدل شبکه عصبی پیچشی

پیشنهادی

مدل شبکه عصبی پیچشی پیشنهادی، از دو لایه پیچشی دوبعدی و سه لایه کاملاً متصل تشکیل شده است. شکل (۴) معماری شبکه پیشنهادی را نمایش می‌دهد.



شکل ۴: معماری شبکه پیشنهادی

در نهایت ویژگی‌ها توسط لایه مسطح ساز^۲، تبدیل به یک بردار ویژگی می‌شوند تا در اختیار لایه کاملاً متصل قرار گیرند. در مرحله اول ۳۵٪ نرخ کاهش^۳ برای جلوگیری از هرگونه تطبیق بیش‌برازش در طول عملیات آموزشی اعمال می‌شود و پس از آن سه لایه کاملاً متصل بکار گرفته می‌شود که برای دو لایه اول تابع فعال سازی غیرخطی رلو و برای لایه آخر تابع فعال سازی بیشینه هموار^۴ استفاده شده است.

۳- منطقه مورد مطالعه

در این قسمت انواع مجموعه داده‌های مورد استفاده در این تحقیق مورد بررسی قرار می‌گیرد.

۳-۱- آیندین پاینس^۵

این مجموعه داده‌ها توسط سنجنده هوابرد آویریس از

کاهش یافته، پنجره T را می‌توان یک تنسور سه‌بعدی با ابعاد $P \times P \times Br$ در نظر گرفت. بنابراین کل مکعب داده ابرطیفی را می‌توان به مجموعه‌ای به شکل $S = \{T_{h,w}\}$ نوشت که از تقسیم پنجره‌های مربعی با پارامترهای $w = 1, 2, \dots, W$ و $h = 1, 2, \dots, H$ تشکیل شده است، در حالی که S در مجموع N تعداد پنجره دارد.

لایه‌های پیچشی با لایه‌های ادغام بیشینه دنبال می‌شوند. در اولین لایه پیچشی دوبعدی از ۳۲ فیلتر^۳ که بیانگر اندازه هسته می‌باشد، استفاده شده است و همچنین پدینگ برای حفظ اندازه خروجی سطح بعدی مورد استفاده قرار گرفته است. پس از اولین لایه پیچشی دوبعدی، لایه ادغام بیشینه 2×2 با گام ۲، برای نمونه‌برداری از نقشه‌های ویژگی در امتداد ابعاد مکانی برای کاهش مقدار پارامترها و محاسبات در شبکه و در نتیجه کنترل بیش‌برازش بکار گرفته شده است. در ادامه یک لایه پیچشی دوبعدی دیگر با ۶۴ فیلتر^۳ اعمال می‌شود و به دنبال آن لایه ادغام بیشینه 2×2 با گام ۲ مجدداً اعمال می‌شود. لازم به ذکر است که در هر دو لایه پیچشی دوبعدی، تابع فعال سازی غیرخطی رلو^۱ ($ReLU$) استفاده شده است.

^۲ Flatten

^۳ Drop Rate

^۴ Softmax

^۵ Salinas

^۱ Rectified Linear Activation Function

کشاورزی و یک سوم آن را جنگل و گیاه در بر گرفته است.

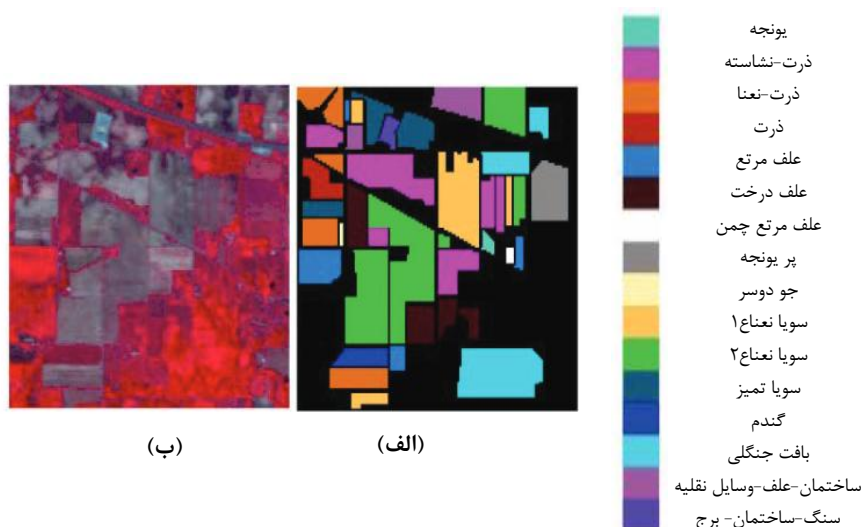
بطور کلی طبقه‌بندی این داده به علت وجود کلاس‌هایی با رفتار مشابه همواره چالش برانگیز بوده است. یک ترکیب رنگی کاذب (*False color*) و نقشه حقیقت زمین (*Ground Truth (GT)*) این داده توسط آزمایشگاه سنجش از دور دانشگاه پردو تهیه شده است (شکل (۵)).

در جدول (۱) نیز انواع کلاس‌های این مجموعه داده‌ها و تعداد نمونه‌های آن نمایش داده شده است.

منطقه ایندین پاین در شمال غربی ایندیانا در ایالات متحده آمریکا جمع آوری شده است. این تصویر با قدرت تفکیک مکانی ۲۰ متر و ابعاد 145×145 پیکسل و ۲۲۴ باند طیفی در محدوده $0.4 - 2.5$ میکرومتر می‌باشد. از آنجایی که ۲۴ باند طیفی توسط نویز آلوده شده‌است، داده‌های تصحیح شده شامل ۲۰۰ باند طیفی پس از حذف باندهای نویزی می‌باشند. این مجموعه داده‌ها از ۱۰۲۴۹ نمونه برچسب‌دار تشکیل شده است، که شامل ۱۶ کلاس از گیاهانی نظیر ذرت، سویا، علف، گندم و غیره است. حدود دو سوم این تصویر را مناطق

جدول ۱: کلاس‌های ایندین پاینس

شماره کلاس	نام کلاس	تعداد نمونه‌ها
۱	یونجه	۴۶
۲	ذرت-نشاسته	۱۴۲۸
۳	ذرت-نعنا	۸۳۰
۴	ذرت	۲۷۳
۵	علف مرتع	۴۸۳
۶	علف درخت	۷۳۰
۷	علف مرتع چمن	۲۸
۸	پر یونجه	۴۷۸
۹	جو دوسر	۲۰
۱۰	سویا نعنای ۱	۹۷۲
۱۱	سویا نعنای ۲	۲۴۵۵
۱۲	سویا تمیز	۵۹۳
۱۳	گندم	۲۰۵
۱۴	بافت جنگلی	۱۲۶۵
۱۵	ساختمان-علف-وسایل نقلیه	۳۸۶
۱۶	سنگ-ساختمان-برج	۹۳

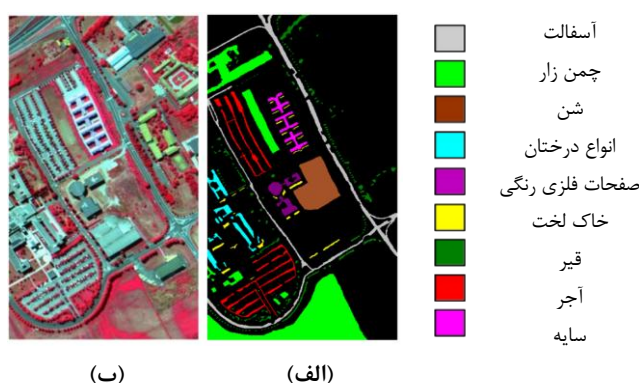


شکل ۵: مجموعه داده ایندین پائینس. الف) نقشه حقیقت زمین، ب) تصویر رنگی کاذب [۲۶]

داده‌های تصحیح شده با اندازه 610×340 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 (۶) تصویر این مجموعه داده‌ها قابل مشاهده می‌باشند. در جدول (۲) نیز انواع کلاس‌های این مجموعه داده‌ها و تعداد نمونه‌های آن نمایش داده شده است.

۳-۲- دانشگاه پاویا (Pavia University)

این تصاویر توسط سنجنده هواپرد روسیس (Reflective Optics System Imaging Spectrometer) طی یک پرواز بر فراز دانشگاه پاویا در شمال ایتالیا، در سال ۲۰۰۱ اخذ شده است. داده‌های تصحیح نشده شامل ۱۱۵ باند طیفی در محدوده طول موج‌های $0.43-0.86$ میکرومتر و دارای قدرت تفکیک مکانی $1/3$ متر هستند. از آنجایی که ۱۲ باند طیفی توسط نویز آلوده شده‌است،



شکل ۶: مجموعه داده‌های دانشگاه پاویا [۲۱]. الف) نقشه حقیقت زمین، ب) تصویر رنگی کاذب

جدول ۲: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های دانشگاه پاویا

شماره کلاس	نام کلاس	تعداد نمونه‌ها
۱	آسفالت ^۱	۶۶۳۱
۲	چمن زار ^۲	۱۸۶۴۹
۳	شن ^۳	۲۰۹۹
۴	انواع درختان ^۴	۳۰۶۴
۵	صفحات فلزی رنگی ^۵	۱۳۴۵
۶	خاک لخت ^۶	۵۰۲۹
۷	قیر ^۷	۱۳۳۰
۸	آجر ^۸	۳۶۸۲
۹	سایه ^۹	۹۴۷

۳-۳- سالیناس (Salinas)

مجموعه داده‌های سالیناس شامل تصاویر ابرطیفی با ۵۱۲×۲۱۷ پیکسل و قدرت تفکیک مکانی ۳/۷ متر است که توسط سنجنده هواپرد آویریس (Airborne Visible Infrared Imaging Spectrometer) در دره سالیناس در کالیفرنیا گرفته شده است. تصاویر طیفی به‌طور پیوسته ۲۲۴ بانده طیفی در محدوده طول موج‌های ۰/۲-۴۲ میکرومتر را تصویر می‌کند. از آنجایی

که ۲۰ باند طیفی توسط نویز آلوده شده است، داده‌های تصحیح شده شامل ۲۰۴ باند طیفی پس از حذف باندهای نویزی می‌باشند. این مجموعه داده‌های شامل ۵۴۱۲۹ نمونه برچسب‌گذاری شده در ۱۶ کلاس می‌باشد. در شکل (۷) تصویر حقیقت زمین و ترکیب رنگی کاذب نشان داده شده است. در جدول (۳) نیز انواع کلاس‌های این مجموعه داده‌ها و تعداد نمونه‌ها نمایش داده شده است.

¹ Asphalt

² Meadows

³ Gravel

⁴ Trees

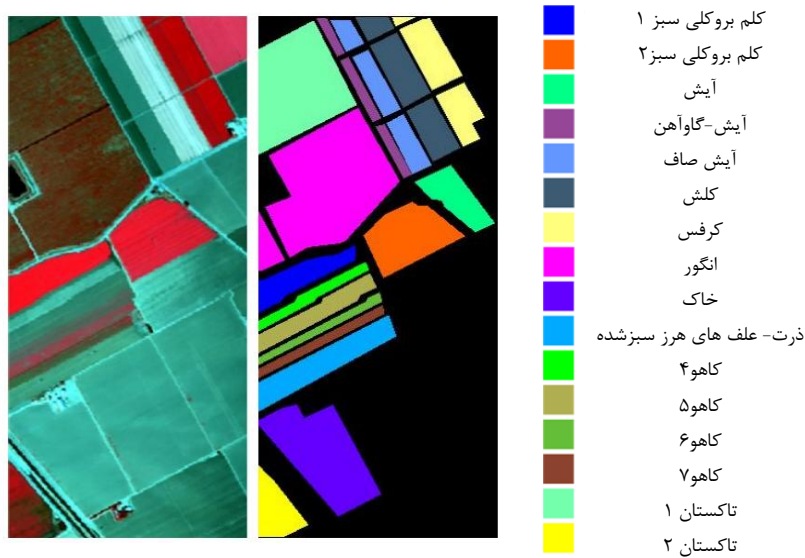
⁵ Painted metal sheets

⁶ Bare Soil

⁷ Bitumen

⁸ Self-Blocking Bricks

⁹ Shadows



(ب)

(الف)

شکل ۷: مجموعه داده‌های سالیانس. الف) نقشه حقیقت زمین، ب) تصویر رنگی کاذب [۲۱]

جدول ۳: کلاس‌های سالیانس

شماره کلاس	نام کلاس	تعداد نمونه‌ها
۱	کلم بروکلی سبز ۱	۲۰۰۹
۲	کلم بروکلی سبز ۲	۳۷۲۶
۳	آیش	۱۹۷۶
۴	آیش-گاواهن	۱۳۹۴
۵	آیش صاف	۲۶۷۸
۶	کلش	۳۹۵۹
۷	کرفس	۳۵۷۹
۸	انگور	۱۱۲۷۱
۹	خاک	۶۲۰۳
۱۰	ذرت- علف های هرز سبز شده	۳۲۷۸
۱۱	کاهو ۴	۱۰۶۸
۱۲	کاهو ۵	۱۹۲۷
۱۳	کاهو ۶	۹۱۶
۱۴	کاهو ۷	۱۰۷۰
۱۵	تاکستان ۱	۷۲۶۸
۱۶	تاکستان ۲	۱۸۰۷

۴- نتایج و بحث

این قسمت به بررسی مشخصات سیستم مورد استفاده، تعیین مناسب ترین فرآپارامترها برای دستیابی به بالاترین درصد دقت، پیاده سازی مدل شبکه و مقایسه روش پیشنهادی با دیگر روش ها می پردازد.

۴-۱- مشخصات سیستم مورد استفاده

سیستم مورد استفاده برای اجرای شبیه سازی الگوریتم موردنظر در این پژوهش، دارای مشخصات پردازنده اینتل (آر) با هسته (تی ام) آی ۷-۶۵۰۰^۱ و کارت گرافیک اینتل (آر) اچ دی گرافیک ۲۵۲۰ با رام ۸ گیگابایت می باشد. همچنین کدنویسی در محیط برنامه نویسی پایتون، نسخه ۳/۸ در بستر تنسورفلو^۲ و کتابخانه های سطح بالای کراس^۳ صورت گرفته است.

۴-۲- انتخاب فرآپارامترها

در مدل سازی شبکه، انتخاب فرآپارامترها از مهم ترین اقدامات اولیه می باشد، که این امر با انجام آزمون و خطا و سنجیدن تمام حالت های ممکن بر اساس تجربیات محققان در مقالات و مستندات [۱۴، ۱۵، ۲۲، ۲۳ و ۲۴] به دست آمده است. یکی از پارامترهای مهم تعیین کننده خروجی شبکه، اندازه پنجره می باشد.

۴-۲-۱- انتخاب اندازه پنجره بهینه

هدف اول، انتخاب مناسب ترین مقدار اندازه پنجره می باشد که این مرحله با در نظر گرفتن پارامترهای آورده شده در جدول (۴) صورت گرفته است. لازم به ذکر است که در تمام مراحل، جهت یافتن پارامترهای بهینه، نسبت مجموعه داده ها بدین گونه بوده است: از کل مجموعه داده ها، ۷۵٪ جهت یادگیری مدل آموزشی اختصاص داده شد (۵۰٪ داده ها جهت آموزش، ۲۵٪ جهت اعتبارسنجی) و ۲۵٪ جهت آزمایش در نظر گرفته شده است.

جدول ۴: فرآپارامترهای فرض شده، جهت انتخاب اندازه پنجره

اندازه دسته ها	۱۲۸
تابع فعال سازی	رلو
نرخ یادگیری	۰/۰۰۱
الگوریتم بهینه سازی	آدام ^۵
تعداد دوره ها	۲۰

با در نظر گرفتن فرضیات جدول بالا، شروع به آزمون و خطا برای یافتن مناسب ترین اندازه پنجره می شود که در جدول (۵) نتایج حاصل از آزمایش مقادیر مختلف پنجره بر روی مجموعه داده های مختلف ذکر شده است. بدین صورت که به ازای هر کدام از اندازه پنجره ها بر روی داده های مربوطه، شبکه پنج بار اجرا شده است و مقادیر دقت کلی و خطا بر اساس میانگین این پنج بار آموزش شبکه گزارش شده است.

همان طور که در جدول (۵) مشاهده می شود، مدل شبکه عصبی پیچشی پیشنهادی، بیشترین درصد دقت کلی و کمترین میزان خطا را به ازای اندازه پنجره ۱۹×۱۹ نشان داده است.

^۱ Intel(R) core(TM) i7-6500

^۲ HD Graphics 520 Intel(R)

^۳ Tensorflow

^۴ Keras

^۵ Adam

جدول ۵: محاسبه پارامتر اندازه پنجره، از نظر معیار دقت کلی و خطا

اندازه پنجره	دانشگاه پاویا		سالیانس		ایندین پاینس	
	% دقت کلی	خطا	% دقت کلی	خطا	% دقت کلی	خطا
۵×۵	۹۹/۴۲	۰/۰۴۲	۹۷/۶۰	۰/۰۸۸	۹۹/۴۱	۰/۰۹۴
۷×۷	۹۹/۹۱	۰/۰۲۶	۹۹/۴۰	۰/۰۴۵	۹۹/۶۱	۰/۰۸۰
۹×۹	۹۹/۹۳	۰/۰۳۱	۹۹/۶۷	۰/۰۳۱	۹۹/۶۱	۰/۰۷۱
۱۱×۱۱	۹۹/۶۴	۰/۰۲۳	۹۹/۹۲	۰/۰۲۳	۹۹/۹۴	۰/۰۶۵
۱۳×۱۳	۹۹/۹۸	۰/۰۱۳	۹۹/۹۹	۰/۰۱۷	۹۹/۷۶	۰/۰۴۹
۱۵×۱۵	۹۹/۹۷	۰/۰۱۶	۹۹/۹۹	۰/۰۱۶	۹۹/۹۲	۰/۰۶۷
۱۷×۱۷	۹۹/۹۹	۰/۰۱۵	۹۹/۹۷	۰/۰۱۸	۹۹/۲۶	۰/۰۸۳
۱۹×۱۹	۱۰۰/۰۰	۰/۰۱۲	۱۰۰/۰۰	۰/۰۰۷	۹۹/۹۶	۰/۰۴۸
۲۱×۲۱	۹۹/۹۶	۰/۰۱۲	۱۰۰/۰۰	۰/۰۱۱	۹۹/۳۷	۰/۰۶۲
۲۳×۲۳	۹۹/۹۱	۰/۰۱۷	۹۹/۹۴	۰/۰۱۳	۹۹/۹۴	۰/۰۴۹
۲۵×۲۵	۹۹/۹۸	۰/۰۰۸	۹۹/۹۸	۰/۰۰۸	۹۹/۹۶	۰/۰۵۶

۴-۲-۲- صحت‌سنجی مقادیر فراپارامترهای انتخاب شده

در تمامی مراحل این بخش، هریک از فراپارامترها با ثابت در نظر گرفتن دیگر فراپارامترهای تعیین شده در جدول (۴) محاسبه خواهد شد.

۴-۲-۱- صحت‌سنجی فراپارامتر اندازه دسته

در بخش اول برای بررسی درستی فرضیات انتخاب شده که با استفاده از آزمون و خطا و همچنین بر اساس مستندات مندرج در مقالات [۱۴، ۱۵، ۲۲، ۲۳، ۲۴]

بدست آمده است، صحت‌سنجی فراپارامتر اندازه دسته بررسی می‌شود که نتایج آن در جدول (۶) مشاهده می‌شود.

همان‌طور که در جدول (۶) مشاهده می‌شود بیشترین درصد دقت کلی و کمترین میزان خطا برای مدل شبکه عصبی پیچشی پیشنهادی، درحالتی که اندازه دسته برابر ۱۲۸ در نظر گرفته شده، به دست آمده است که این مقدار برابر با مقدار فرض شده در جدول (۴) می‌باشد.

جدول ۶: مقایسه معیار دقت و خطا، از نظر فراپارامتر اندازه دسته

اندازه دسته	دانشگاه پاویا		سالیانس		ایندین پاینس	
	% دقت کلی	خطا	% دقت کلی	خطا	% دقت کلی	خطا
۱۲۸	۱۰۰/۰۰	۰/۰۱۲	۱۰۰/۰۰	۰/۰۰۷	۹۹/۹۶	۰/۰۴۸
۲۵۶	۱۰۰/۰۰	۰/۰۱۴	۹۹/۸۶	۰/۰۲۰	۹۹/۸۰	۰/۰۶۳

در بخش دوم صحت‌سنجی تابع فعال‌سازی به ازای دو تابع رلو و سیگموئید بررسی شده است که نتایج آن در جدول (۷) مشاهده می‌شود. همان‌طور که در جدول (۷) مشاهده می‌شود بیشترین درصد دقت کلی، و کمترین

میزان خطا (به غیر از مجموعه داده‌های دانشگاه پاپویا و ایندین پاینس) برای مدل شبکه عصبی پیچشی پیشنهادی، درحالتی است که تابع فعال‌سازی رلو در نظر گرفته شده که این تابع فعال‌سازی همان تابع فعال‌سازی فرض شده در جدول (۴) می‌باشد.

این‌دین پائینس		سالی‌ن‌اس		دانشگاه پ‌ا‌ی‌ا		تابع فعال‌سازی
خطا	% دقت کلی	خطا	% دقت کلی	خطا	% دقت کلی	
۰/۰۴۸	۹۹/۹۶	۰/۰۰۷	۱۰۰/۰۰	۰/۰۱۲	۱۰۰/۰۰	رلو
۰/۰۳۲	۹۹/۷۶	۰/۱۶۱	۹۶/۳۹	۰/۰۰۸	۹۹/۹۳	سیگ‌موتید

در بخش سوم فرارپارامتر نرخ یادگیری، به ازای چهار مقدار مختلف که از متداول‌ترین مقادیر نرخ یادگیری در آموزش شبکه‌های پیچشی می‌باشند صحت‌سنجی شده است که نتایج آن در جدول (۸) مشاهده می‌شود.

همانطور که در جدول (۸) مشاهده می‌شود بیشترین درصد دقت کلی و کمترین میزان خطا برای مدل شبکه عصبی پیچشی پیشنهادی، درحالتی است که میزان نرخ یادگیری برابر ۰/۰۰۱ در نظر گرفته شده است که این مقدار نیز برابر با مقدار فرض شده در جدول (۴) می‌باشد.

ایندین پاینس		سالیناس		دانشگاه پابویا		نرخ یادگیری
خطا	% دقت کلی	خطا	% دقت کلی	خطا	% دقت کلی	
۵٫۸۱	۲۱٫۹۰	۶٫۸۶	۶٫۰۵	۴٫۹۰	۴۱٫۵۹	۰٫۱
۲٫۷۳	۲۳٫۹۰	۲٫۵۶	۲۰٫۸۱	۱٫۸۴	۴۳٫۶۹	۰٫۰۱
۰٫۰۴۸	۹۹٫۹۶	۰٫۰۰۷	۱۰۰٫۰۰	۰٫۰۱۲	۱۰۰٫۰۰	۰٫۰۰۱
۰٫۰۹۳	۹۹٫۷۶	۰٫۰۲۲	۱۰۰٫۰۰	۰٫۰۳۳	۹۹٫۹۷	۰٫۰۰۰۱

در بخش چهارم صحت‌سنجی الگوریتم بهینه‌سازی
با در نظر گرفتن دو نوع الگوریتم آدام و گرادیان شیب
نزولی^۱ (SGD) بررسی شده است. که نتایج آن در

جدول (۹) مشاهده می‌شود.

همانطور که در جدول (۹) مشاهده می‌شود، بیشترین درصد دقت کلی و کمترین میزان خطا برای مدل شبکه عصبی پیچشی پیشنهادی، در حالتی است که از الگوریتم بهینه‌سازی آدام استفاده گردیده است که این مقدار، برابر با مقدار فرض شده در جدول (۴) می‌باشد.

¹ *Stochastic Gradient Descent*

جدول ۹: مقایسه معیار دقت و خطا، از نظر الگوریتم بهینه‌سازی

الگوریتم بهینه‌سازی		دانشگاه پاویا		سالیانس		ایندین پائینس	
% دقت کلی		خطا		% دقت کلی		% دقت کلی	
خطا		خطا		خطا		خطا	
آدام		۰/۰۱۲		۱۰۰/۰۰		۹۹/۸۴	
۰/۰۴۸		۰/۰۱۲		۰/۰۰۷		۰/۰۴۸	
گرادیان شیب نزولی		۰/۱۳۲		۵۹/۴۰		۸۵/۵۱	
۸۸/۵۵		۰/۱۳۲		۱/۲۲		۰/۵۹۲	

۴-۲-۵- صحت‌سنجی فراپارامتر تعداد دوره‌ها

در بخش پنجم صحت‌سنجی فراپارامتر تعداد دوره‌ها، با در نظر گرفتن تعداد دوره‌های مختلف، بررسی گردیده است. که نتایج آن در جدول (۱۰) مشاهده می‌شود.

همانطور که در جدول (۱۰) مشاهده می‌شود بیشترین درصد دقت کلی و کمترین میزان خطا برای مدل شبکه

عصبی پیچشی پیشنهادی، درحالتی است که تعداد دوره‌ها برابر ۲۰ در نظر گرفته شده است، با توجه به اینکه شبکه پیشنهادی پس از ۲۰ دوره همگرا می‌شود اولویت برای انتخاب تعداد دوره براساس زمان آموزش شبکه همان ۲۰ دوره می‌باشد، که این تعداد نیز برابر با تعداد فرض شده در جدول (۴) می‌باشد.

جدول ۱۰: مقایسه معیار دقت و خطا از نظر فراپارامتر تعداد دوره‌ها

تعداد دوره‌ها		دانشگاه پاویا		سالیانس		ایندین پائینس	
% دقت کلی		خطا		% دقت کلی		% دقت کلی	
خطا		خطا		خطا		خطا	
۱۵		۰/۰۱۵		۹۹/۷۱		۹۹/۰۶	
۲۰		۰/۰۱۲		۱۰۰/۰۰		۹۹/۹۶	
۲۵		۰/۰۱۳		۱۰۰/۰۰		۹۹/۸۶	
۵۰		۰/۰۱۴		۱۰۰/۰۰		۹۹/۹۲	
۷۵		۰/۰۱۶		۹۹/۹۹		۹۹/۸۴	
۱۰۰		۰/۰۱۳		۱۰۰/۰۰		۹۹/۸۸	
۰/۰۲۹		۰/۰۱۳		۰/۰۰۹		۰/۰۲۹	

۳-۴- پیاده‌سازی مدل شبکه عصبی پیچشی

پیشنهادی

در این بخش براساس فرآپارامترهای به دست آمده در بخش قبلی در جدول (۱۱) و همچنین مقادیر

در نظر گرفته شده در جدول (۱۲)، فرآیند ساخت مدل، آموزش، ارزیابی و تحلیل خروجی‌های مدل به ازای مجموعه داده‌های مختلف بررسی خواهد شد.

جدول ۱۱: فرآپارامترهای انتخاب شده شبکه عصبی پیچشی پیشنهادی

آدام	الگوریتم بهینه‌سازی
۰/۰۰۱	نرخ یادگیری
رلو	تابع فعال‌سازی
۲۰	تعداد دوره‌ها
۱۲۸	اندازه دسته‌ها
۱۹×۱۹	اندازه پنجره

جدول ۱۲: پارامترهای ایجاد مدل پیشنهادی

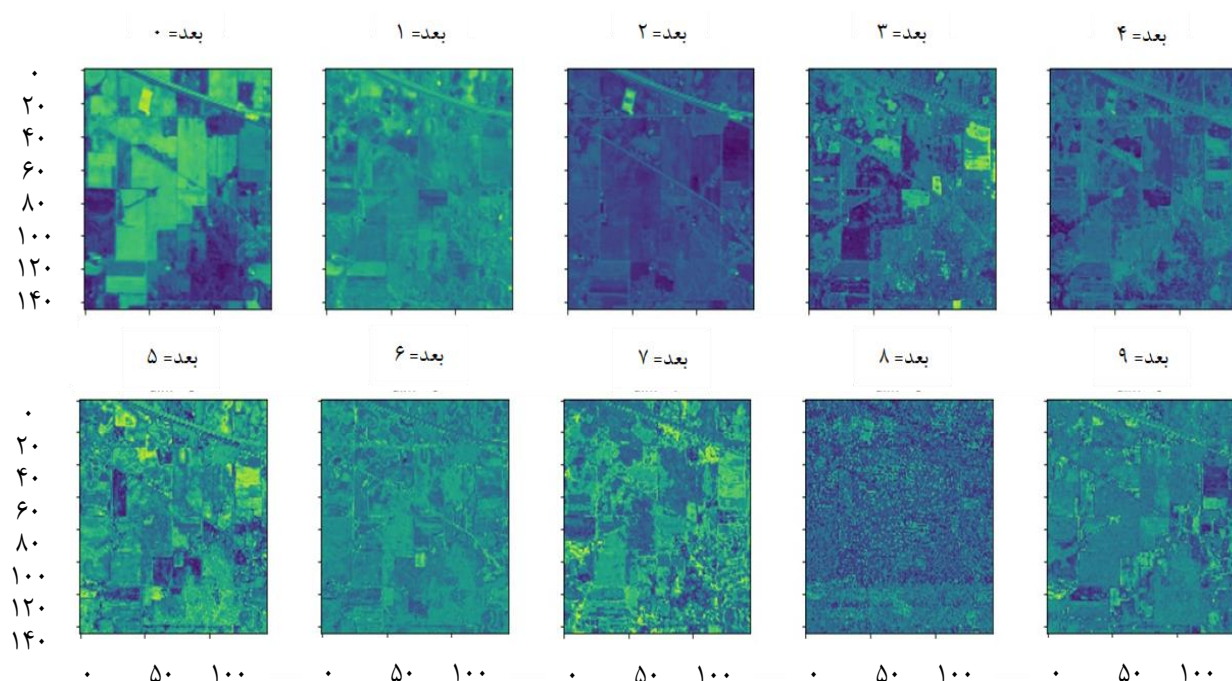
۳×۳	اندازه هسته
۲×۲	اندازه ادغام بیشینه
۰/۳۵	نرخ کاهش
تابع بیشینه هموار	تابع فعال‌سازی لایه خروجی
اختلاف آنتروپی متقاطع	تابع هزینه

۳-۴-۱- پیاده‌سازی مدل شبکه عصبی پیچشی

پیشنهادی با مجموعه داده‌های ایندین پاینس

تعداد داده‌های آموزش و آزمایش برای این مجموعه داده به ترتیب ۷۶۷۹ و ۲۵۶۹ بوده است. پس از اعمال PCA بر روی این مجموعه داده‌ها، تعداد باندهای طیفی از ۲۰۰ به ۶۹ کاهش داده شد. این به آن معناست که ۶۹ باند برتر دارای ۹۹/۹٪ واریانس در داده‌های طیفی می‌باشد. آزمایش‌ها بر روی اندازه‌های مکانی مختلف مانند ۹×۹، ۱۱×۱۱ و ۱۳×۱۳ و ... انجام شد و در نهایت مشاهده شد که اندازه پنجره ۱۹×۱۹ بهترین نتایج

طبقه‌بندی را برای این مجموعه داده‌ها ارائه می‌دهد. در شکل (۸) تصویر ۱۰ باند استخراج شده از مجموع ۶۹ باند کاهش یافته این مجموعه داده‌ها را نشان می‌دهد. مدت زمان آموزش شبکه برای این مجموعه داده‌ها هفت دقیقه و ۲۴ ثانیه به ثبت رسیده است که این زمان برای آموزش شبکه در ۵۰ تکرار دوره بود، درحالی که با حفظ همین مقادیر دقت و خطا در تعداد ۲۰ دوره، این زمان به سه دقیقه و ۱۱ ثانیه کاهش یافت. بنابراین بعد مورد استفاده برای مجموعه داده‌های ایندین پاینس، ۶۹×۱۹×۱۹ است.



شکل ۸: ۱۰ باند از ۶۹ باند کاهش یافته مجموعه داده‌های ایندین پاینس پس از اعمال PCA

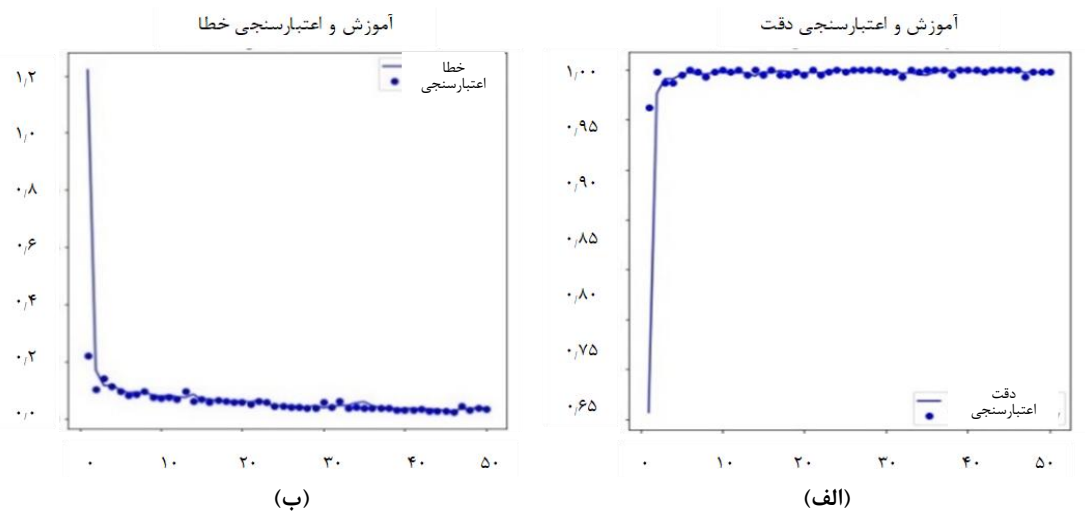
عناصر غیر قطری مجموعه خطاها می‌باشند. در اصل ماتریس درهم ریختگی، پراکندگی توزیع کلاس‌ها را از لحاظ درستی یا نادرستی نمایش می‌دهد. معیارهای $f1$ - $score$ ، صحت^۲، کارایی^۳ و میزان دقت برای هر کلاس، براساس نام و شماره هر کلاس، از مجموعه داده‌های ایندین پاینس در جدول (۱۴) قرار داده شده است. همچنین ستون آخر سمت چپ جدول نشان‌گر تعداد داده‌های موجود در هر کلاس می‌باشد.

شکل (۹) نمودارهای مربوط به دقت و خطا را برحسب داده‌های آموزش و اعتبارسنجی، به ازای ۵۰ دوره نمایش می‌دهد. شکل (۱۰) نقشه پیش‌بینی طبقه‌بندی شده مجموعه داده ایندین پاینس را در مقایسه با تصویر حقیقت زمین آن نمایش می‌دهد. در جدول (۱۳) ماتریس درهم ریختگی^۱ مجموعه داده‌های ایندین پاینس مشاهده می‌شود. تعداد پیکسل‌هایی که درست طبقه‌بندی شده‌اند بر روی قطر اصلی قرار می‌گیرند.

^۲ Precision

^۳ Recall

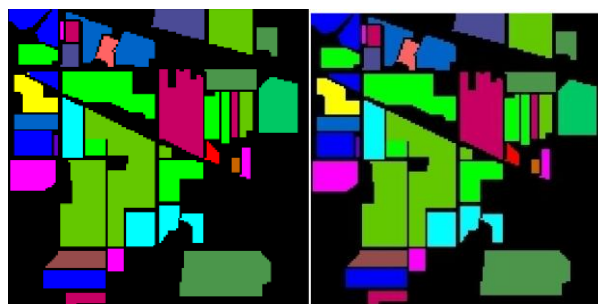
^۱ Confusion Matrix



شکل ۹: نمودارهای میزان دقت و خطا برای مجموعه داده ایندین پاینس، (الف) دقت، (ب) خطا

جدول ۱۳: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های ایندین پاینس

داده‌های آموزشی (برچسب شناخته شده)	۱	۱۲	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	۲	۰	۳۵۷	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	۳	۰	۰	۲۰۸	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	۴	۰	۰	۰	۶۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	۵	۰	۰	۰	۰	۱۲۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	۶	۰	۰	۰	۰	۰	۱۸۳	۰	۰	۰	۰	۰	۰	۰	۰	۰
	۷	۰	۰	۰	۰	۰	۰	۷	۰	۰	۰	۰	۰	۰	۰	۰
	۸	۰	۰	۰	۰	۰	۰	۰	۱۱۹	۰	۰	۰	۰	۱	۰	۰
	۹	۰	۰	۰	۰	۰	۰	۰	۰	۵	۰	۰	۰	۰	۰	۰
	۱۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۲۴۳	۰	۰	۰	۰	۰
	۱۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۶۱۳	۰	۰	۰	۰
	۱۲	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱۴۸	۰	۰	۱
	۱۳	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۵۲	۰	۰
	۱۴	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۳۱۷	۰
	۱۵	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۹۷	۰
	۱۶	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۲۴
کلاس‌ها (برچسب پیش بینی شده)																
		۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵



(ب)

(الف)

شکل ۱۰: نقشه‌های طبقه‌بندی مجموعه داده‌های ایندین پاینس. (الف) تصویر حقیقت زمین [۲۴]، (ب) تصویر روش پیشنهادی

جدول ۱۴: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های ایندین پاینس

شماره کلاس	نام کلاس	صحت	کارایی	<i>f1-score</i>	دقت	تعداد داده
۱	یونجه ۱	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱۲
۲	ذرت-نشاسته	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۳۵۷
۳	ذرت-نعنا	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۲۰۸
۴	ذرت	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۶۰
۵	علف مرتع	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱۲۱
۶	علف درخت	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱۸۳
۷	علف مرتع چمن	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۷
۸	یونجه ۲	۱٫۰۰	۰٫۹۹	۱٫۰۰	۰٫۹۹	۱۲۰
۹	جو دوسر	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۵
۱۰	سویا نعنای ۱	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۲۴۳
۱۱	سویا نعنای ۲	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۶۱۴
۱۲	سویا تمیز	۱٫۰۰	۰٫۹۹	۱٫۰۰	۰٫۹۹	۱۴۹
۱۳	گندم	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۵۲
۱۴	بافت جنگلی	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۳۱۷
۱۵	ساختمان-علف-وسایل نقلیه	۱٫۰۰	۱٫۰۰	۱٫۰۰	۱٫۰۰	۹۷
۱۶	سنگ-ساختمان-برج	۰٫۹۶	۱٫۰۰	۰٫۹۸	۱٫۰۰	۲۴

دقت برای هر کلاس براساس نام و شماره هر کلاس، از مجموعه داده‌های دانشگاه پائوینا در جدول (۱۵) قرار داده شده است. مدت زمان آموزش شبکه برای این مجموعه داده‌های ۱۴ دقیقه و ۵۰ ثانیه، به ثبت رسیده است. که این زمان برای آموزش شبکه در ۵۰ تکرار دوره بود، درحالی که با حفظ همین مقادیر دقت و خطا در تعداد ۲۰ دوره، این زمان به پنج دقیقه و سه ثانیه کاهش یافت.

۴-۳-۲- پیاده‌سازی مدل شبکه عصبی پیچشی

پیشنهادی با مجموعه داده‌های دانشگاه پائوینا

تعداد داده‌های آموزش و آزمایش برای این مجموعه داده‌های به‌ترتیب ۳۲۰۷۷ و ۱۰۶۹۷ می‌باشد. پس از اعمال *PCA* بر روی این مجموعه داده‌ها، تعداد باندهای طیفی از ۱۰۳ به ۱۶ کاهش یافت. بعد مورد استفاده برای مجموعه داده‌های دانشگاه پائوینا، $16 \times 19 \times 19$ است. معیارهای *f1-score*، صحت، کارایی و میزان

جدول ۱۵: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های دانشگاه پابوا

شماره کلاس	نام کلاس	صحت	کارایی	<i>f1-score</i>	دقت	تعداد داده
۱	آسفالت	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۱۶۵۸
۲	چمن زار	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۴۶۶۳
۳	شن	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۵۲۵
۴	انواع درختان	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۷۶۶
۵	صفحات فلزی رنگی	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۳۳۶
۶	خاک لخت	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۱۲۵۸
۷	قیر	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۳۳۳
۸	آجر	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۹۲۱
۹	سایه	۱,۰۰	۱,۰۰	۱,۰۰	۱,۰۰	۲۳۷

در این پژوهش بر اساس دقت کلی^۲ و ضریب کاپا^۳ (برحسب درصد) بررسی شده است. جهت یکسان‌سازی شرایط آموزش برای مقایسه نتایج، مطابق با شرایط مقاله [۱۵]، ۷۰٪ داده‌ها به آموزش و ۳۰٪ به آزمایش اختصاص داده شده است، همانطور که مشاهده می‌شود روش پیشنهادی نسبت به سایر روش‌ها از نظر معیارهای دقت کلی و ضریب کاپا مناسب‌تر می‌باشد. در شکل (۱۱) نتایج نقشه‌های طبقه‌بندی مجموعه داده ایندین پاینس [۱۵]، نمایش داده شده است. در جدول (۱۸) مقایسه‌ای از نظر مدت زمان آموزش و زمان آزمایش داده‌های شبکه با روش *HybridSN* و روش پیشنهادی انجام شده است. جهت یکسان‌سازی شرایط آموزش برای مقایسه نتایج، کدهای مربوط به روش *HybridSN* از [۲۳] استخراج و در محیط یکسان با روش پیشنهادی اجرا شده است همانطور که مشاهده می‌شود روش پیشنهادی نسبت به روش *HybridSN* از نظر زمان سپری شده به صرفه‌تر می‌باشد.

۴-۳-۳- پیاده‌سازی مدل شبکه عصبی پیچشی

پیشنهادی با مجموعه داده‌های سالیانس

تعداد داده‌های آموزش و آزمایش برای این مجموعه داده‌های به‌ترتیب ۴۰۵۹۱ و ۱۳۵۳۶ می‌باشد. پس از اعمال *PCA* بر روی این مجموعه داده‌ها، تعداد باندهای طیفی از ۲۰۴ به شش کاهش یافت. بنابراین بعد مورد استفاده برای مجموعه داده‌های سالیانس، $19 \times 19 \times 6$ می‌باشد. معیارهای *f1-score*، صحت، کارایی و میزان دقت برای هر کلاس در جدول (۱۶) قرار داده شده است. مدت زمان آموزش شبکه برای این مجموعه داده‌ها ۱۶ دقیقه و ۲۲ ثانیه، به ثبت رسیده است که این زمان برای آموزش شبکه در ۵۰ تکرار دوره بود، درحالی که با حفظ همین مقادیر دقت و خطا در تعداد ۲۰ دوره، این زمان به شش دقیقه و ۱۱ ثانیه کاهش یافت.

۴-۴- جمع‌بندی و مقایسه روش پیشنهادی با

دیگر روش‌ها

در جدول (۱۷) نتایج طبقه‌بندی مجموعه داده‌های مختلف با روش‌های *SVM* [۲۴]، *2D-CNN* [۱۴]، *3D-CNN* [۲۵] و *HybridSN* [۱۵] و روش پیشنهادی

² Overall Accuracy

³ Kappa Coefficient

¹ 2 Dimension Convolutional Neural Network

جدول ۱۶: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های سالیانس

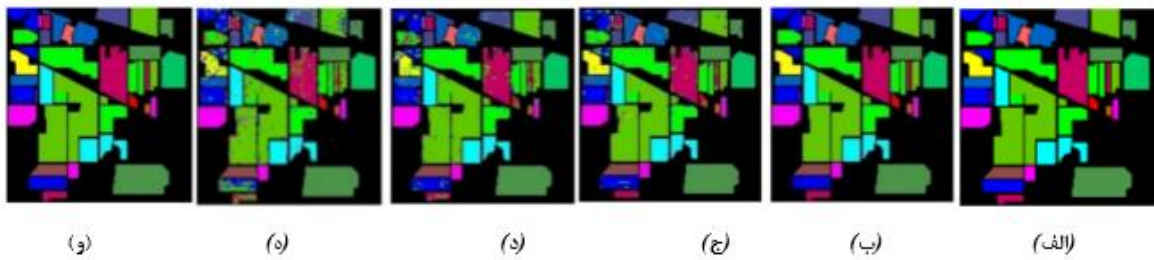
شماره کلاس	نام کلاس	صحت	کارایی	<i>f1-score</i>	دقت	تعداد داده
۱	کلم بروکلی سبز ۱	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۵۰۲
۲	کلم بروکلی سبز ۲	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۹۳۲
۳	آیش	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۴۹۴
۴	آیش-گاوا آهن	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۳۴۹
۵	آیش صاف	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۶۷۰
۶	کلش	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۹۹۰
۷	کرفس	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۸۹۵
۸	انگور	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۲۸۱۸
۹	خاک	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱۵۵۱
۱۰	ذرت- علف های هرز سبز شده	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۸۲۰
۱۱	کاهو ۴	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۲۶۷
۱۲	کاهو ۵	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۴۸۲
۱۳	کاهو ۶	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۲۲۹
۱۴	کاهو ۷	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۲۶۸
۱۵	تاکستان ۱	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۱۸۱۷
۱۶	تاکستان ۲	۱.۰۰	۱.۰۰	۱.۰۰	۱.۰۰	۴۵۲

جدول ۱۷: مقایسه دیگر روش‌ها، از نظر پارامتر میزان دقت کلی و ضریب کاپا با روش پیشنهادی

مجموعه داده	دانشگاه پاویا		سالیانس		ایندین پائینس	
	دقت کلی	کاپا	دقت کلی	کاپا	دقت کلی	کاپا
<i>SVM</i>	94.34 ± 0.2	92.59 ± 0.7	92.95 ± 0.3	92.11 ± 0.2	85.30 ± 2.8	83.10 ± 3.2
<i>2D-CNN</i>	97.86 ± 0.2	97.16 ± 0.5	97.38 ± 0.0	97.08 ± 0.1	89.48 ± 0.2	87.96 ± 0.5
<i>3D-CNN</i>	96.53 ± 0.1	95.51 ± 0.2	93.96 ± 0.2	93.32 ± 0.5	91.10 ± 0.4	89.98 ± 0.5
<i>HybridSN</i>	99.98 ± 0.0	99.98 ± 0.0	100.00 ± 0.0	100.00 ± 0.0	99.75 ± 0.1	99.71 ± 0.1
روش پیشنهادی	100.00 ± 0.0	100.00 ± 0.0	100.00 ± 0.0	100.00 ± 0.0	99.90 ± 0.06	99.89 ± 0.08

جدول ۱۸: مقایسه دیگر روش‌ها از نظر مدت زمان آموزش شبکه با روش پیشنهادی

مجموعه داده	دانشگاه پاویا		سالیانس		ایندین پائینس	
	(بر حسب دقیقه)		(بر حسب دقیقه)		(بر حسب دقیقه)	
<i>2D-CNN</i>	۱.۸		۲.۲		۱.۹	
<i>3D-CNN</i>	۵۸.۰		۷۴.۰		۱۵.۲	
<i>HybridSN</i>	۲۰.۳		۲۵.۵		۱۴.۱	
روش پیشنهادی	۱۳.۸		۱۵.۷		۶.۳	



شکل ۱۱: نتایج نقشه‌های طبقه‌بندی مجموعه داده‌های ایندین پاینس. (الف) حقیقت زمین، (ب) SVM، (ج) 2D-CNN، (د) 3D-CNN، (ه) HybridSN [۲۳]، (و) روش پیشنهادی

۵- نتیجه گیری و پیشنهادات

در این مقاله با توجه به آزمایش‌های صورت گرفته یک معماری جدید دوبعدی از شبکه‌های عصبی پیچشی جهت طبقه‌بندی تصاویر ابرطیفی پیشنهاد شده است که شامل دو لایه پیچشی دوبعدی و سه لایه کاملاً متصل است. از روش پیشنهادی جهت طبقه‌بندی سه مجموعه داده ایندین پاینس، سالیانس و دانشگاه پالویا استفاده شده است. نتایج حاصله نشان از برتری معماری CNN پیشنهادی نسبت به سایر روش‌های بکارگرفته شده مانند SVM و شبکه‌های یادگیری عمیق دوبعدی و سه بعدی موجود بوده است. در مدل پیشنهادی انتخاب فرامترها با انجام آزمون و خطا و سنجیدن تمام حالت‌های ممکن به دست آمده است. جهت آموزش مدل پیشنهادی، ابتدا اطلاعات طیفی و مکانی، در یک مرحله با یکدیگر ترکیب می‌شوند و همزمان ویژگی‌های طیفی- مکانی سطح بالا استخراج می‌گردند. با اعمال PCA، تصویر ابرطیفی گرفته شده به پنجره‌های کوچک‌تر تجزیه می‌شوند که هر کدام حاوی اطلاعات طیفی و مکانی برای یک پیکسل خاص می‌باشند، به گونه‌ای که با کاهش ابعاد ورودی، عملکرد

روش پیشنهادی بهتر از سایر روش‌ها بوده است. جهت ارزیابی عملکرد روش پیشنهادی از جهت دقت طبقه‌بندی و زمان اجرا، نتایج با عملکرد چندین روش طبقه‌بندی دیگر مقایسه شد که حاکی از برتری روش پیشنهادی نسبت به سایر روش‌ها بوده است. اصلی‌ترین رقیب روش پیشنهادی، روش HybridSN بوده است که علاوه بر اینکه دقت طبقه‌بندی روش پیشنهادی بیشتر از آن می‌باشد، در بهترین حالت زمان اجرای الگوریتم (۵۰٪ - ۳۰٪) بهتر بوده است. در تحقیقات آینده می‌توان تنظیم پارامترها و فرامترهای شبکه را بر اساس روش‌های مطرح در حوزه بهینه سازی تکاملی مدنظر داشت تا مقادیر بهینه پارامترها به صورت هدفمند انتخاب شوند و علاوه بر این روش‌های دیگر استخراج ویژگی را مورد توجه قرار داد. همچنین می‌توان نسبت بکارگیری داده‌های آموزش و آزمایش را تغییر داد و از روش‌های تولید داده‌های شبیه‌سازی شده جهت تولید داده‌های آموزشی افزونه استفاده کرد که می‌تواند نیاز به داده‌های آموزش را کاهش دهد و در نهایت مدلی واحد جهت طبقه‌بندی تصاویر ابرطیفی ارائه کرد.

مراجع

- [1] A. Plaza, J. Benediktsson, et al, "Recent advances in techniques for hyperspectral image processing," *Remote sensing of environment*, vol. 113, pp. 110-122, 2009.
- [2] F. Melgani, and L. Bruzzone, "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines," *IEEE transactions on geoscience and remote sensing*, Vol. 42, pp. 1778-1790, 2004.
- [3] 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, pp. 2094–2107, June 2014.
- [4] T. Li, J. Zhang, X. Zhao, and Y. Zhang, *Classification of hyperspectral image based on deep belief networks*, in *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 1–5, Oct 2014.
- [5] W. Zhao and S. Du, *Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach*, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, pp. 4544–4554, Aug 2016.
- [6] X. Zhou, S. Li, F. Tang, K. Qin, S. Hu, and S. Liu, *Deep learning with grouped features for spatial spectral classification of hyperspectral images*, *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 97–101, Jan 2017.
- [7] W. Li, G. Wu, F. Zhang, and Q. Du, *Hyperspectral image classification using deep pixel-pair features*, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 844–853, Feb 2017.
- [8] H. Teffahi, H. Yao, S. Chaib, and N. Belabid, *A novel spectral-spatial classification technique for multispectral images using extended multi-attribute profiles and sparse autoencoder*, *Remote Sensing Letters*, vol. 10, no. 1, pp. 30–38, 2019.
- [9] Z. Lin, Y. Chen, X. Zhao, and G. Wang, *Spectral-spatial classification of hyperspectral image using autoencoders*, in *2013 9th International Conference on Information, Communications Signal Processing*, pp. 1–5, Dec
- for hyperspectral image classification*, " *Journal of Sensors*, vol. 2015, no. 3, p. 12pages, 2015.
- [12] P. H. Swain, S. B. Vardeman, and J. C. Tilton, *Contextual classification of multispectral image data*, *Pattern Recognition*, vol. 13, no. 6, pp. 429 – 441, 1981.
- [13] 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, pp. 6232–6251, Oct 2016.
- [14] K. Makantasis, K. Karantzas, A. Doulamis, and N. Doulamis, *Deep supervised learning for hyperspectral data classification through convolutional neural networks*, in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4959–4962, July 2015.
- [15] S. K. Roy, G. Krishna, S. R. Dubey and B. B. Chaudhuri, *HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification*, in *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, Feb. 2020, doi: 10.1109/LGRS.2019.2918719.
- [16] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [17] C. Szegedy, Wei L., Yangqing J., P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. *Going deeper with convolutions*. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, 2015
- [18] K. Sohn, H. Lee, X. Yan, *Learning structured output representation using deep conditional generative models*, in: *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, 2015, pp. 3483–
- [19] K. O'Shea and R. Nash. *An introduction to*

- convolutional neural networks."CoRR, abs/1511.08458, 2015.
- [20] Yanming Guo, Yu Liu, Ard Oerlemans, Songyang Lao, Song Wu and Michael S. Lew, *Deep learning for visual understanding: A review, Neurocomputing*," <http://dx.doi.org/10.1016/j.neucom.2015.09.116>
- [21] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi and J. A. Benediktsson, "Deep Learning for Hyperspectral Image Classification: An Overview," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690-6709, Sept. 2019, doi: 10.1109/TGRS.2019.2907932.
- [22] A. N. Abbasi and M. He, "Convolutional Neural Network with PCA and Batch Normalization for Hyperspectral Image Classification," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 959-962, doi: 10.1109/IGARSS.2019.8899329.
- [23] A Deep Learning approach to Hyperspectral Image Classification using an improved Hybrid 3D-2D Convolutional Neural Network September 2020 Pages 85–92.
<https://doi.org/10.1145/3411408.3411462>
- [24] S. Ghaderizadeh, D. Abbasi-Moghadam, A. Sharifi, N. Zhao and A. Tariq, "Hyperspectral Image Classification Using a Hybrid 3D-2D Convolutional Neural Networks," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7570-7588, 2021, doi: 10.1109/JSTARS.2021.3099118.
- [25] A. Ben Hamida et al. 3-D Deep Learning Approach for Remote Sensing Image Classification." In: *IEEE Transactions on Geoscience and Remote Sensing* 56.8 (2018), pp. 4420–44 2013.



Improving the Classification Accuracy of Hyperspectral Images Using Convolutional Neural Networks

Mahsa Tekyeh-Nejad¹, Ata Ebrahimzadeh^{2*}, Maliheh Ahmadi³

1- Master's degree in Telecommunication system - Faculty of Electrical and Computer Engineering - Babol Noshirvani University of Technology

2- Professor of Department of Telecommunication System - Faculty of Electrical and Computer Engineering - Babol Noshirvani University of Technology

3- PhD in Telecommunication Systems - Faculty of Electrical Engineering - Shiraz University of Technology

Abstract

Hyperspectral image classification is a crucial aspect of remote sensing image analysis. Deep learning methods have been successfully used to classify the remote sensing data. In recent years, the convolutional neural networks (CNNs) have been significantly used in hyperspectral image classification, which has tried to overcome the computational and processing challenges of hyperspectral data. By increasing the number of parameters and layers of convolutional neural networks, their efficiency in solving complex problems decreases. For this reason, in this article, a new architecture of convolutional neural networks has been introduced which has a good performance and reduces the computing time.

In this paper, we introduce a novel CNN that utilizes spectral-spatial information as input and employs the principal component analysis (PCA) to reduce the spectral bands. To prevent overfitting, we combine batch normalization and dropout techniques. Our two-dimensional CNN includes convolutional layers, pooling layers, and fully connected layers. We also incorporate PCA and patch selection to enhance the accuracy of our model. To evaluate the effectiveness of our proposed model, we conducted experiments on three datasets: Indian Pines, Pavia University, and Salinas. Our simulation results demonstrate that our model achieves a classification accuracy of 100%, with less training time and complexity than the existing models.

Key words: Hyperspectral image classification; Convolutional neural network; Principal component analysis; Choose the right patches.