نشریه علمی مهندسی فناوری اطلاعات مکانی

سال یازدهم، شماره نخست، بهار ۱۴۰۲

*Vol.11, No.1, Spring 2023*

۸۲- ۵۹

مقاله پژوهشی

DOR: 20.1001.1.20089635.1402.11.1.4.5

# 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 داشته اند 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. به‌همین 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 به‌صورت 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 کاملا 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 ایندین‌پاینس، دانشگاه پاویا و سالیناس صورت گرفته است. نتایج شبیه‌سازی نشان می‌دهد که روش پیشنهادی علاوه بر داشتن درصد دقت بالای موفقیت (۱۰۰درصد) در طبقه‌بندی و مدت زمان کمتر، دارای پیچیدگی کمتری نسبت به مدل‌های موجود می‌باشد.

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

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

تلفن: ۰۱۱۳۲۳۳۹۲۱۴

*Email: ezadehata@gmail.com*

[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]

[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]

[ DOI: 10.61186/jgit.11.1.59 ]

## ۱- مقدمه

طیف‌سنجی تصاویر[1] 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 هوابرد یا فضابرد ماهواره‌ای گرفته می‌شود [۱]. این تکنیک نسبتا جدید در اوایل دهه ۱۹۸۰ با تولید یک سنجنده طیف‌سنج هوابرد که بعدها "آویریس(*Aviris*)[2]" 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 به‌دست 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*)[3] 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[4] برای این نوع از تصاویر طراحی و معرفی شده‌اند. روش‌های طبقه‌بندی به سه روش کلی از جمله طبقه‌بندی تحت نظارت[5]، بدون نظارت[6] و نیمه نظارتی[7] 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 پیش بینی کند. برخی روش‌های طبقه‌بندی نظارت‌شده شامل روش ماشین بردارپشتیبان(*SVM*)[8]، شبکه‌های عصبی مصنوعی(*ANN*)[9] و طبقه‌بند حداقل فاصله (*MDC*)[10] 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 قرارگرفته 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 پشته ای[11] (*SAE*) به‌عنوان 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[12](*LR*) در شبکه قرار می‌گیرد [۳].

تانگ لی و همکاران (۲۰۱۴)، مدلی را پیشنهاد دادند که در آن شبکه باور عمیق[13] (*DBN*) با دو لایه پنهان ویژگی‌های طیفی-مکانی پیکسل‌های مکعب ابرطیفی را استخراج می‌کند به طوری که ابتدا ابعاد طیفی مکعب

[1] *Imaging Spectroscopy*
[2] *Airborne Visible Infrared Imaging Spectrometer*
[3] *Hyper Spectral Images*
[4] *Classification*
[5] *Supervision Classification*
[6] *Unsupervised Classification*
[7] *Semi-supervised Classification*
[8] *Support Vector Machine*
[9] *Artificial Neural Networks*
[10] *Minimum Distance Classifier*
[11] *Stacked Autoencoder*
[12] *Logistic Regression*
[13] *Deep Belief Network*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

ابرطیفی، با استفاده از آنالیز مولفه اصلی (*PCA*)[14] کاهش می‌یابد و تنها سه جزء اصلی اول حفظ می‌شوند. در مرحله بعد، یک پنجره سه‌بعدی به اندازه ۳×۷×۷ در اطراف هر نمونه آموزشی (پیکسل) تشکیل می‌شود که در ادامه تبدیل به بردارشده و به عنوان ورودی به شبکه داده می‌شود. طرح طبقه‌بندی نشان داده شده در مقاله‌ی ذکر شده دارای چهار لایه ی ورودی، دو ماشین بولتزمن محدود[15] (*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[16] (*CNN*) معرفی شده‌اند و بسیار مورد توجه قرار گرفته‌اند. *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 [۴،۵،۶،۷،۸،۹ و 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 این مطالعه شامل یک لایه پیچشی، یک لایه ادغام بیشینه[17]، یک لایه کاملا متصل (*FC*)[18] 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 (*BLDE*)[19] 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 (*۳D-CNN*)[20] طراحی کرده اند. 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

[14] *Principle Component Analysis*
[15] *Restricted Boltzmann Machine*
[16] *Convolutional Neural Network*
[17] *Max Pooling*
[18] *Fully Connected*
[19] *Balanced Local Discriminant Embedding*
[20] *3 Dimension Convolutional Neural Network*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

و چن و همکارانش در بررسی قبلی،مدلی پیشنهاد کردند که در آن، برای استخراج اطلاعات طیفی و مکانی به‌طور همزمان از مکعب ابرطیفی داده استفاده شده است. به عبارت دیگر، یک پنجره سه‌بعدی در اطراف هر پیکسل هدف تشکیل می‌شود که حاوی اطلاعات طیفی و مکانی است. با این حال، بر خلاف مدل یوشی چن و همکارانش، که این پنجره سه‌بعدی را به عنوان ورودی مستقیم به *CNN* می‌دهد، در این مدل (*R-PCA*)[21] روی هر پنجره سه‌بعدی برای کاهش ابعاد داده‌های ورودی اعمال می‌شود [۱۴].

سوالپا کومار روی و همکاران (۲۰۲۰)، یک مدل پیچشی که ترکیبی از شبکه عصبی پیچشی دوبعدی و سه‌بعدی می‌باشد را برای استفاده از مزایای قابلیت یادگیری خودکار ویژگی‌های دو بعدی و سه‌بعدی ارائه دادند، که این روش سعی در غلبه برکاستی‌های مدل های قبلی دارد. لایه‌های پیچشی دوبعدی و سه‌بعدی برای مدل پیشنهادی به گونه‌ای تدوین شده‌اند که از ویژگی‌های 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 کاملا متصل می‌باشد که مدل "هیبرید اس ان" (*HybridSN*)[22] نامگذاری شده است. روش پیشنهادی با روش‌های دیگر از نظر دقت طبقه‌بندی مقایسه شد و مشاهده گردید که این روش، از روش‌های دیگر نتایج بهتری ارائه می‌دهد [۱۵].

بررسی‌ها نشان می‌دهد که اکثر تحقیقات ارائه شده از لایه‌های پیچشی سه بعدی برای طبقه‌بندی استفاده می‌کنند که این مسئله پیچیدگی کار را زیاد می‌کند. انتخاب پنجره مناسب می‌تواند دقت طبقه‌بند را افزایش دهد. همچنین کاهش تعداد باندهای طیفی می‌تواند به کاهش پیچدگی سیستم نهایی منجر شود. روش پشنهادی این مقاله به دنبال تحقق این دو هدف می‌باشد. ساختار تحقیق بدین صورت است که در قسمت دوم به روش شناسی پرداخته می‌شود که در آن شبکه‌های عصبی پیچشی و انواع لایه‌های آن و همچنین معماری شبکه پیشنهادی بررسی می‌گردد، درقسمت سوم منطقه های مورد مطالعه بررسی می‌شود، در قسمت چهارم نتایج و بحث ارائه شده است و در قسمت پایانی، نتیجه گیری و پیشنهادات ارائه خواهد شد.

## ۲- روش شناسی

دراین بخش روش‌ها و ابزارهای تحقیق معرفی می‌شود. ابزارهای تحقیق شامل شبکه‌های عصبی پیچشی و لایه‌های مختلف آن و همچنین توصیف روش پیشنهادی می‌باشد.

### ۲-۱- شبکه‌های عصبی پیچشی

در میان انواع مختلف شبکه‌های عصبی عمیق، شبکه‌های عصبی پیچشی بیشتر مورد مطالعه قرار گرفته‌اند. *CNN* یک معماری یادگیری عمیق شناخته شده است که از مکانیسم ادراک بصری طبیعی موجودات زنده الهام گرفته است. این شبکه‌ها به سرعت برای حل مشکلات بینایی رایانه مانند طبقه‌بندی تصویر و تشخیص اشیا در سال‌های اخیر توسعه یافته‌اند. *CNN* یک شبکه عصبی است که از یک عملیات ریاضی به نام کانولوشن در لایه‌های خود استفاده می‌کند [۱۶]. عملیات کانولوشن به‌صورت زیر در رابطه (۱) تعریف می‌شود:

رابطه(۱)

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

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

رابطه(۲)

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

[21] *Random Principle Component Analysis*
[22] *Hybrid Spectral Convolutional Neural Network*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

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:

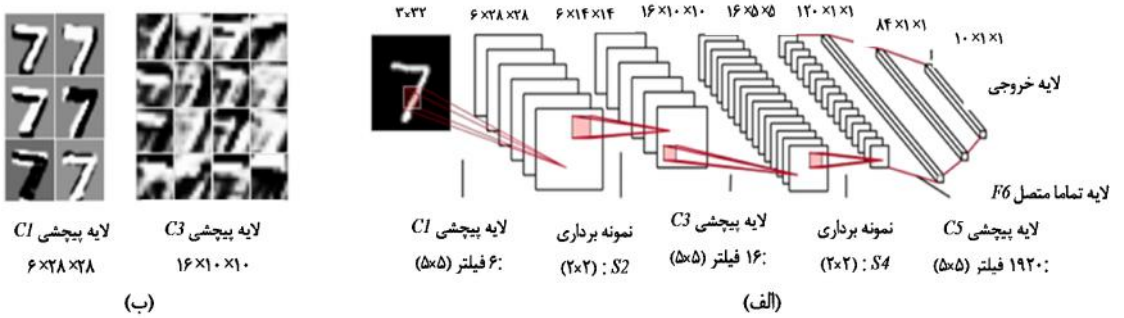
رابطه(۳)

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

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 استخراج‌شده پیچیده‌تر و در نتیجه عمل‌کرد بهتری به دست می‌آید [۱۷]. با این حال لازم به ذکر است که با افزایش لایه‌ها پیچیدگی شبکه نیز افزایش می‌یابد.

### 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.




**شکل ۱: (الف) معماری شبکه *LeNet*-۵، (ب) تجسم ویژگی‌ها در شبکه *LeNet*-۵ [۱۸]**

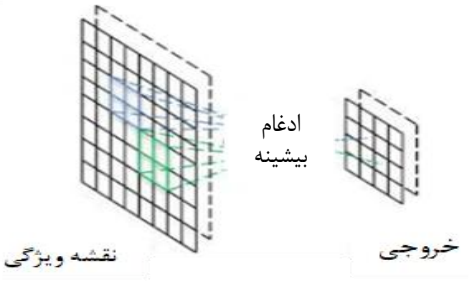
### ۲-۱-۲- لایه ادغام

این لایه برای نمونه‌برداری از نقشه‌های ویژگی در امتداد ابعاد مکانی شامل عرض و ارتفاع، برای کاهش مقدار پارامترها و محاسبات در شبکه و درنتیجه، 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 ویژگی مد نظر، ابعاد هسته آن انتخاب می‌شود. لایه ادغام دارای پارامتر قابل آموزش نمی‌باشد و صرفا یک نمونه‌برداری 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، معمولا مربع‌هایی با اندازه ۴×۴ انتخاب می‌شوند و روی

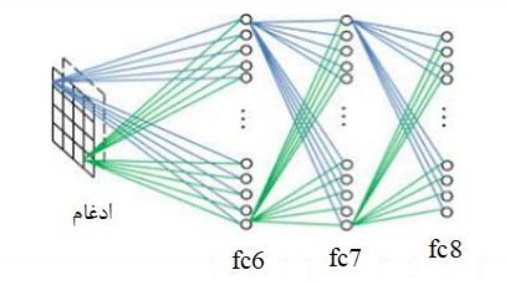
تصویر حرکت داده می‌شوند و در هر چهار پیکسل، مقدار بیشینه آن‌ها انتخاب شده و به لایه بعدی انتقال داده می‌شود. در شکل (۲) یک لایه ادغام مشاهده می‌شود که بر روی صفحه ۸×۸ اعمال شده و اندازه فیلتر ادغام آن ۲×۲ است. در نتیجه، ابعاد صفحه به ۴×۴ کاهش پیدا کرده است. با مقایسه بین این دو مدل ادغام، ثابت شده است که ادغام بیشینه موجب همگرایی سریع‌تر و انتخاب ویژگی‌های متمایزتر می‌شود.

**۲-۱-۳- لایه کاملا متصل**

لایه‌های کاملا متصل هر نورون در یک لایه را، به یک نورون در لایه دیگر متصل می‌کند. لایه کاملا 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 کاملا متصل نشان داده شده است [۱۹].




**شکل ۳: عملیات لایه کاملا متصل [۲۰]**




**شکل ۲: عملیات لایه ادغام [۲۰]**

## 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، درحالی‌که 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 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 ۹۹٫۹٪ 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

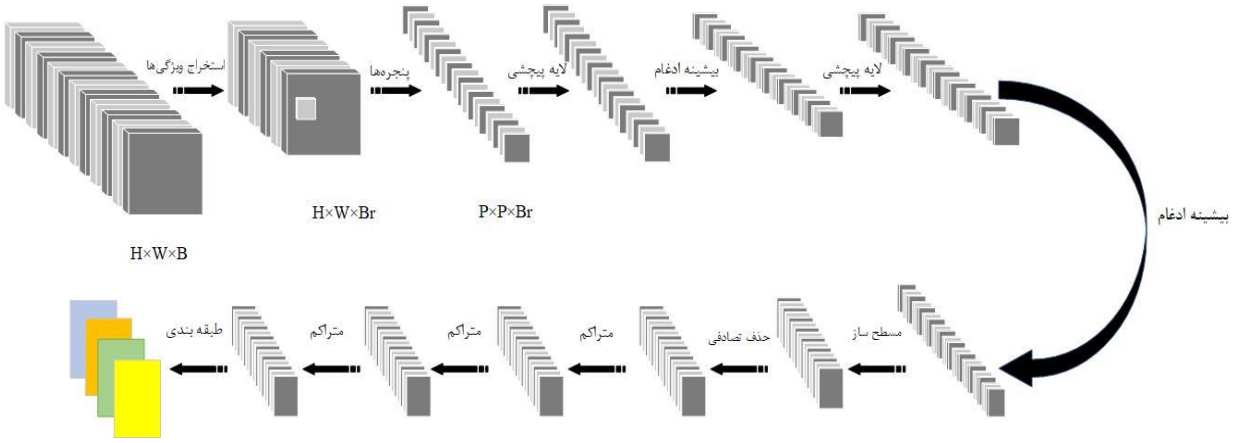


[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

کاهش یافته، پنجره $T$ را می‌توان یک تنسور سه‌بعدی با ابعاد $P\times P\times B_r$ 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 $h = 1,2,...H$ و $w = 1,2,...W$ تشکیل شده است، درحالی که $S$ در مجموع $N$ تعداد پنجره دارد.

### ۲-۲-۱- 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**

لایه‌های پیچشی با لایه‌های ادغام بیشینه دنبال می‌شوند. در اولین لایه پیچشی دوبعدی از ۳۲ فیلتر۳×۳ 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 ۲×۲ 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 فیلتر۳ ×۳ اعمال می‌شود و به دنبال آن لایه ادغام بیشینه ۲×۲ با گام ۲ مجددا اعمال می‌شود. لازم به ذکر است که در هردو لایه پیچشی دوبعدی، تابع فعال‌سازی غیرخطی رلو($ReLU$)[1] استفاده شده است.

درنهایت ویژگی‌ها توسط لایه مسطح ساز[2]، تبدیل به یک بردار ویژگی می‌شوند تا در اختیار لایه کاملا متصل قرار گیرند. در مرحله اول ۳۵ ٪ نرخ کاهش[3] برای جلوگیری از هرگونه تطبیق بیش‌برازش در طول عملیات آموزشی اعمال می‌شود و پس از آن سه لایه کاملا متصل بکار گرفته می‌شود که برای دو لایه اول تابع فعال‌سازی غیرخطی رلو و برای لایه آخر تابع فعال‌سازی بیشینه هموار[4] استفاده شده است.

## ۳- منطقه مورد مطالعه

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

### ۳-۱- ایندین‌پاینس[5]

این مجموعه داده‌ها توسط سنجنده هوابرد آویریس از

[1] *Rectified Linear Activation Function*

[2] *Flatten*
[3] *Drop Rate*
[4] *Softmax*
[5] *Salinas*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

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 ۱۴۵× ۱۴۵پیکسل و ۲۲۴ باند طیفی در محدوده ۲/۵- ۰/۴ میکرومتر می‌باشد. از آنجایی که ۲۴ باندطیفی 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 کشاورزی و یک سوم آن را جنگل و گیاه در برگرفته است.

بطورکلی 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)).

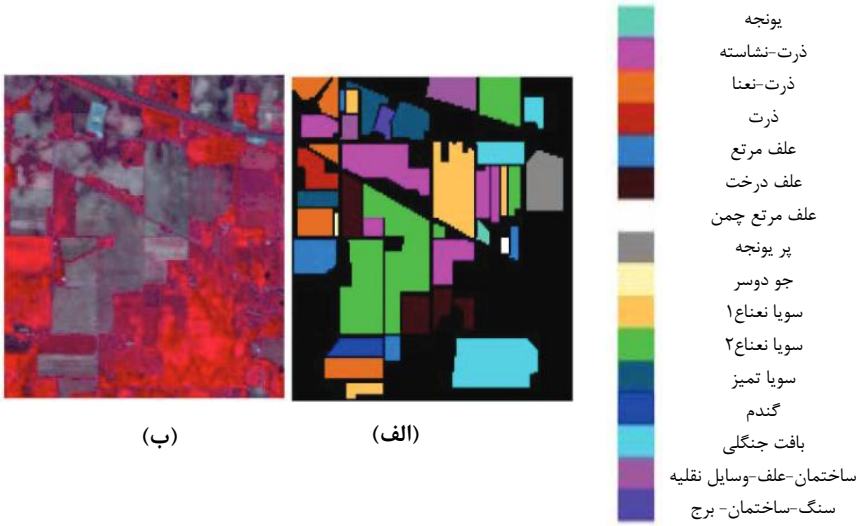
در جدول (۱) نیز انواع کلاس‌های این مجموعه داده‌ها و تعداد نمونه‌های آن نمایش داده شده است.

**جدول ۱: کلاس‌های ایندین‌پاینس**

| 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 نعناع۱ | ۹۷۲ |
| ۱۱ | سویا نعناع۲ | ۲۴۵۵ |
| ۱۲ | سویا تمیز | ۵۹۳ |
| ۱۳ | گندم | ۲۰۵ |
| ۱۴ | بافت جنگلی | ۱۲۶۵ |
| ۱۵ | ساختمان-علف-وسایل نقلیه | ۳۸۶ |
| ۱۶ | سنگ-ساختمان- برج | ۹۳ |



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

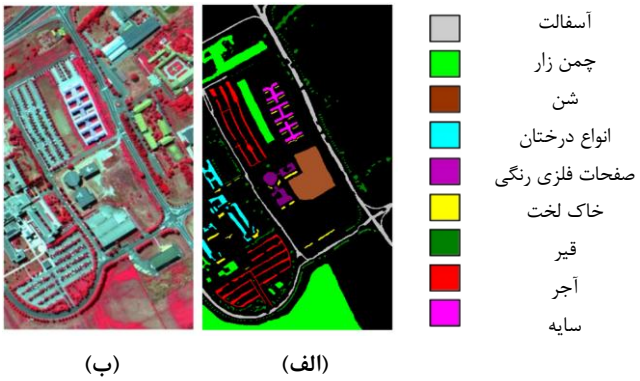



**شکل ۵: مجموعه داده ایندین‌پاینس. الف) نقشه حقیقت زمین، ب) تصویر رنگی کاذب [۲۶]**

### ۳-۲- دانشگاه پاویا (*Pavia University*)

این تصاویر توسط سنجنده هوابرد 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 ۰٫۸۶-۰٫۴۳ میکرومتر و دارای قدرت تفکیک مکانی ۱٫۳ متر هستند. از آنجایی که ۱۲ باندطیفی توسط نویز آلوده شده‌است، داده‌های تصحیح شده با اندازه ۳۴۰ ×۶۱۰ 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.




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



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

جدول۲: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های دانشگاه پاویا

| شماره کلاس | نام کلاس | تعداد نمونه‌ها |
|---|---|---|
| ۱ | آسفالت[1] | ۶۶۳۱ |
| ۲ | چمن زار[2] | ۱۸۶۴۹ |
| ۳ | شن[3] | ۲۰۹۹ |
| ۴ | انواع درختان[4] | ۳۰۶۴ |
| ۵ | صفحات فلزی رنگی[5] | ۱۳۴۵ |
| ۶ | خاک لخت[6] | ۵۰۲۹ |
| ۷ | قیر[7] | ۱۳۳۰ |
| ۸ | آجر[8] | ۳۶۸۲ |
| ۹ | سایه[9] | ۹۴۷ |

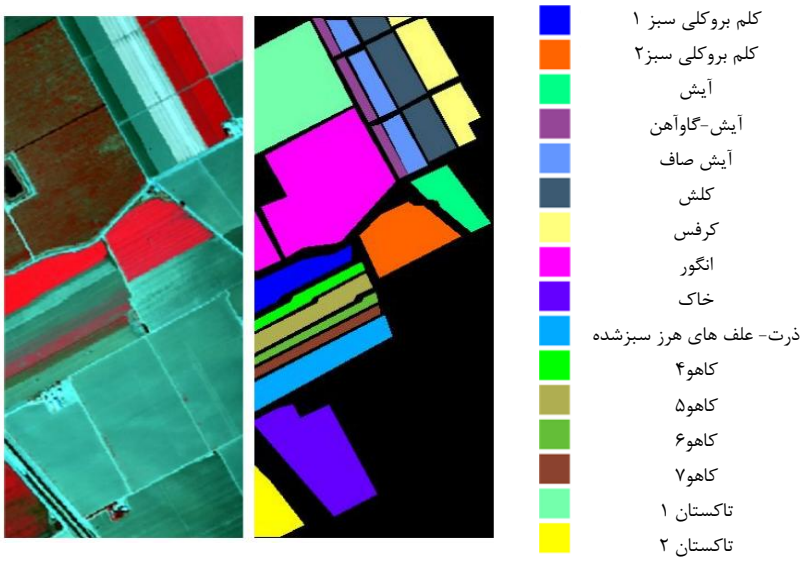
**۳-۳- سالیناس** (*Salinas*)

مجموعه داده‌های سالیناس شامل تصاویر ابرطیفی با ۵۱۲×۲۱۷ پیکسل و قدرت تفکیک مکانی ۳٫۷ متر است که توسط سنجنده هوابرد آویریس (*Airborne Visible Infrared Imaging Spectrometer*) در دره سالیناس در کالیفرنیا گرفته شده است. تصاویر طیفی به‌طور پیوسته ۲۲۴ باندطیفی در محدوده طول موج‌های ۴۲-۰٫۲ 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*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]




(ب) (الف)

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

**جدول ۳: کلاس‌های سالیناس**

| شماره کلاس | نام کلاس | تعداد نمونه‌ها |
|---|---|---|
| ۱ | کلم بروکلی سبز ۱ | ۲۰۰۹ |
| ۲ | کلم بروکلی سبز۲ | ۳۷۲۶ |
| ۳ | آیش | ۱۹۷۶ |
| ۴ | آیش-گاوآهن | ۱۳۹۴ |
| ۵ | آیش صاف | ۲۶۷۸ |
| ۶ | کلش | ۳۹۵۹ |
| ۷ | کرفس | ۳۵۷۹ |
| ۸ | انگور | ۱۱۲۷۱ |
| ۹ | خاک | ۶۲۰۳ |
| ۱۰ | ذرت- علف های هرز سبزشده | ۳۲۷۸ |
| ۱۱ | کاهو۴ | ۱۰۶۸ |
| ۱۲ | کاهو۵ | ۱۹۲۷ |
| ۱۳ | کاهو۶ | ۹۱۶ |
| ۱۴ | کاهو۷ | ۱۰۷۰ |
| ۱۵ | تاکستان ۱ | ۷۲۶۸ |
| ۱۶ | تاکستان ۲ | ۱۸۰۷ |



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

## ۴- نتایج و بحث

این قسمت به بررسی مشخصات سیستم مورد استفاده، تعیین مناسب ترین فراپارامترها برای دست‌یابی به بالاترین درصد دقت، پیاده‌سازی مدل شبکه و مقایسه روش پیشنهادی با دیگر روش‌ها می‌پردازد.

### 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) آی۷-۶۵۰۰[1] و کارت گرافیک اینتل (آر) اچ دی گرافیک ۵۲۰[2] با رام ۸ گیگابایت می‌باشد. همچنین کدنویسی در محیط برنامه‌نویسی پایتون، نسخه ۳٫۸ در بستر تنسورفلو[3] و کتابخانه‌های سطح بالای کراس[4] صورت گرفته است.

### ۴-۲- انتخاب فراپارامترها

در مدل‌سازی شبکه، انتخاب فراپارامترها از مهم‌ترین اقدامات اولیه می‌باشد، که این امر با انجام آزمون و خطا و سنجیدن تمام حالت‌های ممکن بر اساس تجربیات محققان در مقالات و مستندات [۱۴،۱۵،۲۲،۲۳و۲۴] به دست آمده است. یکی از پارامترهای مهم تعیین‌کننده خروجی شبکه، اندازه پنجره می‌باشد.

#### ۴-۲-۱- انتخاب اندازه پنجره بهینه

هدف اول، انتخاب مناسب‌ترین مقدار اندازه پنجره می‌باشد 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 داده‌ها بدین گونه بوده است: از کل مجموعه داده‌ها، ۷۵٪ جهت یادگیری مدل آموزشی اختصاص داده شد (۵۰٪ داده‌ها جهت آموزش، ۲۵٪ جهت اعتبارسنجی) و ۲۵٪ جهت آزمایش در نظر گرفته شده است.

**جدول۴: فراپارامترهای فرض شده، جهت انتخاب اندازه پنجره**

| | |
|---|---|
| **اندازه دسته‌ها** | ۱۲۸ |
| **تابع فعال‌سازی** | رلو |
| **نرخ یادگیری** | ۰٫۰۰۱ |
| **الگوریتم بهینه‌سازی** | آدام[5] |
| **تعداد دوره‌ها** | ۲۰ |

با در نظر گرفتن فرضیات جدول بالا، شروع به آزمون و خطا برای یافتن مناسب‌ترین اندازه پنجره می‌شود که در جدول (۵) نتایج حاصل از آزمایش مقادیر مختلف پنجره بر روی مجموعه داده‌های مختلف ذکر شده است. بدین صورت که به ازای هرکدام از اندازه پنجره‌ها بر روی داده‌های 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 می‌شود، مدل شبکه عصبی پیچشی پیشنهادی، بیشترین درصد دقت کلی و کمترین میزان خطا را به ازای اندازه پنجره ۱۹×۱۹ نشان داده است.

[1] *Intel(R) core(TM) i7-6500*
[2] *HD Graphics 520 Intel(R)*
[3] *Tensorflow*
[4] *Keras*
[5] *Adam*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

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

| اندازه پنجره | دانشگاه پاویا | | سالیناس | | ایندین‌پاینس | |
|---|---|---|---|---|---|---|
| | ٪ 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 درنظر گرفتن دیگر فراپارامترهای تعیین شده در جدول(۴) محاسبه خواهد شد.

### ۴-۲-۲-۱- 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 | | سالیناس | | ایندین‌پاینس | |
|---|---|---|---|---|---|---|
| | ٪ 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 |



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

### ۴-۲-۲-۲- صحت‌سنجی تابع فعال‌سازی

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 پاویا 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.

**جدول۷: مقایسه معیار دقت و خطا، از نظر فراپارامتر تابع فعال‌سازی**

| تابع فعال‌سازی | دانشگاه پاویا | | سالیناس | | ایندین‌پاینس | |
|---|---|---|---|---|---|---|
| | ٪ دقت کلی | خطا | ٪ دقت کلی | خطا | ٪ دقت کلی | خطا |
| رلو | **۱۰۰٫۰۰** | ۰٫۰۱۲ | **۱۰۰٫۰۰** | **۰٫۰۰۷** | **۹۹٫۹۶** | **۰٫۰۴۸** |
| سیگموئید | ۹۹٫۹۳ | **۰٫۰۰۸** | ۹۶٫۳۹ | ۰٫۱۶۱ | ۹۹٫۷۶ | **۰٫۰۳۲** |

### ۴-۲-۲-۳- صحت‌سنجی فراپارامتر نرخ یادگیری

در بخش سوم فراپارامتر 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.

**جدول۸: مقایسه معیار دقت و خطا، از نظر فراپارامتر نرخ یادگیری**

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

### ۴-۲-۲-۴- صحت‌سنجی الگوریتم بهینه‌سازی

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، درحالتی است که از الگوریتم بهینه‌سازی آدام استفاده گردیده است که این مقدار برابر با مقدار فرض شده در جدول(۴) می‌باشد.

[۱] *Stochastic Gradient Descent*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

**جدول۹: مقایسه معیار دقت و خطا، از نظر الگوریتم بهینه‌سازی**

| الگوریتم بهینه‌سازی | دانشگاه پاویا | | سالیناس | | ایندین‌پاینس | |
|---|---|---|---|---|---|---|
| | ٪ دقت کلی | خطا | ٪ دقت کلی | خطا | ٪ دقت کلی | خطا |
| آدام | **۱۰۰٫۰۰** | **۰٫۰۱۲** | **۱۰۰٫۰۰** | **۰٫۰۰۷** | **۹۹٫۸۴** | **۰٫۰۴۸** |
| گرادیان شیب نزولی | ۸۸٫۵۵ | ۰٫۱۳۲ | ۵۹٫۴۰ | ۱٫۲۲ | ۸۵٫۵۱ | ۰٫۵۹۲ |

### ۴-۲-۲-۵- 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.

**جدول۱۰: مقایسه معیار دقت و خطا از نظر فراپارامتر تعداد دوره‌ها**

| تعداد دوره‌ها | دانشگاه پاویا | | سالیناس | | ایندین‌پاینس | |
|---|---|---|---|---|---|---|
| | ٪ دقت کلی | خطا | ٪ دقت کلی | خطا | ٪ دقت کلی | خطا |
| ۱۵ | ۹۹٫۹۹ | ۰٫۰۱۵ | ۹۹٫۷۱ | ۰٫۰۱۴ | ۹۹٫۰۶ | ۰٫۱۱۵ |
| **۲۰** | **۱۰۰٫۰۰** | **۰٫۰۱۲** | **۱۰۰٫۰۰** | **۰٫۰۰۷** | **۹۹٫۹۶** | **۰٫۰۴۸** |
| ۲۵ | **۱۰۰٫۰۰** | ۰٫۰۱۳ | **۱۰۰٫۰۰** | ۰٫۰۰۸ | ۹۹٫۸۶ | ۰٫۰۴۱ |
| ۵۰ | **۱۰۰٫۰۰** | ۰٫۰۱۴ | **۱۰۰٫۰۰** | ۰٫۰۱۶ | ۹۹٫۹۲ | ۰٫۰۴۰ |
| ۷۵ | **۱۰۰٫۰۰** | ۰٫۰۱۶ | ۹۹٫۹۹ | ۰٫۰۰۷ | ۹۹٫۸۴ | ۰٫۰۳۳ |
| ۱۰۰ | **۱۰۰٫۰۰** | ۰٫۰۱۳ | **۱۰۰٫۰۰** | ۰٫۰۰۹ | ۹۹٫۸۸ | ۰٫۰۲۹ |



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

### ۴-۳- پیاده‌سازی مدل شبکه عصبی پیچشی پیشنهادی

در این بخش براساس فراپارامترهای به دست آمده در بخش قبلی در جدول (۱۱) و همچنین مقادیر درنظرگرفته شده در جدول (۱۲)، فرآیند ساخت مدل، آموزش، ارزیابی و تحلیل خروجی‌های مدل به ازای مجموعه داده‌های مختلف بررسی خواهد شد.

**جدول ۱۱: فراپارامترهای انتخاب شده شبکه عصبی پیچشی پیشنهادی**

| | |
|---|---|
| **الگوریتم بهینه‌سازی** | آدام |
| **نرخ یادگیری** | ۰٫۰۰۱ |
| **تابع فعال‌سازی** | رلو |
| **تعداد دوره‌ها** | ۲۰ |
| **اندازه دسته‌ها** | ۱۲۸ |
| **اندازه پنجره** | ۱۹×۱۹ |

**جدول ۱۲: پارمترهای ایجاد مدل پیشنهادی**

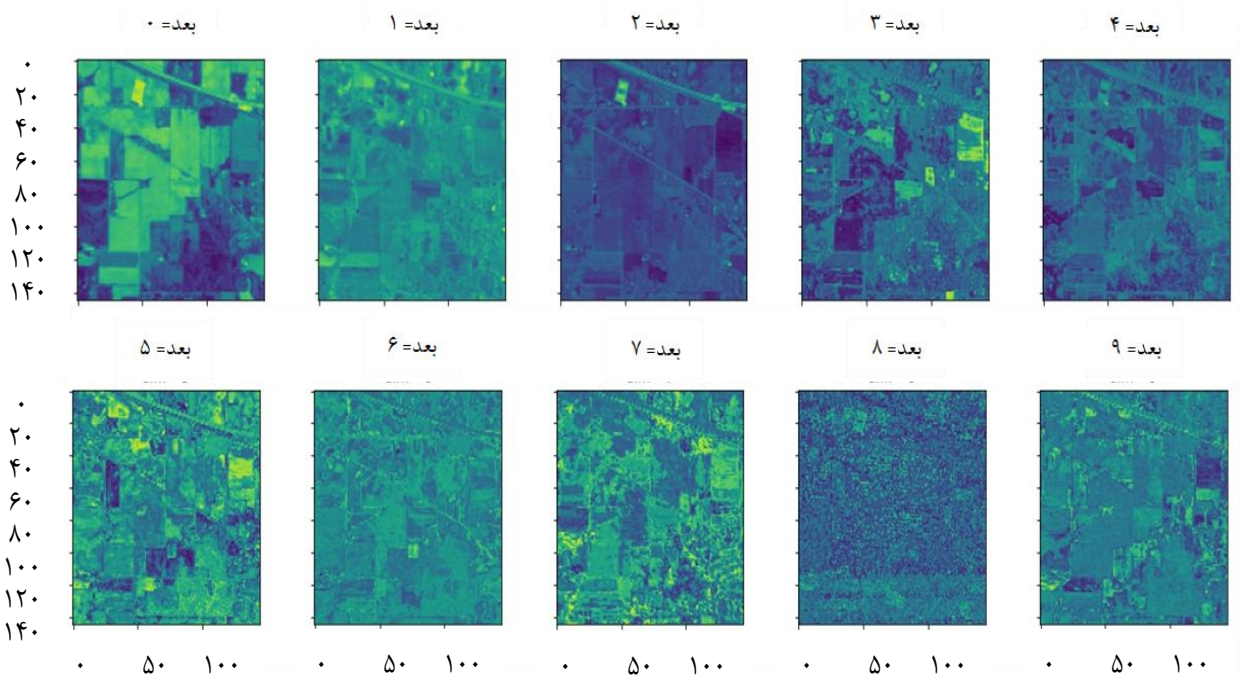
| | |
|---|---|
| **اندازه هسته** | ۳×۳ |
| **اندازه ادغام بیشینه** | ۲×۲ |
| **نرخ کاهش** | ۰٫۳۵ |
| **تابع فعال‌سازی لایه خروجی** | تابع بیشینه هموار |
| **تابع هزینه** | اختلاف آنتروپی متقاطع |

#### ۴-۳-۱- پیاده‌سازی مدل شبکه عصبی پیچشی پیشنهادی با مجموعه داده‌های ایندین‌پاینس

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 درنهایت 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 باندکاهش 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.



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]




**شکل۸: ۱۰ باند از ۶۹ باند کاهش یافته مجموعه داده‌های ایندین‌پاینس پس از اعمال *PCA***

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 ایندین‌پاینس را در مقایسه با تصویر حقیقت زمین آن نمایش می‌دهد.

در جدول (۱۳) ماتریس درهم ریختگی[1] مجموعه داده‌های ایندین‌پاینس مشاهده می‌شود. تعداد پیکسل‌هایی که درست طبقه‌بندی شده‌اند بر روی قطر اصلی قرار می‌گیرند.

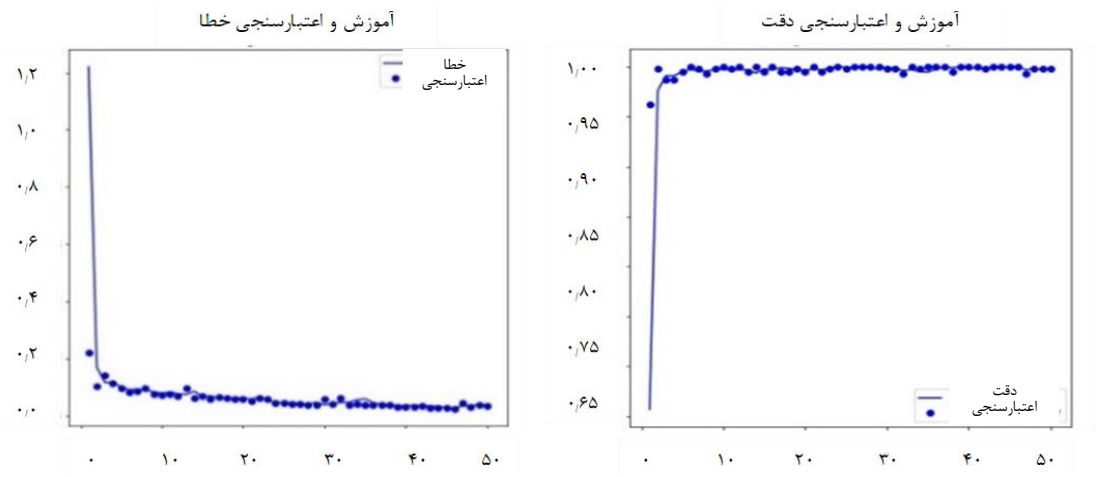
عناصر غیر قطری مجموعه خطاها می‌باشند. در اصل ماتریس درهم ریختگی، پراکندگی توزیع کلاس‌ها را از لحاظ درستی یا نادرستی نمایش می‌دهد. معیارهای *f1-score*، صحت[2]، کارایی[3] و میزان دقت برای هرکلاس، براساس نام و شماره هر کلاس، از مجموعه داده‌های ایندین‌پاینس در جدول (۱۴) قرار داده شده است. همچنین ستون آخر سمت چپ جدول نشان‌گر تعداد داده‌های موجود در هر کلاس می‌باشد.

[1] *Confusion Matrix*

[2] *Precision*
[3] *Recall*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]



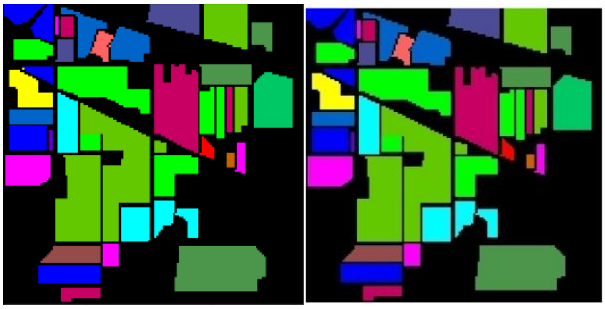

**شکل ۹: نمودارهای میزان دقت و خطا برای مجموعه داده ایندین‌پاینس، الف) دقت، ب) خطا**

**جدول ۱۳: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های ایندین‌پاینس**

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



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]



(ب) (الف)

**شکل ۱۰: نقشه‌های طبقه‌بندی مجموعه داده‌های ایندین‌پاینس. الف) تصویر حقیت زمین [۲۴]، ب) تصویر روش پیشنهادی**

**جدول۱۴: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های ایندین‌پاینس**

| شماره کلاس | نام کلاس | صحت | کارایی | *f1-score* | دقت | تعداد داده |
|---|---|---|---|---|---|---|
| ۱ | یونجه۱ | 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 | یونجه۲ | ۱٫۰۰ | ۰٫۹۹ | ۱٫۰۰ | ۰٫۹۹ | ۱۲۰ |
| ۹ | جو دوسر | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۵ |
| ۱۰ | سویا نعناع۱ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۲۴۳ |
| ۱۱ | سویا نعناع۲ | 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 |

### ۴-۳-۲- پیاده‌سازی مدل شبکه عصبی پیچشی پیشنهادی با مجموعه داده‌های دانشگاه پاویا

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 پاویا، ۱۶× ۱۹× ۱۹ است. معیارهای *f1-score* ، صحت، کارایی و میزان دقت برای هر کلاس براساس نام و شماره هر کلاس، از مجموعه داده‌های دانشگاه پاویا 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.



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

جدول۱۵: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های دانشگاه پاویا

| شماره کلاس | نام کلاس | صحت | کارایی | *f1-score* | 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 |

### ۴-۳-۳- پیاده‌سازی مدل شبکه عصبی پیچشی پیشنهادی با مجموعه داده‌های سالیناس

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 سالیناس، ۶× ۱۹× ۱۹ 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*[1] [۱۴]، *3D-CNN* [۲۵]، *HybridSN* [۱۵] و روش پیشنهادی در این پژوهش بر اساس دقت کلی[2] و ضریب کاپا[3] (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 شده‌است، 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 اینـدین‌پاینس [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 شده‌است. 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.

[1] *2 Dimension Convolutional Neural Network*

[2] *Overall Accuracy*
[3] *Kappa Coefficient*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

**جدول۱۶: معیارهای ارزیابی برای هر کلاس از مجموعه داده‌های سالیناس**

| شماره کلاس | نام کلاس | صحت | کارایی | *f1-score* | دقت | تعداد داده |
|---|---|---|---|---|---|---|
| ۱ | کلم بروکلی سبز ۱ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۵۰۲ |
| ۲ | کلم بروکلی سبز۲ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۹۳۲ |
| ۳ | آیش | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۴۹۴ |
| ۴ | آیش-گاوآهن | 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 سبزشده | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۸۲۰ |
| ۱۱ | کاهو۴ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۲۶۷ |
| ۱۲ | کاهو۵ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۴۸۲ |
| ۱۳ | کاهو۶ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۱٫۰۰ | ۲۲۹ |
| ۱۴ | کاهو۷ | 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 |

**جدول۱۷: مقایسه دیگر روش‌ها، از نظر پارامتر میزان دقت کلی وضریب کاپا با روش پیشنهادی**

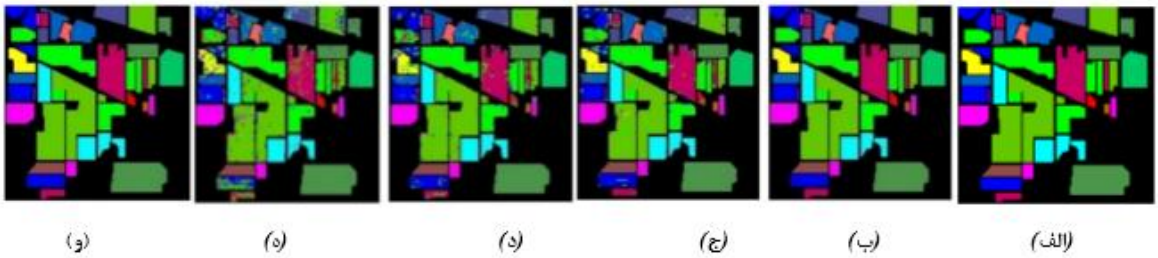
| مجموعه داده | دانشگاه پاویا | | سالیناس | | ایندین‌پاینس | |
|---|---|---|---|---|---|---|
| روش | دقت کلی | کاپا | دقت کلی | کاپا | دقت کلی | کاپا |
| *SVM* | ۹۴٫۳۴ ± ۰٫۲ | ۹۲٫۵۹ ± ۰٫۷ | ۹۲٫۹۵ ± ۰٫۳ | ۹۲٫۱۱ ± ۰٫۲ | ۸۵٫۳۰ ± ۲٫۸ | ۸۳٫۱۰ ± ۳٫۲ |
| *2D-CNN* | ۹۷٫۸۶ ± ۰٫۲ | ۹۷٫۱۶ ± ۰٫۵ | ۹۷٫۳۸ ± ۰٫۰ | ۹۷٫۰۸ ± ۰٫۱ | ۸۹٫۴۸ ± ۰٫۲ | ۸۷٫۹۶ ± ۰٫۵ |
| *3D-CNN* | ۹۶٫۵۳ ± ۰٫۱ | ۹۵٫۵۱ ± ۰٫۲ | ۹۳٫۹۶ ± ۰٫۲ | ۹۳٫۳۲ ± ۰٫۵ | ۹۱٫۱۰ ± ۰٫۴ | ۸۹٫۹۸ ± ۰٫۵ |
| *HybridSN* | ۹۹٫۹۸ ± ۰٫۰ | ۹۹٫۹۸ ± ۰٫۰ | **۱۰۰٫۰۰ ± ۰٫۰** | **۱۰۰٫۰۰ ± ۰٫۰** | ۹۹٫۷۵ ± ۰٫۱ | ۹۹٫۷۱ ± ۰٫۱ |
| روش پیشنهادی | **۱۰۰٫۰۰ ± ۰٫۰** | **۱۰۰٫۰۰ ± ۰٫۰** | **۱۰۰٫۰۰ ± ۰٫۰** | **۱۰۰٫۰۰ ± ۰٫۰** | **۹۹٫۹۰ ± ۰٫۰۶** | **۹۹٫۸۹ ± ۰٫۰۸** |

**جدول۱۸: مقایسه دیگر روش‌ها از نظر مدت زمان آموزش شبکه با روش پیشنهادی**

| مجموعه داده | دانشگاه پاویا | سالیناس | ایندین‌پاینس |
|---|---|---|---|
| روش | (برحسب دقیقه) | (برحسب دقیقه) | (برحسب دقیقه) |
| *2D-CNN* | ۱٫۸ | ۲٫۲ | ۱٫۹ |
| *3D-CNN* | ۵۸٫۰ | ۷۴٫۰ | ۱۵٫۲ |
| *HybridSN* | ۲۰٫۳ | ۲۵٫۵ | ۱۴٫۱ |
| روش پیشنهادی | **۱۳٫۸** | **۱۵٫۷** | **۶٫۳** |



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]




**شکل ۱۱: نتایج نقشه‌های طبقه‌بندی مجموعه داده‌های ایندین‌پاینس. (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، سالیناس و دانشگاه پاویا 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 فراپارامترها 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، به‌گونه‌ای 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. اصلی ترین 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 فراپارامترهای شبکه را بر اساس روش‌های مطرح در حوزه بهینه سازی تکاملی مدنظر داشت تا مقادیر بهینه پارامترها به‌صورت هدفمند انتخاب شوند و علاوه بر این روش‌های دیگر استخراج ویژگی را مورد توجه قرار داد. همچنین می‌توان نسبت بکارگیری داده های آموزش و آزمایش را تغییر داد و از روش های تولید داده‌های شبیه‌سازی شده جهت تولید داده‌های آموزشی افزونه استفاده کرد که می‌تواند نیاز به داده های آموزش را کاهش دهد و درنهایت 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 enviroment, 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*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

*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*

*[10] L. Shu, K. McIsaac, and G. R. Osinski, "Hyperspectral image classification with stacking spectral patches and convolutional neural networks, "IEEE Transactions on Geoscience and Remote Sensing, pp. 1–10, 2018.*

*[11] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural netoworks 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. 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 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*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

*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.*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]
[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]

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

*Journal of Geospatial Information Technology*
**Vol.11, No.1, Spring 2023**

***Research Paper***

# *Improving the Classification Accuracy of Hyperspectral Images Using Convolutional Neural Networks*

***Mahsa Tekyeh-Nejad [1], Ata Ebrahimzadeh [2]*, Maliheh Ahmadi [3]***

*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.*

***Correspondence Address****: Mazandaran- Babol- Shariati St.- Noshirvani University of Technology, Babol- Faculty of Electrical and Computer Engineering - Department of Telecommunication System.*
***Tel****: +981132339214*
***Email:*** *ezadehata@gmail.com*



[ Downloaded from jgit.kntu.ac.ir on 2026-10-11 ]

[ DOR: 20.1001.1.20089635.1402.11.1.4.5 ]
[ DOI: 10.61186/jgit.11.1.59 ]