طبقه‌بندی عمیق تصویر Hyperspectral با CNN متنی

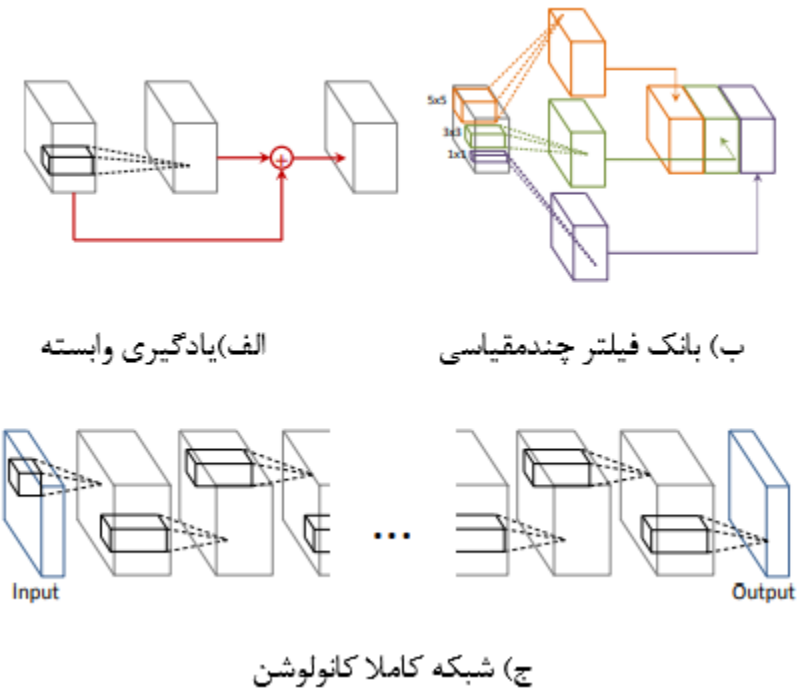
چکیده

در این مقاله که یک شبکه عصبی کانولوشن عمیق جدید (CNN) که عمیق‌تر و گسترده‌تر از شبکه‌های عمیق موجود برای طبقه‌بندی تصویر Hyperspectral است ارائه شده است. برخلاف روش‌های فعلی پیشرفته در طبقه‌بندی تصویر Hyperspectral مبتنی بر CNN، شبکه پیشنهاد شده به نام CNN عمیق متنی، می‌تواند به‌طور مطلوب تعاملات متقابل محتوا را با بهره‌برداری از روابط فضایی-طیفی محلی از بردارهای پیکسل همسایه، بررسی کند. بهره‌برداری مشترک اطلاعات spatio-spectral توسط فیلترکانولوشن چند مقیاسی که به عنوان جزء اولیه خط لوله پیشنهادی CNN مورد استفاده قرار می‌گیرد، به دست می‌آید. ویژگی‌های اولیه فضایی و طیفی نگاشت‌های حاصل از فیلترکانولوشن چند مقیاسی را با هم ترکیب می کنند تا یک ویژگی مشترک فضایی طیفی ایجاد کنند. ویژگی مشترک نشان‌دهنده ویژگی‌های طیفی و فضایی غنی از تصویر Hyperspectral است و سپس از طریق یک شبکه کاملا متقارن تغذیه می‌شود که در نهایت برچسب مربوطه هر pixelvector را پیش‌بینی می‌کند. مجموعه داده‌های استفاده شده در روش پیشنهادی: مجموعه داده‌های Pines هند، مجموعه داده‌های Salinas و مجموعه داده‌های دانشگاه Pavia. مقایسهٔ عملکرد نشان می‌دهد که عملکرد سازگاری پیشرفته در رویکرد پیشنهادی بر روی وضعیت فعلی در سه مجموعه داده نشان داده شده است.

کلمات کلیدی: شبکه عصبی کانولوشن (CNN)، طبقه‌بندی تصویر Hyperspectral، یادگیری وابسته، بانک فیلتر چندمقیاسی، شبکه‌ی کاملا کانولوشن (FCN)

1. مقدمه

اخیرا، شبکه‌های عصبی کانولوشن عمیق (DCNN) برای طیف گسترده‌ای از وظایف ادراکی بصری مانند تشخیص / طبقه‌بندی شی، تشخیص عمل / فعالیت و غیره مورد استفاده قرار گرفته است. به دنبال موفقیت قابل توجه DCNN در نمایش تصویر / ویدیو، قابلیت‌های منحصر به فرد خود را از استخراج زیر ساختارهای غیرخطی از داده‌های تصویری و همچنین شناخت مقوله‌های محتوای معنایی با بهینه‌سازی پارامترهای چند لایه استخراج می‌کند. اخیرا تلاش‌های بیشتری برای استفاده از روش‌های مبتنی بر یادگیری عمیق برای طبقه‌بندی HEX (HIPS) صورت گرفته است [1] - [8]. با این حال، در حال حاضر مجموعه داده‌های HSI در مقیاس بزرگ در دسترس نیستند، که منجر به فراگیری بهینه DCNN با تعداد پارامترهای زیاد بنا به عدم وجود نمونه‌های آموزش دیده می‌گردد. دسترسی محدود به داده‌های گسترده، رویکردهای مبتنی بر CNN برای طبقه‌بندی [1] - [6] HSI را از استفاده‌ی شبکه‌های عمیق‌تر و گسترده‌تر منع می‌کند که می‌تواند به‌طور بالقوه بهتر از اطلاعات طیفی و فضایی بسیار غنی موجود در تصاویر hypersepctral استفاده کند.



شکل 1. مولفه‌های کلیدی شبکه پیشنهادی

از این رو، رویکردهای مدرن و پیشرفته مبتنی بر CNN، بیشتر به استفاده از شبکه‌های کوچک مقیاس با تعداد لایه‌ها و گره‌های نسبتا کمتر در هر لایه برای کاهش هزینه عملکرد تمرکز می‌کنند. عمیق‌تر و گسترده‌تر به معنای استفاده از تعداد نسبتا بیشتری لایه (عمق) و گره در هر لایه (عرض) است. به این ترتیب، کاهش ابعاد طیفی تصویربرداری hypersepctral به طور کلی از طریق تطبیق با تکنیک‌های کم عمق، مانند تجزیه و تحلیل مولفه‌های اصلی (PCA)، تشخیص اختیاری موضعی [3] (BLDE)، تجزیه و تحلیل اختلال محدودیت زوج و اختلاف ناپیوستگی (PCDA-NSD) [10] و غیره است. با این حال، بهره‌برداری از شبکه‌های بزرگ در مقیاس بزرگ، هنوز هم مطلوب است تا به‌طور مشترک از زیرساخت‌های غیرخطی ساختار طیفی و فضایی داده‌های Hyperspectral ساکن در فضای ویژگی‌های چند بعدی استفاده کند. در روش پیشنهادی، قصد داریم یک شبکه عمیق‌تر و وسیع‌تر با توجه به مقادیر محدود داده‌های Hyper-Terra بسازیم که بتواند به‌طور مشترک از اطلاعات طیفی و مکانی هم بهره بگیرد. برای مقابله با مسائل مربوط به آموزش شبکه بزرگ مقیاس در مقدار محدودی از داده‌ها، یک مفهوم به تازگی معرفی شده از "یادگیری وابسته" را به اثبات می‌رسانیم که نشان‌دهنده توانایی قابل توجه برای افزایش قابلیت‌های شبکه‌های بزرگ است. یادگیری وابسته [11] اساس یادگیری زیرگروههای لایه‌ها به نام ماژول‌ها است به‌طوری‌که هر یک از ماژول‌ها توسط سیگنال وابسته، که تفاوت بین خروجی مورد نظر و ورودی ماژول است بهینه می‌شود، همانطور که در شکل a1 نشان داده شده است، ساختار وابسته از شبکه‌ها باعث افزایش قابل توجهی در عمق و عرض شبکه می‌شود که منجر به افزایش یادگیری و در نهایت بهبود عملکرد تولید می‌شود. بنابراین، شبکه پیشنهاد شده نیاز به پیش پردازش برای کاهش ابعاد داده‌های ورودی ندارد.

برای دستیابی به عملکرد بهتری برای طبقه‌بندی HSI، ضروری است که ویژگی‌های طیفی و فضایی به صورت مشترک بهره‌برداری شوند. همانطور که در [1] - [3]، [7]، [8] مشاهده می‌شود، روش‌های حاضر برای طبقه‌بندی مبتنی بر HSI به طور کامل از اطلاعات طیفی و فضایی استفاده می‌کنند. دو نوع متفاوت از اطلاعات، طیفی و فضایی، به‌طور جداگانه از پیش پردازش به دست می‌آیند و سپس برای استخراج ویژگی‌ها و طبقه‌بندی در [1]، [7] پردازش می‌شوند. هو و همکارانش [2] به طور مشترک اطلاعات طیفی و مکانی را تنها با استفاده از بردارهای پیکسل طیفی منفرد به

عنوان ورودی به CNN پردازش کرده‌اند. در این مقاله با الهام از [12]، یک رویکرد مبتنی بر یادگیری عمیق ارائه می‌دهیم که از لایه‌های کانولوشن کامل [13] (FCN)برای استفاده بهتر از اطلاعات طیفی و فضایی از داده‌های hypersepctral به کار می‌رود. در مرحله اولیه CNN عمیق پیشنهاد شده، یک فیلترینگ کانولوشن با مقیاس چندگانه شبیه به "ماژول آغازین" در [12] به طور همزمان توسط نگاشت‌های محلی و گرافیکی طیفی ایجاد می‌شود. بانک فیلتر چندگانه اساسا برای بهره برداری از ساختارهای مختلف فضایی محلی و همچنین همبستگی‌های طیفی محلی استفاده می‌شود. سپس نگاشت‌های فضایی و طیفی اولیه تولید شده با استفاده از بانک فلیتر با هم ترکیب می‌شوند تا یک نگاشت ویژگی spatio-spectral مشترک ایجاد شود که ویژگی‌های spatio-spectral غنی از بردارهای پیکسل hypersepctral را تشکیل دهد. نگاشت مشترک ویژگی به نوبه خود به عنوان ورودی برای لایه‌های بعدی مورد استفاده قرار می‌گیرد که به طور پیش‌فرض برچسب‌های بردارهای پیکسل hypersepctral مربوطه را پیش‌بینی می‌کنند.

شبکه ارائه شده یک شبکه end-to-end است که به‌طور کامل بهینه شده و نیازی به پیش‌پردازش و پس‌پردازش ندارد. شبکه پیشنهادی یک شبکه کاملا متقارن [13] (FCN)(شکل c1) برای گرفتن تصاویر hypersepctral با اندازه دلخواه به عنوان ورودی است و از هیچ لایه زیرنمونه‌ای (pooling) استفاده نمی‌کند که در غیر این صورت خروجی با اندازه‌های مختلفی از ورودی را منجر می‌شود؛ به این معنی که شبکه می‌تواند تصاویر hypersepctral را با اندازه دلخواه پردازش کند. در این روش، شبکه پیشنهادی را در سه مجموعه داده‌ی معیار با اندازه‌های مختلف (145 × 145 پیکسل برای مجموعه داده‌های Pines هند، 640 × 340 پیکسل برای مجموعه داده‌های دانشگاه پاویا و 217 × 512 برای مجموعه داده Salinas) ارزیابی می‌کنیم. شبکه پیشنهاد شده از سه مولفه اصلی تشکیل شده است. یک شبکه کاملا مجازی، یک بانک چندرسانه‌ای و یادگیری وابسته همانطور که در شکل 1 نشان داده شده است. مقایسه عملکرد طبقه‌بندی پیشرفته‌ی، شبکه پیشنهادی را بر حسب حالت فعلی در سه مجموعه داده نشان می‌دهد.

موارد مطرح در مقاله به شرح زیر است:

• شبکه‌ی عمیق‌تر و وسیع‌تر با کمک "یادگیری وابسته" برای غلبه بر کمبود بهینه‌گی در عملکرد شبکه که عمدتا ناشی از مقدار محدود نمونه‌های آموزشی است معرفی می‌کنیم.

• یک معماری CNN عمیق ارائه می‌کنیم که می‌تواند به‌طور مشترک اطلاعات طیفی و فضایی تصاویر hypersepctral را بهینه‌سازی کند.

• روش پیشنهادی یکی از اولین تلاش‌ها برای استفاده از یک شبکه عصبی کاملا پیچیده برای طبقه‌بندی hypersepctral است.

ادامه‌ی مقاله به شرح زیر سازماندهی شده است. در بخش دوم، کارهای مربوطه شرح داده شده است. جزئیات شبکه پیشنهادی در بخش 3، بیان شده است. مقایسه‌ی عملکرد در میان شبکه‌های پیشنهادی و رویکردهای فعلی موجود در بخش 4 شرح داده شده است. در بخش پنجم مقاله با نتیجه‌گیری به پایان رسیده است.

2. کارهای مرتبط

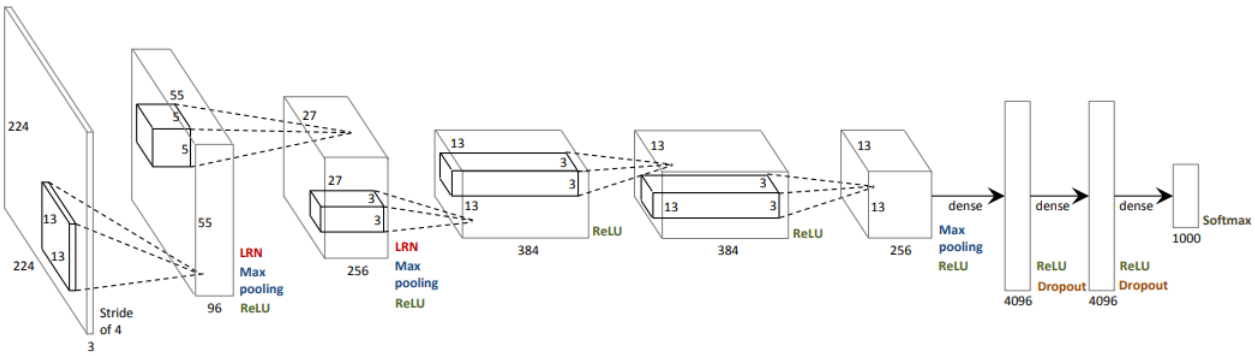
A. ارائه‌ی مسیری عمیق‌تر با CNN برای شناسایی اشیا / طبقه‌بندی

LeCun و همکارانش اولین CNN عمیق به نام LeNet5 [15] را که شامل دو لایه کانولوشن، دو لایه‌ی کاملا متصل و یک لایه اتصال Gaussian با لایه‌های اضافی برای جمع آوری بود، معرفی کرده است. با ظهور پایگاه‌های تصویری در مقیاس وسیع و تکنولوژی پیشرفته محاسباتی، شبکه‌های نسبتا عمیق‌تر و گسترده‌تر مانند AlexNet [16] در مجموعه داده‌های تصویری وسیع مانند ImageNet [17] ساخته شدند. AlexNet از پنج لایه کانولوشن با سه لایه کاملا متصل استفاده می‌کند. Simonyan و Zisserman [18] عمق CNN را با VGG-16، با 16 لایه کانولوشن، به طور قابل توجهی افزایش داده‌اند. Szegedy و همکارانش [12] یک شبکه 22 لایه عمیق را به نام GoogLeNet با استفاده از پردازش چندمرحله‌ای معرفی کرده‌اند که با استفاده از مفهوم "ماژول آغازگر" به دست می‌آید. He و همکارانش [11] یک شبکه عمیق‌تر از آنچه قبلا استفاده کرده بودند با استفاده از یک روش یادگیری جدید به نام «یادگیری وابسته» ساختند که می‌تواند به طور قابل توجهی بهبود کارایی آموزش شبکه‌های عمیق را افزایش دهد.

B. CNN عمیق برای طبقه‌بندی تصویر

تعدادی از روش‌های ارائه شده برای مرتفع کردن مسائل طبقه‌بندی HSI [4]، [19] - [42] ارائه شده‌اند. به‌تازگی، روش‌های هسته‌ای، مانند یادگیری چندهسته‌ای [19] - [25]، به‌طور گسترده مورد استفاده قرار گرفته‌اند، زیرا آنها می‌توانند یک کلاس را قادر به یادگیری تکاملی محدود به یک پارامتر کنند. این مرز با طراحی داده‌ها بر روی یک فضای هسته‌ای چندبعدی هیلبرت ساخته شده است [43]. این کار برای استفاده از مجموعه داده‌ها با نمونه‌های آموزش محدود مناسب است. بااین حال، پیشرفت اخیر در روش‌های مبتنی بر یادگیری عمیق، به‌دلیل قابلیت‌های آن، می‌تواند از طریق ساختارهای غیرخطی و پیچیده تصاویر با استفاده از لایه‌های بسیاری از فیلترهای کانولوشن، بهبود چشمگیر عملکرد را نشان دهد. تا به امروز، چندین روش مبتنی بر یادگیری عمیق [1] - [6] برای طبقه‌بندی HSI توسعه داده شده است. اما بعضی از آنها به دلیل پیشرفت ناشی از عدم وجود نمونه‌های آموزش کافی و استفاده از شبکه‌های نسبتا کوچک مقیاس به پیشرفت‌های فراوانی دست یافتند.

رویکردهای یادگیری عمیق به طور معمول نیاز به مجموعه‌ای بزرگ مقیاس دارند که اندازه آنها باید متناسب با تعداد پارامترهای استفاده شده توسط شبکه برای جلوگیری از بیش برازش در یادگیری شبکه باشد.



تصویر 2. AlexNet [16]. شبکه‌ای شامل پنج لایه کانولوشن و سه لایه کاملا متصل است. در تصویر، مکعب‌ها و جعبه‌ها نشانگرهای داده را نشان می‌دهند. چندین تابع غیرخطی نیز در شبکه استفاده می‌شود. توابع غیرخطی در کنار بلوک‌های خروجی هر لایه قرار می‌گیرند.

Chen و همکارانش [1] از stacked autoencoders (SAE) برای یادگیری ویژگی‌های عمیق از hypersepctral در حالت بدون نظارت همراه با رگرسیون لجستیک مورد استفاده برای طبقه‌بندی ویژگی‌های عمیق

استخراج شده به دسته‌های مناسب استفاده می‌کنند. هر دو بردار پیکسل طیفی و بردار فضایی مربوطه‌ی به دست آمده از تجزیه‌وتحلیل مولفه‌های اصلی (PCA) به داده‌های hypersepctral بر روی ابعاد طیفی جداگانه از یک منطقه محلی به دست می‌آیند و سپس به طور مشترک به عنوان ورودی به SAE مورد استفاده وارد می‌شوند. در [7]، chen و همکارانش SAE را با یک شبکه عمیق (DBN) جایگزین کردند، که شبیه شبکه عصبی کانولوشن عمیق برای طبقه‌بندی HSI است. li و همکارانش [8] از یک DBL دو لایه استفاده کردند اما از کاهش اندازه اولیه استفاده نکرد، که به ناچار باعث از بین رفتن اطلاعات بحرانی تصاویر hypersepctral بود. Hu و همکارانش [2] از بردارهای پیکسل طیفی به‌صورت جداگانه از طریق CNN ساده استفاده کردند که در آن فیلترهای کانولوشن محلی به بردارهای طیفی برای استخراج ویژگی‌های طیفی محلی اعمال می‌شود. سپس نگاشت‌های ویژگی تولید شده پس از جمع‌آوری حداکثری به عنوان ورودی به مرحله طبقه‌بندی کاملا متصل برای طبقه‌بندی مورد استفاده قرار می‌گیرند. Chen و همکارانش [4] همچنین از شبکه عصبی کانولوشن عمیق با پنج لایه‌ی کانولوشن و یک لایه کاملا متصل برای طبقه‌بندی hyperspectral استفاده کردند.

بر خلاف روش‌های مبتنی بر یادگیری عمیق، ابتدا در تلاش برای ساختن شبکه عمیق‌تر و وسیع‌تر با استفاده از مقادیر نسبتا کم نمونه‌های آموزشی بودیم. انتظار می‌رود که شبکه به طور موثر بهینه‌سازی شده و عملکرد پیشرفته‌ای را در شبکه‌های نسبتا جزئی و باریک ارائه دهد.

3. شبکه‌های عصبی عمیق

در این بخش ابتدا مدل CNN که به طور گسترده مورد استفاده قرار می‌گیرید با عنوان AlexNet شرح داده می‌شود و سپس معماری کل شبکه پیشنهادی مورد بحث قرار می‌گیرد. دو عنصر کلیدی شبکه پیشنهادی، «بانک فیلتر چندمقیاسی کانولوشن» و «یادگیری وابسته» توضیح داده می‌شود. در نهایت روند یادگیری شبکه در انتهای بخش بحث می‌شود.

A. شبکه عصبی مصنوعی عمیق

مدل CNN عمیق که استفاده شده شامل چندین لایه از نورون‌ها است، که هر کدام از آنها یک سطح متفاوتی از ویژگی‌های غیرخطی را از ورودی که از ویژگی‌های سطح پایین به بالا است، استخراج می‌کند. غیرخطی بودن در هر لایه با استفاده از یک کارکرد غیرخطی برای تولید لایه‌ها در هر لایه انجام می‌شود. شبکه پیشنهادی اساسا یک شبکه عصبی کانولوشن با یک تابع فعال غیرخطی است که در [16] استفاده شده است. در این بخش ابتدا معماری AlexNet، یک مدل CNN عمیق، همانطور که در شکل 2 نشان داده شده است، را توصیف می‌کنیم تا پایه‌ای برای درک معماری شبکه پیشنهادی ارائه شود. AlexNet شامل پنج لایه کانولوشن و سه لایه کاملا متصل است. هر لایه کاملا متصل دارای وزن خطی WFC است که رابطه بین ورودی x و خروجی y را مرتبط می‌کند:

$$y = W_{FC} \cdot x, \qquad (1)$$

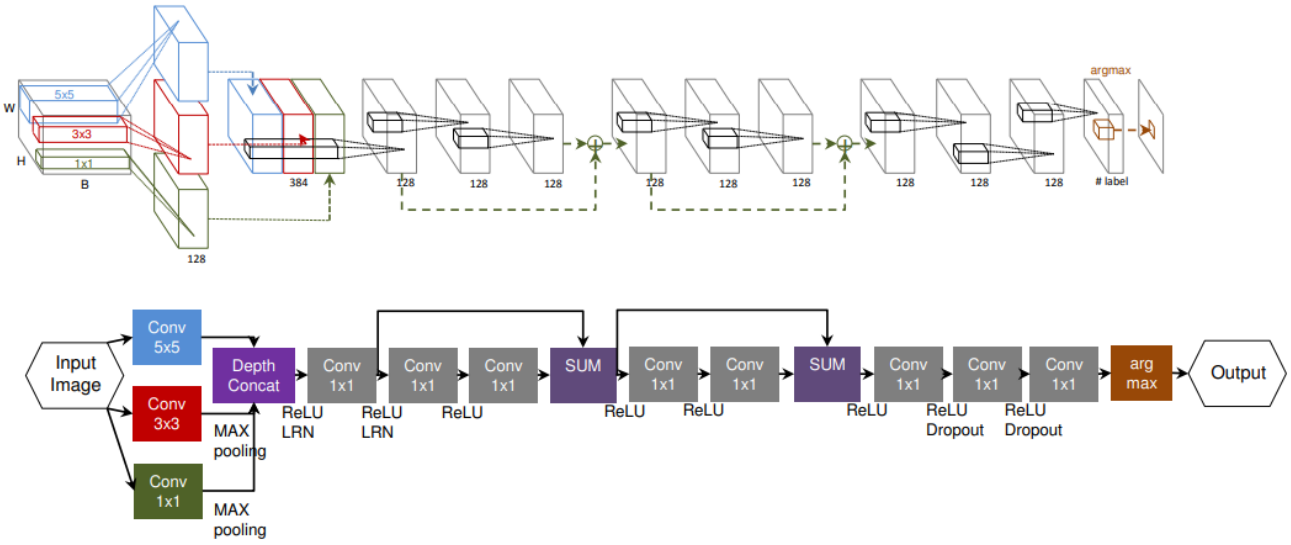
x و y نشان‌دهنده بردارهای ورودی و خروجی هستند. یک لایه کانولوشن با N فیلتر محلی $\{W_{C,i}\}_{i=1,2,\ldots,N}$، ویژگی‌های غیرخطی محلی را از ورودی استخراج می‌کنند که به صورت زیر بیان می‌شود:

$$y = \{W_{C,i} * x\}_{i=1,2,\ldots,N}, \qquad (2)$$

که * یک کنولوشن را نشان می‌دهد. اندازه فیلتر تمام $\{W_{C,i}\}_{i=1,2,\ldots,N}$ به دقت تعیین می‌شود که بسیار کوچک‌تر از اندازه WFC است.

در [16]، چندین عامل غیرخطی مانند نرمال‌سازی پاسخ محلی (LRN)، حداکثر جمع شدن، واحد خطی تصحیح شده (ReLU)، خروج از برنامه و softmax مورد استفاده قرار می‌گیرد. LRN هر فعال‌سازی ai را بر روی فعالیت‌های محلی n فیلتر مجاور که در موقعیت (px، py) مرکزی قرار دارند، نرمال می‌کند و هدف آن تعمیم پاسخ‌های فیلتر است.

$$a_i^*(p_x, p_y) = a_i(p_x, p_y) / \left(k + \alpha \sum_{j=i-n/2}^{i+n/2} \left(a_j(p_x, p_y) \right)^2 \right)^{\beta}, \qquad (3)$$



شکل 3: تصویر معماری شبکه‌ی پیشنهاد شده. ردیف اول لایه‌های ورودی و خروجی لایه‌های کانولوشن و اتصالات آنها را نشان می‌دهد. تعداد فیلترهای هر لایه کانولوشن در زیر خروجی آن نشان داده شده است. ردیف دوم یک نمودار جریان شبکه را نشان می‌دهد.

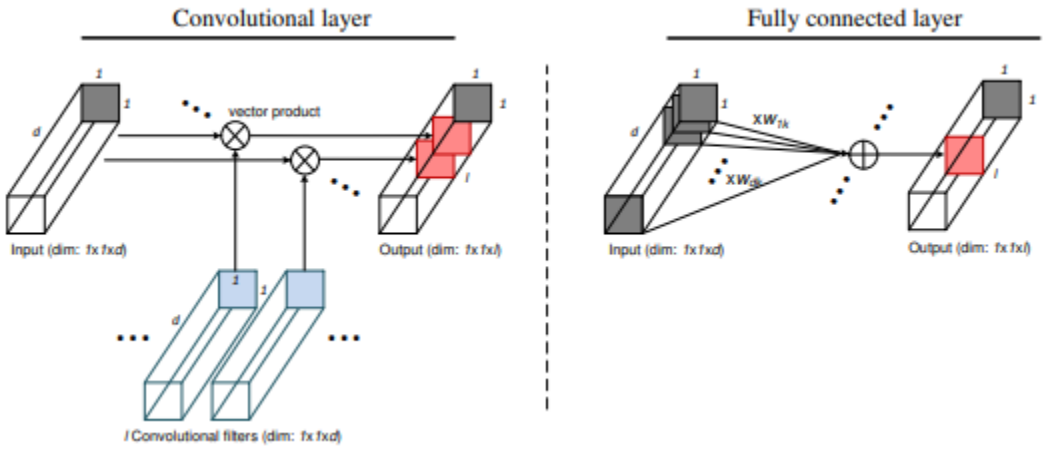
جایی که k، n، α، و β دارای پارامترهای هیدرولیکی هستند. خروجی لایه‌ها با استفاده از جایگزینی یک زیر منطقه خروجی با حداکثر مقدار، که معمولا برای کاهش ابعاد در CNN استفاده می‌شود، نمونه‌ای از لایه‌ها را می‌سازد. ReLU مقادیر منفی را به صفر می‌رساند و برای شبکه و در جهت یادگیری پارامترها با استفاده از فعال‌سازی مثبت استفاده می‌شود. ReLU اساسا عملکرد تابع sigmoid را که معمولا برای شبکه‌های عصبی دیگر استفاده می‌شود، جایگزین می‌کند، عمدتا به این دلیل که یادگیری CNN عمیق با ReLU چندین بار سریعتر از شبکه با سایر توابع فعال غیرخطی مانند tanh است. Dropout یک تابع است که خروجی گره‌های جداگانه هر لایه را با احتمال با یک آستانه مشخص صفر می‌کند و هر مقداری را در داخل (0، 1) قرار می‌دهد. در این کار، از آستانه 0.5 استفاده کردیم. Dropout با جلوگیری از هماهنگی چندگانه داده‌های آموزشی همزمان (که به عنوان "سازگاری های پیچیده" شناخته می‌شود) از بین می‌رود. Softmax یک تعمیم از تابع لجستیک است، که به عنوان نرمالگر گرادیان-log از توزیع احتمالی طبقه‌بندی شده تعریف شده است:

$$P(y=j|x,\{f_k\}_{k=1,2,\cdots,K}) = \frac{e^{f_j(x)}}{\sum_{k=1}^{K} e^{f_k(x)}}, \qquad (4)$$

که fj یک تابع طبقه‌بندی برای کلاس j است که ورودی و خروجی آن x و y است. بنابراین، softmax برای طبقه‌بندی احتمالاتی چندطبقه‌ای از جمله طبقه‌بندی HSI مفید است.

B. معماری شبکه پیشنهادی

یک شبکه کاملا مجتمع جدید [13] (FCN)با تعدادی لایه کانولوشن برای طبقه‌بندی HSI ارائه می‌دهیم، همانطور که در شکل 3 نشان داده شده است. قسمت اول شبکه، یک "بانک چندمقیاسی" همراه دو بلوک از لایه‌های کانولوشن مرتبط با یادگیری وابسته است. سه لایه کانولوشن بعدی به نحوی مشابه با لایه‌های کاملا متصل برای طبقه‌بندی AlexNet عمل می‌کنند که طبقه‌بندی را با استفاده از ویژگی‌های محلی انجام می‌دهند.



شکل 4. مدل سازگار. برای طبقه‌بندی پیکسل، یک لایه کانولوشن می‌تواند همان اثر را به عنوان لایه کاملا متصل با همان تعداد وزن به دست آورد. در تصویر بالا، لایه کنوولاسیون از فیلترهای کانولوشنی استفاده می‌کند که ابعاد آن $d \times 1 \times 1$ است و وزن لایه کاملا متصل $\{w_{i,j}\}_{i=1,\ldots,d\ j=1,\ldots,l}$ است، هر دو لایه کنولوشن و لایه متصل از وزن $l \times$ استفاده می‌کنند.

مشابه AlexNet، لایه‌های کانولوشن هفتم و هشتم در آموزش دیده می‌شوند. ReLU پس از بانک فلیتر چندمقیاسی، لایه‌های کانولوشن دوم، سوم، پنجم، هفتم، هشتم و دو ماژول یادگیری وابسته استفاده می‌شود. خروجی اولین لایه کانولوشن توسط LRN نرمال می‌شود. توجه داشته باشید که ارتفاع و عرض تمام بلوک‌های داده در معماری یکسان هستند و تنها عمق آنها تغییر می‌کند. کاهش حجم در طول پردازش FCN انجام نمی‌شود.

توجه داشته باشید که برگرداندن 1 × 1 × d با فیلترهای 1 × 1 × 1 می‌تواند همان اثر را برگرداند که به‌عنوان کاملا متصل 1 × 1 × d گره ورودی به خروجی، همانطور که در شکل 4 نشان داده شده داشته باشد. با توجه به مدل convolutionalized، FCN می‌تواند برای طبقه‌بندی پیکسل، مانند تقسیم معنایی، طبقه‌بندی HSI، و غیره استفاده شود. از آنجا که شبکه ما براساس FCN است، شبکه پیشنهاد شده بر روی 5 × 5 پیکسل در محدوده بردارهای پیکسل‌های فردی و برای تست کل تصویر یاد می‌گیرد.

جدول 1. مقایسه متغیرهای شبکه مختلف CNNS برای طبقه‌بندی تصویر و HSI.

Method	# of Layer	param	data size	param/data
AlexNet [16]	8	59.3M	12M	4.94
VGG16 [18]	16	135.1M	12M	11.26
GoogLeNet [12]	22	6.8M	12M	0.57
ResNet152 [11]	152	56.0M	12M	4.66
[2]-Indian Pines	3	79.5K	1.6K	49.69
[2]-Salinas	3	80.3K	3.1K	25.90
[2]-U. of Pavia	3	59.8K	1.8K	33.22
The Proposed-Indian Pines	9	1122.5K	6.4K	175.39
The Proposed-Salinas	9	1875.8K	12.4K	151.27
The Proposed-U. of Pavia	9	610.6K	7.2K	84.81

شبکه‌های پیشنهادی چقدر عمیق هستند؟ شبکه پیشنهاد شده شامل 9 لایه است که خیلی بیشتر از سایر CNN ها برای طبقه‌بندی HSI آموزش داده شده بر روی مجموعه داده‌های مشابه مناسب است [2]. با این حال، عمق 9 هنوز به نظر نمی‌رسد به اندازه کافی بزرگ باشد، به ویژه در مقایسه با CNN های پیشرفته موجود برای طبقه‌بندی تصویر، مانند ResNet [11]. این مسئله به این دلیل است که CNN های مبتنی بر HSI باید در مقادیر بسیار کمتر

از نمونه‌های آموزشی تربیت شوند تا از CNN های طبقه‌بندی تصویر که به طور عمده در پایگاه داده‌های بزرگ مانند ImageNet (1.2 M) آموزش داده می‌شود [17] آموزش داده شوند. استراتژی عمیق پیشنهاد شده محدود به داده‌های بسیار محدود آموزش HSI، با استفاده از تعداد بسیار زیادی لایه برای اجتناب از overfitting است. با این حال، هنوز هم تعداد بسیار زیادی از لایه‌ها از سایر CNN های مبتنی بر HSI استفاده می‌کنند. جدول 1 مقایسه CNN های مختلف برای طبقه‌بندی تصویر و HSI با متغیرهای شبکه، مانند تعداد لایه‌ها و پارامترها، اندازه داده‌های آموزشی و نسبت تعدادی از پارامترها و اندازه‌های داده را نشان می‌دهد.

بنا به افزایش اطلاعات استفاده شده در CNN های طبقه‌بندی تصویر، شبکه پیشنهاد شده از یک استراتژی تقویت اطلاعات توصیف شده در بخش III-E استفاده می‌کند. همانطور که در جدول 1 نشان داده شده است، شبکه پیشنهاد شده نسبت تعداد پارامترها و داده‌های آموزش را نسبت به مقادیر پایه [2] برای یک مجموعه داده‌های آموزشی مشابه، نسبتا بزرگتر است. همچنین پارامتر نسبت داده‌ها نسبت به شبکه‌های پیشنهادی تقریبا هشت برابر بزرگتر از هر کدام از CNN های طبقه‌بندی تصویر است. این مسئله نشان می‌دهد که معماری شبکه پیشنهادی طراحی شده است تا اطمینان حاصل شود که عمق مناسب لایه‌ها به طور کامل از داده‌های آموزشی بهره‌برداری می‌کند.

C. بانک فیلتر چندمقیاسی

اولین لایه کانولوشن که به تصویر Hyperspectral ورودی اعمال شده است از یک بانک فیلتر چندبعدی استفاده می‌کند که به صورت محلی تصویر ورودی را با سه فیلتر کانولوشن با ابعاد مختلف (1 × 1 × B، 3 × 3 × B، و 5 × 5 × B که در آن B تعداد نوارهای طیفی است) پردازش می‌کند. فیلترهای 3 × 3 × B و 5 × 5 × B برای بهره‌برداری از همبستگی‌های فضایی محلی تصویر ورودی استفاده می‌شود در حالی که فیلتر 1 × 1 × B برای استفاده از ارتباطات طیفی استفاده می‌شود. خروجی اولین لایه کانولوشن ترکیبی برای ایجاد یک نگاشت ویژگی مشترک فضایی طیفی استفاده شده به عنوان ورودی به لایه‌های کانولوش بعدی است.

با این حال، از آنجا که اندازه نگاشت‌های ویژگی از سه فیلتر کانولوشن متفاوت است، استراتژی برای تنظیم اندازه نگاشت‌های ویژگی با ترکیب آنها برای نگاشت ویژگی مشترک موردنیاز است. ابتدا، یک فضا با عرض دو پیکسل که با صفرها در اطراف تصویر ورودی پوشیده شده است، به‌طوری‌که اندازه ویژگی‌ها از فیلترهای (1 × 1، 3 × 3 و 5 × 5 H + 4)، (W +4، H + 2)، (W +2 و (W ،H) تشکیل شده است. H و W ارتفاع و عرض تصویر ورودی است. اندازه تمام نگاشت‌های ویژگی (W ،H) پس از 5 × 5 و 3 × 3 حداکثر به طور کامل به نگاشت‌های ویژگی فیلتر 1 × 1 و 3 × 3 اعمال می‌شود.

کانولوشن 3 × 3 و 5 × 5 با تعداد زیادی نوارهای طیفی می‌تواند گران باشد و ادغام خروجی از کانال فیلتر باعث افزایش اندازه شبکه می‌شود که ناگزیر به پیچیدگی محاسباتی بالا می‌شود. همانطور که اندازه شبکه افزایش می‌یابد، بهینه‌سازی شبکه با تعداد کمی نمونه آموزشی با بیش‌برازش و اختلاف روبرو خواهد شد. بنابراین، باید استراتژی برای رسیدگی به مسائل فوق مورد استفاده قرار گیرد. برای مقابله با مسائل، از تقویت اطلاعات آموزشی و ماژول‌های یادگیری وابسته در بخش‌های د-3 و ه-4 استفاده می‌کنیم.

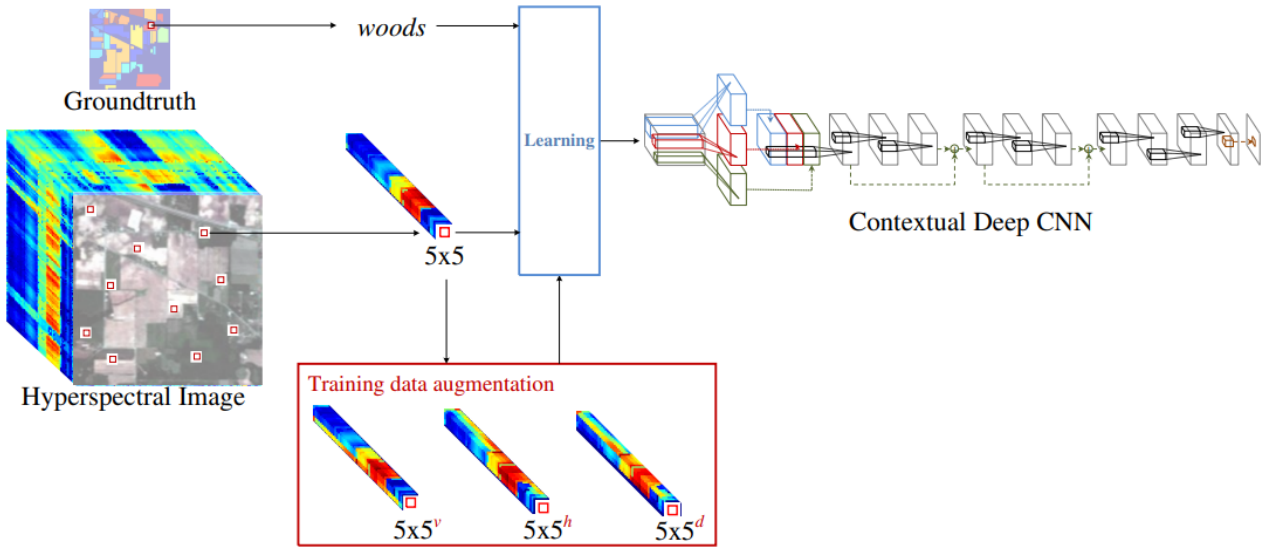
قابلیت بانک فیلتر چند مقیاسی . بانک فیلتر چندمقیاسی با مفهومی شبیه به ماژول اولیه در [12] برای بهینه‌سازی بهره‌برداری از ساختارهای متنوع و مختلف تصویر ورودی سازگار است. [12] اثربخشی ماژول اولیه را نشان می‌دهد که باعث می‌شود شبکه عمیق‌تر شود و همچنین از ساختارهای محلی تصویر ورودی به منظور دستیابی به عملکرد state-of-the-art در طبقه‌بندی تصویر استفاده شود. بانک فیلتر چندگانه در شبکه پیشنهاد شده به شیوه‌ای متفاوت به کار می‌رود که هدف آن ساختارهای فضایی محلی مشترک در ارتباط با همبستگی‌های طیفی محلی در مرحله اولیه ساختار پیشنهادی است.

D. آموزش وابسته

لایه‌های کانولوشن بعد از فیلترهای 1 × 1 × B برای استخراج ویژگی‌های غیرخطی از نگاشت ویژگی طیف مشترک استفاده می‌شود. بنابراین از دو ماژول «یادگیری وابسته» استفاده می‌کنیم [11]، که به طور قابل ملاحظه‌ای برای بهبود کارایی آموزش شبکه‌های عمیق نشان داده شده است. یادگیری وابسته یادگیری لایه‌ها با اشاره به ورودی لایه با استفاده از فرمول زیر است:

$$y = \mathcal{F}(x, \{W_i\}) + x, \qquad (5)$$

که در آن x و y بردارهای ورودی و خروجی لایه‌های در نظر گرفته شده است. تابع F: = y - x نگاشت وابسته از ورودی به خروجی وابسته y - x با استفاده از فیلترهای کانولوشن Wi است. [11] ثابت کرده است که بهینه‌سازی Wi با نگاشت وابسته آسان‌تر از بهینه‌سازی این وزن‌ها با نگاشت نامربوط است. در شبکه پیشنهادی، دو لایه کانولوشن برای نگاشت وابسته وجود دارد، که "اتصالات میانبر" نامیده می‌شود. یادگیری وابسته در عمل بسیار موثر است، که در [11] اثبات شده است. ReLU تابعی است که اولین لایه را در ماژول غیرخطی ایجاد می‌کند.



شکل 5. روند یادگیری شبکه پیشنهاد شده. در تصویر Hyperspectral، 1*1 پیکسل آموزشی و 5 × 5 پیکسل همسایه به ترتیب با یک مستطیل قرمز و سفید نشان داده شده است. در جعبه قرمز نشان داده شده از داده‌های

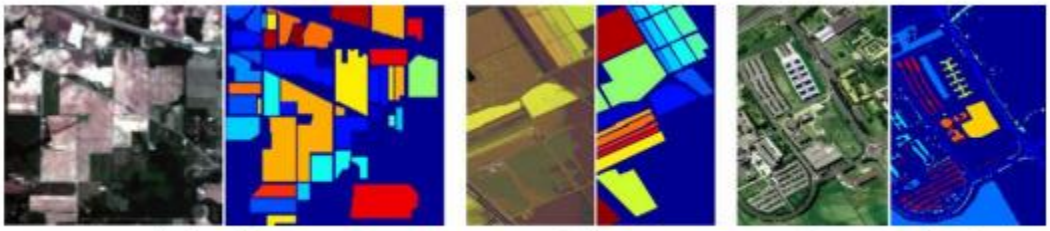
آموزش پیشرفته، 5 × V5، 5 5H ×، و 5 × d5 نمونه‌های آموزشی هستند که به ترتیب در محور افقی، عمودی و مورب نشان داده شده است.

توجه داشته باشید که هر دو بانک فیلتر چندمقیاسی و یادگیری وابسته در افزایش عمق و عرض شبکه در حالی که محدودیت محاسبات وجود دارد موثر است [11]، [12]. این مسئله کمک می‌کند تا به طور موثر آموزش شبکه عمیق با تعداد کمی از نمونه‌های آموزشی صورت گیرد.

E. یادگیری شبکه پیشنهادی

ما به صورت تصادفی یک تعداد مشخصی از پیکسل‌ها را از تصویر Hyperspectral به منظور ارزیابی عملکرد شبکه پیشنهاد شده نمونه برداری می‌کنیم. برای هر پیکسل آموزشی، حدود 5 × 5 پیکسل همسایه را برای یادگیری لایه‌های کانولوشه در نظر می‌گیریم. شبکه پیشنهاد شده حاوی حدود K1000 پارامتر است که از چند صدها پیکسل آموزشی از هر دسته مشخص می‌شود. برای جلوگیری از بیش برازش، تعداد نمونه‌های آموزشی را چهار بار توسط بازتاب نمونه‌های آموزش در محورهای افقی، عمودی و مورب افزایش می‌دهیم. شکل 5 روند یادگیری شبکه پیشنهاد شده را نشان می‌دهد.

برای یادگیری شبکه پیشنهاد شده، گرادیان تصادفی (SGD) با اندازه دسته‌ای 10 نمونه با K100 تکرار، حرکت 0.9، افتادگی 0.0005 و گاما 0.1 استفاده می‌شود. در ابتدا نرخ یادگیری پایه را به 0.001 تنظیم کردیم. نرخ یادگیری پایه پس از 33333 تکرار به 000001 کاهش می‌یابد و بعد از 66 666 تکرار به 0000001 کاهش می‌یابد. برای یادگیری شبکه، آخرین لایه argmax با یک لایه softmax که معمولا برای یادگیری لایه‌های کانولوشن استفاده می‌شود، جایگزین می‌شود. لایه اول، دوم و نهم کانولوشن از یک توزیع گاوسی با میانگین صفر با انحراف معیار 0.01 و لایه‌های وابسته با انحراف استاندارد 0.005 شروع می‌شوند. بی‌نظمی از تمام لایه‌های کانولوشن به جز آخرین لایه به یک مقدار اولیه و آخرین لایه صفر شروع می‌شود.



شکل 6. سه مجموعه داده Pines.HSI هندی، سالیناس و مجموعه داده‌های دانشگاه پاویا. برای هر مجموعه داده، تصویر ترکیبی رنگی در سمت چپ داده شده و تصویر واقعی زمین در سمت راست نشان داده شده است. در تصویر واقعی، پیکسل متعلق به همان کلاس با همان رنگ تصویر می‌باشد.

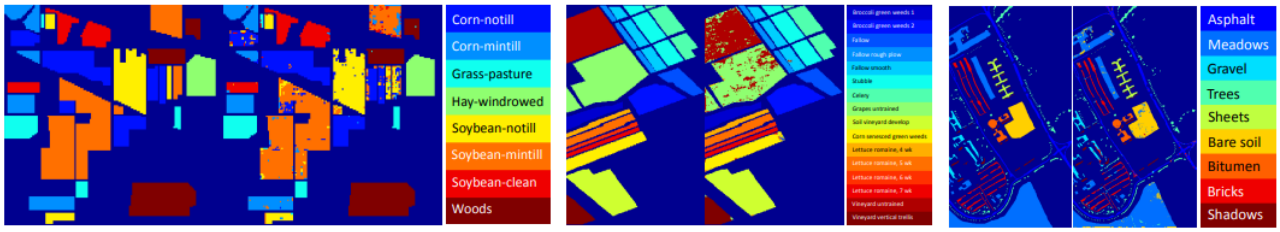
جدول 2 کلاس‌های انتخابی برای ارزیابی و تعداد نمونه‌های آموزش و نمونه‌های مورد استفاده از داده‌های هندسی PINES

No	Class	Training	Test
1	Corn-notill	200	1228
2	Corn-mintill	200	630
3	Grass-pasture	200	283
4	Hay-windrowed	200	278
5	Soybean-notill	200	772
6	Soybean-mintill	200	2255
7	Soybean-clean	200	393
8	Woods	200	1065
	Total	1600	6904

4. نتایج تجربی

A. مجموعه داده‌ها و پایه‌ها

عملکرد طبقه‌بندی HSI بر روی مجموعه داده‌های پیشنهاد شده ارزیابی شد: مجموعه داده‌های هند، مجموعه داده‌های سالیناس و مجموعه داده‌های دانشگاه پاویا، همانطور که در شکل 6 نشان داده شده است.



(a) Indian Pines (b) Salinas (c) University of Pavia

شکل 7. نقشه‌های ترکیبی RGB زمین (چپ) و داده‌های طبقه‌بندی (مرکز) از شبکه پیشنهاد شده برای مجموعه داده‌ها.

جدول 3 کلاس‌های انتخابی برای ارزیابی و تعداد نمونه‌های آموزش و آزمون که از داده‌های SALINAS استفاده می‌کند

No	Class	Training	Test
1	Broccoli green weeds 1	200	1809
2	Broccoli green weeds 2	200	3526
3	Fallow	200	1776
4	Fallow rough plow	200	1194
5	Fallow smooth	200	2478
6	Stubble	200	3759
7	Celery	200	3379
8	Grapes untrained	200	11071
9	Soil vineyard develop	200	6003
10	Corn senesced green weeds	200	3078
11	Lettuce romaines, 4 wk	200	868
12	Lettuce romaines, 5 wk	200	1727
13	Lettuce romaines, 6 wk	200	716
14	Lettuce romaines, 7 wk	200	870
15	Vineyard untrained	200	7068
16	Vineyard vertical trellis	200	1607
	Total	3200	50929

145 × 145 پیکسل و 220 نوار طیفی انعکاسی محدوده‌ی 0.4 تا 2.5 میکرومتر را با وضوح فضایی 20 متر پوشش می‌دهد. داده‌های دانشگاه هند در ابتدا دارای 16 کلاس است، اما ما فقط 8 کلاس را با تعداد نسبتا زیادی از نمونه‌ها استفاده می‌کنیم. مجموعه داده Salinas حاوی 216 × 512 پیکسل و 224 نوار طیفی است. که شامل 16 کلاس است و با وضوح بالای فضایی 3.7 متر مشخص می‌شود. مجموعه داده‌های دانشگاه پاویا حاوی 610 × 340 پیکسل با 103 نوار طیفی است که محدوده طیفی از 0.43 تا 0.86 میکرومتر با وضوح فضایی 1.3 متر را پوشش می‌دهد.

9 کلاس در مجموعه داده وجود دارد. برای مجموعه داده‌های Salinas و مجموعه داده‌های دانشگاه پاویا، ما از همه کلاس‌ها استفاده می‌کنیم چون هر دو مجموعه داده شامل کلاس‌هایی با تعداد نسبتا کم نمونه نیستند.

عملکرد شبکه پیشنهاد شده در [2] مقایسه شده است که از معماری عمیق CNN و SVM مبتنی بر هسته RBF در سه مجموعه داده‌ی hyperspectral استفاده شده است. CNN عمیق استفاده شده در [2] شامل دو لایه کانولوشن و دو لایه کاملا متصل است که خیلی کمتر از شبکه پیشنهاد شده ما با 9 لایه کانولوشن است.

جدول 4. کلاس‌های انتخاب شده برای ارزیابی و تعداد نمونه‌های آموزش و تست مورد استفاده در داده‌های دانشگاه پاویا

No	Class	Training	Test
1	Asphalt	200	6431
2	Meadows	200	18449
3	Gravel	200	1899
4	Trees	200	2864
5	Sheets	200	1145
6	Bare soils	200	4829
7	Bitumen	200	1130
8	Bricks	200	2482
9	Shadows	200	747
	Total	1800	40976

در حال حاضر، مجموعه داده‌های Pines هند و دانشگاه پاویا، یک رویکرد با استفاده از شبکه‌های عمیق (D-DBN) [6] برای طبقه‌بندی HSI ارائه می‌دهد [2]. ما همچنین از D-DBN به‌عنوان پایه‌ای در این کار استفاده می‌کنیم. برای مجموعه داده‌های Pines هند، از سه نوع شبکه‌ی عصبی که در [2] مورد ارزیابی قرار گرفته است، استفاده می‌کنیم: یک شبکه عصبی دو لایه کاملا متصل (NN دو لایه)، یک شبکه عصبی کاملا متصل با یک لایه مخفی (سه لایه NN) و LeNet-5 کلاسیک [15].

برای مقایسه عادلانه، 200 نمونه از هر کلاس انتخاب می‌کنیم و از آنها به‌عنوان نمونه‌های آموزشی استفاده می‌کنیم [2]. بقیه کلاس‌ها برای آزمایش شبکه پیشنهاد شده مورد استفاده قرار می‌گیرند. کلاس‌های انتخاب شده و تعداد نمونه‌های آموزش و آزمایش از سه مجموعه داده در جدول‌های 2، 3 و 4 ذکر شده است. در کارهای گذشته‌ی مربوط

به طبقه‌بندی HSI، پارتیشن‌های مختلف داده‌های آموزش / آزمون برای ارزیابی رویکردهای آنها مورد استفاده قرار می‌گیرند. در میان آنها، پارتیشن‌بندی مجموعه داده‌ها با استفاده از 200 نمونه آموزشی، دارای دو مزیت در ارزیابی شبکه پیشنهاد شده است. 1) ارزیابی با این پارتیشن می‌تواند همکاری ما را تأیید کند، که یک شبکه عمیق‌تر و وسیع‌تر با تعداد کمی نمونه‌ی آموزشی ایجاد می‌کند و 2) [2] با استفاده از این پارتیشن می‌توان عملکرد معقول و نسبتا خوب پایه‌های اولیه مانند RBF-SVM و CNN را بهبود داد. برای تمام آزمایشات، پارتیشن‌بندی تصادفی آموزش / تست 20 بار انجام می‌شود و میانگین و انحراف معیار دقت طبقه‌بندی (OA) را نشان می‌دهد. ما تمام آزمایشات را بر روی فریمورک Caffe [44] با یک X GPU Titan انجام داده‌ایم.

جدول 5. مقایسه عملکرد طبقه‌بندی نفوذپذیر در شبکه پیشنهاد شده و پایه‌ها بر روی سه مجموعه داده. بهترین عملکرد در بین 20 پارتیشن آموزش / تست در Parentheses نشان داده شده است. بهترین عملکرد در میان تمام روش‌ها در FOLD BOLD مشخص می‌شود.

Method	Performance		
	Indian Pines	Salinas	University of Pavia
Two-layer NN [2]	86.49	·	·
RBF-SVM [2]	87.60	91.66	90.52
Three-layer NN [1], [2]	87.93	·	·
LeNet-5 [2], [15]	88.27	·	·
Shallower CNN [2]	90.16	92.60	92.56
D-DBN [6]	91.03 ± 0.12	·	93.11 ± 0.06
The proposed network	**93.61 ± 0.56 (94.24)**	**95.07 ± 0.23 (95.42)**	**95.97 ± 0.46 (96.73)**

جدول 6. مقایسه عملکرد شبکه پیشنهاد شده در درصد W.R.T. و عرض‌های مختلف (تعداد کرنل‌ها در هر لایه).

Dataset	64	128	192	256
Indian Pines	80.38 ± 14.20	**93.61 ± 0.56**	93.47 ± 0.41	92.79 ± 0.81
Salinas	91.35 ± 3.62	93.60 ± 0.58	**95.07 ± 0.23**	94.10 ± 0.55
University of Pavia	94.77 ± 0.83	**95.97 ± 0.46**	95.86 ± 0.50	95.78 ± 0.52

جدول 7. زمان آموزش (در دومین) شبکه پیشنهاد شده W.R.T. و در عرض‌های مختلف (تعداد کرنل‌ها در هر لایه).

Dataset	64	128	192	256
Indian Pines	351	482	576	738
Salinas	428	598	696	896
University of Pavia	349	474	597	751

B. طبقه‌بندی HSI

جدول 5 یک مقایسه عملکردی بین شبکه پیشنهاد شده و مقدمات پایه در مجموعه داده‌ها را نشان می‌دهد. Hu و همکارانش [2] تنها یک نمونه از عملکرد طبقه‌بندی را بدون نشان دادن اینکه آیا ارزش بهترین یا متوسط دقت ارزیابی‌های چندگانه چیست، گزارش می‌دهد. شبکه پیشنهاد شده عملکرد بهتری در تمام سطوح پایه و در تمام مجموعه داده‌ها را فراهم کرد. میانگین عملکرد طبقه‌بندی شبکه پیشنهاد شده به ترتیب 2.58٪، 2.47٪ و 2.86٪ برای بهترین داده‌های رده‌بندی پایه در داده‌های Pines هندی، مجموعه داده‌های Salinas و داده‌های دانشگاه پاویا است. این افزایش عملکرد به طور عمده با ایجاد یک شبکه عمیق‌تر و گسترده‌تر و همچنین بهره‌برداری از اطلاعات فضایی طیفی داده‌های Hyperspectral به دست آمده است. یادگیری وابسته نیز به بهبود عملکرد با بهینه‌سازی توانایی آموزش در تعداد نسبتا کمی از نمونه‌ها کمک می‌کند. نقشه جغرافیایی (سمت چپ) و نقشه طبقه‌بندی (راست) که توسط شبکه پیشنهاد شده برای تمام مجموعه داده‌ها به دست آمده است نیز در شکل 7 نشان داده شده است. نقشه‌ی طبقه‌بندی از یک پارتیشن آموزش / تست دلخواه در میان 20 نمونه کشیده شده است.

جدول .8 مقایسه عملکرد شبکه پیشنهادی در .W.R.T های مختلف (تعداد ماژول‌های آموزش).

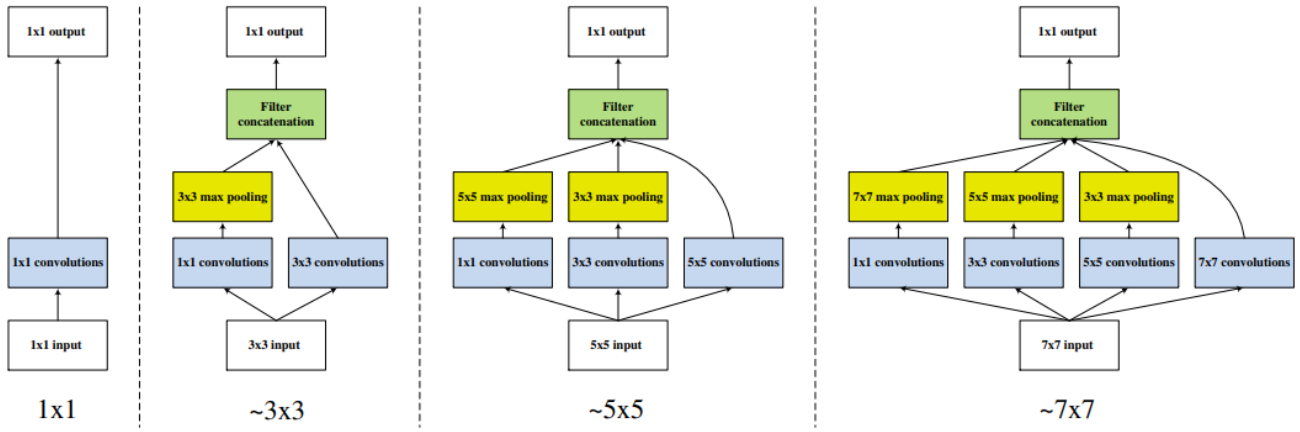
Dataset	1	2	3
Indian Pines	92.74 ± 0.69	**93.61 ± 0.56**	92.63 ± 0.84
Salinas	94.06 ± 0.26	**95.07 ± 0.23**	94.01 ± 0.47
University of Pavia	95.63 ± 0.50	**95.97 ± 0.46**	95.66 ± 0.59

جدول 9 زمان آموزش (در دومین) شبکه پیشنهاد شده در .W.R.T های مختلف (تعداد ماژول‌های آموزش).

Dataset	1	2	3
Indian Pines	431	482	549
Salinas	616	696	777
University of Pavia	426	474	544

C. پیدا کردن عمق و عرض مطلوب شبکه

برای پیدا کردن عرض بهینه شبکه پیشنهاد شده، شبکه را با تغییر تعداد فیلترهای کانولوشه (یعنی تعداد هسته‌ها) ارزیابی می‌کنیم: 64، 128، 192، 256 برای هر سه مجموعه داده. جدول 6 عملکرد شبکه پیشنهاد شده با تعداد مختلف هسته (عرض شبکه) را نشان می‌دهد در حالی‌که جدول 7 زمان آموزش برای همه موارد را نشان می‌دهد. برای مجموعه داده‌های Pine هندی و مجموعه داده‌های دانشگاه پاویا، 128 پهنای بهینه برای بهترین عملکرد است، در حالی‌که 192 بهترین برای مجموعه داده‌های Salinas است. از آنجاییکه مجموعه داده‌های Salinas حاوی نمونه‌های آموزشی بیشتری با تعداد بیشتری از کلاس‌ها نسبت به سایر مجموعه‌های داده است، برای دستیابی به عملکرد مطلوب، وزن بیشتری لازم است.



شکل 8: معماری بانک‌های فیلتر مختلف چندمقیاسی.

جدول 10. مقایسه عملکرد شبکه پیشنهادی .W.R.T بانک فیلتر چندمقیاسی با اندازه‌های مختلف. ~ 7 × 7 به معنای فیلترینگ چندگانه است که شامل 1 × 1، 3 × 3، 5 × 5، و 7 × 7 فیلتر انباشته است.

Dataset	1×1	~3×3	~5×5	~7×7
Indian Pines	53.67 ± 16.63	87.37 ± 4.12	**93.61 ± 0.56**	93.47 ± 0.77
Salinas	50.62 ± 30.87	92.08 ± 0.77	**95.07 ± 0.23**	94.20 ± 0.43
University of Pavia	65.62 ± 8.18	93.59 ± 1.35	**95.97 ± 0.46**	95.91 ± 0.50

همانطور که در جدول 6 و 7 نشان داده شده است، اضافه کردن فیلترهای بیشتر به شبکه بهینه، نه تنها موجب کاهش عملکرد می‌شود بلکه باعث افزایش هزینه‌های محاسباتی نیز می‌گردد.

همچنین شبکه پیشنهاد شده با عمق‌های مختلف را برای به دست آوردن عمق مطلوب ارزیابی می‌کنیم. عمق می‌تواند با استفاده از اعداد مختلف از ماژول‌های یادگیری وابسته متغیر باشد. در جدول 8 نشان داده شده است. جدول 9 زمان آموزش برای همه موارد را نشان می‌دهد. برای هر سه مجموعه داده، با استفاده از دو ماژول یادگیری وابسته، بهترین عملکرد را در میان تمام تغییرات به دست می‌آورد. با استفاده از سه ماژول یادگیری وابسته ممکن است یک مسئله با بیش برازش روبه رو شود، که منجر به تخریب عملکرد می‌شود. همچنین در جدول 9 نشان داده شده است که با استفاده از سه ماژول یادگیری وابسته، محاسبات بسیار پر هزینه می‌شود.

براساس این ارزیابی‌ها، شبکه با دو ماژول یادگیری وابسته و عرض 128 برای هر لایه و برای هر دو مجموعه داده‌ی Pines هند و مجموعه داده‌ی دانشگاه Pavia انتخاب می‌شود. برای مجموعه داده‌های Salinas، شبکه با دو ماژول یادگیری وابسته و عرض 192 برای هر لایه انتخاب شده است.

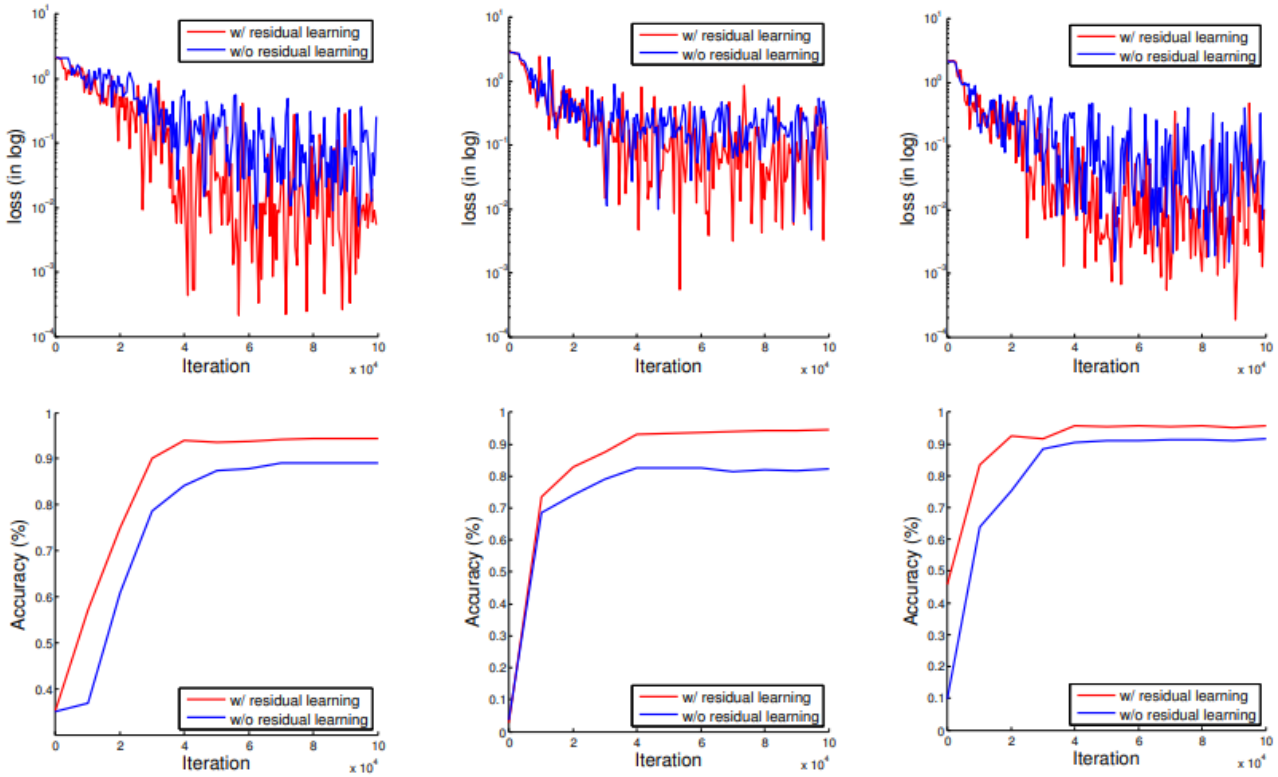
D. اثربخشی بانک فیلتر چندمقیاسی

برای تأیید اثربخشی بانک فیلتر چندمقیاسی که به طور مشترک از اطلاعات فضایی و زمانی استفاده می‌کند، شبکه پیشنهاد شده را با شبکه بدون بانک چندمقیاسی مقایسه می‌کنیم که از یک فیلتر 1×1 در لایه اول استفاده می‌کند. ما همچنین شبکه را با فیلتر چندمقیاسی با تنظیمات مختلف مقایسه می‌کنیم: 1×1، 3×3، 5×5 و 7×7. شکل 8 معماری تمام بانکهای مختلف چندمقیاسی را نشان می‌دهد. همانطور که در جدول 12 نشان داده شده است، فیلتر چندمقیاسی به‌طور قابل توجهی بهتر از شبکه بدون آن (فقط x11) برای هر سه مجموعه داده است (39.94٪ برای مجموعه داده‌های Pines هند، 44.45٪ برای مجموعه داده Salinas و 30.35٪ برای دانشگاه Pavia در میانگین عملکرد طبقه‌بندی). تضعیف عملکرد عمدتا به دو دلیل انجام می‌شود؛ 1) هیچ بهره‌برداری مشترکی از اطلاعات spatiospectral انجام نمی‌گیرد و 2) تقویت داده‌ها با آینه‌سازی مناطق محلی نمی‌تواند به دلیل عدم وجود فضا استفاده شود.

ما همچنین شبکه پیشنهاد شده را با بانک‌های چندمقیاسی با مقادیر مختلف مقایسه می‌کنیم. همانطور که در جدول 12 نشان داده شده است، تضعیف عملکرد با استفاده از بانک فلیتر multiscale با تمام فیلترهای 7 × 7 با 7 ~ 7 نشان داده شده است که به دلیل "سرریز" در نزدیکی مرزهای کلاس ناشی از استفاده از فیلتر فضایی 7 × 7 است. بنابراین، انتخاب می‌کنیم از یک بانک فیلترینگ چندبعدی 1 × 1، 3 × 3 و 5 × 5 برای شبکه پیشنهاد شده استفاده کنیم.

E. اثربخشی یادگیری وابسته

برای تأیید اثربخشی یادگیری وابسته، عملکرد شبکه پیشنهاد شده را باغ یک شبکه مشابه با اولین ماژول وابسته با دو لایه کانولوشن معمولی مقایسه می‌کنیم، همانطور که در جدول 11 نشان داده شده است. هر دو شبکه‌ها بر روی تعداد لایه‌های کانولوشن یکسانی ساخته می‌شوند که 9 است. این یافته‌ها نشان می‌دهد که شبکه بدون استفاده از ماژول‌های یادگیری وابسته، به دلیل تمرکز زیاد در آموزش و ناکام مانده است.



(a) Indian Pines (b) Salinas (c) University of Pavia

شکل 9: ارزیابی اثربخشی یادگیری وابسته. از دست دادن آموزش (بالا) و دقت طبقه‌بندی (پایین) در سه مجموعه داده با شبکه پیشنهاد شده و شبکه با اولین ماژول یادگیری وابسته با دو لایه کانولوشن به عنوان تابع تکرار آموزش ارائه شده است. توجه داشته باشید که "یادگیری وابسته" معماری پیشنهادی است و "یادگیری وابسته" معماری اصلاح شده جایگزین اولین ماژول یادگیری وابسته با دو لایه غیرخطی منظم به عنوان دو لایه پیچیده با لایه‌های غیرخطی یکسان است.

این شبکه با اولین ماژول یادگیری وابسته جایگزین با دو لایه کانولوشن نیز نتوانست منجر به بهینه‌سازی مطلوب عملکرد پارامترهای شبکه شود، همانطور که در جدول 11 نشان داده شده است. در شکل 9، مقایسه از دست دادن آموزش و دقت طبقه‌بندی به عنوان تابع تکرار آموزش برای دو شبکه نشان داده شده است که از یک پارتیشن آموزش / آزمون دلخواه محاسبه می‌شود. از دست دادن آموزش در ردیف اول شکل 9، نشان می‌دهد که شبکه پیشنهاد شده از دست دادن کمتری در طول یادگیری و در پایان تکرارها نسبت به شبکه دیگر دارد. ردیف دوم شکل 9 نیز نشان می‌دهد که کاهش کمتر در طول یادگیری منجر به بهبود دقت طبقه‌بندی می‌شود. این مشاهدات حاکی از آن است که یادگیری وابسته به میزان قابل توجهی اثربخشی یادگیری را افزایش می‌دهد که در نتیجه موجب کاهش آموزش و افزایش دقت طبقه‌بندی می‌گردد.

F. تغییرات عملکرد با توجه به حجم مجموعه آموزش

برای تجزیه و تحلیل تاثیر اندازه مجموعه داده‌های آموزشی در یادگیری شبکه پیشنهاد شده، عملکرد شبکه پیشنهاد شده با تغییر اندازه داده‌ها مقایسه می‌شود: 50، 100، 200، 400، یا 800 نمونه در هر کلاس. جدول 12 دقت طبقه‌بندی شبکه پیشنهادی .w.r.tرا نشان می‌دهد.

جدول 11. مقایسه عملکرد طبقه‌بندی شبکه پیشنهاد شده و شبکه با اولین ماژول آموزشی وابسته با جایگزینی با لایه‌های غیرمنظم (در درصد).

Dataset	w/ conv. layer	w/ residual learning
Indian Pines	49.73 ± 24.58	**93.61 ± 0.56**
Salinas	46.75 ± 25.98	**95.07 ± 0.23**
University of Pavia	50.23 ± 27.78	**95.97 ± 0.46**

برای داده‌های Pines هند، یادگیری با 800 نمونه در هر یک از کلاس‌ها را انجام نمی‌دهیم، زیرا چندین کلاس با نمونه‌های نامناسب دارند (مثلا 483 برای Grass-pasture، 478 برای Hay-windrowed، 593 برای Soybean-clean).

همانطور که انتظار می‌رود، دقت طبقه‌بندی شبکه پیشنهاد شده به‌عنوان افزایش حجم داده آموزش به طور یکنواخت افزایش می‌یابد. همچنین خاطرنشان می‌کنیم که حتی برای اندازه‌های کوچک‌تر آموزش داده شده مانند 50 و 100، شبکه پیشنهاد شده دقت بیشتری نسبت به طبقه‌بندی HSI مبتنی بر یادگیری چندگانه [20] (MKL) ارائه می‌دهد، همانطور که در جدول 12 نشان داده شده است.

جدول 12. مقایسه عملکرد شبکه پیشنهادی (در درصد) W.R.T. با تعداد نمونه‌های آموزشی در یک کلاس.

Dataset	Method	50	100	200	400	800
Indian Pines	MKL [20]	77.40 ± 1.78	80.63 ± 0.99	·	·	·
	The proposed network	80.50 ± 3.93	87.39 ± 0.88	93.61 ± 0.56	94.68 ± 0.47	·
Salinas	MKL [20]	89.33 ± 0.44	90.60 ± 0.43	·	·	·
	The proposed network	91.36 ± 1.11	93.15 ± 0.43	95.07 ± 0.23	96.55 ± 0.29	97.14 ± 0.53
University of Pavia	MKL [20]	91.52 ± 0.98	92.72 ± 0.33	·	·	·
	The proposed network	91.39 ± 0.80	93.10 ± 0.45	95.97 ± 0.46	96.81 ± 0.25	97.31 ± 0.26

G. تجزیه و تحلیل مثبت‌های کاذب

جدول 13 ماتریس‌های confuse را برای سه مجموعه داده، که از یک پارتیشن آموزش / تست دلخواه محاسبه می‌شوند، نشان می‌دهد. برای داده‌های Pines هندی، شبکه پیشنهاد شده عملکرد زیر 95٪ را فقط در دو کلاس از میان هشت کلاس نشان می‌دهد. همانطور که در جدول 2 نشان داده شده است، دو کلاس نمونه‌هایی هستند که نمونه‌های بسیار بیشتری از نمونه‌های دیگر دارند. یادگیری شبکه با داده‌های آموزشی نسبتا کوچک به نظر می‌رسد

نتواند مشخصات کلی طیفی کلاس‌ها را نشان دهد. به‌طور مشابه، تقریبا 5٪ مثبت کاذب هر یک از دو کلاس به‌عنوان کلاس دیگری برچسب‌گذاری می‌شوند، زیرا توزیع طیفی دو کلاس بیشتر از مابقی است. گرایش مشابهی برای مجموعه داده Salinas نشان داده شده است. شبکه پیشنهاد شده برای بدست آوردن اطلاعات بیشتر در دو کلاس، بدتر عمل می‌کند. همانطور که در جدول 3 نشان داده شده است: 83/4 درصد برای grapes بدون آموزش و 89/4 درصد برای vineyaedبدون آموزش. بیشترین اثرات کاذب دو کلاس، طبقه‌بندی به عنوان کلاس دیگری از دو کلاس است. برای مجموعه داده‌های دانشگاه پاویا، عملکرد طبقه‌بندی کلاس Brick به طور قابل توجهی بدتر شده است، که کمتر از 90 درصد است. بیشترین مثبت‌های کاذب کلاس Brick به‌عنوان gravel طبقه‌بندی می‌شوند.

برای ارزیابی اینکه شبکه پیشنهاد شده برای پیکسل‌ها در نزدیکی مرزهای بین کلاس‌های مختلف انجام می‌شود، ما تمام پیکسل‌ها را با توجه به فاصله پیکسل تا مرز طبقه‌بندی کردیم. پیکسل در مرز به عنوان صفر نامگذاری شده است. به طور مشابه، پیکسل‌های نزدیک لبه به مقدار یک مقداردهی می‌شوندو مابقی به صورت ≤ 2 برچسب‌گذاری شده‌اند. توجه داشته باشید که از 5 × 5 پیکسل همسایه برای بهره‌برداری از اطلاعات فضایی هر پیکسل استفاده می‌کنیم. برای پیکسل‌هایی که به‌عنوان ≤ 2 برچسب‌گذاری می‌شوند، 5 × 5 پیکسل همسایه از همان کلاس هستند. جدول 14 تعداد مثبت‌های غلط در مقایسه با تمام داده‌های آزمون را در هر دسته پیکسل برای تمام سه مجموعه داده نشان می‌دهد. برای تمام مجموعه داده‌ها، مشاهده می‌شود که بخش‌های بزرگتر از مثبت کاذب، انتظار می‌رود در نزدیکی مرزهای تولید شده است. مثبت کاذب نزدیک مرزهای کلاس یکی از عوامل مهم برای تخریب عملکرد شبکه پیشنهاد شده است. پیکسل‌های دور از مرزها با فاصله بیش از یک پیکسل، تحت تاثیر قرار نمی‌گیرند و به همین دلیل کمتر به اشتباه طبقه بندی می‌شوند.

5. نتیجه‌گیری

در روش پیشنهادی، یک شبکه عصبی کانولوشن کامل با مجموع 9 لایه ایجاد کرده‌ایم که بسیار عمیق‌تر از دیگر شبکه‌های کانولوشن موجود برای طبقه بندی HSI است. به‌طورکلی به خوبی شناخته شده است که شبکه‌ی عمیق

بهینه می‌تواند منجر به بهبود عملکرد در شبکه‌های کم عمق شود. برای افزایش بهره‌وری یادگیری، شبکه‌ی پیشنهادی آموزش دیده در نمونه‌های آموزشی نسبتا ضعیف، یک روش یادگیری جدید معرفی شده که با نام یادگیری وابسته مورد استفاده قرار گرفته است. برای استفاده از اطلاعات طیفی و فضایی تعبیه شده در تصاویر Hyperspectral، شبکه‌ی پیشنهاد شده به طور مشترک از تعامل فضایی محلی با استفاده از یک بانک فلیتر multiscale در مرحله اولیه شبکه استفاده می‌کند. بانک فیلتر چندرسانه‌ای شامل سه فیلتر کانولوشن با اندازه‌های مختلف می‌باشد: دو فیلتر (3 × 3 و 5 × 5) برای بهره‌برداری از همبستگی فضایی محلی استفاده می‌شود در حالی که 1 × 1 برای پاسخ دادن به همبستگی‌های طیفی استفاده می‌شود.

همانطور که در نتایج تجربی پشتیبانی می‌شود، شبکه‌ی پیشنهاد شده عملکرد پیشرفته‌ی طبقه‌بندی در سه مجموعه داده‌ی معیار را براساس رویکردهای حاضر با استفاده از معماری‌های مختلف CNN ارائه می‌دهد. عملکرد بهبود یافته عمدتا از 1) استفاده از یک شبکه عمیق‌تر با آموزش‌های پیشرفته و 2) بهره‌برداری مشترک از اطلاعات فضایی طیفی است. عمق (تعداد لایه‌ها) و عرض (تعداد هسته‌های مورد استفاده در هر لایه) از شبکه پیشنهاد شده و همچنین تعدادی از ماژول‌های یادگیری وابسته توسط اعتبار متقابل تعیین می‌شود. عملکرد طبقه‌بندی نیز نشان می‌دهد که شبکه‌ی پیشنهادی با دو ماژول یادگیری وابسته بهتر از آن است که دارای تنها یک ماژول باشد که از اثربخشی یادگیری وابسته در شبکه پیشنهادی پشتیبانی می‌کند.

جدول 13.ماتریس confusion. برچسب‌های طبقه متوسط و کلاس‌های طبقه‌بندی شده به واسطه X و Y. اعداد در امتداد محور مربوط به شماره‌های کلاس در جدول های 2، 3و 4 برای سه مجموعه داده.

	1	2	3	4	5	6	7	8
1	**90.1** %	1.5 %	0.0 %	0.0 %	1.0 %	4.9 %	2.2 %	0.3 %
2	1.8 %	**97.1** %	0.0 %	0.0 %	0.0 %	1.1 %	0.0 %	0.0 %
3	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
4	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %
5	1.3 %	0.0 %	0.1 %	0.0 %	**95.9** %	2.2 %	0.5 %	0.0 %
6	5.5 %	3.7 %	0.0 %	0.0 %	3.1 %	**87.1** %	0.7 %	0.0 %
7	2.0 %	0.8 %	0.0 %	0.0 %	0.0 %	0.8 %	**96.4** %	0.0 %
8	0.0 %	0.0 %	0.6 %	0.0 %	0.0 %	0.0 %	0.0 %	**99.4** %

(a) Indian Pines

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
2	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
3	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
4	0.0 %	0.0 %	0.0 %	**99.3** %	0.7 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
5	0.0 %	0.0 %	0.0 %	0.5 %	**98.5** %	0.0 %	0.0 %	0.0 %	0.0 %	0.2 %	0.0 %	0.2 %	0.0 %	0.0 %	0.6 %	0.0 %
6	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
7	0.2 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**99.8** %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
8	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**83.4** %	0.0 %	0.9 %	0.0 %	0.0 %	0.0 %	0.3 %	15.5 %	0.0 %
9	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**99.6** %	0.0 %	0.4 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %
10	0.0 %	0.0 %	1.0 %	0.0 %	0.0 %	0.2 %	0.0 %	0.3 %	0.3 %	**94.6** %	1.6 %	1.0 %	0.0 %	0.6 %	0.4 %	0.0 %
11	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**99.3** %	0.7 %	0.0 %	0.0 %	0.0 %	0.0 %
12	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %
13	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %
14	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %
15	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %
16	0.0 %	0.0 %	0.0 %	0.1 %	0.1 %	0.0 %	0.5 %	1.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.2 %	0.1 %	**98.0** %

(b) Salinas

	1	2	3	4	5	6	7	8	9
1	**94.6** %	0.0 %	1.2 %	0.0 %	0.0 %	0.0 %	2.8 %	1.2 %	0.0 &
2	0.0 %	**96.0** %	0.0 %	1.7 %	0.0 %	2.3 %	0.0 %	0.0 %	0.0 %
3	0.5 %	0.0 %	**95.5** %	0.0 %	0.0 %	0.3 %	0.0 %	4.7 %	0.0 %
4	0.0 %	3.1 %	0.0 %	**95.9** %	0.0 %	0.9 %	0.0 %	0.0 %	0.0 %
5	0.0 %	0.0 %	0.0 %	0.0 %	**100.0** %	0.0 %	0.0 %	0.0 %	0.0 %
6	0.0 %	4.4 %	0.0 %	0.2 %	0.0 %	**94.1** %	0.0 %	1.2 %	0.0 %
7	2.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**97.5** %	0.4 %	0.0 %
8	1.7 %	0.1 %	8.9 %	0.0 %	0.0 %	0.5 %	0.0 %	**88.8** %	0.0 %
9	0.1 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.4 %	0.0 %	**99.5** %

(c) University of Pavia

جدول 14. طبقه‌بندی مثبت کاذب

Dataset	# of FP / # of test data			Percentage		
	0	1	≥ 2	0	1	≥ 2
Indian pines	93 / 717	80 / 109	310 / 5478	12.97 %	11.28 %	5.66 %
Salinas	94 / 1093	81 / 1082	2688 / 48754	8.60 %	7.49 %	5.51 %
University of Pavia	254 / 3455	299 / 4135	1737 / 33386	7.35 %	7.23 %	4.30 %

REFERENCES

[1] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), vol. 7, no. 6, pp. 2094–2107, 2014.

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

[3] 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 (TGARS), vol. 54, no. 8, 2016.

[4] 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 (TGARS), vol. 54, no. 10, pp. 6232–6251, 2016.

[5] P. Liu, H. Zhang, and K. Eom, "Active deep learning for classification of hyperspectral images," IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), no. 10, pp. 712– 724, 2017.

[6] P. Zhong, Z. Gong, S. Li, and C.-B. Sch´´ onlieb, "Learning to diversify deep belief networks for hyperspectral image classification," IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), no. 99, pp. 1–15, 2017.

[7] Y. Chen, X. Zhao, and X. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network," IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), vol. 8, no. 6, pp. 2381–2392, 2015.

[8] T. Li, J. Zhang, and Y. Zhang, "Classification of hyperspectral image based on deep belief networks," in IEEE Conference on Image Processing (ICIP), 2014. [9] K. Pearson, "On lines and planes of closest fit to systems of points in space," Philosophical Magazine, vol. 2, no. 11, pp. 559–572, 1901.

[10] X. Wang, Y. Kong, Y. Gao, and Y. Cheng, "Dimensionality reduction for hyperspectral data based on pairwise constraint discriminative analysis and nonnegative sparse divergence," IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), no. 10, pp. 1552–1562, 2017.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in IEEE conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[12] C. Szegedy, W. Liu, Y. Jia, 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), 2015.

[13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in IEEE conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[14] H. Lee and H. Kwon, "Contextual deep cnn based hyperspectral classification," in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016.

[15] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel, "Backpropagation applied to handwritten zip code recognition," Nerual Computation, vol. 1, pp. 541–551, 1989.

[16] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in Conference on Neural Information Processing Systems (NIPS), 2012.

[17] J. Deng, W. Dong, L. J. J. R. Socher, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in IEEE conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in International Conference on Learning Representations (ICLR), 2015.

[19] P. Gurram and H. Kwon, "Sparse kernel-based ensemble learning with fully optimized kernel parameters for hyperspectral classification problems," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 51, pp. 787–802, 2013.

[20] Y. Gu, T. Liu, X. Jia, J. A. Benediktsson, and J. Chanussot, "Nonlinear multiple kernel learning with multiple-structure-element extended morphological profiles for hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 3235–3247, 2016.
[21] F. de Morsier, M. Borgeaud, V. Gass, J.-P. Thiran, and D. Tuia, "Kernel low-rank and sparse graph for unsupervised and semi-supervised classification of hyperspectral images," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 3410–3420, 2016.
[22] J. Liu, Z. Wu, J. Li, A. Plaza, and Y. Yuan, "Probabilistic-kernel collaborative representation for spatial-spectral hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 2371–2384, 2016.
[23] Q. Wang, Y. Gu, and D. Tuia, "Discriminative multiple kernel learning for hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 3912–3927, 2016.
[24] B. Guo, S. R. Gunn, R. I. Demper, and J. D. B. Nelson, "Customizing kernel functions for SVM-based hyperspectral image classification," IEEE Transactions on Image Processing (TIP), vol. 17, pp. 622–629, 2008.
[25] L. Yang, M. Wang, S. Yang, R. Zhang, and P. Zhang, "Sparse spatiospectral lapSVM with semisupervised kernel propagation for hyperspectral image classification," IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), no. 99, pp. 1–9, 2017.
[26] R. Roscher and B. Waske, "Shapelet-based sparse representation for landcover classification of hyperspectral images," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 1623–1634, 2016.
[27] J. Liu and W. Lu, "A probabilistic framework for spectral-spatial classification of hyperspectral images," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 5375–5384, 2016.
[28] A. Zehtabian and H. Ghassemian, "Automatic object-based hyperspectral image classification using complex diffusions and a new distance metric," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 4106–4114, 2016.
[29] S. Jia, J. Hu, Y. Xie, L. Shen, X. Jia, and Q. Li, "Gabor cube selection based multitask joint sparse representation for hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 3174–3187, 2016.
[30] J. Xia, J. Chanussot, P. Du, and X. He, "Rotation-based support vector machine ensemble in classification of hyperspectral data with limited training samples," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 1519–1531, 2016.
[31] Z. Zhong, B. Fan, K. Ding, H. Li, S. Xiang, and C. Pan, "Efficient multple feature fusion with hashing for hyperspectral imagery classification: A comparative study," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 4461–4478, 2016.
[32] J. Xia, L. Bombrun, T. Adali, Y. Berthoumieu, and C. Germain, "Spectral-spatial classification of hyperspectral images using ica and edge-preserving filter via an ensemble strategy," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 4971–4982, 2016.
[33] H. Yang and M. Crawford, "Spectral and spatial proximity-based manifold alignment for multitemporal hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 51–64, 2016.
[34] M. Toksoz and ¨ ´I. Ulusoy, "Hyperspectral image classification via basic thresholding classifier," IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 54, pp. 4039–4051, 2016.
[35] P. Zhong and R. Wang, "Learning conditional random fields for classification of hyperspectral images," IEEE Transactions on Image Processing (TIP), vol. 19, pp. 1890–1907, 2010.
[36] K. Bernard, Y. Tarabaika, J. Angulo, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach," IEEE Transactions on Image Processing (TIP), vol. 21, pp. 2008–2021, 2012.
[37] Y. Gao, R. Ji, P. Cui, Q. Dai, and G. Hua, "Hyperspectral image classification through bilayer graph-based learning," IEEE Transactions on Image Processing (TIP), vol. 23, pp. 2769–2778, 2014.

[38] M. Brell, K. Segl, L. Guanter, and B. Bookhagen, “Hyperspectral and lidar intensity data fusion: A framework for the rigorous correction of illumination, anisotropic effects, and cross calibration,” IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 55, pp. 2799– 2810, 2017.
[39] S. Jia, J. Hu, J. Zhu, X. gJia, and Q. Li, “Three-dimensional local binary patterns for hyperspectral imagery classification,” IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 55, pp. 2399–2413, 2017.
[40] S. Jia, B. Deng, J. Zhu, and Q. Li, “Superpixel-based multitask learning framework for hyperspectral image classification,” IEEE Transactions on Geoscience and Remote Sensing (TGARS), vol. 55, pp. 2575–2588, 2017.
[41] S. Mei, Q. Bi, J. Ji, J. Hou, and Q. Du, “Hyperspectral image classification by exploring low-rank property in spectral or/and spatial domain,” IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), no. 99, pp. 1–12, 2017.
[42] H. Su, Y. Cai, and Q. Du, “Firefly-algorithm-inspired framework with band selection and extreme learning machine for hyperspectral image classification,” IEEE Journal of Selected Topics in applied Earth Observations and Remote Sensing (J-STARS), no. 10, pp. 309–320, 2017.
[43] E. Strobl and S. Visweswaran, “Deep multiple kernel learning,” in IEEE International Conference on Machine Learning and Applications (ICMLA), 2013.
[44] Y. Jia*, E. Shelhamer*, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” in ACM Multimedia (ACMMM), 2014.