نشریه علمی مهندسی فناوری اطلاعات مکانی

سال نهم، شماره دوم، تابستان ۱۴۰۰

Vol.9, No.2, Summer 2021

۱ -۲۷

مقاله پژوهشی

DOR: 20.1001.1.20089635.1400.9.2.3.2

طبقه‌بندی تصاویر ابرطیفی با استفاده از ادغام ویژگی‌های طیفی و مکانی در شبکه‌های عصبی پیچشی

عبید شریفی[1*]، بهنام اصغری بیرامی[2]، مهدی مختارزاده[3]

۱- دانشجوی کارشناسی ارشد سنجش از دور – دانشکده مهندسی نقشه برداری – دانشگاه صنعتی خواجه نصیرالدین طوسی

۲- دانشجوی دکتری سنجش از دور – دانشکده مهندسی نقشه برداری – دانشگاه صنعتی خواجه نصیرالدین طوسی

۳- دانشیار گروه آموزشی فتوگرامتری و سنجش از دور - دانشکده مهندسی نقشه برداری – دانشگاه صنعتی خواجه نصیرالدین طوسی

تاریخ دریافت مقاله: ۱۳۹۸/۰۵/۰۷ تاریخ پذیرش مقاله: ۱۳۹۸/۱۰/۰۸

چکیده

سنجنده‌های ابرطیفی به واسطه اخذ تعداد زیادی از باندهای طیفی همواره دارای اهمیت خاصی در پایش پدیده‌های سطح زمین می‌باشند. طبقه‌بندی تصاویر ابرطیفی مهم‌ترین روش پردازش داده‌های ابرطیفی است که تا به حال تلاش‌های زیادی برای افزایش دقت آن صورت گرفته است. شبکه‌های عصبی پیچشی و ویژگی‌های مکانی در سال‌های اخیر جایگاه مهمی در بهبود دقت طبقه‌بندی تصاویر ابرطیفی داشته‌اند. در تحقیقات پیشین توجه زیادی به استفاده همزمان از قابلیت‌های روش‌های استخراج ویژگی مکانی سطح پایین در شبکه‌های عصبی پیچشی نشده است. به همین دلیل در مقاله حاضر یک معماری جدید از شبکه‌های عصبی پیچشی برای طبقه‌بندی تصاویر ابرطیفی معرفی شده است که به عنوان ورودی شبکه از بردار طیفی_مکانی حاصل از ترکیبات مختلف ویژگی‌های مکانی سطح پایین شامل پروفایل‌های مورفولوژی، بانک فیلترگابور و الگوی باینری محلی با ویژگی‌های طیفی استخراج شده از روش تبدیل مولفه اصلی استفاده می‌کند. آزمایش‌های این مقاله که بر روی دو تصویر ابرطیفی حقیقی از دو منطقه کشاورزی و شهری صورت گرفته است، نشان از برتری روش پیشنهادی دارد. نتایج نهایی نشان می‌دهد که دقت کلی طبقه‌بندی با روش پیشنهادی می‌تواند در بهترین حالت ۲٫۵ درصد از روش‌های رقیب بهتر باشد.

کلید واژه‌ها : طبقه‌بندی تصاویر ابرطیفی، شبکه‌های عصبی پیچشی، پروفایل‌های مورفولوژی، بانک فیلتر گابور، الگوی باینری محلی.

* نویسنده مکاتبه کننده: گروه فتوگرامتری و سنجش‌ازدور، دانشکده مهندسی نقشه‌برداری، پردیس دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران.

تلفن: ۰۲۱۸۸۸۷۷۰۷۱

Email: O.sharifi@kntu.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]

[DOR: 20.1001.1.20089635.1400.9.2.3.2]

[DOI: 10.52547/jgit.9.2.1]

۱- مقدمه

سنجنده‌های ابرطیفی با اخذ تعداد زیادی از باندهای طیفی از نواحی فرابنفش نزدیک تا مادون قرمز بازتابی به ابزار بسیار سودمندی برای مطالعه پدیده‌های سطح زمین تبدیل شده‌اند. طبقه‌بندی تصاویر که یکی از مهم‌ترین روش‌های استخراج اطلاعات از تصاویر ابرطیفی[۱] (*HSI*) می‌باشد، به معنی در یک گروه قرار دادن پیکسل‌های مشابه و تعیین برچسب برای هر پیکسل است. روش‌های طبقه‌بندی تصاویر به طور کلی به دو دسته نظارت‌شده و نظارت‌نشده دسته‌بندی می‌شوند. در روش‌های نظارت‌نشده، طبقه‌بندی‌کننده بدون نیاز به اطلاعات عامل خارجی و با جست‌وجوی فضای ویژگی سعی در دسته‌بندی پیکسل‌های مشابه بر اساس معیارهای گوناگون دارد، اما در روش‌های نظارت‌شده کاربر با انتخاب پیکسل‌های نمونه برای هر کلاس سعی می‌کند تا طبقه‌بندی‌کننده را طوری تربیت کند که در مواجه شدن با پیکسل‌های جدید برچسب آنها را به درستی تشخیص دهد. از جمله طبقه‌بندی‌های نظارت‌شده سنتی می‌توان به *K* همسایه نزدیک[۲]، حداقل فاصله و رگرسیون لجستیک[۳] اشاره کرد[۱]. در تحقیقات بعدی، ماشین بردار پشتیبان[۴] (*SVM*) به عنوان یک روش کارآمد و پایدار برای طبقه‌بندی *HSI*، به ویژه برای حالت‌های که تعداد نمونه‌های آموزشی کم می‌باشد مورد استفاده قرار گرفتند[۲و۳].

در سال‌های اخیر شبکه‌های عصبی[۵] (*NN*) وارد مباحث سنجش از دوری شدند و شبکه‌هایی مانند شبکه‌های پرسپترون چندلایه[۶] [۴]، ماشین بولتزمن[7] و تابع پایه شعاعی[۸] [۵] برای طبقه‌بندی داده‌های سنجش از دور مورد استفاده قرار گرفته‌اند. تحقیقات گسترده اخیر در مورد شبکه‌های عصبی نشان داده است که شبکه‌های عصبی جایگزین امیدوارکننده‌ای برای روش‌های مختلف طبقه‌بندی سنتی هستند. جامع‌ترین آنها کار میشی و همکاران است[۹]. در این پروژه سه طبقه‌بندی کلی شبکه‌های عصبی، طبقه‌بندی‌های آماری و یادگیری ماشین با ۲۳ روش مقایسه شده‌اند و بیش از ۲۰ مجموعه داده مختلف مورد استفاده قرار گرفتند. نتیجه کلی آنها این است که هیچ طبقه‌بندی تنها برای تمام مجموعه داده‌ها مناسب نیست؛ هرچند شبکه‌های عصبی عملکرد خوبی را نسبت به بقیه روش‌ها دارند. از جمله مزیت‌های شبکه‌های عصبی می‌توان به این موارد اشاره کرد: ۱) توانایی سازگاری با داده‌های مختلف که دارای توزیع آماری مشخصی نیستند ۲) شبکه‌های عصبی توانایی تقریب هر تابع با دقت دلخواه را دارا هستند [۶و۷]. ۳) شبکه‌های عصبی چون مدل‌های غیرخطی هستند در مدل‌سازی روابط پیچیده دارای انعطاف‌پذیری بالایی هستند. ۴) شبکه‌های عصبی قادر به برآورد احتمال اولیه هستند که مبنای ایجاد قانون طبقه‌بندی و انجام تحلیل‌های آماری است [۸]. هینتون و سالاخوتدینف (۲۰۰۶)، یک معماری عمیق‌تر از شبکه‌های عصبی که شامل چند لایه رمزگذار خودکار بوده و مدل قدرتمندی برای طبقه‌بندی داده‌های سنجش از دوری است مورد استفاده قرار گرفت[۱۱]. نتایج تحقیق نشان داد که این معماری عمیق‌تر دارای عملکردی قابل رقابت با *SVM* هست و این موضوع باعث شروع تحقیقات بر روی شبکه‌های عصبی عمیق[۹] (*DNN*) شد.

یافته‌های علم اعصاب نشان می‌دهد سیستم بصری انسان توسط یک دنباله‌ای از سطوح مختلف پردازشی تشکیل شده است که در این سیستم از ویژگی‌های سطحی می‌توان به ویژگی‌های عمیق که توانایی مغز در شناسایی شیء را بهبود می‌دهند، رسید. معماری

[1] *Hyperspectral image (HSI)*
[2] *K-nearest neighbor (KNN)*
[3] *Logistic regression (LR)*
[4] *Support Vector machine (SVM)*
[5] *Neural networks (NN)*
[6] *Multilayer perceptron (MLP)*
[7] *Restricted Boltzmann machine (RBM)*
[8] *Radial basis function (RBF)*
[9] *Deep neural networks (DNNs)*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

شبکه‌های عصبی عمیق نیز که از روند شبکیه به قشر مغز شبیه‌سازی شده شامل چندین لایه برای استخراج ویژگی‌های عمیق می‌باشند. یادگیری عمیق در بسیاری از زمینه‌های تحقیقاتی، از جمله مباحث سنجش از دوری مانند طبقه‌بندی تصاویر ابرطیفی و تشخیص هدف مورد استفاده قرار گرفتند[۱۲]. سال ۲۰۰۶ اولین روش یادگیری عمیق تحت عنوان شبکه‌های باور عمیق[۱] (*DBN*) برای طبقه‌بندی تصاویر ابرطیفی مورد استفاده قرار گرفت[۱۰]. در *DBN* از ترکیب چند لایه ماشین بولتزمن برای استخراج ویژگی‌های عمیق استفاده می‌شود[۱۳]. چن و همکاران (۲۰۱۴) نیز یک روش یادگیری عمیق دیگر برای طبقه‌بندی تصاویر ابرطیفی معرفی کردند که در این روش از ترکیب چندین رمزگذار خودکار[2] (*SAE*) برای استخراج ویژگی‌های عمیق استفاده می‌شود[۱۴].

در ساختار هر دو روش *DBNs* و *SAEs* با توجه به وجود اتصالات کامل بین لایه‌های مختلف تعداد زیادی 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 شود و در یک بردار همراه اطلاعات طیفی وارد شبکه شوند. به منظور حل این مشکلات در سال‌های اخیر شبکه‌های عصبی پیچشی[۳] (*CNNs*) معرفی شده‌اند و بسیار مورد توجه قرار گرفتند. *CNNs* به کمک در نظر گرفتن ارتباطات محلی می‌توانند اطلاعات مکانی را استخراج کنند. همچنین در این شبکه‌ها استراتژی اشتراک وزن باعث کاهش چشم‌گیری در پارامترهای قابل آموزش شبکه می‌شود. از سال ۲۰۱۲، *CNNs* توجه زیادی را در موضوعاتی نظیر طبقه‌بندی تصویر [۱۶]، قطعه‌بندی [۱۷]، تشخیص شیء [۱۸]، تحلیل ویدئو [۱۹] و غیره به خود جلب کرده است. در ادامه چندین کار تحقیقاتی مهم در ارتباط با طبقه‌بندی تصاویر ابرطیفی به کمک شبکه‌های عصبی پیچشی بررسی می‌شود.

هو و همکاران (۲۰۱۵) از یک شبکه عصبی پیچشی برای طبقه‌بندی تصاویر ابرطیفی استفاده کردند که در معماری *CNN* مورد استفاده از یک لایه کانولوشن، یک لایه ادغام[۴] با عملگر ماکزیمم و دو لایه با اتصال کامل استفاده شده است[۲۰]. لایه کاملاً متصل نهایی با استفاده از تابع حداکثر نرم‌کنندگی[۵] میزان تعلق پیکسل مورد نظر را به کلاس‌های مختلف مشخص می‌کند. در این شبکه ورودی یک بردار $n\times 1$ می‌باشد که n تعداد باندهای طیفی تصویر ورودی هست بنابراین طبقه‌بندی فقط براساس اطلاعات طیفی انجام می‌شود. پردازش‌ها که بر روی سه داده ابرطیفی انجام شده، نشان از برتری *CNN* نسبت به طبقه‌بندی کننده *SVM* با استفاده از کرنل *RBF* دارد.

در تحقیقات بعدی چن و همکاران (۲۰۱۶) یک ساختار سه‌بعدی جدید از *CNN* برای طبقه‌بندی تصاویر ابر طیفی ارائه داده‌اند که در این رویکرد از چندین لایه کانولوشن و لایه ادغام استفاده شده است[۲۱]. ورودی این شبکه به صورت $d\times d\times n$ می‌باشد که d ابعاد کرنل مکانی و n تعداد باندهای طیفی است. *CNN* پیشنهاد شده به علت ورودی سه‌بعدی توانایی استخراج ویژگی‌های مکانی را نیز دارد. در این تحقیق برای حل مسئله بیش برازش[۶] از تکنیک حذف تصادفی[۷] و تولید نمونه مجازی استفاده می‌شود. دو نتیجه مهم پیاده-

[1] *Deep belief networks (DBNs)*
[2] *Stacked Autoencoder (SAE)*
[3] *Convolutional neural networks (CNNs)*
[4] *Pooling*
[5] *Softmax*
[6] *Overfitting*
[7] *Dropout*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

سازی‌های این تحقیق عبارت‌اند از: ۱) دقت طبقه‌بندی با استفاده از ویژگی‌های عمیق استخراج شده از *CNN* نسبت به روش‌های دیگر استخراج ویژگی مانند آنالیز مولفه اصلی*(PCA)* [1] و تجزیه و تحلیل عوامل*(FA)* [2] بهتر است. ۲) در شبکه‌های *3D-CNN* چون علاوه بر ویژگی‌های طیفی، ویژگی‌های مکانی نیز استخراج می‌شود، نتایج خیلی بهتری نسبت به *1D-CNN* و *2D-CNN* تولید می‌کند.

از سوی دیگر استفاده از ویژگی‌های طیفی به تنهایی در پروسه طبقه‌بندی تصاویر سنجش از دوری هنگامی که عوارض موجود رفتارهای طیفی نزدیک دارند عموماً منجر به دقت‌های بالایی نمی‌شود[۲۳و۲۴]. با توجه به پیشرفت‌های اخیر در فن‌آوری‌های سنجش از دور و بهبود قدرت تفکیک مکانی سنجنده‌ها یک راهکار مناسب استفاده از اطلاعات مکانی است، زیرا عوارض همسایه با احتمال بالا مربوط به یک کلاس می‌باشند و مکمل ارزشمندی برای ویژگی‌های طیفی در بحث طبقه‌بندی می‌باشند. در یک نگاه کلی ویژگی‌های مکانی سه نوع هستند:۱) ویژگی‌های مکانی سطح پایین (پروفایل مورفولوژی[۲۷]) ۲) ویژگی‌های مکانی سطح میانه (کیف کلمات بصری[3][۴۷]) ۳) ویژگی‌های مکانی عمیق. برای استخراج ویژگی‌های مکانی سطح پایین روش‌های زیادی پیشنهاد شده که از پرکاربردترین آنها می‌توان به ویژگی‌های مستخرج از ماتریس هم رخداد، ویژگی‌های بانک فیلتر گابور، پروفایل‌های مورفولوژی[4] و روش الگوی باینری محلی[5] *(LBP)* اشاره کرد. نتایج تحقیقاتی نظیر فراری و همکاران (۲۰۰۱) که از بانک فیلتر گابور در جهت‌ها و مقیاس‌های متفاوت استفاده شده است[۲۹]، میرزاپور و قاسمیان (۲۰۱۳) که از مفهوم پروفایل مورفولوژی استفاده شده[۲۷] و تحقیق اوجالا و همکاران (۲۰۰۲) که از روش *LBP* برای تولید ویژگی‌های مکانی استفاده کرده است[۳۰]، نشان می‌دهد استفاده از ویژگی‌های مکانی سطح پایین در کنار ویژگی‌های طیفی باعث بهبود دقت طبقه‌بندی تصاویر سنجش از دوری می‌شود. در تحقیقات مذکور از قابلیت‌های روش‌های یادگیری عمیق استفاده نشده است در حالی که در سال‌های اخیر روش یادگیری عمیق *CNN* بسیار مورد توجه قرار گرفتند. در این راستا، اپتوئل و همکاران (۲۰۱۶) در ورودی *CNN* علاوه بر ویژگی‌های طیفی مستخرج از تبدیل *PCA* از ویژگی‌های مکانی استخراج شده از پروفایل‌های خصوصیت[6] استفاده کردند[۲۳]. در این مقاله پس از استخراج مولفه‌های *PCA*، برای هر باند *PCA*، پروفایل‌های خصوصیت استخراج شدند و این دو ویژگی در کنار هم قرار گرفته و همزمان وارد *CNN* شدند. نتایج پیاده‌سازی بر روی دو داده ابرطیفی نشان می‌دهد که استفاده از پروفایل‌های ویژگی در کنار ویژگی‌های طیفی *PCA* در شبکه های عمیق سبب بهبود هر چه بیشتر دقت طبقه‌بندی نسبت به موقعی می‌شود که در ورودی شبکه فقط از ویژگی‌های طیفی استفاده شود. همچنین، سن تارا و همکاران (۲۰۱۷) یک روش یادگیری عمیق برای طبقه‌بندی طیفی_مکانی ارائه داده اند که شامل سه مرحله است: ۱) انتخاب ویژگی‌های طیفی ۲) در مرحله دوم این تحقیق به اندازه تعداد باندهای طیفی انتخاب شده در مرحله اول شبکه‌های عصبی پیچشی طراحی شده و این شبکه‌ها برای استخراج ویژگی‌های مکانی به صورت موازی بر روی باندهای انتخاب شده اعمال می‌شوند ۳) در مرحله آخر ویژگی‌های طیفی و مکانی تولید شده از مرحله قبل در کنار هم قرار می‌گیرند و وارد طبقه‌بندی‌کننده حداکثر نرم‌کنندگی [۲۰] می‌شوند[۴۱]. نتایج تحقیق نشان می‌دهد که روش ارائه شده باعث بهبود دقت طبقه‌بندی نسبت به روش‌های دیگر شده است. در تحقیق روی و

[1] *Principal component analysis*
[2] *Factor analysis*
[3] *bag of visual words*
[4] *Morphological Profiles (MP)*
[5] *Local binary patterns*
[6] *Attribute profiles*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

همکاران (۲۰۱۹) ابتدا به کمک *2D-CNN* ویژگی‌های مکانی عمیق استخراج می‌شوند و سپس این ویژگی‌ها در کنار ویژگی‌های طیفی وارد *3D-CNN* شده و طبقه-بندی انجام می‌شود. نتایج این تحقیق نیز نشان می‌دهد که استفاده از ویژگی‌های مکانی عمیق در ورودی *CNN* باعث بهبود دقت طبقه‌بندی تصاویر ابرطیفی می-شود[۴۰].

از مرور تحقیقات پیشین می‌توان به جایگاه مهم شبکه-های عصبی پیچشی و ویژگی‌های مکانی در بهبود دقت طبقه‌بندی پی برد. از آنجایی که ویژگی‌های سطح پایین مختلف دارای اطلاعات گوناگون مکانی می‌باشند به نظر می‌رسد که ادغام این ویژگی‌ها با ویژگی‌های طیفی در شبکه‌های عصبی عمیق موجب تولید ویژگی-های سطح عمیق باکیفیتی می‌شود. از این‌رو در مقاله حاضر یک ساختار جدید *3D-CNN* برای طبقه‌بندی تصاویر ابرطیفی معرفی شده، که برای بررسی تاثیر ویژگی‌های مکانی سطح پایین در نتایج *CNN* که پیش-تر مورد استفاده نبوده است در ورودی شبکه از ویژگی-های مکانی گوناگونی نظیر پروفایل مورفولوژی، فیلتر گابور و *LBP* در کنار ویژگی‌های طیفی استفاده می-شود. این مسئله از این جهت دارای اهمیت است که شبکه پیشنهادی توانایی تولید ویژگی‌های طیفی_مکانی عمیقی را دارا است که در طبقه‌بندی به سطوح دقت بالایی می‌رسد (بالای ۹۶ درصد). این موضوع حتی در زمانی که داده آموزشی اندک در اختیار باشد نیز صادق است. همچنین در معماری شبکه پیشنهادی از تکنیک-های پیشرفته‌ای مانند حذف شدن تصادفی[۴۴]، تابع فعال‌ساز واحد خطی اصلاح کننده *(ReLU)*[1] [۲۲]، و نرمال‌سازی دسته‌ای[2] *(BN)* [۴۵] که پیش‌تر نیز مورد استفاده قرار گرفته‌اند استفاده می‌شود.

ساختار تحقیق به این صورت است که در قسمت دوم به مبانی نظری تحقیق پرداخته می‌شود که در آن به معرفی شبکه‌های عصبی پیچشی و طرز استخراج ویژگی‌ها پرداخته می‌شود. در قسمت سوم طرح کلی تحقیق و روش پیشنهادی شرح داده می‌شود. در قسمت چهارم پس از معرفی دو داده ابرطیفی حقیقی مورد استفاده، نتایج پیاده‌سازی‌ها ارائه شده است. در قسمت آخر نیز نتیجه‌گیری و جمع‌بندی نهایی صورت می‌گیرد.

۲- مبانی نظری تحقیق

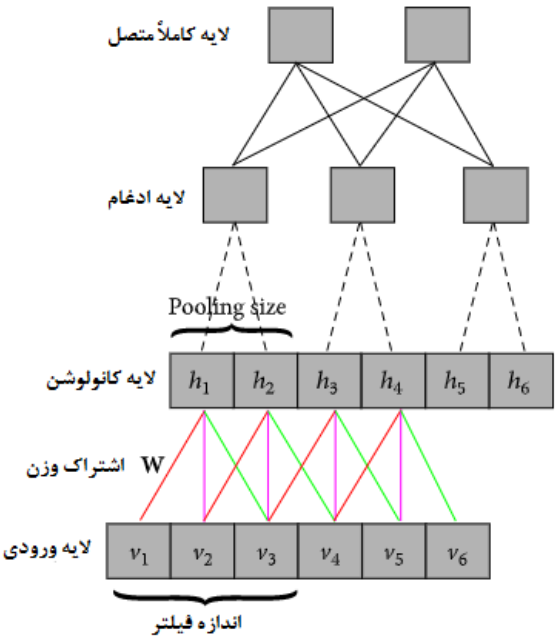
در این قسمت ابتدا به بررسی شبکه‌های عصبی پیچشی و لایه‌های آن می‌پردازیم و در ادامه روش‌های استخراج ویژگی مکانی را بررسی می‌کنیم.

۲-۱- شبکه‌های عصبی پیچشی

CNNs شبکه‌های عصبی عمیقی هستند که از تکنیک آموزش به جلو[3] استفاده می‌کنند و شامل ترکیبی از لایه‌های کانولوشن، ادغام و کاملاً متصل هستند. *CNNs* به علت استفاده از اطلاعات مکانی پیکسل‌های همسایه علاوه بر ویژگی‌های طیفی توانایی استخراج ویژگی‌های مکانی عمیق را نیز دارا هستند که این ویژگی‌ها باعث بهبود دقت طبقه‌بندی تصاویر ابرطیفی می‌شود.[۲۰]. لایه کانولوشن برای استخراج ویژگی‌ها از یک سری کرنل خطی استفاده می‌کند. اولین لایه کانولوشن مسئول تولید ویژگی‌های سطح پایین مانند لبه‌ها، جهت گرادیان و دیگر موارد است. با افزایش تعداد لایه-های کانولوشن، معماری با ویژگی‌های سطح بالا نیز سازگار می‌شود[۲۱]. بعد از لایه کانولوشن معمولاً یک لایه ادغام قرار می‌گیرد. لایه ادغام خروجی چندین نورون لایه‌ی کانولوشن را از طریق یک تابع که معمولاً از تابع ماکزیمم استفاده می‌شود با هم ادغام می‌کند. ادغام خروجی نورون‌ها باعث تولید ویژگی‌های غیر قابل تغییر از محل که ویژگی‌های مناسب برای طبقه‌بندی تصاویر ابرطیفی می‌باشند می‌شود[۳۱]. یک معماری متداول شبکه‌های عصبی در شکل (۱) نشان داده شده است.

[1] *Rectifier Linear Unit*
[2] *Batch Normalization*
[3] *Feed forward*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]



شکل ۱: معماری متداول *CNN* که متشکل از لایه‌های کانولوشن، ادغام و لایه کاملاً متصل است[۲۰].

CNNs همانطور بیان شد دارای دو ویژگی خاص هستند: ۱) اتصالات محلی ۲) اشتراک وزن. در *CNN* نورون‌ها در لایه کانولوشن فقط به نورونی در لایه بعدی متصل هستند که دارای ارتباط مکانی باشند، پس مقدار ورودی تابع فعال‌ساز با ضرب یک ورودی کوچک محلی در وزن W محاسبه می‌شود. سپس وزن W در کل فضای ورودی به اشتراک گذاشته می‌شود. همانطور که در شکل(۱) نشان داده شده است، خطوطی که رنگ یکسان دارند وزن آنها یکسان است. مزیت‌های استفاده از اشتراک وزن در *CNN* عبارت‌اند از [۲۰]:

۱. به کمک استفاده از بحث اشتراک وزن تعداد پارامترهای قابل آموزش شبکه کاهش پیدا می‌کند، پس به تعداد نمونه آموزشی کمتری برای آموزش شبکه احتیاج است.

۲. باتوجه به شکل(۱) در این حالت چون پارامترهای وزن از چندین ورودی برای یادگیری استفاده می‌کند، مثلا برای پارامتر w_1 که به علت بحث اشتراک وزن در تعیین مقدارهای خروجی h_1، h_2، h_3 و h_4 نقش دارد از چهار وروی v_1، v_2، v_3 و v_4 برای آموزش آن استفاده می‌شود. به همین دلیل اگر یکی از ورودی‌ها نویزی باشد روی پارامتر وزن تاثیر زیادی نمی‌گذارد.

۳. اشتراک وزن در *CNN* باعث افزایش قدرت تعمیم‌پذیری شبکه و حل مشکل بیش برازش می‌شود.

۴. برای اجرای *CNN* به حافظه کمتری احتیاج است.

در فرایند طبقه‌بندی با *CNN* چون پردازش شبکه زمان‌بر و سنگین است، ابتدا نمونه‌های آموزشی به‌طور تصادفی به چند دسته تقسیم می‌شوند و هر گروه شامل تعداد برابری از نمونه‌های آموزشی می باشد. برای هر تکرار فقط یک دسته از نمونه‌ها برای آموزش به شبکه فرستاده می‌شوند و پس از انجام عملیات هرکدام از لایه‌ها مقدار خروجی تعیین می‌شود. پس از تعیین مقدار خروجی، با روش تصادفی گرادیان نزولی با شتاب *(sgdm)*[۱] و استفاده از تابع هزینه آنتروپی متقاطع[۲] شبکه آموزش داده می‌شود. هنگامی که تمام گروه‌های کوچک تولید شده از نمونه‌های آموزشی وارد

[1] *Stochastic gradient descent with momentum*

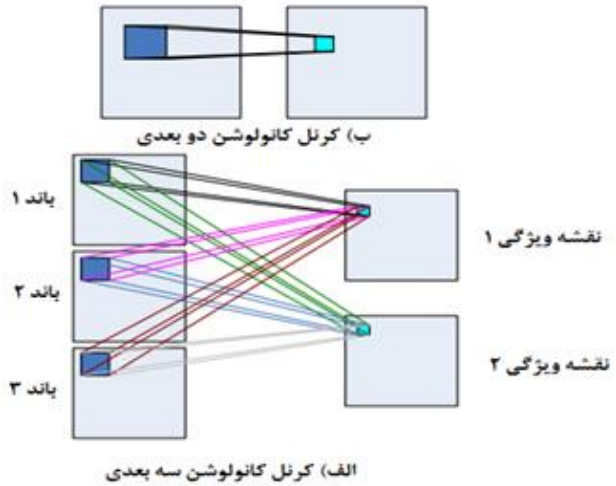
[2] *Cross Entropy*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

شبکه شدند، یک اپک به اتمام می‌رسد. روند آموزش تا زمانی که به حداکثر تعداد اپک نرسد متوقف نخواهد شد. در ادامه ساختار لایه‌های *CNN* بررسی می‌شود.

۲-۱-۱- لایه کانولوشن

لایه کانولوشن برای استخراج ویژگی از فیلترهای کانولوشن خطی با ابعاد $k_w \times k_h \times d$ استفاده می‌کند، که k_w طول فیلتر، k_h عرض فیلتر و d ارتفاع فیلتر که برابر تعداد باندهای تصویر ورودی است. کرنل‌های کانولوشن با اندازه $1\times1\times d$ تنها می‌تواند ویژگی‌های طیفی را استخراج کنند و برای استخراج ویژگی‌های مکانی کاربرد ندارند. هنگامی که هدف طبقه‌بندی فقط براساس اطلاعات مکانی باشد بعد سوم که نشان‌دهنده تعداد باندهای ورودی می‌باشد یک است و عملاً کرنل مورد استفاده یک کرنل دو بعدی است و در حالت طبقه‌بندی طیفی_مکانی از کرنل سه بعدی استفاده می‌شود. در شکل(۲) نحوه عملیات کرنل کانولوشن دو بعدی و سه بعدی نشان داده شده است.



شکل ۲: نحوه عملکرد کرنل کانولوشن دوبعدی و سه‌بعدی. اندازه کرنل کانولوشن در محدوده طیفی ۳ است[۲۲].

پس از اعمال هر کرنل یک نقشه ویژگی ایجاد می‌شود. در این مقاله چون از کرنل سه‌بعدی که در شکل(۲b) نشان داده شده استفاده می‌شود در نتیجه طبق رابطه(۱) مقدار نورون لایه iام در نقشه ویژگی jام برای پیکسلی با موقعیت (x,y,z) که x و y مربوط به شماره سطر و ستون پیکسل و z نشان دهنده باند مورد استفاده می‌باشد به دست می آید[۲۲]. در رابطه (۱)، k_w طول فیلتر، k_h عرض فیلتر، d تعداد باندهای تصویر ورودی، m تعداد نقشه‌های ویژگی در لایه $(i\text{-}1)$، w_{ijm}^{pqr} وزن موقعیت (p,q,r) مربوط به نقشه ویژگی m است، b_{ij} بایاس نقشه ویژگی jام در لایه iام است، $g(x)$ تابع فعال‌ساز و $v_{(i-1)m}$ مقدار خروجی نورون‌های لایه $(i-1)$ در موقعیت‌های مختلف است. وزن کرنل‌ها پارامترهایی هستند که به کمک داده‌های آموزشی به‌روزرسانی می‌شوند.

$$V_{ij}^{xyz} = g(b_{ij} + \sum_m \sum_{p=0}^{k_w-1} \sum_{q=0}^{k_h-1} \sum_{r=0}^{d} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}) \qquad \text{رابطه (۱)}$$

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

۲-۱-۲- لایه ادغام

در *CNNs* بعد از لایه کانولوشن معمولاً یک لایه ادغام وجود دارد، به نحوی که تعداد نقشه‌های ویژگی در این لایه با لایه کانولوشن برابر است. در لایه ادغام از یک کرنل با سایز $p \times q$ ، که معمولاً $p,q \in \{۲,۳,۴,۵\}$ انتخاب می‌شوند، استفاده می‌شود[۴۶]. هدف از لایه ادغام، کاهش قدرت محاسباتی مورد نیاز برای پردازش داده‌ها از طریق کاهش ابعاد است. علاوه بر این، لایه ادغام به علت ترکیب نتایج خروجی چندین نورون لایه کانولوشن، باعث تولید ویژگی‌هایی می‌شود که در مقابل دوران و تغییر موقعیت ثابت هستند. این ویژگی‌ها که تحت عنوان ویژگی‌های غالب[1] شناخته می‌شوند سرعت همگرایی شبکه را افزایش داده و باعث آموزش بهتر شبکه می‌شوند[۳۱].

در لایه ادغام می‌توان از توابع مختلف استفاده کرد که پرکاربردترین آن‌ها ماکزیمم و میانگین هستند. نحوه عملکرد این توابع بر روی داده ورودی در شکل(۳) نشان داده شده است. تحقیقات نشان داده که تابع ماکزیمم نتایج بهتری از تابع میانگین ایجاد می‌کند، به همین دلیل در لایه ادغام عموما از تابع ماکزیمم استفاده می‌شود[۲۱].

در لایه ادغام پارامتری برای آموزش وجود ندارد ولی نکته‌ای که وجود دارد این است که در حالت انتشار رو به عقب در لایه ادغام که از تابع ماکزیمم استفاده می‌شود باید خطا را به نورونی برگردانیم که بیشترین مقدار را داشته است.

۲-۱-۳- لایه کاملاً متصل

بعد از چندین لایه کانولوشن و ادغام، لایه‌های کاملاً متصل قرار می‌گیرند. این لایه ویژگی‌های طیفی و مکانی را که در لایه‌های قبلی استخراج شده‌اند برای طبقه‌بندی داده ورودی باهم ترکیب می‌کند. همانطور که از نام این لایه معلوم است، برای اینکه دو لایه متوالی به طور کامل متصل شوند، تمام نورون‌ها در لایه قبلی باید به تمام نورون‌ها در لایه بعدی متصل شوند. مقدار خروجی لایه کاملاً متصل را با استفاده از مقادیر وزن و بایاس می‌توان از رابطه(۲) محاسبه کرد[۳۲]:

$$o^{(k)} = \left(o^{(k-1)}\right)^T w^{(k)} + b^{(k)} \qquad \text{رابطه (۲)}$$

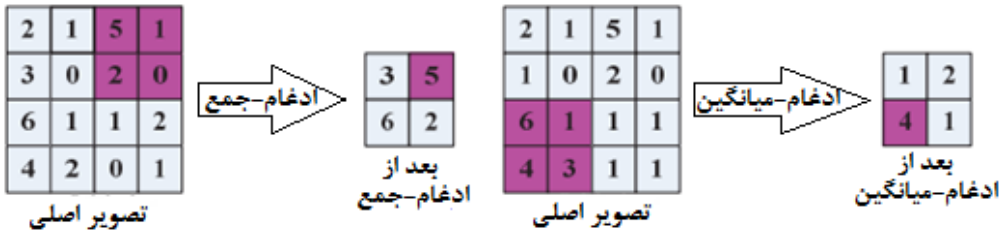
در بحث طبقه‌بندی با *CNN* تعداد نورون‌های آخرین لایه کاملاً متصل به اندازه تعداد کلاس‌ها انتخاب می‌شوند و با استفاده از تابع حداکثر نرم_کنندگی که از رابطه (۳) محاسبه می‌شود[۲۰]، یک خروجی احتمالی که مشخص‌کننده احتمال تعلق نمونه به هرکدام از کلاس‌ها است تولید می‌شود. نمونه به کلاسی تعلق دارد که بیشترین احتمال را داشته باشد.

$$y = \frac{1}{\sum_{K=1}^{M} e^{W_{L,K}^T X_L + b_{L,K}}} \begin{bmatrix} e^{W_{L,1}^T X_L + b_{L,1}} \\ \vdots \\ e^{W_{L,M}^T X_L + b_{L,M}} \end{bmatrix} \qquad \text{رابطه (۳)}$$

در رابطه(۳)، M تعداد کلاس‌ها و عبارت مخرج کسر $\sum_{K=1}^{M} e^{W_{L,K}^T X_L + b_{L,K}}$ فقط برای نرمال‌سازی تابع *softmax* است.

[1] *Dominant*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]



شکل ۳: نحوه عملکرد ادغام-جمع و ادغام-میانگین [۳۴]

۲-۲- روش‌های تولید ویژگی مکانی

در این تحقیق برای تولید ویژگی‌های مکانی سطح پایین از سه روش پروفایل مورفولوژی، فیلتر گابور و الگویی دوتایی محلی استفاده می‌شود که در ادامه به بررسی این سه روش می‌پردازیم.

۲-۲-۱- پروفایل‌های مورفولوژی

مورفولوژی یک تکنیک پردازشی غیرخطی برای آنالیز روابط مکانی بین پیکسل‌ها و ابزاری برای استخراج اجزا تصویر است. مورفولوژی در ابتدا برای تصاویر باینری تعریف شد و سپس در مورد تصاویر با درجات خاکستری و همچنین تصاویر رنگی و ابرطیفی توسعه پیدا کرد [۲۶]. اپراتورهای پایه در مورفولوژی سایش[1] و اتساع[2] هستند[۲۷]. اتساع، پیکسل‌های را به مرز عوارض در تصویر اضافه می‌کند، در حالی که سایش پیکسل‌هایی را از مرز عوارض حذف می‌کند. تعداد پیکسل‌های حذف یا اضافه شده، بستگی به سایز و شکل المان ساختاری استفاده شده در پردازش تصویر دارد. ترکیب اتساع و سایش دو اپراتور باز و بسته را تولید می‌کند. اپراتور باز یک سایش است که با اتساع دنبال می‌شود و اپراتور بسته، یک اتساع است که با یک سایش دنبال می‌شود.

به‌منظور توسعه اپراتورهای مورفولوژی به تصاویر رنگی و ابر طیفی چندین متدولوژی ارائه شده است. پروفایل مورفولوژی یکی از این روش‌ها می‌باشد که در این روش ابتدا مولفه‌های اصلی[3] PCA استخراج می‌شوند، سپس روی هریک از مولفه‌ها که تصاویر خاکستری است اپراتورهای باز و بسته را با سایزهای مختلف المان ساختاری اعمال می‌کنیم. هر مولفه اصلی به همراه تصاویر حاصل از n اپراتور باز و n اپراتور بسته ، برداری $2n+1$ بعدی تشکیل می‌دهند که آن را پروفایل مورفولوژی می‌نامند و از رابطه (۴) محاسبه می‌شود[۳۳].

رابطه(۴)

$$MP(x)=\left[CP_n(x),\ldots,I(x),\ldots,OP_n(x)\right]$$

در رابطه (۴)، $OP_n(x)$ و $CP_n(x)$ به ترتیب پروفایل اپراتوز باز و بسته در پیکسل x و n تعداد تکرار اپراتورهای باز و بسته به ازای هریک از المان‌هاس ساختاری است. مجموعه‌ی پروفایل‌های مورفولوژی حاصل از همه مولفه‌های اصلی، پروفایل مورفولوژی گسترده شده[4] را تشکیل می‌دهند و از رابطه (۵) محاسبه می‌شود[۳۳]:

رابطه(۵)

$$EMP(x)=\left\{MP_{PC^1}(x),\ldots,MP_{PC^m}(x)\right\}$$

در رابطه(۵)، m تعداد ویژگی‌های منتخب اولیه و منظور از MP، پروفایل مورفولوژی هریک از ویژگی‌های منتخب اولیه است. خصوصیات مکانی پروفایل‌های مورفولوژی وابسته به المان ساختاری مورد استفاده

[1] *Erosion*
[2] *Dilation*
[3] *Principal component (PC)*
[4] *Expanded Morphological Profiles*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

برای ایجاد این پروفایل‌ها است. برای مثال اگر یک خط با n پیکسل در جهتی خاص به عنوان المان ساختاری استفاده شود، این المان می‌تواند ساختارهای خطی در آن جهت را حفظ کند. از نقطه نظر طبقه‌بندی، استفاده از یک المان ساختاری همه‌جانبه (ایزوتروپیک) اطلاعات بیشتری تولید می‌کند به همین دلیل در این تحقیق از المان ساختاری دیسکی که یک المان ساختاری همه‌جانبه است استفاده می‌شود.

۲-۲-۲- گابور

فیلترهای گابور در زمینه‌های مختلف پردازش تصویر از قبیل طبقه‌بندی بافت، تشخیص لبه، شناسایی اثر انگشت و کدگذاری تصاویر استفاده شده‌اند[۲۸]. فیلتر گابور به دلیل خاصیت تفکیک‌پذیری خوب و چندگانه در حوزه مکان و فرکانس برای استخراج ویژگی‌های بافت از تصاویر مورد استفاده قرار می‌گیرد. ایده فیلترهای گابور براساس استخراج ساختارهای با جهت خطی در یک تصویر است[۲۹]. فیلتر گابور در واقع همان فیلتر گوسی است که به یک فرکانس مرکزی (U,V) انتقال یافته است. موجک‌های گابور یک سری فیلترها هستند که از یک فیلتر مادر ساخته می‌شوند. حال اگر این فیلتر مادر را به صورت رابطه (۶) در نظر بگیریم[۲۸]، از این فیلتر مادر برای تولید یک سری فیلترهای دیگر $h_{s,d}(x,y)$ در مقیاس‌ها و جهت‌های مختلف استفاده می‌شود (رابطه(۷) و (۸)).

رابطه (۶)
$$\varphi(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jU_h x\right\}$$

رابطه (۷)
$$h_{s,d}(x,y)=a^{-s}\varphi(X,Y)$$

رابطه (۸)
$$a=\left(\frac{U_h}{U_1}\right)^{\frac{1}{N_S-1}}$$

رابطه (۹)
$$\sigma_x=\frac{(a+1)\sqrt{ln4}}{2\pi(a-1)U_h}$$

رابطه (۱۰)
$$\sigma_y=\frac{\sqrt{ln4-\left(\frac{ln4}{2\pi U_h\sigma_x}\right)^2}}{2\pi tan\left(\frac{\pi}{2N_d}\right)\left[U_h-\frac{ln4}{U_h(2\pi\sigma_x)^2}\right]}$$

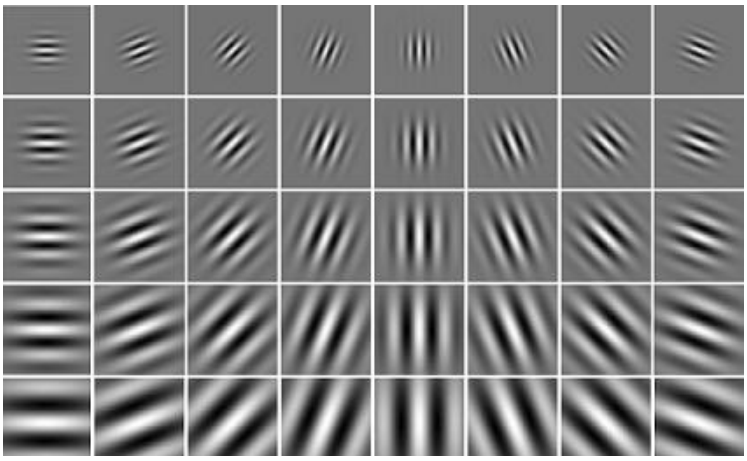
رابطه (۱۱)
$$X=a^{-s}\left[(x-x_.)\cos\left(\frac{\pi d}{N_d}\right)+(y-y_.)\sin\left(\frac{\pi d}{N_d}\right)\right]$$

رابطه (۱۲)
$$Y=a^{-s}\left[-(x-x_.)\sin\left(\frac{\pi d}{N_d}\right)+(y-y_.)\cos\left(\frac{\pi d}{N_d}\right)\right]$$

در رابطه‌های بالا $s\in\{1,\dots N_s\}$ و $d\in\{1,\dots N_d\}$ شاخص‌های مقیاس و جهت هستند، U_h و U_1 فرکانس‌های مرکزی حداقل و حداکثر موردنظر و مرکز فیلتر مکانی در مختصات $x_.$ و $y_.$ می‌باشد. ضریب مقیاس

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

a^{-s} نیز برای مستقل شدن انرژی فیلتر از مقیاس N_S است. در شکل (۴) فیلترهای گابور ایجاد شده برای پنج مقیاس و چهار جهت در حوزه مکان نشان داده شده است.



شکل ۴: فیلترهای گابور ایجاد شده برای پنج مقیاس و هشت جهت[۲۸]

اعمال فیلترهای گابور بر روی یک تصویر تک باند، برای هر پیکسل $\mathrm{N_d} \times \mathrm{N_s}$ ویژگی تولید می‌کند و از آنجا که مقادیر خروجی این فیلترها مختلط است، معمولاً از مقدار بزرگی[1] آنها استفاده می‌شود. علاوه بر این، برای کاهش واریانس درون کلاسی و در نتیجه کاهش اشتباهات طبقه‌بندی، فیلتر پایین گذر گاوسی به مقادیر بزرگ اعمال می‌شود[۲۸].

۲-۲-۳- الگویی دوتایی محلی

یک مشکل عمده در استخراج بافت از تصاویر این است که بافت‌های موجود در دنیای واقعی اغلب از نظر تغییرات در جهت‌گیری، مقیاس و شکل ظاهری یکنواخت نیستند[۳۱]. تغییر مقیاس خاکستری در مقابل نوردهی نامنظم نیز یک بحث مهم دیگر است[۳۱]. علاوه بر این، میزان پیچیدگی محاسباتی بسیاری از روش‌های بافت پیشنهاد شده بسیار بالا است.

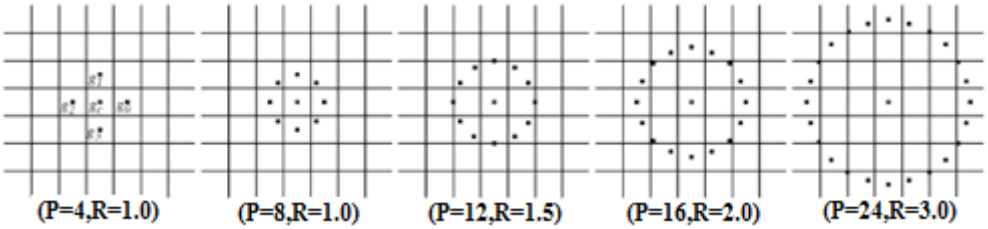
روش *LBP* که برای استخراج بافت توسط اوجلا و همکاران (۲۰۰۲) معرفی شد، یک روش استخراج ویژگی غیر پارامتریک و بسیار قدرتمند است که به علت سادگی در محاسبات و قدرت تفکیک‌پذیری بالا، بسیار مورد توجه است[۳۰]. در این رویکرد از توزیع مقادیر خاکستری پیکسل‌های همسایه که در یک محدوده دایره‌ای از پیکسل مرکزی قرار گرفتند، یک پارامتر محاسبه می‌شود که در مقابل تغییرات مقیاس خاکستری ثابت است. حالت مستقل از دوران نیز با ایجاد یک مجموعه ثابت از الگوها که در مقابل دوران ثابت هستند و اپراتور *LBP* قابلیت تشخیص این الگوها را در تصویر دارد ایجاد می‌شود[۳۱].

برای محاسبه اپراتور *LBP* یک پیکسل با مقدار درجه خاکستری g_C، ابتدا باید تعداد پیکسل‌های همسایه و شعاع دایره‌ای که پیکسل‌های همسایه بر روی آن قرار می‌گیرند تعیین شود. بعد از تعیین دو پارامتر P و R با در نظر گرفتن مختصات (۰،۰) برای پیکسل مرکزی مختصات دیگر پیکسل‌های همسایه از طریق رابطه(۱۳) تعیین می‌شود. در شکل (۵) نحوه قرارگیری پیکسل-های همسایه برای حالت‌های مختلف نشان داده شده است.

$$\left(-\mathrm{R}\sin\left(2\pi p/P\right),\ R\cos\left(2\pi p/P\right)\right) \qquad \text{رابطه(۱۳)}$$

[1] *Magnitude*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]



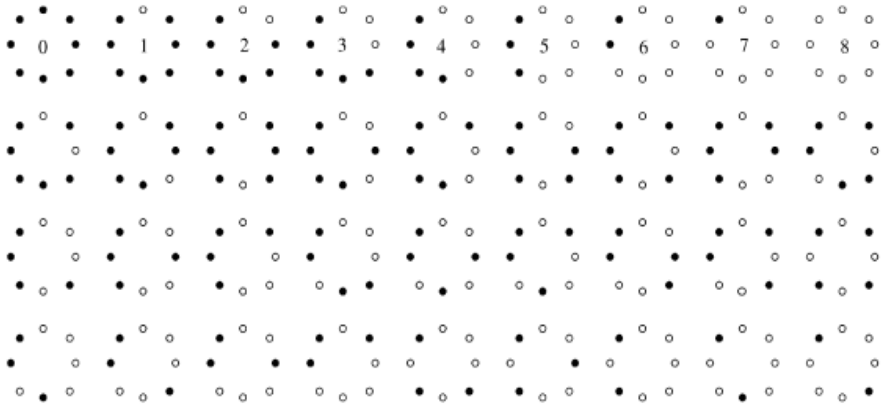
شکل ۵: نحوه قرارگیری پیکسل‌های همسایه در روش *LBP* برای حالت‌های مختلف *(P,R)* [۳۰]

بعد از تعیین موقعیت پیکسل‌های همسایه برای ایجاد حالت تغییرناپذیری در مقابل مقیاس خاکستری، مقدار درجه خاکستری پیکسل‌های همسایه $g_p\left(p=۰,\ldots,P-۱\right)$ از پیکسل مرکزی کم می‌شود و براساس اختلاف به دست آمده یک برچسب صفر و یک برای هرکدام از پیکسل‌های همسایه در نظر گرفته می‌شود. اپراتور *LBP* که در مقابل مقیاس خاکستری ثابت هست به صورت رابطه(۱۴) تعریف می‌شود[۳۰]:

$$LBP_{P,R}=\sum_{p=۰}^{P-۱}S\left(g_p-g_C\right)۲^p$$ رابطه(۱۴)

که:

$$s(x)=\begin{cases}۱, & x\geq ۰\\ ۰, & x<۰\end{cases}$$ رابطه(۱۵)



شکل ۶: ۳۶ الگوی تغییرناپذیر در مقابل دوران. دایره‌های سیاه و سفید به ترتیب نشان‌دهنده برچسب صفر و یک می‌باشند[۳۰].

برای تغییرناپذیری در مقابل دوران نیز از یک سری الگوها که در شکل (۶) نشان داده شده‌اند، استفاده می‌شود. مشاهدات نشان داده که این الگوها بیش از ۹۰ درصد الگوهایی بافت موجود در تصاویر را شامل می‌شوند[۳۰]. این الگوهای بنیادی که در سطر اول شکل(۶) نشان داده شده‌اند تحت عنوان الگوهای

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

یکنواخت (یونیفرم) معرفی شدند زیرا همه آنها دارای ساختار دایره‌ای یکنواخت با کمترین انتقال مکانی هستند[۳۰]. براساس این ویژگی مشترک الگوهای یکنواخت، یک معیار یکنواختی U معرفی شده است که تعداد انتقال‌های مکانی از صفر به یک و بلعکس را برای الگوهای به دست آمده از رابطه (۱۴) محاسبه می‌کند[۳۰]. به عنوان مثال برای دو الگو "۰۰۰۰۰۰۰۰" و "۱۱۱۱۱۱۱۱" مقدار U برابر صفر است همچنین مقدار U برای هفت الگو دیگر سطر اول کمتر یا مساوی ۲ است، اما در ۲۷ حالت دیگر مقدار U حداقل چهار است. بر این اساس، الگوهای از تصویر که مقدار U برای آنها کمتر یا مساوی دو باشد الگوهای یکنواخت در نظر گرفته می‌شوند.

اپراتور *LBP* که در مقابل مقیاس و دوران تغییرناپذیر باشد به صورت رابطه(۱۶) تعریف می‌شود[۳۰]. بالانویس $riu۲$ مربوط به استفاده از روش الگوهای یکنواخت برای ایجاد تغییرناپذیری در مقابل دوران است.

رابطه (۱۶)

$$LBP_{P,R}^{riu۲}=\begin{cases}\sum_{p=۰}^{P-۱} s(g_p-g_c) & U(LBP_{P,R})\le ۲\\ P+۱ & Otherwise\end{cases}$$

که:

رابطه (۱۷)

$$U\left(LBP_{P,R}\right)=\left|s\left(g_{P-۱}-g_C\right)-s\left(g_{۰}-g_C\right)\right|+\sum_{p=۱}^{P-۱}\left|s\left(g_p-g_C\right)-s\left(g_{p-۱}-g_C\right)\right|$$

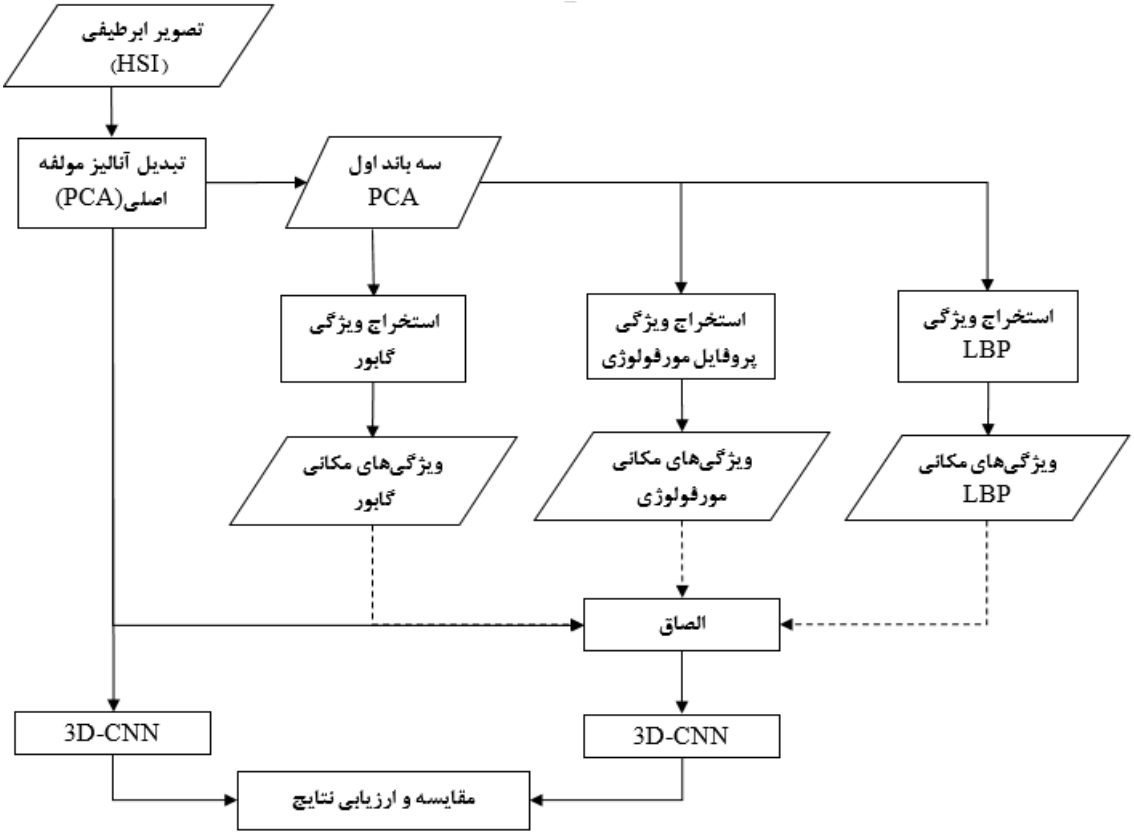
۳- روش تحقیق

فلوچارت کلی تحقیق در شکل (۷) نشان داده شده است که در ادامه به توضیح هرکدام می‌پردازیم.

در تصاویر ابرطیفی با افزایش باندها تا یک تعداد مشخصی دقت تفکیک‌پذیری کلاس‌ها افزایش پیدا می‌کند، ولی از یک حدی به بعد به علت محدودیت نمونه‌های آموزشی و کاهش دقت برآورد پارامترهای آماری دقت طبقه‌بندی کاهش پیدا می‌کند به همین دلیل از روش‌های استخراج ویژگی برای کاهش بعد تصاویر ابرطیفی استفاده می‌شود. *PCA* که یک روش نظارت‌نشده می‌باشد یکی از ساده‌ترین و در عین حال مهم‌ترین روش‌های استخراج ویژگی است. در این مقاله برای استخراج ویژگی‌های طیفی از *PCA* استفاده شده است. پس از اعمال تبدیل *PCA* از *N* مولفه اول که دارای ۹۹٫۹ درصد واریانس هستند به عنوان ویژگی‌های طیفی استفاده شده است. در این مقاله برای تولید ویژگی‌ها مکانی سطح پایین از سه روش مورفولوژی، گابور و *LBP* استفاده می‌شود. از آنجا که تولید این سه ویژگی مکانی از تمام باندهای *PCA* باعث افزایش بیش از حد تعداد باندها می‌شود و مشکلاتی نظیر زمان‌بر شدن آموزش شبکه می‌شود، به همین دلیل از سه مولفه اول *PCA* برای استخراج ویژگی‌های مکانی استفاده می‌شود. در شکل(۸) میزان وابستگی بین ویژگی‌های طیفی و مکانی استخراج شده برای دو داده ابرطیفی نشان داده شده است. در این شکل همانطور که مشاهده می‌شود برای مناطق خارج از محدوده‌های مشخص شده که میزان وابستگی بین ویژگی‌های یک روش با روش‌های دیگری را نشان می‌دهد میزان وابستگی بسیار کم است در نتیجه ویژگی‌های طیفی و مکانی تولید شده از هم متمایز می‌باشند؛ اما برای چهار محدوده مشخص شده که به ترتیب میزان وابستگی بین خود ویژگی‌های تولید شده از روش‌های *PCA*، مورفولوژی، گابور و *LBP* را نشان می‌دهد در بین ویژگی‌های *LBP* میزان وابستگی بالا است و از آنجا که روش کاهش ویژگی *PCA* بر روی ویژگی‌های *LBP* باعث کاهش دقت طبقه‌بندی شد در تحقیق از کل

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

ویژگی‌های *LBP* برای طبقه‌بندی استفاده شد. پس از تولید ویژگی‌های مکانی، این ویژگی‌ها در ترکیبات گوناگون کنار ویژگی‌های مستخرج از تبدیل *PCA* قرار می‌گیرند و وارد شبکه *3D-CNN* می‌شوند. خط‌چین‌ها به این معنی است که هریک از ویژگی‌های مکانی هم می‌توانند در بردار ادغام حضور داشته باشند هم خیر. در مرحله دیگر پیاده‌سازی ویژگی‌های طیفی به تنهایی وارد شبکه *3D-CNN* می‌شوند و نتایج این قسمت با حالت‌های متفاوت طیفی_مکانی مقایسه می‌شود.

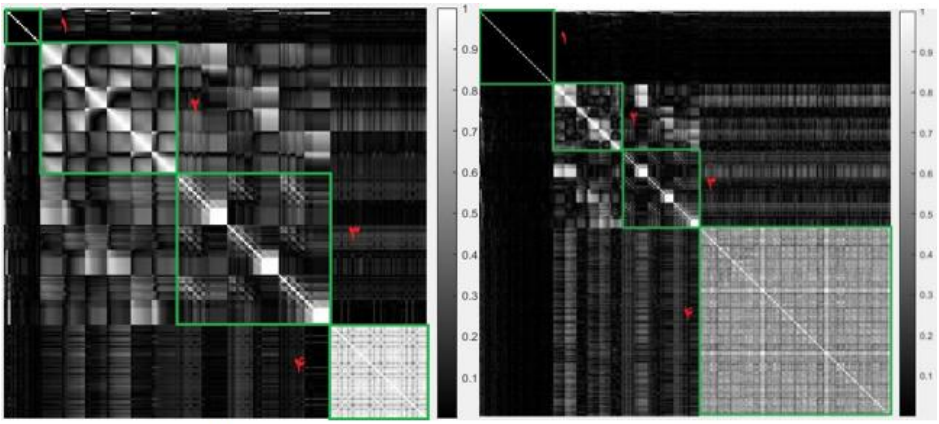


شکل ۷: فلوچارت کلی تحقیق

یک موضوع مهم در استفاده از شبکه‌های عصبی معماری شبکه مورد استفاده است. در طراحی معماری *CNN* سه پارامتر مهم عبارت‌اند از: ۱) عمق شبکه: منظور تعداد لایه‌های کانولوشن مورد استفاده در معماری شبکه می‌باشد. ۲) تعداد نقشه‌های ویژگی[1] لایه کانولوشن ۳) اندازه داده ورودی. در این تحقیق باتوجه به آزمایش‌های انجام شده برای تعیین این سه پارامتر، از ساختار نشان داده شده در شکل(۹) استفاده می‌کنیم. طبق شکل(۹) شبکه مورد استفاده دارای سه لایه کانولوشن، لایه نرمال‌سازی و تابع فعال‌ساز *ReLU*، دو لایه ادغام با عملگر ماکزیمم و یک لایه کاملاً متصل می‌باشد.

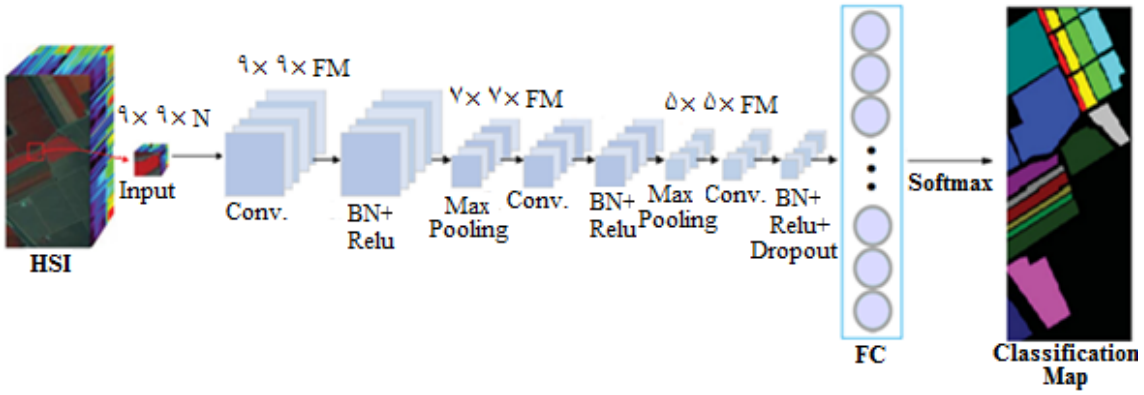
[1] *Feature maps*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]



الف ب

شکل ۸: میزان وابستگی بین ویژگی‌های طیفی و مکانی استخراج شده برای: الف) دانشگاه پاویا ب) ایندین پاین. محدوده‌های مشخص شده به ترتیب مربوط به ویژگی‌های *PCA*، مورفولوژی، گابور و *LBP* است.



شکل ۹: ساختار *CNN* مورد استفاده در این تحقیق

در این معماری داده ورودی به شبکه دارای ابعاد $N \times ۹ \times ۹$ است که N تعداد باندهای تصویر ورودی می‌باشد. ابعاد کرنل مورد استفاده در تمامی لایه‌های کانولوشن و ادغام، ۳ × ۳ است. در لایه کانولوشن برای جلوگیری از دست رفتن لبه‌ها از تکنیک Padding استفاده شد. در این تکنیک یک سری سطر و ستون صفر متناسب با اندازه کرنل مورد استفاده به مرز تصویر اضافه می‌شود، اما در لایه‌های ادغام از تکنیک Padding استفاده نشد در نتیجه سایز داده‌های خروجی پس از اعمال کرنل لایه ادغام از رابطه(۱۸) محاسبه می‌شود[۴۱].در رابطه (۱۸)، stride که نشان‌دهنده گام حرکتی کرنل می‌باشد در هر دو جهت برابر یک درنظر گرفته شد. p و q سایز کرنل مورد استفاده برای لایه ادغام، m و n هم اندازه داده ورودی به لایه ادغام می‌باشند. در شکل(۱۰)، FM تعداد نقشه ویژگی‌های تولید شده به وسیله لایه کانولوشن می‌باشد که از سه برابر تعداد باندهای ورودی به عنوان تعداد نقشه‌های ویژگی در لایه‌های کانولوشن استفاده شده است. همچنین تعداد نورون‌های موجود در لایه کاملاً متصل را برابر تعداد کلاس‌ها در نظر گرفتیم.

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

رابطه (۱۸)
$$\left(\frac{\left(m-p+۲\times Padding\right)}{Stride}+۱\right)*\left(\frac{\left(n-q+۲\times Padding\right)}{Stride}+۱\right)$$

۴- نتایج تجربی و بحث

در این قسمت ابتدا داده‌های مورد استفاده در این تحقیق معرفی شده و در ادامه نتایج پیاده‌سازی ارائه می‌شود. در بحش پایانی نیز به مقایسه نتایج این تحقیق با چند تحقیق جدید و پیشرفته پرداخته می‌شود.

۴-۱- داده‌های مورد استفاده

اولین داده‌ی مورد استفاده در این پژوهش تصویر اخذ شده توسط سنجنده‌ی هوابرد *AVIRIS* از منطقه ایندین پاینس در شمال غربی منطقه ایندیانا در ایالات متحده آمریکا در تاریخ ۱۲ ژوئن ۱۹۹۲ است. این تصویر با حد تفکیک مکانی ۲۰ متر دارای ابعاد ۱۴۵ × ۱۴۵پیکسل و ۲۲۴ باند طیفی در محدوده ۰/۴ تا ۲/۵ میکرومتر با پهنای باند ۱۰ نانومتر است. حدود دوسوم این تصویر را مناطق کشاورزی و یک‌سوم آن را جنگل و گیاه در برگرفته است، همچنین در این تصویر نواحی ساختمانی، خط ریل و دو اتوبان دو لایه و جاده‌های کوچکی نیز به چشم می‌خورد. با حذف ۲۴ باند در محدوده‌های جذبی بخار آب از بقیه ۲۰۰ باند باقی‌مانده در پردازش‌های این پژوهش استفاده شده است. این منطقه شامل شانزده کلاس از گیاهانی نظیر ذرت، سویا، علف، گندم و غیره است. در این تصویر گونه‌های مختلف گیاهی با امضاهای طیفی نزدیک به هم از کلاس‌های ذرت و سویا وجود دارد که کار طبقه‌بندی را مشکل می‌کند. این تصویر به طور کلی ابعاد بزرگی ندارد و فاصله مکانی پوشش‌های مختلف زمینی بسیار کم است. علاوه بر این در برخی کلاس‌ها مانند کلاس یونجه[1] تعداد پیکسل‌ها بسیار کم و در برخی کلاس‌ها مانند کلاس‌های سویا تعداد پیکسل‌ها بسیار زیاد است. با توجه به نکاتی که ذکر شد طبقه‌بندی این داده در سنجش از دور همیشه چالش‌برانگیز بوده است و علت انتخاب داده ایندین در این تحقیق بررسی کارایی *CNN* در افزایش دقت طبقه‌بندی این داده است. یک ترکیب رنگی کاذب و نقشه حقیقت زمینی این داده که توسط آزمایشگاه سنجش از دور دانشگاه پردو تهیه شده است در شکل (۱۰) نشان داده شده است.

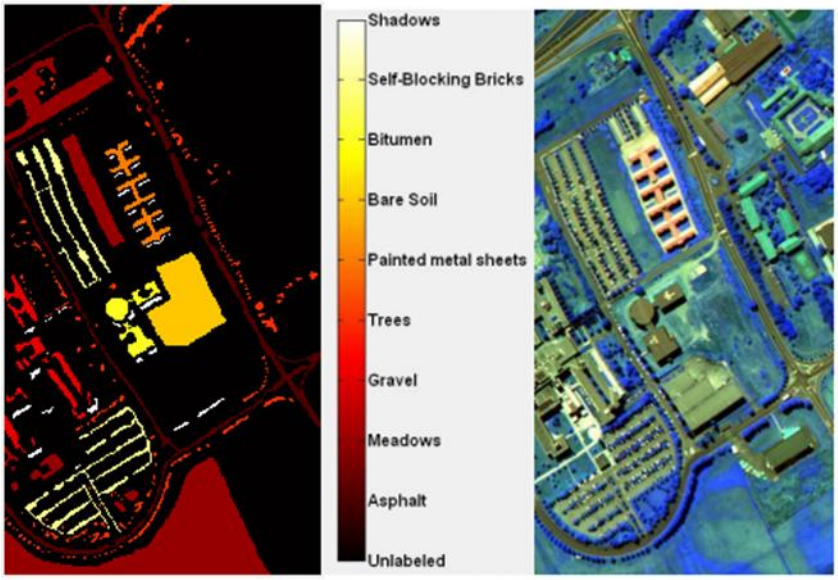
تصویر دوم توسط سنجنده *Rosis 3* از شمال ایتالیا و دانشگاه پاویا در ابعاد ۳۴۰ × ۶۱۰ پیکسل اخذ شده است. سنجنده‌ای مذکور دارای ۱۱۵ باند طیفی در محدوده ۴۳۰ تا ۸۶۰ نانومتر است. پس از حذف ۱۲ باند نویزی از ۱۰۳ باند باقی‌مانده در این تحقیق استفاده شده است. حد تفکیک مکانی این تصویر ۱/۳ متر و دارای ۹ کلاس اطلاعاتی مانند آسفالت، سایه، خاک لخت و غیره می‌باشد. طبقه‌بندی این تصویر به علت وجود کلاس‌های پیچیده شهری و نواحی تودرتوی مکانی چالش‌برانگیز است. تصویر با ترکیب رنگی کاذب و نقشه حقیقت زمینی این داده که نشان‌دهنده پراکندگی کلاس‌ها هست در شکل (۱۱) نشان داده شده است.

[1] *Alfalfa*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]



شکل ۱۰: تصویر رنگی کاذب و نقشه حقیقت زمینی داده ایندین پاینس



شکل ۱۱: تصویر تصویر رنگی کاذب و نقشه حقیقت زمینی داده دانشگاه پاویا

۴-۲- ملاحظات پیاده‌سازی

کلیه پیاده‌سازی‌های این مقاله در نرم‌افزار متلب *2018b* در سیستم کامپیوتری با مشخصات سی پی یو *Intel® Core™ i5* و رم ۸ گیگابایت انجام شد. برای بخش *CNN* از کدهای موجود در تولباکس *Deep-learning* متلب استفاده شده است. به علت تاثیر تعداد نمونه‌های آموزشی در نتایج شبکه‌های عصبی در این مقاله ۵٪ و ۱۰٪ حقیقت زمینی برای تصویر ایندین، ۱٪ و ۵٪ حقیقت زمینی هم برای تصویر پاویا به طور تصادفی به عنوان نمونه آموزشی هر کلاس انتخاب شدند و الباقی پیکسل‌های حقیقت زمینی به عنوان نمونه‌های تست برای ارزیابی نهایی دقت طبقه‌بندی انتخاب شدند. از دو

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

پارامتر آماری دقت کلی[1] (*OA*) و ضریب کاپا[2] (K) که از ماتریس ابهام استخراج می‌شوند برای ارزیابی عملکرد شبکه استفاده شده است (به منظور بررسی روابط این روش‌ها رجوع شود به[۳۵]).

در این مقاله ورودی شبکه *CNN*، بردارهای طیفی_مکانی می‌باشد که برای استخراج این ویژگی‌های طیفی و مکانی از پارامترهای زیر استفاده می‌شود.

۱) ویژگی‌های طیفی: همانطور که پیش‌تر عنوان شد به دلیل وجود باندهای همبسته و اضافی در تصاویر ابرطیفی، از تبدیل *PCA* برای کاهش ابعاد استفاده شد و n مولفه اول آن، که شامل ۹۹٫۹ درصد واریانس تجمعی بودند به عنوان ویژگی‌های طیفی انتخاب شدند. در تصویر ایندین پاینس ابعاد ویژگی طیفی *PCA*، ۶۹ باند و در تصویر دانشگاه پاویا ۱۵ باند می‌باشد.

۲) پروفایل‌های مورفولوژی: برای تولید پروفایل‌ها در هر دو داده از عملگرهای مورفولوژی باز کردن و بستن با بازسازی المان ساختاری دیسک شکل با اندازه‌های $r=[۳,۶,۹,\ldots,۳۰]$ استفاده می‌شود. این مقادیر بر اساس تجربه و خطا تعیین شده است. تعداد ویژگی‌های تولید شده در این روش برای سه باند اول *PCA* هر دو داده، ۶۳ باند است.

۳) ویژگی‌های گابور: به منظور تولید ویژگی‌های گابور با توجه به تحقیقات انجام شده در [۲۷]، از $N_d = ۴$، $N_s = ۶$، $U_h = ۰٫۴۹$ و $U_1 = ۰٫۰۱$ استفاده می‌شود. تعداد ویژگی‌های گابور تولید شده برای سه باند اول *PCA* هر دو داده، ۷۲ باند است.

۴) ویژگی‌های *LBP*: در روش استخراج ویژگی مکانی *LBP* ما از حالت یکنواخت استفاده می‌کنیم. در روش *LBP*، شعاع و تعداد پیکسل‌های همسایه دو پارامتر مهم می‌باشند. این پارامترها براساس تجربه و خطا تعیین شدند که بر این اساس برای داده ایندین از شعاع یک و هشت پیکسل همسایه و برای داده دانشگاه پاویا نیز از شعاع یک و چهار پیکسل همسایه استفاده می‌شود. ابعاد ویژگی‌های مکانی استخراج شده از روش *LBP* به ازای سه باند اول *PCA* برای داده ایندین پاینس ۱۷۷ باند و برای داده پاویا ۴۵ باند است.

۴-۳- آنالیز نتایج

همانطور که اشاره شد هدف از این مقاله بررسی قابلیت *CNN* پیشنهاد شده در طبقه‌بندی تصاویر ابرطیفی و همچنین بررسی تاثیر ویژگی‌های مکانی سطح پایین استخراج شده از چندین روش تولید ویژگی مکانی در نتایج *CNN* پیشنهاد شده می‌باشد. به همین منظور، پس از تعیین معماری شبکه و استخراج ویژگی‌های مکانی، این ویژگی‌ها در ترکیبات مختلف کنار ویژگی طیفی مستخرج از تبدیل *PCA* قرار می‌گیرند و بردار طیفی_مکانی حاصل وارد *CNN* می‌شود. نتایج حاصل از طبقه‌بندی برای داده ایندین پاین و دانشگاه پاویا در جداول (۱) و (۲) نشان داده شده است.

[1] *Overall accuracy (OA)*
[2] *Kappa coefficient*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

جدول ۱: دقت طبقه‌بندی برای داده ایندین پاین

داده آموزشی	ویژگی‌های ورودی	SVM		KNN		CNN	
		OA	K	OA	K	OA	K
٪۵	LDA	۷۲٫۹۹	۶۹٫۳۴	۷۳٫۶۵	۶۹٫۴۰	**۹۴٫۶۳**	**۹۳٫۹۲**
	PCA	۷۳٫۷۰	۶۹٫۸۲	۶۱٫۹۶	۵۵٫۳۶	**۹۳٫۶۵**	**۹۲٫۷۸**
	PCA + MP	۹۵٫۶۳	۹۵٫۱۲	۸۹٫۰۹	۸۷٫۵۵	**۹۵٫۷۸**	**۹۵٫۱۹**
	PCA + Gabor	۷۳٫۷۲	۶۹٫۸۲	۶۱٫۹۶	۵۵٫۳۶	**۹۳٫۰۹**	**۹۲٫۱۵**
	PCA + LBP	۷۴٫۶۷	۷۰٫۹۴	۶۴٫۰۴	۵۵٫۴۵	**۹۶٫۱۶**	**۹۵٫۶۳**
	PCA + MP + Gabor	**۹۵٫۶۳**	**۹۵٫۱۲**	۸۹٫۰۹	۸۷٫۵۵	۹۴٫۳۸	۹۳٫۵۹
	PCA + MP + LBP	۹۵٫۶۳	۹۵٫۱۲	۸۹٫۰۹	۸۷٫۵۵	**۹۶٫۳۷**	**۹۵٫۸۷**
	PCA + Gabor + LBP	۷۴٫۶۷	۷۰٫۹۴	۶۴٫۰۴	۵۵٫۴۵	**۹۶٫۳۶**	**۹۵٫۸۶**
	PCA + MP + Gabor + LBP	۹۵٫۶۳	۹۵٫۱۲	۸۹٫۰۹	۸۷٫۵۵	**۹۶٫۸۹**	**۹۶٫۴۵**
٪۱۰	LDA	۷۳٫۷۰	۶۹٫۸۲	۷۴٫۷۴	۷۰٫۸۳	**۹۸٫۳۷**	**۹۸٫۱۵**
	PCA	۷۳٫۸۹	۷۰٫۰۴	۶۱٫۹۶	۵۵٫۳۶	**۹۸٫۲۴**	**۹۸٫۰۰**
	PCA + MP	۹۸٫۰۵	۹۷٫۷۸	۹۲٫۵۰	۹۲٫۵۷	**۹۸٫۶۱**	**۹۸٫۴۱**
	PCA + Gabor	۸۰٫۶۳	۷۷٫۹۲	۶۶٫۵۵	۶۱٫۳۸	**۹۷٫۹۵**	**۹۷٫۶۶**
	PCA + LBP	۸۲٫۹۰	۸۰٫۵۱	۶۶٫۶۲	۶۱٫۴۵	**۹۸٫۴۴**	**۹۸٫۲۲**
	PCA + MP + Gabor	۹۸٫۰۵	۹۷٫۷۸	۹۳٫۵۰	۹۲٫۵۷	**۹۸٫۵۸**	**۹۸٫۳۸**
	PCA + MP + LBP	۹۸٫۰۵	۹۷٫۷۸	۹۳٫۵۰	۹۲٫۵۷	**۹۸٫۹۳**	**۹۸٫۷۸**
	PCA + Gabor + LBP	۸۲٫۹۰	۸۰٫۵۱	۶۶٫۶۲	۶۱٫۴۵	**۹۸٫۹۵**	**۹۸٫۸۰**
	PCA + MP + Gabor + LBP	۹۸٫۰۶	۹۷٫۷۹	۹۳٫۵۰	۹۲٫۵۷	**۹۹٫۱۹**	**۹۹٫۰۷**

جدول۲ : دقت طبقه‌بندی برای داده دانشگاه پاویا

داده آموزشی	ویژگی‌های ورودی	SVM		KNN		CNN	
		OA	K	OA	K	OA	K
٪۱	LDA	۸۵٫۸۲	۸۰٫۹۵	۸۴٫۵۵	۷۸٫۹۳	**۹۲٫۱۳**	**۸۹٫۴۸**
	PCA	۸۹٫۷۱	۸۶٫۱۵	۷۹٫۳۹	۷۱٫۴۶	**۹۵٫۵۸**	**۹۴٫۱۲**
	PCA + MP	۹۷٫۳۴	۹۶٫۶۹	۹۴٫۳۲	۹۲٫۴۵	**۹۷٫۶۶**	**۹۶٫۹۱**
	PCA + Gabor	۸۹٫۷۱	۸۶٫۱۵	۷۹٫۳۹	۷۱٫۴۶	**۹۳٫۹۹**	**۹۲٫۰۴**
	PCA + LBP	۹۰٫۴۰	۸۷٫۱۰	۷۹٫۴۳	۷۱٫۵۳	**۹۵٫۱۷**	**۹۳٫۶۰**
	PCA + MP + Gabor	**۹۷٫۳۴**	**۹۶٫۶۹**	۹۴٫۳۲	۹۲٫۴۵	۹۷٫۲۳	۹۶٫۳۳
	PCA + MP + LBP	۹۷٫۳۴	۹۶٫۶۹	۹۴٫۳۲	۹۲٫۴۵	**۹۷٫۹۲**	**۹۷٫۲۸**
	PCA + Gabor + LBP	۹۰٫۴۰	۸۷٫۱۰	۷۹٫۴۳	۷۱٫۵۳	**۹۵٫۱۴**	**۹۳٫۶۲**
	PCA + MP + Gabor + LBP	۹۷٫۳۴	۹۶٫۶۹	۹۴٫۳۲	۹۲٫۴۵	**۹۸٫۱۰**	**۹۷٫۴۸**
	LDA	۸۶٫۲۴	۶۹٫۸۲	۸۷٫۲۹	۸۲٫۸۴	**۹۷٫۲۵**	**۹۶٫۳۶**
	PCA	۹۳٫۰۰	۹۰٫۶۶	۸۳٫۷۲	۷۷٫۷۰	**۹۹٫۰۹**	**۹۸٫۰۸**
	PCA + MP	۹۸٫۵۷	۹۸٫۴۳	۹۸٫۰۸	۹۷٫۴۶	**۹۹٫۶۶**	**۹۹٫۵۵**
	PCA + Gabor	۹۳٫۰۰	۹۰٫۶۶	۸۳٫۷۲	۷۷٫۷۰	**۹۹٫۳۰**	**۹۹٫۰۷**

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

٪۵	PCA + LBP	۹۴٫۳۹	۹۲٫۵۲	۸۳٫۹۳	۷۷٫۹۹	**۹۹٫۵۷**	**۹۹٫۴۳**
	PCA + MP + Gabor	۹۸٫۵۷	۹۸٫۴۳	۹۸٫۰۸	۹۷٫۴۶	**۹۹٫۷۲**	**۹۹٫۶۳**
	PCA + MP + LBP	۹۸٫۵۷	۹۸٫۴۳	۹۸٫۰۸	۹۷٫۴۶	**۹۹٫۸۵**	**۹۹٫۸۰**
	PCA + Gabor + LBP	۹۴٫۳۹	۹۲٫۵۲	۸۳٫۹۳	۷۷٫۹۹	**۹۹٫۶۸**	**۹۹٫۵۸**
	PCA + MP + Gabor + LBP	۹۸٫۵۷	۹۸٫۴۳	۹۸٫۰۸	۹۷٫۴۶	**۹۹٫۹۰**	**۹۹٫۸۶**

با توجه به دقت‌های به دست آمده در جداول (۱) و (۲) برای دو داده ایندین پاین و دانشگاه پاویا می‌توان به نتایج زیر رسید:

۱. اولین موضوع مربوط به دقت‌های بالای حاصل شده از *CNN* پیشنهاد شده نسبت به دو طبقه‌بندی‌کننده مرسوم *SVM* و *KNN* می‌باشد که لزوم استفاده از این روش را در طبقه‌بندی تصاویر ابرطیفی نشان می‌دهد. همانطور که از دقت‌های تولید شده در جدول (۱) و (۲) می‌بینیم *CNN* مورد استفاده در همه حالت‌های طیفی و طیفی_مکانی به دقت‌های بسیار خوبی منجر شده است اما دو روش *SVM* و *KNN* بجز حالتی که از ویژگی مکانی مورفولوژی استفاده شده است در بقیه حالت‌ها نسبت به *CNN* پیشنهاد شده، دقت‌های بسیار پایین‌تری تولید کرده‌اند.

۲. نکته دیگر مربوط به بررسی تاثیر تعداد نمونه‌های آموزشی بر نتایج طبقه‌بندی *CNN* است. نتایج تولید شده برای دو حالت نمونه آموزشی متفاوت نشان می‌دهد که روش *CNN* همانند دو روش *SVM* و *KNN* وابسته به تعداد نمونه‌های آموزشی است و با افزایش تعداد نمونه‌های آموزشی به طور قابل توجهی دقت آن افزایش پیدا می‌کند.

۳. یکی از موضوعات اصلی که در این تحقیق مورد بررسی قرار گرفت تاثیر استفاده از اطلاعات مکانی گوناگون در کنار اطلاعات طیفی بود. برای بررسی این موضوع از ویژگی‌های مکانی سطح پایین استخراج شده از سه روش مورفولوژی، گابور و *LBP* در ترکیبات مختلف کنار ویژگی طیفی *PCA* استفاده شد. همانظور که از دقت‌های تولید شده برای حالت طیفی و حالت‌های متنوع طیفی_مکانی می‌بینیم، با افزودن ویژگی‌های مکانی سطح پایین به ویژگی‌های طیفی، دقت کلی طبقه‌بندی با *CNN* پیشنهاد شده بهبود قابل توجهی پیدا می‌کند. در این قسمت برای بررسی بهتر این موضوع از آماره Z که برای مقایسه نتایج طبقه‌بندی بین دو روش می‌باشد استفاده شد. آماره Z به صورت مقابل تعریف می‌شود: $Z_{ij} = \left(n_{ij} - n_{ji}\right) / \left(n_{ij} + n_{ji}\right)^{1/2}$، در این رابطه n_{ij} تعداد نمونه‌های تستی می‌باشد که در روش i درست طبقه‌بندی شده‌اند ولی در j اشتباه طبقه‌بندی شده‌اند و بلعکس. در آمار Z اگر $Z_{ij} >$ 0 باشد نتایج طبقه‌بندی i بهتر از j است و در حالتی که $|Z_{ij}| >$ ۱٫۹۶ اختلاف بین دو روش قابل توجه[1] می‌باشد[۴۳]. مقادیر محاسبه شده برای آماره Z بین دو ویژگی طیفی *PCA* و آنالیز مولفه خطی *(LDA)*[2] با حالت طیفی_مکانی که از کل ویژگی‌های مکانی استخراج شده استفاده می‌شود در جدول (۳) نشان داده شده است. همانطور که از مقادیر محاسبه شده می‌بینیم در هر دو حالت نمونه آموزشی، اختلاف بین نتایج طبقه‌بندی طیفی_مکانی محاسبه شده از روش *CNN* با نتایج طبقه‌بندی *CNN* برای ویژگی‌های طیفی *PCA* و

[1] *Significant*
[2] *Linear Discriminant Analysis*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

LDA بسیار قابل توجه می‌باشد و این موضوع لزوم استفاده از ویژگی‌های مکانی سطح پایین در کنار ویژگی‌های طیفی، در *CNN* را نشان می‌دهد.

۴. نکته دیگر در نتایج جداول (۱) و (۲) این است که در داده ایندین پاین ویژگی مکانی *LBP* و در داده دانشگاه پاویا ویژگی مکانی مورفولوژی ویژگی قوی‌تری برای طبقه‌بندی این داده‌ها با استفاده از *CNN* هستند. ازآنجایی که یکی از علل اصلی افزایش زمان پردازش در طبقه‌بندی *CNN* ابعاد داده ورودی می‌باشد، می‌توان تنها با استفاده از این ویژگی‌ها در کنار ویژگی طیفی *PCA* علاوه بر کاهش چشم‌گیر زمان پردازش به نتایج قابل رقابت با حالتی که از تمام ویژگی‌های مکانی در کنار ویژگی‌های طیفی *PCA* استفاده شود، رسید.

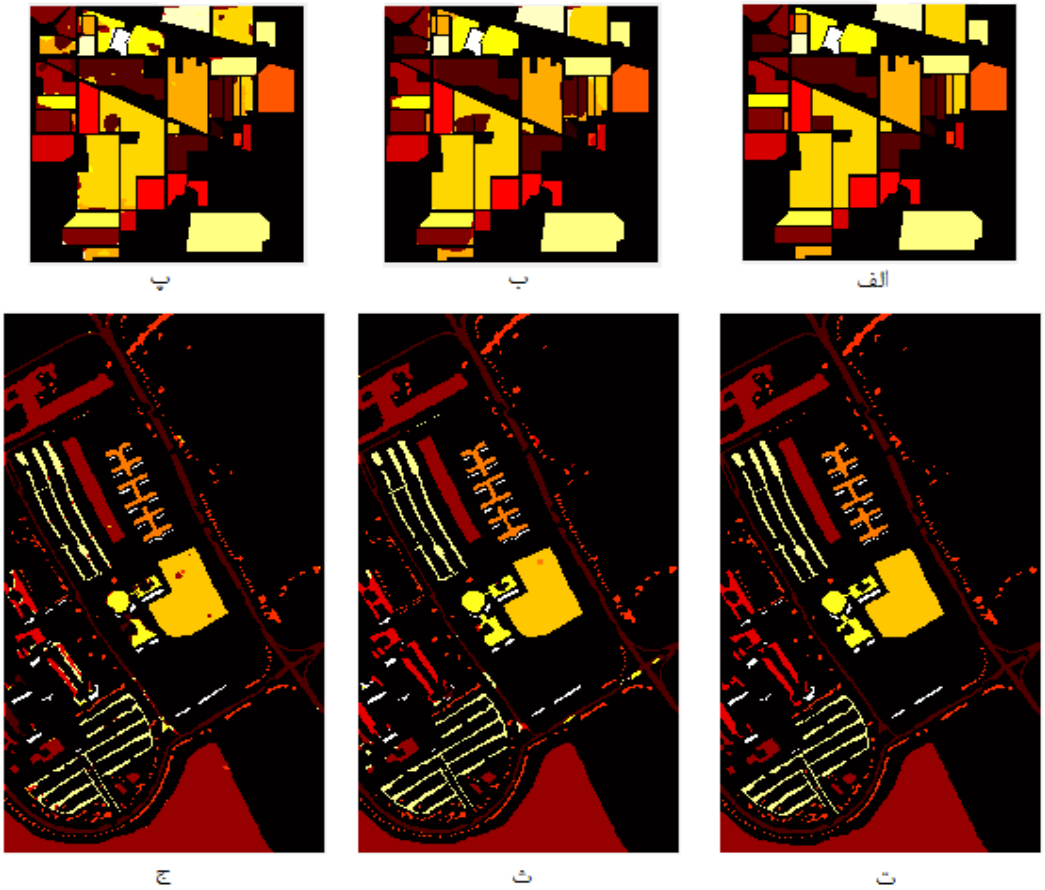
۵. یک موضوع مهم دیگر نتایج نسبتاً ضعیف تولید شده از ویژگی گابور نسبت به دو ویژگی مورفولوژی و *LBP* است. همانطور که در بحث تولید ویژگی‌های مکانی اشاره شد برای تولید ویژگی‌های گابور از یک فیلتر مادر استفاده می‌شود که این فیلتر هر بار در یک جهت و مقیاس متفاوت بر کل تصویر اعمال شده و می‌تواند ساختارهای موجود در آن جهت را حفظ کند و اطلاعات ضعیفی از دیگر مقیاس‌ها و جهت‌ها تولید می‌کند. چون در *3D-CNN* روابط مکانی بین پیکسل‌ها نیز مهم می‌باشد، در نتیجه نتایج طبقه‌بندی حاصل از ویژگی گابور نسبت به دو روش مورفولوژی و *LBP* که ویژگی‌های مکانی را در همه جهت‌ها استخراج می‌کنند، ضعیف‌تر است.

جدول ۳: محاسبه آماره Z بین نتایج طبقه‌بندی حاصل از *CNN* برای ویژگی‌های طیفی *PCA* و *LDA* با حالت طیفی_مکانی که از کل ویژگی‌های مکانی استفاده می‌کند

داده ابرطیفی	داده آموزشی	*Z-Statistics*	
		LDA	*PCA*
ایندین پاین	٪۵	۷٫۴۲	۹٫۳۱
	٪۱۰	۶٫۷۱	۸٫۴۶
دانشگاه پاویا	٪۱	۴۱٫۴۰	۲۲٫۸۳
	٪۵	۳۱٫۷۰	۱۶٫۵۹

در شکل(۱۲) نقشه‌های حاصل از طبقه‌بندی در حالت نمونه آموزشی کمتر برای هر دو داده ابرطیفی در کنار تصاویر حقیقت زمینی نشان داده شده است. همانطور که از نقشه‌های طبقه‌بندی شده مشهود است با افزودن اطلاعات مکانی سطح پایین استخراج شده از سه روش پروفایل‌های مورفولوژی، گابور و *LBP* به اطلاعات طیفی مستخرج از تبدیل *PCA*، نقشه‌های طبقه‌بندی حاصل دارای نویز کمتری نسبت به موقعی می‌باشند که فقط از ویژگی‌های طیفی در طبقه‌بندی تصاویر استفاده شود.

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]



شکل ۱۲- داده ایندین پاین – (الف) تصویر حقیقت زمینی – (ب) طبقه‌بندی با ویژگی‌های طیفی_مکانی (روش پیشنهادی) – (پ) طبقه‌بندی با ویژگی‌های طیفی – دانشگاه پاویا – (ت) تصویر حقیقت زمینی – (ث) طبقه‌بندی با ویژگی‌های طیفی_مکانی (روش پیشنهادی) – (ج) طبقه‌بندی با ویژگی‌های طیفی

۴-۴- مقایسه با روش‌های دیگر

برای تشخیص جایگاه روش پیشنهادی در این تحقیق، نتایج حاصل شده با چند روش یادگیری عمیق دیگر مقایسه شده است. این روش‌ها عبارت‌اند از:

۱) *EPF*[1]: در این روش ابتدا تصویر ابرطیفی به کمک یک طبقه‌بندی کننده مانند *SVM* طبقه‌بندی شده، سپس بر روی نقشه‌های طبقه‌بندی احتمالی به دست آمده یک فیلترینگ حفظ حاشیه مانند فیلتر گاید اعمال می‌شود. در نهایت، طبق نقشه احتمالات فیلتر شده، کلاس هر پیکسل بر اساس حداکثر احتمال انتخاب می‌شود[۳۶].

۲) *R-VCANet*: این روش برای استخراج ویژگی‌های عمیق از تصاویر ابرطیفی طراحی شده است. *R-VCANet* دارای ساختار شبکه‌ای بسیار ساده‌تر نسبت به دیگر روش‌های عمیق می‌باشد و چون پارامترهای کرنل کانولوشن از طریق *VCA*[2] به

[1] *Edge-preserving filters*

[2] *Vertex component analysis*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

دست می‌آیند پس به تعداد نمونه آموزشی کمتری برای آموزش احتیاج دارند[۳۷].

۳) *GFDN*: در روش *GFDN*[1] ابتدا ویژگی‌های گابور برای سه مولفه اول *PCA* استخراج شده سپس این ویژگی‌ها در کنار ویژگی‌های طیفی قرار گرفته و بردار طیفی_مکانی حاصل وارد شبکه عمیق *SAE* می‌شود[۳۸].

۴) *RPNet*: روش *RPNet* یک روش جدید یادگیری عمیق می‌باشد که در این شبکه، تکه‌های تصادفی انتخاب شده از تصویر و بدون هیچ آموزشی به عنوان کرنل‌های کانولوشن در نظر گرفته می‌شوند و این کرنل‌ها برای استخراج ویژگی بر روی تصاویر اعمال می‌شوند[۳۹].

۵) *HybridSN*: روش *HybridSN* از ترکیب *2D-CNN* با *3D-CNN* استفاده می‌کند. در این روش ابتدا به کمک *2D-CNN* ویژگی‌های مکانی عمیق استخراج می‌شوند و سپس این ویژگی‌ها در کنار ویژگی‌های طیفی وارد *3D-CNN* شده و طبقه-بندی انجام می‌شود.[۴۰].

نتایج مقایسه میان این تحقیقات و روش پیشنهادی در جدول (۴) نمایش داده شده است. باید توجه داشت که نتایج گزارش شده در این جدول برای سه روش *R-VCANet*، *GFDN* و *HybridSN* از مقالات بالا استخراج شده‌اند و جاهای خالی به این علت می‌باشد که در این تعداد نمونه دقت طبقه‌بندی در مقالات گزارش نشده است، اما دو روش *EPF* و *RPNet* با توجه به مقالات مربوط پیاده‌سازی شده‌اند.

جدول(۴) نشان می‌دهد که نتایج الگوریتم پیشنهادی دارای برتری نسبت به دیگر روش‌ها می‌باشد. علت این برتری می‌تواند به علت معماری جدید در نظر گرفته شده برای *CNN* و همچنین ادغام ویژگی‌های مکانی گوناگون با اطلاعات مکمل و ویژگی‌های طیفی در بردار ورودی باشد. یک پارامتر مهم دیگر برای مقایسه بین روش‌های طبقه‌بندی، زمان اجرا پردازش می‌باشد. به عنوان مثال زمان سپری شده برای انجام طبقه‌بندی داده ایندین با ۱۰٪ نمونه آموزشی و روش پیشنهاد شده در این تحقیق تقریباً ۱۳۰۰ ثانیه می‌باشد ولی برای دو روش *EPF* و *RPNet* که در سیستمی مشابه پیاده‌سازی شدند به ترتیب ۵۲۵ و ۹۰۶ ثانیه است. در نتیجه این دو روش در زمان کمتری اجرا می‌شوند ولی نکته مهمی که وجود دارد تفاوت معنی‌دار ۱٫۲۹ درصدی دقت طبقه‌بندی روش پیشنهادی با روش *RPNet* و ۷٫۸۶ درصدی با روش *EPF* است که باعث می‌شود روش پیشنهادی همچنان پیشرو باشد. البته باید به این نکته نیز اشاره شود که یکی از علل اصلی افزایش زمان در روش پیشنهادی تعداد ویژگی‌های ورودی است به همین دلیل می‌توان مثلاً برای داده ایندین فقط از ویژگی مورفولوژی که در کنار ویژگی طیفی به دقت کلی ۹۸٫۶۱ می‌رسد و بهتر از دقت روش‌های *EPF* و *RPNet* می‌باشد، استفاده کرد و زمان پردازش را به ۲۴۷ ثانیه کاهش داد.

[1] *Gabor filtering and deep network*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

جدول ۴: مقایسه روش مورد استفاده در این پژوهش با روش‌های رقیب

داده ابرطیفی	داده آموزشی	نتایج طبقه‌بندی					
		EPF	*R-VCANet*	*GFDN*	*RPNet*	*HybridSN*	*Proposed*
ایندین پاین	٪۵	۸۵٫۲۲	۹۵٫۹۸	---	۹۵٫۴۶	---	**۹۶٫۸۹**
	٪۱۰	۹۱٫۳۳	۹۷٫۹۰	---	۹۷٫۷۶	۹۸٫۳۹	**۹۹٫۱۹**
دانشگاه پاویا	٪۱	۹۵٫۵۲	۹۶٫۷۷	۹۲٫۳۲	۹۵٫۷۹	---	**۹۸٫۱۰**
	٪۵	۹۸	---	۹۴٫۷۰	۹۸٫۸۸	---	**۹۹٫۹۰**

۵- نتیجه‌گیری و پیشنهادات

در این مقاله باتوجه به آزمایش‌ها انجام شده یک معماری سه‌بعدی جدید از شبکه‌های عصبی پیچشی برای طبقه‌بندی تصاویر ابرطیفی ارائه شده است که شامل سه لایه کانولوشن، دو لایه ادغام و یک لایه کاملاً متصل است. از روش پیشنهادی در طبقه‌بندی دو تصویر ابرطیفی واقعی یکی از منطقه کشاورزی با حد تفکیک مکانی متوسط و دیگری از یک منطقه شهری با حد تفکیک بالا استفاده شد. نتایج طبقه‌بندی به دست آمده از *CNN* پیشنهاد شده برای این دو داده ابرطیفی نشان از برتری روش *CNN* نسبت به دیگر روش‌های طبقه‌بندی مرسوم مانند *KNN* و *SVM* دارد. در بخش دیگر تحقیق برای بررسی تاثیر ویژگی‌های مکانی سطح پایین در نتایج طبقه‌بندی، در ورودی *CNN* پیشنهاد شده علاوه بر ویژگی‌های طیفی از ویژگی‌های مکانی استخراج شده از سه روش مورفولوژی، فیلتر گابور و *LBP* نیز استفاده شد. برای مقایسه نتایج به دست آمده از ورودی‌های طیفی با ورودی‌های متنوع طیفی_مکانی از آماره Z استفاده شد. خروجی این پارامتر نشان می‌دهد که ورودی طیفی_مکانی باعث بهبود قابل توجه نتایج طبقه‌بندی *CNN* پیشنهاد شده نسبت به حالتی می‌شوند که در ورودی فقط از ویژگی طیفی استفاده شود. همچنین عملکرد روش پیشنهادی در افزایش دقت طبقه‌بندی با چندین روش طبقه‌بندی دیگر مقایسه شد که نتایج نشان از برتری روش پیشنهادی نسبت به دیگر روش‌های طبقه‌بندی می‌باشد. جدی‌ترین رقیب روش پیشنهادی را می‌توان روش *RPNet* در نظر گرفت که دقت طبقه‌بندی با روش پیشنهادی در بهترین حالت ۲٫۵ درصد از آن بیشتر است. در تحقیقات آتی می‌توان روش‌های دیگر استخراج ویژگی‌های طیفی، مکانی و ترکیب این ویژگی‌ها را به منظور رسیدن به دقت‌های بالاتر در *CNN* مدنظر داشت. علاوه بر این چون شبکه‌های عصبی پیچشی یک روش نوین در بحث طبقه‌بندی تصاویر ابرطیفی می‌باشد، می‌تواند تحقیقات گسترده‌ای در موارد مختلف مانند ارائه روشی‌های نوین در بحث آموزش شبکه، حل مشکل بیش‌برازش و موارد مختلف انجام شود.

مراجع

[1] *G. Foody, M. Giles, and A. Mathur, " A relative evaluation of multiclass image classification by support vector machines " , IEEE Transactions on geoscience and remote sensing 42, 2004.*

[2] *F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines", IEEE Transactions on geoscience and remote sensing 42, 2004.*

[3] *JA. Gualtieri and S. Chettri, " Support vector machines for classification of hyperspectral data", In IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium, vol. 2, pp. 813-815. IEEE, 2000.*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

[4] PM. Atkinson and ARL.Tatnall, "Introduction neural networks in remote sensing", International Journal of remote sensing 18, pp. 699-709, 1997.

[5] L. Bruzzone, and DF. Prieto, "A technique for the selection of kernel-function parameters in RBF neural networks for classification of remote-sensing images", IEEE transactions on geoscience and remote sensing 37, pp. 1179-1184, 1999.

[6] G. Cybenko, "Approximation by superpositions of a sigmoidal function", Mathematics of control, signals and systems 2, pp. 303-314, 1989.

[7] K. Hornik, "Approximation capabilities of multilayer feedforward networks", Neural networks 4, pp. 251-257, 1991.

[8] MD. Richard, and RP. Lippmann, "Neural network classifiers estimate Bayesiana posterioriprobabilities", Neural computation 3, pp. 461-483, 1991.

[9] D.Michie, D.J.Spiegelhalter, and C.C. Taylor, " Machine Learning, Neural, and Statistical Classification" London, U.K.:Ellis Horwood, 1994.

[10] GE. Hinton and RR. Salakhutdinov, "Reducing the dimensionality of data with neural networks" science 313, pp. 504-507, 2006.

[11] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, no. 5786, pp. 504–507, 2006.

[12] NL. Roux, and Y. Bengio, "Deep belief networks are compact universal approximators", Neural computation 22, pp. 2192-2207, 2010.

[13] R. Salakhutdinov, Ruslan, and G. Hinton, "Deep boltzmann machines", In Artificial intelligence and statistics, pp. 448-455, 2009.

[14] Y. Chen, L. Zhouhan, X. Zhao, G. Wang, Y. Gu, "Deep learning-based classification of hyperspectral data", IEEE Journal of Selected topics in applied earth observations and remote sensing 7, pp. 2094-2107, 2014.

[15] A. Krizhevsky, I. Sutskever, and GE. Hinton, "Imagenet classification with deep convolutional neural networks", In Advances in neural information processing systems, pp. 1097-1105, 2012.

[16] X. Liang, X. Chunyan, X. Shen, J. Yang, S. Liu, J. Tang, L. Lin, and S. Yan, "Human parsing with contextualized convolutional neural network", In Proceedings of the IEEE International Conference on Computer Vision, pp. 1386-1394, 2015.

[17] C. Xu, C. Lu, X. Liang, J. Gao, W. Zheng, T. Wang, and S. Yan, "Multi-loss regularized deep neural network", IEEE Transactions on Circuits and Systems for Video Technology 26, pp. 2273-2283, 2015.

[18] C. Gan, N. Wang, Y. Yang, DY. Yeung, and AG. Hauptmann, "Devnet: A deep event network for multimedia event detection and evidence recounting", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2568-2577, 2015.

[19] JA. Benediktsson, JA. Palmason, and JR. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles", IEEE Transactions on Geoscience and Remote Sensing 43, pp. 480-491, 2005.

[20] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification", Journal of Sensors , 2015.

[21] 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 54, pp. 6232-6251, 2016.

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

[22] *S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification", Neurocomputing 219,pp. 88-98, 2017,*

[23] *E. Aptoula, MC. Ozdemir, and B. Yanikoglu, "Deep learning with attribute profiles for hyperspectral image classification", IEEE Geoscience and Remote Sensing Letters13, pp. 1970-1974, 2016.*

[24] *B. Kumar, and O. Dikshit, "Texture based hyperspectral image classification", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 8, 2014.*

[25] *M. Zortea, and A. Plaza, "Spatial preprocessing for endmember extraction", IEEE Transactions on Geoscience and Remote Sensing 47, pp. 2679-2693, 2009.*

[26] *J. Nagi, F. Ducatelle, GA. Di Caro, D. Cireşan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, and LM. Gambardella, "Max-pooling convolutional neural networks for vision-based hand gesture recognition", In 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 342-347. IEEE, 2011.*

[27] *F. Mirzapour, and H. Ghassemian, "Using GLCM and Gabor filters for classification of PAN images", In 2013 21st Iranian Conference on Electrical Engineering (Icee), pp. 1-6. IEEE, 2013.*

[28] *F. Mirzapour, and H. Ghassemian, "Improving hyperspectral image classification by combining spectral, texture, and shape features", International Journal of Remote Sensing 36, pp. 1070-1096, 2015.*

[29] *RJ. Ferrari, RM. Rangayyan, JL. Desautels, and AF. Frère, "Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets", IEEE Transactions on Medical Imaging 20, pp. 953-964, 2001.*

[30] *T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", IEEE Transactions on Pattern Analysis & Machine Intelligence 7, pp. 971-987, 2002.*

[31] *Z. Zuo, B. Shuai, G. Wang, X. Liu, X. Wang, B. Wang, and Y. Chen, "Learning contextual dependence with convolutional hierarchical recurrent neural networks", IEEE Transactions on Image Processing 25, pp. 2983-2996, 2016.*

[32] *N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting", The journal of machine learning research 15, pp. 1929-1958, 2014.*

[33] *M. Fauvel, "Spectral and spatial methods for the classification of urban remote sensing data", PhD diss., 2007.*

[34] *X. Liu, Q. Sun, B. Liu, B. Huang, and M. Fu, "Hyperspectral image classification based on convolutional neural network and dimension reduction", In 2017 Chinese Automation Congress (CAC), pp. 1686-1690. IEEE, 2017.*

[35] *BA. Beirami, and M. Mokhtarzade, "SVM classification of hyperspectral images using the combination of spectral bands and Moran's I features" In 2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP), pp. 139-144. IEEE, 2017.*

[36] *X. Kang, S. Li, and JA. Benediktsson, "Spectral–spatial hyperspectral image classification with edge-preserving filtering", IEEE transactions on geoscience and remote sensing 52, pp. 2666-2677, 2013.*

[37] *B. Pan, Z. Shi, and X. Xu, "R-VCANet: A new deep-learning-based hyperspectral image classification method", IEEE Journal of selected topics in applied earth observations and remote sensing 10, pp. 1975-1986, 2017.*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

[38] *X. Kang, C. Li, S.Li, and H. Lin, "Classification of hyperspectral images by Gabor filtering based deep network", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, pp. 1166-1178, 2017.*

[39] *Y. Xu, B. Du, F. Zhang, and L. Zhang, "Hyperspectral image classification via a random patches network", ISPRS journal of photogrammetry and remote sensing 142, pp. 344-357, 2018.*

[40] *A. Santara, K. Mani, P. Hatwar, A. Singh, A. Garg, K. Padia, and P. Mitra, "BASS net: Band-adaptive spectral-spatial feature learning neural network for hyperspectral image classification", IEEE Transactions on Geoscience and Remote Sensing, pp.5293-5301,2017.*

[41] *F.Palsson, J.R.Sveinsson, and M.O.Ulfarsson, "Multispectral and hyperspectral image fusion using a 3-D-convolutional neural network", IEEE Geoscience and Remote Sensing Letters, pp.639-643, (2017).*

[42] *S.K. Roy, G. Krishna, S.R. Dubey, and B.B. Chaudhuri, "HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification", Published In IEEE Geoscience And Remote Sensing Letters, 2019.*

[43] *F. Mirzapour, and H. Ghassemian, "Moment-based feature extraction from high spatial resolution hyperspectral images", International Journal of Remote Sensing, pp.1349-1361, 2016.*

[44] *N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R.Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting", The journal of machine learning research, pp.1929-1958, 2014.*

[45] *S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift", 2015.*

[46] *M. Thoma, "Analysis and optimization of convolutional neural network architectures", Master Thesis, Department of Computer Science Institute for Anthropomatics and FZI Research Center for Information Technology, 2017.*

[47] *N. Singhal, N. Singhal, and V. Kalaichelvi,"Image classification using bag of visual words model with FAST and FREAK", In 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT) IEEE, 2017.*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]
[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]

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

Journal of Geospatial Information Technology
Vol.9 No.2, Summer 2021

Research Paper

Classification of hyperspectral images by fusion of spectral and spatial features in convolutional neural networks

Obeid Sharifii[1*], Behnam Asghari Beirami[2], Mahdi Mokhtarzadeh[3]

1- Ms.c student of remote sensing in Department of Geomatics, College of Engineering, K.N.Toosi University of Technology
2- Ph.D. student of remote sensing in Department of Geomatics, College of Engineering, K.N.Toosi University of Technology
3- Associate Professor in Department of Geomatics, College of Engineering, K.N.Toosi University of Technology

Abstract

Hyperspectral images are useful in monitoring the Earth surface phenomena due to the acquisition of large number of spectral bands. Hyperspectral image classification is the most important field of the hyperspectral data processing, and so far, there have been many attempts to increase its accuracy. Convolutional neural networks (CNNs) and spatial features have had a great role in improving the accuracy of the hyperspectral image classification in recent years. In the previous researches not much attention has been paid to the simultaneous use of the capabilities of the low spatial feature deriving methods in convolutional neural networks. For this reason, in the present paper, a new architecture of convolutional neural networks is introduced for the classification of hyperspectral images that uses the different combinations of spectral features and spatial features which are derived from morphological profiles, Gabor filter and local binary pattern (LBP) as input vectors to the proposed CNN. The experiments which are conducted on two real hyperspectral images from agricultural and urban areas, show the superiority of the proposed method (about 2.5%) in comparison to some recent spatial-spectral classification methods.

Key words: *Hyperspectral image classification, convolutional neural networks, morphological profiles, Gabor filter, local binary pattern.*

Correspondence Address *: Geomatics Engineerig Faculty. K.N. Toosi University of Technology. ValiAsr Street, Mirdamad cross.*
Tel *: +98 87 35186331*
Email*: O.sharifi@kntu.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-10-30]

[DOR: 20.1001.1.20089635.1400.9.2.3.2]
[DOI: 10.52547/jgit.9.2.1]