

بهبود تشخیص صحنه در سنجش از راه دور با



استفاده از یادگیری عمیق و انتخاب گر ویژگی

سید علی جهانی^۱، کیوان محبی^{۲*}، فرساد زمانی بروجنی^۳

کارشناس ارشد رشته مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد اصفهان (خوراسگان)، اصفهان، ایران^۱

استادیار گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد اصفهان (خوراسگان)، اصفهان، ایران^۲

استادیار گروه مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران^۳

چکیده

تصاویر سنجش از دور به عنوان منبع داده با ارزش می تواند در اندازه گیری و مشاهده ساختارهای دقیق در سطح زمین کمک کند. هدف این پژوهش ارائه راهکار تشخیص صحنه در سنجش از راه دور با به کارگیری روش های یادگیری عمیق است. برای رفع محدودیت های روش های پیشین، یک رویکرد ترکیبی استخراج ویژگی مطرح شده است که در آن دو نوع ویژگی عمیق محلی و سراسری و یک نوع محلی دستی با یکدیگر ترکیب شده اند. برای استخراج ویژگی ها یک شبکه کانولوشن پیش آموزش دیده با بیست لایه تمام متصل پیشنهاد شده است؛ همچنین یک مرحله انتخاب ویژگی با استفاده از دو دسته الگوریتم های پالایه و بسته بند در این مدل تعبیه شده است؛ در نهایت با استفاده از الگوریتم های طبقه بندی مختلف تشخیص صحنه انجام می شود. ارزیابی راهکار پیشنهادی برای مجموعه داده های RSSCN7، AID، UCM و NWPU-RESISC45 به ترتیب دقت ۹۹/۲۷٪، ۹۷/۹۱٪، ۹۹/۰۹٪ و ۹۳/۰۹٪ را کسب کرده است.

واژگان کلیدی: سنجش از راه دور، یادگیری عمیق، ویژگی عمیق، یادگیری ترکیبی، از پیش آموزش دیده.

Improving Scene Recognition in Remote Sensing Using Deep Learning and Feature Selector

Seyyed Ali Jahani¹, Keyvan Mohebbi^{1*}, Farsad Zamani Boroujeni²

Master, Department of Computer Engineering, Isfahan (Khorasgan) Branch,
Islamic Azad University, Isfahan, Iran¹

Assistant Professor, Department of Computer Engineering, Isfahan (Khorasgan) Branch,
Islamic Azad University, Isfahan, Iran²

Assistant Professor, Department of Computer Engineering, Science and Research Branch,
Islamic Azad University, Tehran, Iran³

Abstract

Remote sensing images as a valuable data source in Earth observation can help in measuring and observing detailed structures on the Earth's surface. Scene detection in remote sensing images has many applications in various fields such as urban planning, natural hazard detection, environmental monitoring, vegetation mapping, and geographic object detection. One of the key problems in the interpretation of remote sensing images is the scene classification of remote sensing images. Feature extraction is very important in scene detection and classification. Convolutional neural networks are one of the deep learning methods that have significantly increased the performance of tasks such as object recognition and scene classification, but their performance is highly dependent on the number of labeled images available, which are not available enough especially in the field of remote sensing. Recently, transfer learning, especially for the fine-tuning of pre-trained convolutional neural networks, has attracted more attention from researchers as a practical strategy for scene classification in remote sensing. However, the lack of use of local features and global deep model that is trained on the target data set is one of the limitations of current methods. Also, if these networks are not deep enough and the images do not pass through multiple filters, they cannot extract more semantic information, and the

* Corresponding author

* نویسنده عهده دار مکاتبات

extracted features do not have high discrimination power, and as a result, scene recognition is not performed well. On the other hand, the features extracted through local features are very large, and not using feature selector methods reduces the accuracy of the model. In this research, to solve the mentioned limitations, a hybrid approach of feature extraction has been proposed in which three types of features including two types of deep local and global features and one type of manual local feature are combined with each other. To extract deep features, pre-trained convolutional networks have been used. The pre-trained networks used are: ResNet, InceptionNet, GoogleNet and EfficientNet_b0. In order to extract as much information as possible from the images, a convolutional network with 20 fully connected layers is proposed. Also, a combined feature selection stage consisting of two categories of filtering and packing algorithms is included in this model. Finally, scene detection is performed using several different classification algorithms. The different structure of pre-trained convolutional networks and their appropriate depth can be effective in improving the extraction of deep features. In addition, the combination of three categories of different features can provide a more comprehensive knowledge of images. The evaluation of the proposed solution on the UCM, AID, RSSCN7 and NWPU-RESISC45 datasets has obtained the accuracy of 99.27%, 97.91%, 99.09% and 93.09% respectively in identifying and classifying images. As a result, this solution has shown a better performance compared to the models that used the manual extraction of features, as well as the methods that use normal convolutional models.

Keywords: Remote Sensing, Deep Learning, Deep Feature, Hybrid Learning, Pre-Trained.

۱- مقدمه

تصاویر سنجش از دور به عنوان یک منبع داده بسیار ارزشمند در رصد زمین نقش بسیار مهمی ایفا می کنند؛ این تصاویر، ما را قادر می سازند تا به دقت ساختارها و الگوهای زمینی را بررسی و اندازه گیری کنیم. با توجه به رشد روزافزون فناوری در این زمینه، حجم تصاویر سنجش از دور به شدت افزایش یافته است [۱،۲]. به دلیل این افزایش، نیاز به روش ها و رویکردهای نوین برای تفسیر و بهره برداری این تصاویر بیش از پیش احساس می شود. تفسیر تصاویر سنجش از دور به منظور طبقه بندی صحنه ها به یک حوزه پژوهشی فوق العاده مهم و پرچالش تبدیل شده است؛ در این حوزه، هدف اصلی آن است که به هر تصویر سنجش از دور، برچسبی مناسب با طبقه های معنایی از پیش تعریف شده اختصاص دهیم. این کاربردها در زمینه های متنوعی از جمله برنامه ریزی شهری [۳]، برای اینکه بتوان برنامه ریزی کرد که کدام قسمت از شهر را چه چیزی تشکیل داده و مدیریت و برنامه ریزی انجام داد، تشخیص مخاطرات طبیعی [۴، ۵]، از طریق پایش گسل ها، پایش محیط [۶]، برای اینکه بدانیم در هر قسمت چه اتفاقاتی در حال رخ دادن است، نقشه برداری پوشش گیاهی و تشخیص اشیای جغرافیایی [۷] تأثیر می گذارند.

با وجود توسعه روزافزون تصویربرداری سنجش از دور، چالش هایی نیز وجود دارد؛ یکی از مهم ترین این چالش ها، طبقه بندی صحنه های پیچیده در تصاویر سنجش از دور است. در تصاویر با وضوح پایین و متوسط، ممکن است اشیای مختلف منحنی پاسخ طیفی مشابهی داشته باشند و یا یک شیء منحنی های متفاوتی داشته باشد؛ بنابراین، روش های طبقه بندی در سطح پیکسل اشیاء،

محدودیت های زیادی دارند. در مقابل، استفاده از بلوک تصویر (برای مثال صحنه) به عنوان واحد اساسی برای تفسیر تصویر می تواند از اطلاعات محتوایی برای رفع ابهام تفسیر استفاده مؤثری کند [۸]. یکی از محدودیت هایی که در این بین مطرح بوده و در مطالعات پیشین مورد توجه قرار نگرفته، تنوع در ویژگی های قابل استخراج است [۹، ۱۰]؛ به عبارت ساده مطالعات موجود فقط متمرکز بر یک نوع ویژگی مثل ویژگی های عمیق و یا ویژگی های محلی بوده اند و هم زمان از مزایای استخراج هر دو بهره نبرده اند [۱۱]. در پژوهش حاضر، به دنبال یک راه کار ترکیبی جهت استخراج ویژگی های تمایز دهنده برای تفسیر تصاویر سنجش از دور هستیم؛ در این رویکرد، ما از سه نوع ویژگی، شامل دو نوع ویژگی سراسری، همچنین یک نوع از ویژگی های محلی دستی بهره برده و به منظور بهبود کارایی تشخیص صحنه ها، آن ها را با هم ترکیب می کنیم.

ادامه این مقاله زیربخش بندی شده است. بخش دو به تعریف اصول و مفاهیم مورد استفاده در این پژوهش می پردازد. بخش سه، تعدادی از کارهای انجام شده پیشین در زمینه سنجش از دور را مرور می کند. بخش چهار، راه کار پیشنهادی در این زمینه را مطرح کرده و به توضیح مراحل مختلف آن اختصاص دارد. بخش پنج، به ارزیابی راه کار ارائه شده می پردازد؛ همچنین مقایسه این راه کار با روش های پیشین مرتبط در بخش شش انجام و در نهایت، نتیجه گیری و پیشنهاد کارهای آینده در بخش هفت مطرح شده است.

۲- ادبیات پژوهش

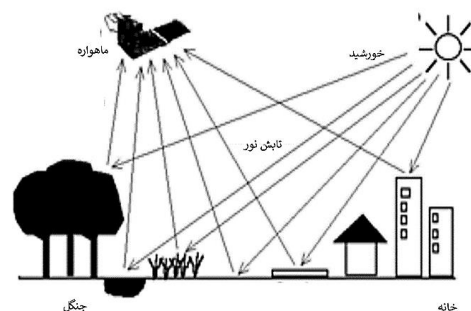
این بخش شامل تعریف مفاهیم پایه مورد استفاده در پژوهش حاضر است.

۲-۱- سنجش از دور

سنجش از دور (RS) که رصد زمین نیز نامیده می‌شود، به جمع‌آوری اطلاعات در مورد اشیاء یا مناطق در سطح زمین بدون تماس مستقیم با جسم یا منطقه اشاره دارد. سنجش از دور روشی برای مشاهده سطح زمین یا جو از خارج از فضا با استفاده از ماهواره و یا هواپیماهای هوابرد است. سنجش از دور از یک یا چند بخش از طیف الکترومغناطیسی استفاده کرده و انرژی الکترومغناطیسی منعکس شده یا ساطع شده از سطح زمین را ثبت می‌کند. مقدار تابش از یک شیء، تحت تأثیر ویژگی‌ها و تابش به آن شیء است. چشم انسان نور خورشیدی بازتاب شده از این اشیاء را ثبت و مغز ما رنگ‌ها، تن‌های خاکستری و تغییرات شدت را تفسیر می‌کند. در سنجش از دور انواع مختلفی از ابزارها و وسایل برای ایجاد تابش الکترومغناطیسی در خارج از این محدوده از چهارصد تا هفتصد نانومتر برای چشم انسان، به‌ویژه مادون قرمز نزدیک، مادون قرمز میانی، مادون قرمز حرارتی و مایکروویو استفاده می‌شود.

تصاویر سنجش از دور کاربردهای زیادی در نقشه‌برداری کاربری و پوشش، کشاورزی، نقشه‌برداری خاک، جنگل‌داری، برنامه‌ریزی شهری، پژوهش‌های باستان‌شناسی، مشاهدات نظامی و بررسی ژئومورفولوژیکی، تغییرات پوشش زمین، جنگل‌زدایی، پویایی پوشش گیاهی، پویایی کیفیت آب و رشد شهری دارد [۱۲].

فرایند سنجش از دور در شکل (۱) نشان داده شده است. نور خورشید به سطح زمین می‌تابد و ماهواره انعکاس تابش نور از اشیاء را تصویربرداری می‌کند و بدین‌صورت می‌توان با طبقه‌بندی این تصاویر، سنجش از دور را انجام داد.



(شکل-۱): فرایند سنجش از دور [۵]
(Figure-1): Remote sensing process [5]

۲-۲- شبکه‌های عصبی از پیش آموزش دیده

شبکه‌های عمیق از پیش آموزش دیده به شبکه‌های عصبی عمیق اطلاق می‌شوند که قبل از استفاده در یک وظیفه خاص، روی یک مجموعه داده بزرگ پیش‌آموزش دیده‌اند؛ این مرحله پیش‌آموزش به شبکه امکان می‌دهد تا ویژگی‌های عمومی را یاد بگیرد که سپس برای انتقال یادگیری به وظایف مشابه دیگر مورد استفاده قرار

می‌گیرند [۱۳]. مزایای اصلی استفاده از شبکه‌های عمیق از پیش آموزش دیده عبارت‌اند از:

۱. آموزش سریع‌تر: از آنجا که لایه‌های شبکه قبلاً ویژگی‌های مفیدی را یاد گرفته‌اند، فقط لایه‌های جدید باید برای وظیفه خاص آموزش داده شوند؛ این منجر به آموزش بسیار سریع‌تر نسبت به آموزش شبکه از ابتدا می‌شود.

۲. عملکرد بهتر: ویژگی‌های آموخته شده از یک مجموعه داده بزرگ در طول فرایند پیش‌آموزش به خوبی به وظایف مشابه دیگر منتقل می‌شوند و در نتیجه بهبود عملکرد بیشتری در مقایسه با آموزش از ابتدا دارند، به خصوص زمانی که از مجموعه داده‌های کوچک‌تر برای تنظیم دقیق استفاده می‌شود.

۳. نیاز به داده کمتر: ویژگی‌های از پیش آموزش دیده به شبکه امکان تعمیم بهتر با استفاده از نمونه‌های کمتر در طول تنظیم دقیق برای کار خاص را می‌دهند.

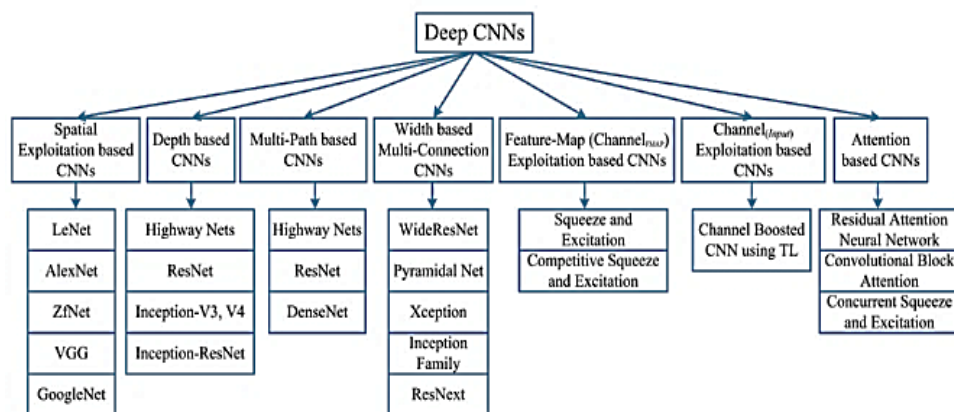
معمول‌ترین شبکه‌های پیش‌آموزش دیده شامل AlexNet, VGGNet, ResNet و Inception هستند که بر روی مجموعه داده ImageNet آموزش دیده‌اند. ویژگی‌های آموخته شده از تصاویر طبیعی در این شبکه‌های عصبی کانولوشنال را می‌توان برای طبقه‌بندی تصاویر برای وظایف بینایی کامپیوتری دیگر استفاده کرد [۱۴].

شبکه‌های از پیش آموزش دیده، با تثبیت لایه‌های قبلی که ویژگی‌های عمومی را ارائه می‌دهند و فقط تطبیق لایه‌های بعدی برای وظایف جدید مورد استفاده قرار می‌گیرند. ویژگی‌های لایه‌های قبلی به عنوان یک استخراج کننده ویژگی عمومی عمل می‌کنند که دوباره مورد استفاده قرار می‌گیرد؛ درحالی که لایه‌های ویژه به مرور در طول تنظیم دقیق تطبیق می‌یابند. این روش نسبت به آموزش شبکه از ابتدا نیاز به داده و زمان آموزش کمتری دارد. تصویری از طبقه‌بندی معماری‌های عمیق کانولوشن در هفت دسته مختلف در شکل (۲) آورده شده است.

۲-۳- استخراج کننده ویژگی محلی

استخراج کننده‌های ویژگی محلی^۱ الگوریتم‌هایی هستند که در تصاویر، نقاط مهم و مناطق برجسته را تشخیص داده و از آن‌ها ویژگی‌های محلی و توصیف‌گرها را استخراج می‌کنند [۱۵]. این ویژگی‌ها به‌ویژه از تغییرات مقیاسی، چرخشی، تغییرات روشنایی و غیره در تصویر نسبت به تبدیل‌های مختلف مستقل‌اند. استفاده از استخراج کننده‌های ویژگی محلی باعث توانایی بالاتر در تشخیص اشیاء، تطبیق تصاویر و انجام وظایف طبقه‌بندی می‌شود؛

¹ Local Feature Extractor



(شکل-۲): طبقه‌بندی معماری‌های عمیق کانولوشن [۱۴]

(Figure-2): Classification of deep convolution architectures [16]

بنابراین، استخراج‌کننده‌های ویژگی محلی به تشخیص نقاط مهم در تصاویر و توصیف ویژگی‌های آن‌ها پرداخته و در وظایفی مانند تطبیق تصاویر و طبقه‌بندی بسیار مفیدند. SIFT و LBP دو مثال از این استخراج‌کننده‌های ویژگی محلی هستند که به‌ترتیب بر اساس تفاوت‌های گرادیان و الگوهای محلی دودویی عمل می‌کنند [۱۴].

۲-۴- الگوریتم‌های انتخاب‌گر ویژگی

الگوریتم‌های انتخاب‌گر ویژگی^۳ یکی از مهم‌ترین مراحل در پردازش داده‌ها و مدل‌سازی ماشینی هستند. این الگوریتم‌ها برای انتخاب ویژگی‌های مهم و معنی‌دار از داده‌ها استفاده می‌کنند تا مدل‌های بهتر و کارآمدتری ایجاد شوند [۱۶]؛ دو الگوریتم معروف در این زمینه، مربع کای^۴ و بسته‌بند هستند. الگوریتم مربع کای به‌طور معمول برای مسائل دسته‌بندی و الگوریتم و بسته‌بند برای مسائل پیش‌بینی (رگرسیون) استفاده می‌شوند. در ادامه توضیح مختصری در مورد هر یک ارائه می‌شود:

الگوریتم مربع کای برای انتخاب ویژگی‌هایی استفاده می‌شود که با متغیر هدف (به‌طور معمول یک متغیر دسته‌بندی) رابطه آماری معنی‌دار دارند. مربع کای به‌طور معمول در مسائل دسته‌بندی و تحلیل متغیرها بر اساس فراوانی‌های مشترک در جداول تراسنشی یا جداول ترکیبی استفاده می‌شود؛ در این الگوریتم ابتدا یک جدول تراسنشی برای ویژگی و متغیر هدف ساخته می‌شود، سپس مقدار کای برای هر ویژگی محاسبه می‌شود. ویژگی‌هایی که مربع کای بزرگ‌تری دارند، به‌عنوان ویژگی‌های مهم انتخاب می‌شوند [۱۴].

در الگوریتم بسته‌بند، ویژگی‌ها به‌صورت گروهی بررسی می‌شوند و هدف انتخاب ویژگی‌هایی است که در

یکی از معروف‌ترین استخراج‌کننده‌های ویژگی محلی، الگوریتم ORB^۱ است، که به‌صورت زیر عمل می‌کند: تشخیص نقاط کلیدی با استفاده از FAST: این الگوریتم ابتدا با استفاده از الگوریتم FAST نقاط کلیدی پتانسیل را در تصویر تشخیص می‌دهد.

محاسبه جهت‌های نقاط کلیدی: برای هر نقطه کلیدی، جهت‌هایی محاسبه می‌شود تا بتوانیم نقاط کلیدی را به‌صورت چرخشی توصیف کنیم.

محاسبه توصیف‌گرهای BRIEF: در این مرحله با استفاده از الگوریتم BRIEF، توصیف‌گرهای دودویی برای نقاط کلیدی محاسبه می‌شود.

توصیف‌گرهای ORB: توصیف‌گرهای BRIEF به‌صورت چرخشی تصویر برای نقاط کلیدی ایجاد می‌شوند و نقاط کلیدی به‌وسیله توصیف‌گرهای ORB توصیف می‌شوند؛ بنابراین، استفاده از ORB به تشخیص نقاط کلیدی و توصیف ویژگی‌های محلی در تصاویر کمک می‌کند. این به‌ویژه در وظایفی مانند تطبیق تصاویر و طبقه‌بندی بسیار مفید است. یکی دیگر از استخراج‌کننده‌های محلی معروف، LBP^۲ است، که به‌صورت زیر عمل می‌کند:

تولید الگوهای دودویی محلی: برای هر پیکسل در تصویر، پیکسل‌های همسایگی آن با توجه به مقدار پیکسل مرکزی در آستانه‌ای تبدیل به یک عدد دودویی می‌شوند. این الگوهای دودویی نمایانگر الگوهای محلی در تصویرند. ساخت هیستوگرام الگوهای محلی: در این مرحله، یک هیستوگرام از الگوهای محلی در یک منطقه محاسبه می‌شود. این توصیف‌گرهای LBP می‌توانند به‌صورت محلی برای نقاط کلیدی یا به‌صورت متراکم در تمام تصویر محاسبه شوند.

^۳ Feature Selection Algorithms

^۴ Chi Squared

^۵ Wrapper

^۱ Oriented FAST and Rotated BRIEF

^۲ Local Binary Patterns

باین حال، تنها حدود هفت میلیون پارامتر دارد که بسیار کمتر از تعداد پارامترهای VGG است. این شبکه می تواند تصاویر ورودی را با اندازه های مختلف بپذیرد؛ علاوه بر این، آن ها به جای لایه به طور کامل متصل، لایه کانولوشن را در شبکه درج کردند تا بتوان ویژگی های خروجی شبکه را بدون عملیات مسطح طبقه بندی کرد، که عملیات طبقه بندی را ساده می کند. روش آن ها عملکرد طبقه بندی را در مجموعه داده UCM، مجموعه داده AID، مجموعه داده OPTIMAL-31 و مجموعه داده NWPU-RESISC45 بهبود قابل توجهی می دهد و به دقتی بین ۹۲٪ تا ۹۵٪ رسیده است [۱۷].

سان و همکاران^۴، یک شبکه دوجته دروازه ای^۵ برای ادغام ویژگی سلسله مراتبی و حذف اطلاعات تداخلی پیشنهاد دادند؛ در این روش ابتدا عملکرد هر یک از ویژگی های کانولوشن تحلیل کمی و ترکیبی برتر از ویژگی های کانولوشن انتخاب می شود؛ سپس، یک اتصال دوطرفه برای تجمع سلسله مراتبی ویژگی های کانولوشن چندلایه پیشنهاد می شود. هم جهت بالابه پایین و هم جهت پایین به بالا برای جمع کردن ویژگی های کانولوشن چندلایه در نظر گرفته می شوند و یک تابع دروازه ای برای حذف اطلاعات تداخل در اتصال دو جهته استفاده می شود. در نهایت، ویژگی های معنایی و ویژگی های ظاهری برای طبقه بندی ادغام می شوند. روش پیشنهادی آن ها در چهار مجموعه داده طبقه بندی صحنه سنجش دور (AID، UC-Merced، WHU-RS19، و OPTIMAL-31) بررسی شده و به دقت ۹۵٪ رسیده است [۱۸].

ژو و همکاران^۶ یک شبکه کانولوشن بدون مقیاس (SF-CNN)^۷ برای طبقه بندی صحنه سنجش از دور معرفی کردند. به طور خاص، در این شبکه لایه های تمام متصل^۸ به لایه های کانولوشنی تبدیل می شوند، که نه تنها اجازه می دهد تصاویر ورودی دارای اندازه های دل خواه باشند؛ بلکه توانایی استخراج ویژگی های متمایز با استفاده از یک استراتژی سنتی مبتنی بر پنجره کشویی^۹ را نیز حفظ می کنند؛ سپس، یک لایه ادغام میانگین سراسری (GAP)^{۱۰} بعد از لایه کانولوشن نهایی اضافه می شود تا تصاویر ورودی با اندازه دل خواه را بتوان برای نمایش ویژگی هایی با اندازه یک نواخت نگاشت کرد؛ در نهایت، از ویژگی های به دست آمده برای ایجاد

کنار هم به بهترین نتیجه برسند؛ مثال بارزی از این الگوریتم، روش های حذف بازگشتی ویژگی^۱ هستند؛ این روش به حذف تدریجی ویژگی های کم اهمیت از مجموعه داده و انتخاب ویژگی های مهم تر برای مدل می پردازد و به طور معمول در مسائل پیش بینی و مدل سازی استفاده می شود؛ به ویژه زمانی که تعداد ویژگی ها زیاد باشد و نیاز به ساده تر کردن مدل و کاهش ابعاد داده وجود داشته باشد. الگوریتم بسته بند از یک مدل ماشینی (به طور معمول مدل دسته بندی یا رگرسیون) شروع می کند، سپس ویژگی ها را به تدریج حذف می کند و عملکرد مدل را مورد ارزیابی قرار می دهد. ویژگی هایی که حذف شده اند و مدل هنوز به خوبی عمل نکند، به تدریج بازگردانده نمی شوند.

در نهایت، انتخاب بین این الگوریتم ها و یا استفاده از الگوریتم های دیگر بستگی به مسئله مورد نظر و خصوصیات داده دارد؛ همچنین، به طور معمول نیاز به تنظیمات و تجربیات مکرر برای انتخاب بهترین ویژگی ها وجود دارد.

۲-۵ شبکه های عصبی گراف (GCN)

شبکه های عصبی گراف^۲ (GNN) یا شبکه های کانولوشن گراف (GCNs) مدل های یادگیری ماشینی هستند که از داده هایی با ساختار نموداری مانند شبکه های اجتماعی یا مولکول ها برای مسائلی که وابستگی های میان نقاط داده مهم اند، بهره می برند؛ این مدل ها به گره ها و ارتباطات بین آن ها در نمودار توجه کرده و تلاش می کنند نمایش های اطلاعاتی از هر گره با در نظر گرفتن اطلاعات از گره های مجاورشان را یاد بگیرند. GNN/GCN به تشخیص الگوها و ویژگی های مهم در داده های گرافیکی کمک می کنند و در بسیاری از وظایف مانند طبقه بندی گره ها، پیش بینی ارتباطات و سایر وظایف مشابه مؤثرند. این مدل ها با استفاده از مراحل مختلفی مانند نمایش ورودی، پیچیدگی نمودار، تجمع چندلایه و لایه خروجی آموزش می بینند.

۳- کارهای انجام شده

ژانگ و همکاران^۳ یک شبکه کانولوشن کامل ساده و کارآمد مبتنی بر DenseNet برای طبقه بندی صحنه های سنجش از دور پیشنهاد دادند؛ در این مدل تعداد زیادی اتصالات متراکم وجود دارد که شبکه را عمیق تر می کند، اما تعداد پارامترها را به میزان قابل توجهی افزایش نمی دهد. شبکه طراحی شده بیش از صد لایه دارد؛

⁴ Sun et al.

⁵ Gated bidirectional network

⁶ Xie et al.

⁷ Scale-free CNN

⁸ Fully Connected

⁹ Sliding-window

¹⁰ Global average pooling (GAP)

¹ Recursive Feature Elimination

² Graph Convolutional Networks

³ Zhang et al.

یک لایه تماممتصل جدید استفاد که برای طبقه‌بندی نهایی به لایه SoftMax تغذیه می‌شود. نتایج تجربی که با استفاده از چندین مجموعه داده واقعی انجام شده، برتری روش پیشنهادی آن‌ها را نسبت به چندین روش طبقه‌بندی معروف، از جمله روش‌های مبتنی بر CNN از پیش آموزش دیده نشان می‌دهد که توانسته در مجموعه داده UC MERCED LAND-USE به دقت ۹۹٪ دست یابد [۱۹].

شی و همکاران^۱ یک روش طبقه‌بندی سبک‌وزن کانولوشن مبتنی بر هم‌جوشی شاخه‌ای ویژگی (LCNN-BFF)^۲ برای طبقه‌بندی صحنه سنچس از دور پیشنهاد دادند؛ برخلاف ساختار کانولوشن معمولی، مدل پیشنهادی آن‌ها ساختار استخراج ویژگی دوخطی^۳ دارد. مدل آن‌ها پیچیدگی را تا حد زیادی کاهش داده است. ارزیابی روش روی چهار مجموعه داده NWPU-RESISC45، RSSCN7، UC و AID نشان داد که مدل به دقتی بین ۹۱٪ تا ۹۷٪ رسیده است [۲۰].

هوانگ و همکاران^۴، یک معماری کانولوشن جدید با نام RSSCNet^۵ با قابلیت تعمیم بالا ارائه دادند؛ در این مدل از سیاست نرخ یادگیری چرخه‌ای دو مرحله‌ای^۶ و روش یادگیری انتقالی بدون انجماد^۷ برای سرعت بخشیدن به آموزش مدل و افزایش دقت استفاده شده است؛ همچنین از الگوریتم t-SNE برای کاهش ابعاد و الگوریتم LIME^۸ برای بهبود نتایج در مواردی که مدل پیش‌بینی‌های اشتباهی انجام داده استفاده شده است. مقایسه نتایج در سه مجموعه داده RSSCN7، UC Merced و WHU-RS19 نشان می‌دهد که مدل آن‌ها توانسته به دقت ۹۸٪ دست یابد [۲۱].

تانگ و همکاران^۹ یک شبکه کانولوشن جدید با نام (ACNet)^{۱۰} پیشنهاد دادند که ساختاری مشابه با شبکه سیامی دارد. ساختار شبکه‌های سیامی از دو زیرشبکه کانولوشن تشکیل شده است که در طی فرایند یادگیری این دو شبکه وزن‌های خود را به اشتراک می‌گذارند و به همین دلیل فرایند یادگیری آن‌ها دارای پیچیدگی محاسباتی کمتری است؛ در این مدل از ساختار سیامی جدید برای استخراج ویژگی‌های دقیق از تصاویر سنچس از دور استفاده شده است. روش آن‌ها بر روی سه

مجموعه داده UCM، AID و NWPU ارزیابی شده و توانسته است به دقتی بین ۹۱٪ تا ۹۹٪ دست یابد [۲۲].

انور و همکاران^{۱۱} ادغام چندین مدل رنگ عمیق را برای طبقه‌بندی صحنه‌های هوایی بررسی کردند. آن‌ها نشان دادند که ترکیب چندین مدل رنگ عمیق عملکرد تشخیص را در مقایسه با استفاده از شبکه RGB به تنهایی بهبود قابل توجهی می‌بخشد. آن‌ها از یک رویکرد فشرده‌سازی نظری اطلاعات^{۱۲} برای مقابله با ابعاد بالای داده استفاده کردند، که منجر به مجموعه‌ای از ویژگی‌های رنگ عمیق فشرده بدون افت قابل توجهی در دقت می‌شود. آزمایش‌های جامع بر روی پنج مجموعه داده صحنه سنچس از دور شامل: UC-Merced با ۲۱ طبقه صحنه، WHU-RS19 با نوزده نوع صحنه، RSSCN7 با هفت دسته، AID با سی طبقه صحنه هوایی، و NWPU-RESISC45 با ۴۵ صحنه انجام شده است. نتایج نشان می‌دهد که ادغام ویژگی‌های رنگ عمیق همیشه عملکرد طبقه‌بندی کلی را در مقایسه با ویژگی‌های عمیق RGB استاندارد بهبود داده و به دقتی بین ۹۰٪ تا ۹۷٪ دست یافته است [۲۳].

تیان و همکاران^{۱۳} یک شبکه متراکم چندمقیاسی با نام SEMSDNet^{۱۴} را پیشنهاد دادند. چارچوب شبکه کانولوشن متراکم چندمقیاسی با طبقه‌بندی کننده‌های متعدد و اتصالات متراکم می‌تواند خودکار به یک شبکه کوچک و یک شبکه عمیق باتوجه به پیچیدگی نمونه‌های آزمایشی و محدودیت منابع محاسباتی تبدیل شود؛ علاوه بر این، به منظور استخراج ویژگی‌های مؤثرتر، شبکه فشرده‌سازی و تحریک^{۱۵} در چارچوب درج شده است تا ویژگی‌های صحنه‌های مختلف را به صورت خودسازگار استخراج کند. نتایج تجربی بر روی مجموعه داده‌های UC Merced، NWPU-RESISC45، AID و SEMSDNet نسبت به روش‌های پیشرفته در عملکرد و کارایی برتری دارد و دقتی در حدود ۹۱٪ تا ۹۸٪ ارائه داده است [۲۴].

بی و همکاران^{۱۶} یک شبکه تجمع توجه چندمقیاسی (MS2AP)^{۱۷} را برای مقابله با محدودیت‌های شبکه‌های کانولوشن ساده پیشنهاد دادند که سه بخش اصلی دارد. نخست، می‌توان آن را به راحتی در مدل‌های کانولوشن فعلی جاسازی کرد تا قابلیت نمایش ویژگی برای

¹¹ Anwer et al

¹² Information theoretic compression

¹³ Tian et al

¹⁴ Multiscale Dense Networks with Squeeze and Excitation Attention

¹⁵ Squeeze-and-Excitation

¹⁶ Bi et al

¹⁷ Multi-Scale Staking Attention Pooling

¹ Shi et al

² Lightweight CNN Classification Method Based on Branch Feature Fusion

³ Bilinear feature extraction

⁴ Hung et al

⁵ Remote Sensing Scene Classification Network

⁶ Two-stage cyclical learning rate policy

⁷ No-freezing transfer learning

⁸ Local Interpretable Model, Agnostic Explanation

⁹ Tang et al

¹⁰ Attention Consistent Network

ژو و همکاران^۵ شبکه عصبی عمیق پیش آموزش دیده کانولوشن^۶ VggNet16 که بر روی ImageNet [۲۸] آموزش دیده است را پیشنهاد دادند. از این مدل از شبکه کانولوشن با عمق کم استفاده کرده است که این مورد از جمله محدودیت های آن محسوب می شود؛ اگر شبکه به اندازه کافی عمیق نباشد، نمی تواند اطلاعات معنایی انتزاعی بیشتری استخراج کند؛ برای استخراج ویژگی از تصاویر سنجش از دور استفاده شده است. ویژگی های تصاویر در لایه های کانولوشن سوم لایه سوم، کانولوشن چهارم لایه سوم و کانولوشن پنجم لایه سوم استخراج و این ویژگی ها بر روی شبکه عصبی گراف آموزش داده شد؛ سپس خروجی های شبکه گراف با خروجی لایه تمام متصل در مدل VggNet16 ادغام و طبقه بندی خطی انجام شده است. این روش بر روی چهار پایگاه داده UCM و NWPU-RESISC45، AID، RSSCN7 به دقتی بین ۸۹٪ تا ۹۴٪ رسیده است [۱۱].

۴- روش پیشنهادی

در این پژوهش، روشی برای سنجش و طبقه بندی تصاویر از راه دور با استفاده از یادگیری عمیق ارائه شده است. برای رفع مشکلاتی همچون عدم به کارگیری انواع مختلف ویژگی ها در تشخیص صحنه و نیز رفع محدودیت های راه کار های پیشین، یک رویکرد ترکیبی برای استخراج ویژگی ارائه شده است؛ در این رویکرد، سه نوع ویژگی شامل دو ویژگی عمیق محلی و سراسری و یک ویژگی محلی دستی، با یکدیگر ترکیب می شوند تا دقت تشخیص صحنه بهبود یابد. روش پیشنهادی شامل مراحل زیر در شکل (۳) نشان داده شده است.

۱. طراحی یک مدل بیست لایه عمیق سراسری و آموزش این مدل با مجموعه داده هدف.
۲. استخراج ویژگی های تصاویر مجموعه داده با استفاده از مدل آموزش داده شده بیست لایه عمیق.
۳. استخراج ویژگی های تصاویر با استفاده از مدل های از پیش آموزش دیده ResNet، InceptionNet، EfficientNet_b0 و GoogleNet.
۴. استخراج ویژگی های تصاویر با استفاده از الگوریتم های استخراج کننده ویژگی های محلی.
۵. ادغام ویژگی های استخراج شده از دو مدل یادگیری عمیق و یک ویژگی های محلی.
۶. استفاده از الگوریتم انتخاب گر ویژگی برای پیدا کردن مهم ترین ویژگی های هر تصویر.
۷. طبقه بندی تصاویر با استفاده از الگوریتم های طبقه بندی کننده ویژگی.
۸. ارزیابی کارایی مدل.

صحنه های سنجش از راه دور را افزایش دهد. دوم، یک ماژول توجه کانال-فضایی^۱ جدید برای استخراج محلی ویژگی های کلیدی در بردار ویژگی پیشنهاد دادند که در مقایسه با ماژول های فعلی، می تواند ویژگی های متمایز از بالا به پایین و ویژگی های پایین به بالا را هم از حوزه کانال و هم از حوزه فضایی ترکیب کند. سوم، یک اپراتور کانولوشن چندمقیاسی پیشنهاد دادند که می تواند نقشه های ویژگی چندمقیاسی را استخراج کند و اندازه های آن ها را ثابت نگه دارد. آزمایش های گسترده نشان می دهد که مدل آن ها در AID، UCM، NWPU به دقتی بین ۹۰٪ تا ۹۹٪ دست می یابد [۲۵].

می و همکاران^۲ یک چارچوب جدید مبتنی بر نمایش پراکنده برای طبقه بندی صحنه در سنجش از راه دور با ترکیب ویژگی های عمیق پیشنهاد دادند؛ در این مدل، ویژگی های چندسطحی ابتدا از لایه های مختلف شبکه کانولوشن استخراج می شوند تا به طور کامل از توانایی یادگیری ویژگی های کانولوشن بهره برداری کنند. در این مدل از شبکه های کانولوشن آموزش دیده موجود برای مثال AlexNet، VGGNet، و ResNet50، برای استخراج ویژگی استفاده شده است، که در آن ها هیچ نمونه برچسب گذاری شده ای مورد نیاز نیست؛ سپس، طبقه بندی مبتنی بر نمایش پراکنده برای ترکیب ویژگی های چندسطحی طراحی می شود، که به ویژه زمانی مؤثر است که فقط تعداد کمی از نمونه های آموزشی در دسترس باشد. نتایج تجربی بر روی دو مجموعه داده UC-Merced و WHU-RS19، نشان داد که روش پیشنهادی آن ها به دقت ۸۷٪ و ۸۸٪ رسیده است [۲۶].

شن و همکاران^۳ یک معماری دوگانه جدید با یک استراتژی گروه بندی-توجه-تلفیقی^۴ برای بهبود عملکرد طبقه بندی صحنه ارائه دادند. به طور خاص، این مدل از دو شبکه عصبی کانولوشن مختلف برای استخراج ویژگی استفاده می کند که در آن استراتژی گروه بندی-توجه-هم جوشی برای ترکیب ویژگی های کانولوشن ها به شیوه ای خوب و چندمقیاسی استفاده می شود؛ به این ترتیب، نمایش ویژگی های حاصل از صحنه افزایش می یابد؛ علاوه بر این، برای پرداختن به مسئله ظواهر مشابه بین صحنه های مختلف، یک تابع هزینه تعریف شده است که تنوع های کوچک درون طبقاتی و فواصل بین طبقاتی بزرگ را افزایش می دهد. آزمایش های گسترده روی چهار مجموعه داده UCM، WHU-RS19، AID و OPTIMAL-31 نشان می دهد که دقتی در حدود ۹۴٪ تا ۹۶٪ حاصل شده است [۲۷].

¹ residual channel-spatial attention module

² Mei et al

³ Shen et al.

⁴ grouping-attention-fusion

⁵ Xu et al.

⁶ Pretrained Convolutional Neural Networks

(جدول ۱-): ساختار شبکه بیست لایه
(Table-1): The structure of 20-layer network

لایه	نوع	تعداد پالایه	اندازه پالایه
۱	Conv2D	۶۴	۳*۳
۲	Conv2D	۶۴	۳*۳
۳	MaxPooling2D	-	۲*۲
۴	Conv2D	۱۲۸	۳*۳
۵	Conv2D	۱۲۸	۳*۳
۶	MaxPooling2D	-	۲*۲
۷	Conv2D	۲۵۶	۳*۳
۸	Conv2D	۲۵۶	۳*۳
۹	Conv2D	۲۵۶	۳*۳
۱۰	MaxPooling2D	-	۲*۲
۱۱	Conv2D	۵۱۲	۳*۳
۱۲	Conv2D	۵۱۲	۳*۳
۱۳	Conv2D	۵۱۲	۳*۳
۱۴	MaxPooling2D	-	۲*۲
۱۵	Conv2D	۵۱۲	۳*۳
۱۶	Conv2D	۵۱۲	۳*۳
۱۷	Conv2D	۵۱۲	۳*۳
۱۸	MaxPooling2D	-	۲*۲
۱۹	Flatten	-	-
۲۰	Dense	۷	-

در مرحله بعدی، ویژگی‌های محلی هر تصویر با الگوریتم‌های ORB و LBP استخراج می‌شود؛ ویژگی‌های استخراج‌شده ترکیب و پس از نرمال‌سازی بین صفر و یک، بهترین ویژگی‌ها برای هر تصویر با استفاده از الگوریتم‌های مربع کای و بسته‌بند انتخاب می‌شود. در نهایت با استفاده از چندین الگوریتم طبقه‌بندی مختلف شامل روش‌های انفرادی مثل ماشین بردار پشتیبان^۱ و روش‌های گروهی مثل جنگل تصادفی^۲ تشخیص صحنه انجام می‌شود.

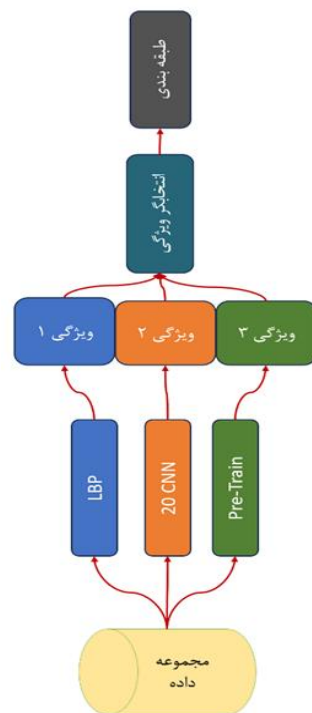
۵- مجموعه داده

مرحله نخست در راه‌کار پیشنهادی دریافت داده‌ها است. در این پژوهش، از چهار مجموعه داده به شرح زیر استفاده شده است:

- مجموعه داده نخست، پایگاه داده RSSNC7 است. این مجموعه داده را دانشگاه ووهان از Google Earth استخراج کرد و در مجموع ۲۸۰۰ تصویر دارد که به هفت دسته تقسیم می‌شوند و ابعاد هر تصویر 400×400 پیکسل است. هر دسته صحنه در چهار مقیاس مختلف (۱:۷۰۰، ۱:۱۳۰۰، ۱:۲۶۰۰ و ۱:۵۲۰۰) نمونه‌برداری می‌شوند. شکل (۴) برخی از تصاویر نمونه از این مجموعه را ارائه می‌دهد.

¹ Support Vector Machine

² Random Forest



(شکل ۳-): روش پیشنهادی
(Figure-3): Proposed approach

در راه‌کار پیشنهادی ابتدا داده‌ها از منبع مربوطه دریافت و مراحل پیش‌پردازش شامل نرمال‌سازی تصاویر و تغییر ابعاد آن‌ها اعمال می‌شود؛ سپس یک مدل کانولوشن بیست لایه تمام‌متصل طراحی می‌شود که هدف از آن، استفاده از یک مدل عمیق تمام‌متصل است که بتواند ویژگی‌های مهم را با توجه به عمیق‌بودن از تصاویر استخراج کند و با مجموعه داده هدف مورد آموزش قرارگیرد که این می‌تواند منجر به افزایش کیفیت ویژگی‌های استخراج‌شده شود. شبکه‌های عمیق از پیش آموزش‌دیده به شبکه‌های عصبی عمیقی اطلاق می‌شوند که پیش از استفاده در یک وظیفه خاص، روی یک مجموعه داده بزرگ آموزش دیده‌اند تا بتوانند به‌طور مؤثری ویژگی‌های مفید برای هر تصویر را استخراج کنند. برای هر پایگاه داده، ویژگی‌های تصاویر با استفاده از مدل‌های پیش‌آموزش‌دیده ResNet، InceptionNet، GoogleNet و EfficientNet_b0 استخراج می‌شود؛ سپس ویژگی‌های هر پایگاه داده با مدل بیست لایه تمام‌متصل که ساختار آن در جدول (۱) آمده است، استخراج می‌شود. در ساختار شبکه بیست لایه، لایه‌های یک تا هجده لایه‌های کانولوشن هستند و هر کدام دارای تعداد مشخصی پالایه است. اندازه پالایه در تمام لایه‌های کانولوشن سه در سه و لایه‌های سه، شش، ده، چهارده و هجده لایه‌های MaxPooling2D هستند. لایه نوزده، لایه Flatten، داده‌ها را به یک بردار تبدیل می‌کند. لایه بیست لایه Dense است که برای طبقه‌بندی استفاده می‌شود.

زمین شناسی ایالات متحده (USGS) جمع آوری کرده و در مجموع شامل ۲۱۰۰ تصویر است که به ۲۱ دسته تقسیم می شوند و هر تصویر دارای ابعاد 256×256 پیکسل با وضوح یک فوت است. شکل (۷) برخی از تصاویر نمونه از این مجموعه داده های معیار را نشان می دهد.



(شکل-۷): نمونه هایی از مجموعه داده UCM
(Figure-7): UCM Dataset Samples

۶- طراحی آزمایش ها

روش پیشنهادی با استفاده از زبان پایتون^۱ و کتابخانه پایتروچ^۲ پیاده سازی شده است؛ همچنین برای استفاده از توصیف گرهای محلی و روش های انتخاب ویژگی از کتابخانه های opencv, scikit learn, numpy, pandas استفاده و تمامی آزمایش ها در محیط Google Colab پیاده سازی شده است.

به منظور ارزیابی کارایی روش پیشنهادی از معیار دقت^۳ که به وسیله رابطه (۱) محاسبه می شود، استفاده شده است:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

(جدول-۲): پارامترهای مدل بیست لایه برای آموزش
(Table-2): The parameters of 20-layer model for training

پارامترها	مقدار
تکرار ^۴	۳۳۰
اندازه دسته ^۵	۶۴
تابع فعال ساز ^۶	Relu
بهینه ساز ^۷	Adam
نرخ یادگیری ^۸	۰.۰۰۱
درصد داده های آموزش ^۹	۸۰

پارامترهای مربوط به مدل مورد استفاده در راه کار پیشنهادی در جدول (۲) آورده شده است.

¹ Python
² Pytorch
³ Accuracy
⁴ epochs
⁵ Batch size
⁶ Activation function
⁷ Optimizer function
⁸ Learning rate
⁹ Training rate



(شکل-۴): نمونه هایی از مجموعه داده RSSNC7
(Figure-4): RSSNC7 Dataset Samples

- مجموعه داده دوم، پایگاه داده NWPU-RESISC45، یک مجموعه داده در مقیاس بزرگ با تنوع تصویر فراوان است؛ این مجموعه داده شامل ۳۱۵۰۰ تصویر رنگی در فضای رنگی RGB است که به ۴۵ دسته معنایی تقسیم شده اند. در NWPU-RESISC45، هر کلاس شامل هفتصد تصویر با اندازه ثابت 256×256 پیکسل و وضوح فضایی از حدود ۰/۲ به سی متر متغیر است. شکل (۵) برخی از تصاویر نمونه از این مجموعه داده را نشان می دهد.
- مجموعه داده سوم، موسوم به پایگاه داده AID را دانشگاه ووهان از Google Earth استخراج کرده و در مجموع شامل ده هزار تصویر و هر تصویر بر روی 600×600 پیکسل با وضوح بین ۰.۵ تا هشت متر ثابت شده است. مجموعه داده AID شامل سی دسته معنایی و تعداد تصاویر در هر کلاس از ۲۲۰ تا ۴۲۰ متغیر است. در شکل (۶) تعدادی از تصاویر این پایگاه داده ارائه شده است.



(شکل-۵): نمونه هایی از مجموعه داده NWPU-RESISC45
(Figure-5): NWPU-RESISC45 Dataset Samples



(شکل-۶): نمونه هایی از مجموعه داده AID
(Figure-6): AID Dataset Samples

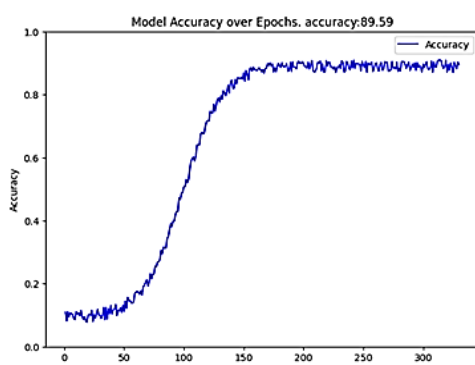
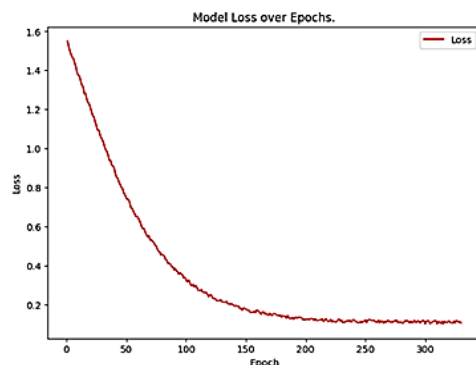
- مجموعه داده چهارم؛ یعنی پایگاه داده UCM را آزمایشگاه بینایی کامپیوتری دانشگاه کالیفرنیا از نقشه ملی سازمان

برای تکرار، عدد ۳۳۰ انتخاب شد؛ زیرا استفاده از تعداد بیشتر برای تکرار آموزش مدل، موجب خطای بیش‌برازش^۱ می‌شود. برای سرعت بالا در آموزش مدل، اندازه دسته ۶۴ انتخاب شد؛ همین‌طور از تابع فعال‌ساز Relu و بهینه‌ساز Adam با نرخ یادگیری ۰/۰۰۱ استفاده شده‌است؛ به این علت که در نرخ ۰/۰۱ مدل نمی‌توانست به‌خوبی آموزش ببیند و همچنین نرخ ۰/۰۰۰۱ منجر به‌کندشدن روند یادگیری شد.

که در هر تکرار چند ویژگی حذف شود تا زمانی که به تعداد بهینه ویژگی‌ها برسد. پارامتر اعتبارسنجی متقابل، تعداد فولدهای اعتبارسنجی متقاطع را برای استفاده در فرایند انتخاب ویژگی مشخص می‌کند. اعتبارسنجی متقابل، روشی برای ارزیابی عملکرد یک مدل یادگیری ماشینی با تقسیم داده‌ها به چندین زیرمجموعه و استفاده از هر زیرمجموعه برای آموزش و آزمایش است.

۶-۱- آزمایش نخست: بررسی اثر استفاده از مدل بیست لایه تمام‌متصل بر روی کارایی تشخیص صحنه ازراه‌دور

در مرحله نخست این آزمایش، مجموعه‌داده RSSNC7 به تعداد ۳۳۰ دور روی مدل آموزش داده می‌شود. همان‌طور که در شکل (۸) نمایش داده شده، ابتدای آموزش دقت مدل از ۱۲٪ شروع و پس از حدود صد دور آموزش به بالای ۸۰ درصد و تابع ضرر نیز پس از صد دور آموزش کمتر شده و پس از ۳۳۰ دور، دقت به ۸۹/۵۰٪ رسیده‌است.



(شکل ۸): مقادیر دقت و ضرر برای مجموعه‌داده RSSNC7
(Figure-8): Accuracy and Loss values for RSSNC7 dataset

در مرحله بعد، مدل با مجموعه‌داده UCM به تعداد ۳۳۰ دور آموزش داده می‌شود. همان‌طور که در شکل (۹) نمایش داده شده، ابتدای آموزش دقت مدل از ۴۱٪ شروع و پس از حدود صد دور آموزش به بالای ۹۰٪ و تابع ضرر پس از صد دور آموزش کمتر شده و پس از ۳۳۰ دور دقت به ۹۷/۲۱٪ رسیده‌است.

(جدول ۳): پارامترهای مربع کای

(Table-3): Chi-squared parameters

تعداد ویژگی برتر	آزمون آماری	نام مدل
۵۰	Chi-squared	SelectKBest

پارامترهای بعدی در جدول (۳) مربوط به انتخاب‌گر ویژگی بر اساس پالایه است که با استفاده از آزمون آماری مربع کای و انتخاب پنجاه ویژگی برتر برای هر تصویر آورده شده‌است.

در جدول (۳) تعداد ویژگی برتر به معنای این است که در نهایت پنجاه ویژگی برتر برای هر تصویر از پایگاه داده از بین ۱۵۰۰ ویژگی برای هر تصویر انتخاب می‌شود؛ همچنین آزمون آماری مورد استفاده در این مدل، آزمون مربع کای وابستگی بین دو متغیر را اندازه‌گیری کرده و فرضیه صفر مبنی بر مستقل بودن دو متغیر را ارزیابی می‌کند.

در زمینه انتخاب ویژگی، آزمون مربع کای برای امتیازدهی به هر ویژگی براساس وابستگی آن به متغیر هدف (یعنی برچسب‌ها) استفاده می‌شود. تابع chi2 در scikit-learn به‌عنوان تابع امتیازدهی برای انتخاب ویژگی با SelectKBest استفاده می‌شود. ورودی این تابع دو آرایه ماتریس ویژگی و متغیر هدف است. ماتریس ویژگی فقط مقادیر غیرمنفی و متغیر هدف حاوی دسته‌بندی است. پارامترهای مربوط به استفاده از روش بسته‌بند برای انتخاب ویژگی در جدول (۴) آورده شده‌است.

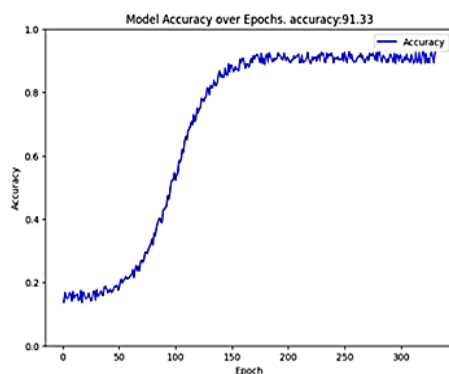
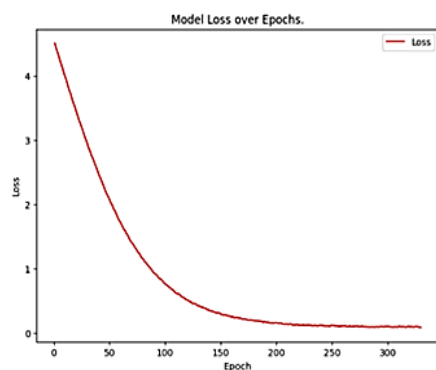
(جدول ۴): پارامترهای الگوریتم بسته‌بند

(Table-4): Wrapper algorithm parameters

اعتبارسنجی متقابل	گام	طبقه‌بندی	انتخاب‌گر
۱۰۰	۳۳	LogisticRegression	RFECV

در این جدول، پارامتر طبقه‌بندی، رگرسیون لجستیک است که دسته‌بندی ویژگی‌ها در انتخاب‌گر را انجام می‌دهد. پارامتر گام^۲ در نمونه RFECV مشخص می‌کند

^۱ overfitting
^۲ Step



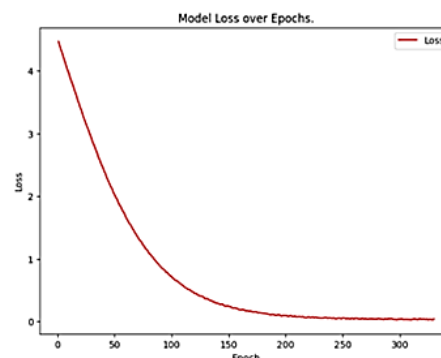
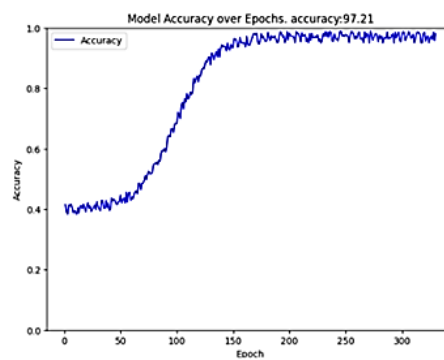
(شکل-۱۱): مقادیر دقت و ضرر برای مجموعه داده AID
(Figure-11): Accuracy and Loss values for AID dataset

در مرحله آخر، مدل با مجموعه داده AID به تعداد ۳۳۰ دور آموزش داده می شود. همان طور که در شکل (۱۱) نمایش داده شده، ابتدای آموزش دقت مدل از ۱۷٪ شروع و پس از حدود صد دور آموزش به بالای ۸۵ درصد رسید. تابع ضرر پس از صد دور آموزش کمتر شده و پس از ۳۳۰ دور، دقت به ۹۱/۳۳٪ رسیده است.

۲-۶- آزمایش دوم: بررسی مدل های از پیش آموزش داده شده بر روی کارایی تشخیص صحنه از راه دور

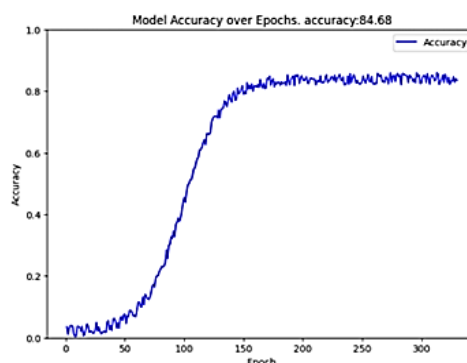
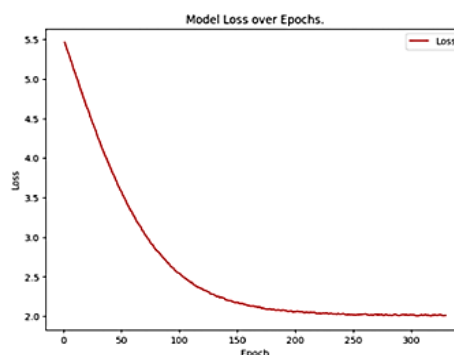
در این آزمایش به بررسی مدل های از پیش آموزش دیده بر روی کارایی تشخیص صحنه می پردازیم. در این مرحله با استفاده از چهار مدل ResNet, GoogleNet, EfficientNet, InceptionNet_b0 تصاویر انجام می شود؛ ابتدا تمام تصاویر مجموعه داده ها به مدل های از پیش آموزش داده شده وارد و برای هر تصویر هزار ویژگی استخراج می شود.

دقت ویژگی های استخراج شده برای مجموعه داده های مختلف با هر مدل در شکل (۱۲) نشان داده شده است. با توجه به نتایج به دست آمده مدل ResNet توانست ویژگی های دقیق تری نسبت به سایر مدل ها استخراج کند و پس از آن InceptionNet_b0 و سپس GoogleNet و در آخر EfficientNet قرار دارد.



(شکل-۹): مقادیر دقت و تابع ضرر برای مجموعه داده UCM
(Figure-9): Accuracy and Loss values for UCM dataset

سپس مدل با مجموعه داده NWPU45 به تعداد ۳۳۰ دور آموزش داده می شود. همان طور که در شکل (۱۰) نمایش داده شده، ابتدای آموزش دقت مدل از ۷٪ شروع و پس از حدود صد دور آموزش به بالای ۷۵٪ و تابع ضرر پس از صد دور آموزش کمتر شده و پس از ۳۳۰ دور، دقت به ۸۴/۶۸٪ رسیده است.

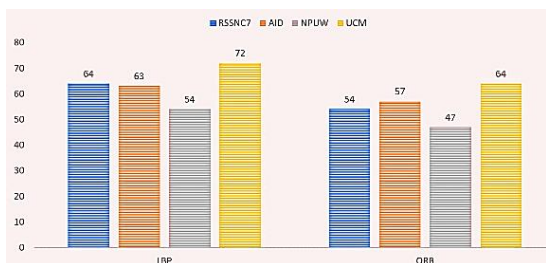


(شکل-۱۰): مقادیر دقت و ضرر برای مجموعه داده NWPU45
(Figure-10): Accuracy and Loss values for NWPU45 dataset

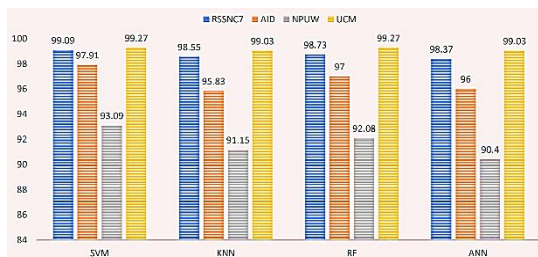
در این آزمایش به بررسی اثر استفاده از الگوریتم‌های بسته‌بند و مربع کای به‌عنوان انتخاب‌گر ویژگی برای ویژگی‌های ترکیب‌شده بر روی کارایی تشخیص صحنه از راه دور می‌پردازیم. دقت و نتایج به‌دست‌آمده برای هر الگوریتم در شکل (۱۴) مشاهده می‌شود. هر دو الگوریتم، پنجاه ویژگی برتر را برای هر تصویر از ۱۵۰۰ ویژگی کلی انتخاب می‌کنند؛ دلیل انتخاب پنجاه ویژگی آن است که این تعداد، بیشترین ویژگی مؤثر را برای هر تصویر دربردارد و با تعداد کمتر یا بیشتر از این مقدار، دقت ویژگی‌های انتخاب‌شده کاهش پیدا می‌کند. برای طبقه‌بندی ویژگی‌ها نیز از روش ماشین بردار پشتیبان استفاده شده‌است. با توجه به نتایج به‌دست‌آمده می‌توان گفت که الگوریتم بسته‌بند، روش بهتری برای انتخاب ویژگی‌های تصویر برای تشخیص صحنه از راه دور است.

۶-۵- آزمایش پنجم: بررسی اثر استفاده از الگوریتم‌های طبقه‌بندی بر کارایی تشخیص صحنه از راه دور

الگوریتم‌های طبقه‌بندی، ویژگی‌های استخراج‌شده به‌وسیله انتخاب‌گر ویژگی را دسته‌بندی می‌کنند. پس از استخراج ویژگی‌ها از دو مدل و توصیف‌گر ویژگی محلی و ترکیب آن‌ها، سپس استفاده از انتخاب‌گر ویژگی و طبقه‌بندی تصاویر، خروجی نهایی بر روی چهار پایگاه داده به‌دست آمده‌است.

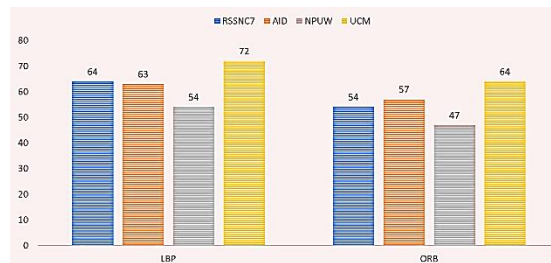


(شکل-۱۴) دقت الگوریتم‌های انتخاب‌گر ویژگی
(Figure-14) Accuracy of feature selection algorithms

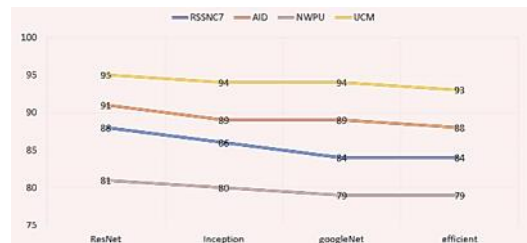


(شکل-۱۵) دقت الگوریتم‌های طبقه‌بندی
(Figure-15): Accuracy of classification algorithms

در این آزمایش به بررسی دقت مدل با استفاده از چهار الگوریتم طبقه‌بندی ماشین بردار پشتیبان^۱، جنگل



(شکل-۱۲) دقت ویژگی‌های مدل‌های از پیش آموزش‌دیده
(Figure-12): Accuracy of features of pre-trained models



(شکل-۱۳) دقت ویژگی‌های توصیف‌گرهای محلی
(Figure-13): Accuracy of features of local descriptors

۶-۳- آزمایش سوم: بررسی اثر استفاده از توصیف‌گرهای محلی بر روی کارایی تشخیص صحنه از راه دور

در این آزمایش به بررسی اثر استفاده از توصیف‌گرهای محلی بر روی کارایی تشخیص صحنه از راه دور می‌پردازیم؛ در این مرحله از دو الگوریتم ORB و LBP استفاده می‌کنیم. هر کدام از این دو الگوریتم، تصاویر مجموعه داده‌ها را دریافت و الگوریتم LBP برای هر تصویر ۴۲۰ و ORB تعداد شش هزار ویژگی استخراج می‌کند؛ پس از آن دقت ویژگی‌های استخراج‌شده با الگوریتم ماشین بردار پشتیبان به‌دست‌آمده که در شکل (۱۳) نمایش داده شده‌است. همان‌طور که مشاهده می‌شود ویژگی‌های استخراج‌شده به‌وسیله الگوریتم LBP دقیق‌ترند.

۶-۴- آزمایش چهارم: بررسی اثر استفاده از بسته‌بند و مربع کای به‌عنوان انتخاب‌گر ویژگی بر روی کارایی تشخیص صحنه از راه دور

برای ادغام ویژگی‌های استخراج‌شده، ابتدا تمام ویژگی‌ها برای هر تصویر به‌صورت یک سطر درآمده‌است و سپس آرایه‌ها به هم متصل می‌شوند. در نهایت برای هر تصویر تمامی ویژگی‌ها در یک سطر قرار می‌گیرند. قبل از استفاده از الگوریتم‌های انتخاب‌گر ویژگی، ویژگی‌های به‌دست‌آمده از مدل بیست لایه و ویژگی‌های مدل‌های از پیش آموزش‌دیده و ویژگی‌های توصیف‌گر محلی با هم ترکیب و برای هر تصویر در یک سطر ذخیره شده‌اند.

¹ Support Vector Machine

(جدول-۶): درصد دقت برای سنجش تصاویر از راه دور

برای پایگاه داده UCM (نسبت تمرین ۸۰٪)

(Table-6): Percentage of accuracy for remote sensing images for UCM database (training ratio 80%)

روش	دقت
BoVW(LBP) [28]	۷۷.۱۲
BoVW(SIFT) [28]	۷۴.۱۲
VGG19+WSPM-CRC [29]	۹۶.۱۴
Resnet+Hybrid-KCRC [30]	۹۷.۲۹
GoogleNet [32]	۹۴.۳۱
VGG-VD-16 [29]	۹۵.۲۱
VGG-M+IFK [28]	۹۶.۹
MLDS [26]	۹۷.۷۸
DFAGCN [14]	۹۸.۴۸
Proposed Approach	۹۹.۲۷

(جدول-۷): درصد دقت برای سنجش تصاویر از راه دور

برای پایگاه داده NWPU (نسبت تمرین ۲۰٪)

(Table-7): Percentage of accuracy for remote sensing images for NWPU database (training ratio 20%)

روش	دقت
BoVW(LBP) [28]	۴۴.۹۷
BoVW(SIFT) [28]	۳۲.۹۶
AlexNet+SAFF [33]	۸۴.۰۰
VGG-VD16+SAFF [33]	۸۷.۸۶
AlexNet [12]	۷۹.۸۵
VGGNET [12]	۷۹.۷۹
GoogleNET [32]	۷۸.۴۸
MIDV-Net_CS [33]	۸۷.۹۹
RADC-Net [34]	۸۷.۶۳
DFAGCN [14]	۸۹.۲۹
Proposed Approach	۹۳.۰۹

(جدول-۸): درصد دقت برای سنجش تصاویر از راه دور

برای پایگاه داده AID (نسبت تمرین ۵۰٪)

(Table-8): Percentage of accuracy for remote sensing images for AID database (training ratio 50%)

روش	دقت
BoVW(LBP) [28]	۶۴.۲۵
BoVW(SIFT) [28]	۶۷.۶۵
VGG19+WSPM-CRC [29]	۹۲.۵۷
Resnet+Hybrid-KCRC [30]	۹۲.۸۷
CaffeNet [31]	۸۹.۵۲
VGG-VD-16 [31]	۸۹.۶۴
GoogleNET [31]	۸۷.۳۹
VGG-VD16+SAFF [33]	۹۳.۸۳
MF2NET [35]	۹۴.۸۴
VGG-16-CapsNet [36]	۹۴.۷۱
DFAGCN [14]	۹۴.۸۸
Proposed Approach	۹۷.۹۱

تصادفی^۱، شبکه عصبی مصنوعی آبشاری^۲ و k نزدیک ترین همسایه^۳ می پردازیم. نتایج ارزیابی در شکل (۱۵) نمایش داده شده است.

به منظور نمایش خروجی روش پیشنهادی، برای هر پایگاه داده، ترکیبی که بیشترین دقت را دارد آورده شده است. طبق نتایج، الگوریتم ماشین بردار پشتیبان، بالاترین دقت را در بین طبقه بندیها دارد.

۷-ارزیابی روش پیشنهادی

به منظور ارزیابی راه کار پیشنهادی، عملکرد آن با تعدادی پژوهش مشابه پیشین مقایسه شد. در بیشتر روش هایی که تاکنون ارائه شده اند، تنها از ویژگی های سراسری عمیق استخراج شده به وسیله شبکه های کانولوشن استفاده شده است؛ در واقع بیشتر پژوهش ها مبتنی بر یک نوع خاص ویژگی مثل سراسری و یا محلی بوده اند و برای رسیدن به دانش کافی از تصاویر، بیشتر از شبکه های کانولوشن پیش آموزش دیده استفاده کرده اند. در این حالت اگر عمق شبکه به اندازه کافی مناسب نباشد ویژگی های استخراج شده قدرت افتراق بالایی ندارند و در نتیجه تشخیص صحنه به خوبی انجام نمی شود؛ از این رو، در این پژوهش از ترکیب ویژگی های عمیق سراسری که از شبکه های کانولوشن استخراج می شود در کنار ویژگی های محلی دستی و عمیق که از توصیف گرهای محلی به دست می آید، استفاده شده است؛ همچنین در این پژوهش، یک مرحله انتخاب ویژگی ترکیبی مبتنی بر روش های مربع کای و بسته بند نیز برای رسیدن به ویژگی های مناسب در مدل وجود دارد. مطالعات نشان می دهد تاکنون از مدل ترکیبی بالا برای تشخیص صحنه در سنجش از دور استفاده نشده است.

(جدول-۵): درصد دقت برای سنجش تصاویر از راه دور برای

پایگاه داده RSSNC7 (نسبت تمرین ۵۰٪)

(Table-5): Percentage of accuracy for remote sensing images for RSSNC7 database (training ratio 50%)

روش	دقت
BoVW(LBP) [28]	۸۱.۶۹
BoVW(SIFT) [28]	۸۱.۳۴
VGG19+WSPM-CRC [29]	۹۱.۳۱
Resnet+Hybrid-KCRC [30]	۹۲.۶۲
Two-stage-deep-feature [31]	۹۲.۳۷
DFAGCN [14]	۹۴.۱۴
Proposed Approach	۹۹.۰۹

نتایج ارزیابی ها روی چهار مجموعه داده UCM، RSSNC7، NWPU و AID به ترتیب در جدول های (۵) تا (۸) آمده است. با توجه به نتایج نشان داده شده، راه کار پیشنهادی برای تمام مجموعه داده ها دقت بالاتری نسبت به روش های مورد مقایسه دارد.

¹ Random Forest

² Cascade ANN

³ KNN

به منظور توسعه پژوهش جاری می توان ترکیب ویژگی های استخراج شده از مدل دستی کانولوشنی بیست لایه و پیش آموزش این مدل با تصاویر مجموعه داده هدف را که باعث استخراج ویژگی های بهتر و دقیق تری از تصاویر می شود، در نظر گرفت.

9-Reference

۹-مراجع

- [1] Q. Hu et al., "Exploring the use of Google Earth imagery and object-based methods in land use/cover mapping," *Remote Sensing*, vol. 5, no. 11, pp. 6026-6042, 2013.
- [2] L. Gómez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, "Multimodal classification of remote sensing images: A review and future directions," *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1560-1584, 2015.
- [3] A. Tayyebi, B. C. Pijanowski, and A. H. Tayyebi, "An urban growth boundary model using neural networks, GIS and radial parameterization: An application to Tehran, Iran," *Landscape and Urban Planning*, vol. 100, no. 1-2, pp. 35-44, 2011.
- [4] Z. Y. Lv, W. Shi, X. Zhang, and J. A. Benediktsson, "Landslide inventory mapping from bitemporal high-resolution remote sensing images using change detection and multiscale segmentation," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 11, no. 5, pp. 1520-1532, 2018.
- [5] T. R. Martha, N. Kerle, C. J. Van Westen, V. Jetten, and K. V. Kumar, "Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis," *IEEE transactions on geoscience and remote sensing*, vol. 49, no. 12, pp. 4928-4943, 2011.
- [6] F. Ghazouani, I. R. Farah, and B. Solaiman, "A multi-level semantic scene interpretation strategy for change interpretation in remote sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8775-8795, 2019.
- [7] Y. Li, Y. Zhang, X. Huang, and A. L. Yuille, "Deep networks under scene-level supervision for multi-class geospatial object detection from remote sensing images," *ISPRS journal of photogrammetry and remote sensing*, vol. 146, pp. 182-196, 2018.
- [8] Y. Gu, Y. Wang, and Y. Li, "A survey on deep learning-driven remote sensing image scene understanding: Scene classification, scene retrieval and scene-guided object detection," *Applied Sciences*, vol. 9, no. 10, p. 2110, 2019.
- [9] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3735-3756, 2020.

در نهایت، مقایسه ای بین استخراج کننده ویژگی محلی LBP، مدل از پیش تمرین داده شده Resnet، شبکه بیست لایه تمام متصل با راه کار پیشنهادی، شامل ترکیب این سه مدل، انجام شده است. نتیجه در جدول (۹) آمده و برای هر چهار مجموعه داده دقت راه کار پیشنهادی مقدار بهتری را نشان می دهد. دلیل این برتری، استفاده از ترکیب ویژگی های محلی و سراسری استخراج شده و معماری مدل بیست لایه از پیش آموزش دیده است؛ همچنین برای استخراج ویژگی ها از مدل های ResNet101, CNN20layer, LBP و برای انتخاب ویژگی های ترکیب شده، از الگوریتم بسته بند استفاده شده و طبقه بندی برعهده ماشین بردار پشتیبان بوده است.

(جدول ۹): مقایسه بین مدل های مختلف

(Table-9): Comparison between different models

پایگاه داده	روش	دقت
RSSNC7	LBP	۶۴.۳۳
	Pre-Train	۸۸.۶۸
	20CNN	۸۹.۵۰
	Proposed Approach	۹۹.۰۹
NWPU	LBP	۵۴.۳۵
	Pre-Train	۸۱.۹۳
	20CNN	۸۴.۶۸
	Proposed Approach	۹۳.۰۹
UCM	LBP	۷۲.۲۵
	Pre-Train	۹۵.۶۳
	20CNN	۹۷.۲۱
	Proposed Approach	۹۹.۲۷
AID	LBP	۶۳.۳۸
	Pre-Train	۹۱.۱۲
	20CNN	۹۱.۳۳
	Proposed Approach	۹۴.۳۱

۸-نتیجه گیری و پیشنهادهای آینده

در حال حاضر، روش هایی برای سنجش تصاویر از راه دور ارائه شده که مبتنی بر استفاده از روش های یادگیری عمیق و شبکه های از پیش آموزش دیده اند. در روش پیشنهادی که بر مبنای ترکیب دو نوع ویژگی عمیق و یک نوع ویژگی محلی است، با استخراج و ترکیب این ویژگی ها و سپس ادغام آنها مهم ترین ویژگی های هر تصویر با استفاده از الگوریتم های انتخاب گر ویژگی انتخاب می شود. ارزیابی این راه کار نشان داد که معماری پیشنهادی به همراه استفاده از مدلی که با مجموعه داده هدف آموزش دیده باشد باعث افزایش عملکرد در تشخیص صحنه در سنجش از راه دور می شود.

- [22] X. Tang, Q. Ma, X. Zhang, F. Liu, J. Ma, and L. Jiao, "Attention consistent network for remote sensing scene classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2030-2045, 2021.
- [23] R. M. Anwer, F. S. Khan, and J. Laaksonen, "Compact deep color features for remote sensing scene classification," *Neural Processing Letters*, vol. 53, no. 2, pp. 1523-1544, 2021.
- [24] T. Tian, L. Li, W. Chen, and H. Zhou, "SEMSDNet: A multiscale dense network with attention for remote sensing scene classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 5501-5514, 2021.
- [25] Q. Bi, H. Zhang, and K. Qin, "Multi-scale stacking attention pooling for remote sensing scene classification," *Neurocomputing*, vol. 436, pp. 147-161, 2021.
- [26] S. Mei, K. Yan, M. Ma, X. Chen, S. Zhang, and Q. Du, "Remote sensing scene classification using sparse representation-based framework with deep feature fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 5867-5878, 2021.
- [27] J. Shen, T. Zhang, Y. Wang, R. Wang, Q. Wang, and M. Qi, "A Dual-Model Architecture with Grouping-Attention-Fusion for Remote Sensing Scene Classification," *Remote Sensing*, vol. 13, no. 3, p. 433, 2021.
- [28] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 248-255: Ieee.
- [29] B.-D. Liu, J. Meng, W.-Y. Xie, S. Shao, Y. Li, and Y. Wang, "Weighted spatial pyramid matching collaborative representation for remote-sensing-image scene classification," *Remote Sensing*, vol. 11, no. 5, p. 518, 2019.
- [30] B.-D. Liu, W.-Y. Xie, J. Meng, Y. Li, and Y. Wang, "Hybrid collaborative representation for remote-sensing image scene classification," *Remote Sensing*, vol. 10, no. 12, p. 1934, 2018.
- [31] Y. Liu, Y. Liu, and L. Ding, "Scene classification based on two-stage deep feature fusion," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 183-186, 2017.
- [32] T. Gong, X. Zheng, and X. Lu, "Remote Sensing Scene Classification with Multi-task Learning," in *Proceedings of the 7th China High Resolution Earth Observation Conference (CHREOC 2020)*, 2022, pp. 403-418: Springer.
- [33] R. Cao, L. Fang, T. Lu, and N. He, "Self-attention-based deep feature fusion for remote sensing scene classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 43-47, 2020.
- [10] W. Wang, Y. Chen, and P. Ghamisi, "Transferring CNN With Adaptive Learning for Remote Sensing Scene Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-18, 2022.
- [11] K. Xu, H. Huang, P. Deng, and Y. Li, "Deep feature aggregation framework driven by graph convolutional network for scene classification in remote sensing," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
- [12] S. Aggarwal, "Principles of remote sensing," *Satellite remote sensing and GIS applications in agricultural meteorology*, vol. 23, no. 2, pp. 23-28, 2004.
- [13] M. a. Ahmadi and R. Dianat, "Introducing a method for extracting features from facial images based on applying transformations to features obtained from convolutional neural networks," (in eng), *Signal and Data Processing, Research* vol. 17, no. 3, pp. 141-156, 2020.
- [14] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial intelligence review*, vol. 53, no. 8, pp. 5455-5516, 2020.
- [15] M. A. Zare Chahooki and z. khalifeh zadeh, "A General Investigation on the Combination of Local and Global Feature Selection Methods for Request Identification on Telegram," (in eng), *Signal and Data Processing, Applicable* vol. 19, no. 2, pp. 175-196, 2022.
- [16] M. Imani and H. Ghassemian, "Supervised Feature Extraction of Face Images for Improvement of Recognition Accuracy," (in eng), *Signal and Data Processing, Research* vol. 16, no. 1, pp. 158-172, 2019.
- [17] J. Zhang, C. Lu, X. Li, H.-J. Kim, and J. Wang, "A full convolutional network based on DenseNet for remote sensing scene classification," *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 3345-3367, 2019.
- [18] H. Sun, S. Li, X. Zheng, and X. Lu, "Remote sensing scene classification by gated bidirectional network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 1, pp. 82-96, 2019.
- [19] J. Xie, N. He, L. Fang, and A. Plaza, "Scale-free convolutional neural network for remote sensing scene classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6916-6928, 2019.
- [20] C. Shi, T. Wang, and L. Wang, "Branch feature fusion convolution network for remote sensing scene classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5194-5210, 2020.
- [21] S.-C. Hung, H.-C. Wu, and M.-H. Tseng, "Remote sensing scene classification and explanation using RSSCNet and LIME," *Applied Sciences*, vol. 10, no. 18, p. 6151, 2020.



- [34] Q. Bi, K. Qin, H. Zhang, Z. Li, and K. Xu, "RADC-Net: A residual attention based convolution network for aerial scene classification," *Neurocomputing*, vol. 377, pp. 345-359, 2020.
- [35] K. Xu, H. Huang, Y. Li, and G. Shi, "Multilayer feature fusion network for scene classification in remote sensing," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 11, pp. 1894-1898, 2020.
- [36] W. Zhang, P. Tang, and L. Zhao, "Remote sensing image scene classification using CNN-CapsNet," *Remote Sensing*, vol. 11, no. 5, p. 494, 2019.



سید علی جهانی دارای مدرک کارشناسی ارشد رشته مهندسی کامپیوتر گرایش هوش مصنوعی است. زمینه اصلی فعالیت و موضوعات پژوهشی مورد علاقه ایشان، هوش مصنوعی، پردازش تصویر و یادگیری عمیق است. نشانی رایانامه ایشان عبارت است از:

Alijahani1919@gmail.com



کیوان محبی استادیار گروه مهندسی کامپیوتر دانشگاه آزاد اسلامی واحد اصفهان (خوراسگان) و دارای مقالات متعددی در کنفرانس‌های ملی و بین‌المللی و نشریات معتبر علمی است و عضویت در کمیته علمی و داوری کنفرانس‌ها و نشریات علمی را در کارنامه خود دارد. موضوعات پژوهشی مورد علاقه ایشان عبارتند از: مهندسی نرم‌افزار، معماری نرم‌افزار، آزمون نرم‌افزار، وب معنایی و شبکه‌های اجتماعی. نشانی رایانامه ایشان عبارت است از:

K.mohebbi@khuif.ac.ir



فرساد زمانی بروجنی در سال ۲۰۱۳ مدرک دکترای خود را در رشته کامپیوتر از دانشگاه پوترا مالزی دریافت کرد. ایشان هم‌اکنون استادیار گروه مهندسی کامپیوتر دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران است. زمینه‌های پژوهشی مورد علاقه ایشان داده‌کاوی، یادگیری ماشین و تجزیه و تحلیل داده‌های حجیم است. نشانی رایانامه ایشان عبارت است از:

Farsad.zamani@srbiau.ac.ir