

بهبود طبقه‌بندی تصاویر ابر طیفی با استفاده از مدل ترکیبی شبکه‌های کپسول و درخت تصمیم تقویتی

چکیده

با توسعه علم سنجش از دور، استفاده از تصاویر هایپراسپکترال روز به روز گسترده تر می شود. طبقه بندی یکی از محبوب ترین موضوعات در سنجش از دور ابرطیفی است. در دو دهه گذشته، تعداد زیادی روش ها برای مقابله با مشکل طبقه بندی داده های هایپراسپکترال پیشنهاد شده است که در پژوهش حاضر از ساختاری مبتنی بر یادگیری شبکه‌های کپسول برای طبقه‌بندی تصاویر ابر طیفی استفاده شده است، به گونه‌ای که ساختار شبکه بتواند با استفاده از یک لایه کانوولوشنی و یک لایه کپسول بهترین حالت تولید ویژگی‌ها را داشته باشد و در عین حال از بیش‌برازش شبکه بر روی نمونه‌های آموزشی جلوگیری کند. نتایج به دست آمده نشان از کیفیت بالای ویژگی‌های تولیدی در ساختار پیشنهادی دارد. در راستای بهبود دقت طبقه‌بندی، رویکرد استخراج ویژگی از طریق شبکه‌ی طراحی‌شده و طبقه‌بندی توسط الگوریتم درخت تقویتی XGBoost، با روش طبقه‌بندی توسط شبکه‌ی عمیق سراسری مقایسه شد تا علاوه بر بررسی و کیفیت‌سنجی ویژگی‌های عمیق برداری تولیدی توسط روش پیشنهادی در طبقه بندی کننده‌های مختلف، میزان توانایی شبکه‌های عمیق سراسری نیز در کاربرد طبقه‌بندی مورد بررسی قرار گیرد. رویکرد کپسول پیشنهادی شامل ۳ لایه اساسی: (۱) Prime caps که کپسول‌هایی به اندازه ۸ و ۳۲ فیلتر 9×9 و گام حرکتی (۲، ۲) Digitcaps دارای ۱۰ کپسول ۱۶ بعدی و (۳) لایه تماماً متصل می باشد. که نتایج حاصل از بررسی دو رویکرد برای شبکه‌ی عمیق و نیز ترکیب شبکه‌های کپسول با الگوریتم درخت تقویتی XGBoost مقایسه شد. رویکرد های نظیر LSTM, GRU, RF-۲۰۰, SVM, GRU-Pretanh و برای مقایسه ی رویکرد پیشنهادی بر اساس پیکربندی هایی که در تحقیقات ان ها اشاره شده بود در نظر گرفته شدند. برای ارزیابی مدل پیشنهادی همچنین مجموعه داده ی Indian Pines که شامل ۱۶ کلاس مختلف می باشد مورد استفاده قرار گرفت. با استفاده از روش پیشنهادی ترکیبی، طبقه‌بندی تصاویر با دقت ۹۹ درصد بر روی داده‌های آموزش و دقت ۹۷/۵ درصد بر روی داده‌های تست انجام می‌شود.

واژگان کلیدی: طبقه‌بندی، تصاویر ابر طیفی، شبکه‌های کپسول، درخت تصمیم تقویتی، مدل ترکیبی.

Improving the classification of hyperspectral images using the combined model of CapsNet and the Extreme Gradient Boosting

Abstract

With the development of remote sensing science, the use of hyperspectral images is becoming more widespread. Classification is one of the most popular topics in hyperspectral remote sensing. In the last two decades, a number of methods have been proposed to address the problem of hyperspectral data classification. In the present study, a structure based on learning capsule networks has been used to classify hyperspectral images, so that the network structure can have the most optimal generation of features by using a convolution layer and a capsule layer, and at the same time avoid overfitting on training data. The obtained results show the high quality of production features in the proposed structure.

In order to improve the classification accuracy, the feature extraction approach through the designed network and the classification by the Extreme Gradient Boosting was compared with the classification method by the global deep network. The proposed capsule approach consists of 3 basic layers: 1) Prime caps, which are capsules of size 8 and 32 with 9×9 filters and movement step 2, 2) Digitcaps with 10 16-dimensional capsules, and 3) fully connected layer. The results of examining two approaches for deep networking as well as combining capsule networks with XGBoost reinforcement tree algorithm were compared. Approaches such as SVM, RF, LSTM, GRU and GRU-Pretanh were considered to compare the proposed approach based on the configurations mentioned in their research.

Up in addition to the study and quality measurement of production vector deep features by the proposed method in different classifiers, the ability of deep global networks in the application of classification should also be examined.

The results of examining two approaches for deep network and also combining CapsNet with XGBoost show that by using the proposed combined method, images are classified with 99% accuracy on training data and 97.5% accuracy on test data.

Keywords:

Classification, Hyperspectral Images, CapsNet, XGBoost ,Fusion Model

سنجنده‌های ابر طیفی^۱ نوع خاصی از سنجنده‌های طیف‌سنجی هستند که پهنای باند موردنظر را به صدها باند باریک مجاور هم تقسیم می‌کنند. در دسترس قرار گرفتن سنجنده‌هایی با توان تفکیک طیفی^۲ بالا و تعداد باند زیاد، دستیابی به جزئیات بیشتر در مورد کلاس‌ها را امکان‌پذیر ساخته و باعث افزایش دقت و اعتبار طبقه‌بندی شده است. با توجه به تعداد زیاد باندها، کاربردهای بسیار زیادی برای تصاویر ابر طیفی وجود دارد (Ghassemian H. et al., ۲۰۰۳). استفاده از تصاویر ابر طیفی در کاربردهایی مانند کشاورزی و برآورد دقیق محصولات، جنگل‌داری و مطالعه پوشش‌های گیاهی، مطالعه منابع آب و شناسایی آلودگی‌ها، اکتشاف منابع طبیعی و معدنی، زمین‌شناسی و نفت، کاربردهای صلح‌آمیز از قبیل شناسایی و پاک‌سازی مناطق جنگی آلوده، کاربردهای نظامی و شناسایی ادوات استتار شده، نتایج خوبی بدست آورده است.

مهم‌ترین چالشی که در پردازش تصاویر ابر طیفی وجود دارد، کم بودن تعداد نقاط آموزشی نسبت به تعداد باندهای تصویر است. در واقع مسئله این است که اگر تعداد نقاط آموزشی نسبت به تعداد باندهای داده کمتر از ۱۰۰ نمونه به ازای هر کلاس باشد، خطای تخمین پارامترهای طبقه‌بندی کننده با افزایش تعداد باندها بیشتر می‌شود (Keshavarz A. & Ghasemiyan H., ۲۰۰۵). با توجه به مسئله ذکر شده، یک راه حل برای بهبود عملکرد طبقه‌بندی کننده‌های آماری، کاهش فضای ویژگی قبل از طبقه‌بندی این تصاویر است که منجر به افزایش نسبت تعداد نقاط آموزشی به تعداد ویژگی می‌شود. برای این منظور دو دسته الگوریتم مورد استفاده قرار می‌گیرد که عبارتند از الگوریتم‌های انتخاب باند^۳ و الگوریتم‌های کاهش ابعاد فضای ویژگی^۴. در سال‌های اخیر، الگوریتم‌های متعددی جهت استخراج ویژگی مطرح و مورد استفاده قرار گرفته‌اند. استخراج ویژگی با مرز تصمیم‌گیری^۵ (Lee C. & Landgrebe DA., ۱۹۹۳)، نگاشت جستجوگر (Jimenez L. & Landgrebe DA., ۱۹۹۹)، تحلیل ممیز و تبدیل موجک^۶ (Kaewpijit S. et al., ۲۰۰۳)، تحلیل مؤلفه‌های اصلی^۷، تحلیل مؤلفه‌های مستقل^۸ (Du P. et al., ۲۰۰۸)، از جمله الگوریتم‌هایی هستند که به منظور کاهش ابعاد فضای ویژگی و بهبود طبقه‌بندی تصاویر ابر طیفی مورد استفاده قرار گرفته‌اند. در بین سال‌های ۱۹۹۸ تا ۲۰۰۰، گوالتری^۹ و همکاران کارایی ماشین‌های بردار پشتیبان^{۱۰} در طبقه‌بندی تصاویر ابر طیفی را بررسی کرده، نشان دادند این روش امکان آنالیز تصاویر ابر طیفی در فضای ویژگی با ابعاد زیاد و بدون نیاز به فرآیند کاهش ویژگی را فراهم می‌کند (Chetri SR ۲۰۰۰). همچنین ملگانی^{۱۱} و همکاران (۲۰۰۴) نیز کارایی ماشین‌های بردار پشتیبان در طبقه‌بندی تصاویر ابر طیفی را مطالعه نمودند (Melgani F. & Bruzzone L., ۲۰۰۴). لیانگ^{۱۲} و همکاران (۲۰۰۵) ماشین بردار رگرسیون^{۱۳} را برای طبقه‌بندی تصاویر ابر طیفی پیشنهاد دادند. پیش‌بینی‌های احتمالی، تخمین پارامترها به صورت اتوماتیک، و امکان انتخاب هسته‌های اختیاری از مزایای RVM نسبت به SVM است (Liyang W. et al., ۲۰۰۵). فوکوناگا^{۱۴} و همکاران (۲۰۰۶) مشکل کم بودن تعداد نمونه‌های آموزشی در طبقه‌بندی تصاویر ابر طیفی را بررسی و نشان دادند تعداد نقاط آموزشی موردنیاز برای طبقه‌بندی کننده‌های خطی متناسب با افزایش تعداد باندها به صورت خطی افزایش می‌یابد

۱. Hyperspectral

۲. Spectral Resolution

۴. Band Selection

۵. Dimension reduction

۸. Feature extraction based on decision boundaries

۹. Automatic wavelet spectral analysis

۶. Principal Component Analysis

۷. Independent Component Analysis

۱۰. Gualteri

۱۱. Support Vector Machin

۱۲. Melgani

۱. Liyang

۱۳. Regression Vector Machin

۱۴. Fukunaga

(Fukunaga A. & Sber G., ۲۰۰۸). بروزونه^۱ (۲۰۰۶) روش TSVM^۲ را برای طبقه‌بندی داده‌هایی که نمونه‌های آموزشی آن‌ها کاملاً قابل اعتماد نیستند ارائه نمود (Bruzzone L. et al., ۲۰۰۶). فاول^۳ و همکاران (۲۰۰۷) استفاده همزمان از ویژگی‌های طیفی و مکانی در طبقه‌بندی تصاویر ابرطیفی به روش SVM را مورد بررسی قرار دادند. نتایج این بررسی روی تصویر ROSIS موید افزایش کم دقت طبقه‌بندی بدون کاهش ابعاد فضای ویژگی و بهبود چشمگیر دقت با استفاده از الگوریتم کاهش ابعاد فضای ویژگی بود (Fauvel M. et al., ۲۰۰۷). اکبری و همکاران (۲۰۱۶) با استخراج اطلاعات از پیکسل‌های همسایگی در فرآیند انتخاب نمونه‌ها، سعی نمودند دقت الگوریتم‌های طبقه‌بندی طیفی-مکانی در تصاویر ابر طیفی را بهبود ببخشند (Akbari D. et al., ۲۰۱۶). غفاری و همکاران (۲۰۱۶) به بررسی انتخاب باندهای بهینه جهت بهبود جداسازی طیفی تصاویر ابر طیفی پرداختند. ایشان روش جدیدی برای انتخاب باند متشکل از اولویت‌بندی‌ها در راستای کاهش اثر تغییرپذیری طیفی و کاهش همبستگی بین باندها ارائه نموده و نشان دادند می‌توان به طور مؤثری به حذف تعداد زیادی از باندهای غیرضروری اقدام نمود (Ghaffari O. et al., ۲۰۱۶).

در چند سال اخیر، شبکه‌های عصبی کانولوشن^۴ به دلیل دارا بودن ساختاری مناسب برای کار با تصاویر، به عنوان روش منتخب بر مبنای یادگیری عمیق^۵ برای تولید ویژگی از تصاویر در نظر گرفته شده است (Ghamisi P. et al., ۲۰۱۷). موضوع مهم در این راستا، تعیین ساختار شبکه متناسب با داده‌های ورودی به شبکه است. لئو^۶ (۲۰۱۸) سعی در استفاده از شبکه‌های کپسول^۷ داشت و یک رویکرد مبتنی بر CNN را با مدل مبتنی بر CapsNet مقایسه کرد. نتایج این مطالعه موید آن بود که روش مبتنی بر CapsNet مزایای مورد انتظار را فراهم نمی‌کند و هیچ پیشنهادی نشان داده نشده است (Luo, Y. et al., ۲۰۱۸). وانگ^۸ و همکاران (۲۰۱۸) یک ساختار CapsNet بر اساس شبکه‌های عصبی بازگشتی^۹ جهت تجزیه و تحلیل احساسات و عواطف ارائه دادند (Wang Y. et al., ۲۰۱۸). در پژوهشی دیگر تعریف شبکه‌های عصبی کپسول این‌گونه بیان شد که نوع کاملاً جدیدی از معماری یادگیری عمیق را نشان می‌دهد که سعی در غلبه بر محدودیت‌ها و اشکالات CNN هم‌چون از دست رفتن اطلاعات در مرحله Max Pooling دارد (Sabour S. et al., ۲۰۱۷). هو^{۱۰} و همکاران (۲۰۱۵) یک شبکه کانولوشنی عصبی پنج لایه برای داده‌های ابر طیفی ارائه دادند، که در این روش از اطلاعات طیفی بدون در نظر گرفتن اطلاعات مکانی استفاده شد (Hu W. et al., ۲۰۱۵). یو^{۱۱} و همکاران (۲۰۱۹) یک شبکه کانولوشنی کپسول را برای تشخیص و سایل نقلیه با استفاده از تصاویر سنجش از دوری پیشنهاد دادند. در رویکرد پیشنهادی، تصویر سنجش از دوری با استفاده از استراتژی قطعه‌بندی سوپر پیکسل^{۱۲} به چندین تکه^{۱۳} تقسیم شد و به عنوان ورودی برای تشخیص وجود وسایل نقلیه مورد استفاده قرار گرفت (Yu Y. et al., ۲۰۱۹). گائو^{۱۴} و همکاران (۲۰۲۰)، رویکرد شبکه عصبی کپسول چند مقیاسه را برای تشخیص تغییرات از تصاویر رادار با دهانه مصنوعی^{۱۵} طراحی کردند. در این مطالعه، مدل کپسول صرفاً برای مشخص کردن رابطه بین ویژگی‌ها استفاده شد و از یک

-
۱. Bruzzone
 ۲. Transductive SVM
 ۳. Fauvel
 ۴. Convolution neural networks
 ۵. Deep Learning
 ۶. Luo, Y.; Zou
 ۷. CapsNet
 ۸. wang
 ۹. Recurrent Neural Network
 ۱۰. Hu, W
 ۱۱. Yu, Y
 ۱۲. superpixel segmentation strategy
 ۱۳. patch
 ۱۴. Gao, Y
 ۱۵. Synthetic aperture radar

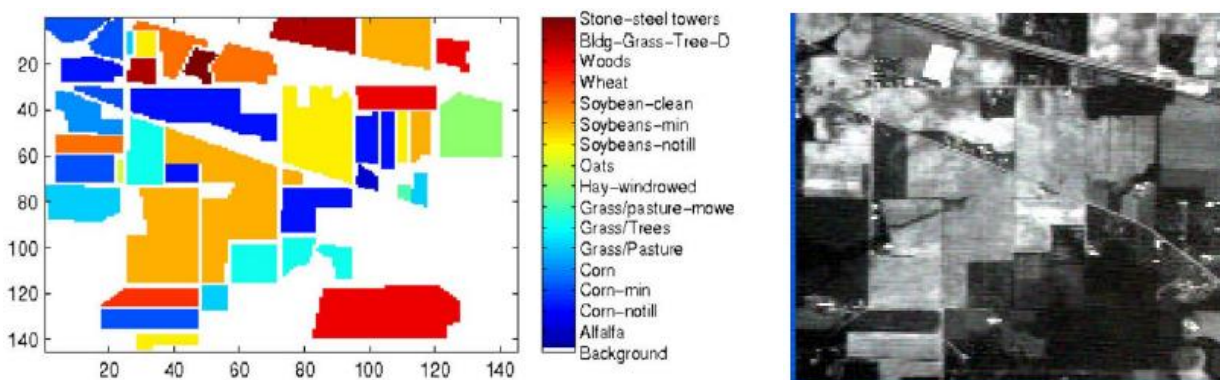
انطباقی کانولوشن برای افزایش قدرت شناسایی رابطه بین ویژگی‌ها استفاده شد (Gao Y. et al., ۲۰۲۰). در ادامه شبکه‌های عصبی کانولوشن سه بعدی (CNN) جهت طبقه‌بندی تصاویر ابر طیفی با استفاده همزمان از ویژگی‌های طیفی و مکانی مورد بحث و بررسی قرار گرفتند و در نهایت مدل شبکه کپسول سه بعدی با استفاده از کپسول‌های طیفی-مکانی جایگزین شبکه‌های کانولوشنی شدند.

اکتاو^۲ برای طبقه‌بندی تصاویر ابر طیفی طراحی کردند. در رویکرد طراحی شده، شبکه کانولوشنی معماری‌های موازی را سازگار کرده و ویژگی‌های تصاویر را استخراج کرد (Xu Q. et al., ۲۰۲۰).

در این پژوهش، از ساختاری مبتنی بر یادگیری عمیق برای طبقه‌بندی تصاویر ابرطیفی استفاده شده است به گونه‌ای که بتوان از ویژگی‌های معماری شبکه کپسول برای برطرف کردن معایب شبکه کانولوشنی جهت طبقه‌بندی تصاویر ابر طیفی و مقایسه نتایج حاصل از این دو شبکه استفاده کرد. در این مطالعه با توجه به این موضوع که خروجی‌های شبکه کپسول به صورت بردار می‌باشد، می‌توان ارتباط بین ویژگی‌های استخراجی را از طریق اندازه و جهت بردار استخراج کرد که هر کدام از المان‌های بردار خروجی نشان‌دهنده‌ی یک توصیف خاص از موجودیت تصویر می‌باشد. رویکرد استخراج ویژگی از طریق شبکه طراحی شده و طبقه‌بندی توسط الگوریتم درخت تقویتی^۳ XGBoost با روش طبقه‌بندی توسط شبکه‌ی عمیق سراسری مقایسه گردیده است تا علاوه بر بررسی و کیفیت‌سنجی ویژگی‌های عمیق تولیدی توسط روش پیشنهادی در طبقه‌بندی کننده‌های متفاوت، میزان توانایی شبکه‌های عمیق سراسری نیز در کاربرد طبقه‌بندی مورد بررسی قرار گیرد.

۲- داده‌ها و معرفی منطقه مورد مطالعه

تصاویر ابرطیفی به طور گسترده در جامعه سنجش از دور به منظور استفاده از ترکیب صدها کانال طیفی در یک صحنه واحد استفاده می‌شوند. طبقه‌بندی تصاویر ابرطیفی به دلیل ماهیت پیچیده صحنه تصویر (به عنوان مثال حجم زیادی از داده‌ها، پیکسل‌های مختلط و نمونه‌های آموزشی محدود) به عنوان یک مسئله پرچالش به شمار می‌آید و به همین دلیل تلاش‌های بسیاری برای پرداختن به این موضوع انجام یافته است. مجموعه داده‌های مورد استفاده در این تحقیق، داده‌های جامع از زمین‌های پاینس در آمریکا است که در ژوئن سال ۱۹۹۲ جمع‌آوری شده‌اند و ۱۶ کلاس که عمدتاً مربوط به پوشش زمینی می‌باشند را شامل می‌شوند (Ghassemian H. et al., ۲۰۰۳). در شکل ۱ این تصویر به همراه نقشه کاربری نشان داده شده است.



^۲. Xu

^۳. Octave convolution

^۴. Xtreme Gradient Boosting

(الف)

(ب)

شکل ۱. (الف) ترکیب رنگی تصویر Indian Pines (ب) نقشه‌ی کاربری تصویر Indian Pines

این مجموعه داده به عنوان داده‌های پایه برای ارائه‌ی مدل‌های طبقه‌بندی در اکثر تحقیقات بررسی و مورد استفاده قرار گرفته‌اند. داده‌های این مجموعه این امکان را فراهم می‌کند تا مقایسات جامعی از رویکردها در تحقیق پیش‌رو انجام شود. ابعاد این تصاویر ابرطیفی $145 \times 145 \times 200$ می‌باشند که اطلاعات هر یک از دسته‌ها در جدول ۱ و پوشش زمینی در شکل ۷ ارائه شده‌است. به منظور جلوگیری از ایجاد مشکل در تحلیل‌های عددی، نرمال‌سازی مقادیر باندها انجام شد. برای نرمال‌سازی داده‌ها از دو روش بهره گرفته شد؛ در روش تمامی باندها، مقدار درجه خاکستری هر پیکسل در هر باند بر جمع مقادیر درجات خاکستری آن پیکسل در تمامی باندها تقسیم می‌گردد و در روش تک باند مقدار درجه خاکستری هر پیکسل در هر باند بر جمع مقادیر درجات خاکستری تمامی پیکسل‌های آن باند تقسیم می‌شود. تأثیر روش نرمالیزه کردن بر نتایج طبقه‌بندی با استفاده از هسته گاوسین بررسی گردیده و نشان داده شد با هر دو روش نتایج مشابهی به دست می‌آید لذا از روش تمامی باندها در این پژوهش استفاده گردید.

جدول ۱. پوشش زمینی کلاس‌های تصویر

شماره کلاس	تعداد پیکسل‌های کلاس در تصویر	پوشش زمینی	شماره کلاس	تعداد پیکسل‌های کلاس در تصویر	پوشش زمینی
۱	۵۴	یونجه	۹	۲۰	جو دوسر
۲	۱۴۳۴	ذرت بدون شخم	۱۰	۹۶۸	سویا بدون شخم
۳	۸۳۴	ذرت کم شخم	۱۱	۲۴۶۸	سویا کم شخم
۴	۲۳۴	ذرت	۱۲	۶۱۴	سویا شخم زده
۵	۴۹۷	سبزه/چمن زار	۱۳	۲۱۲	گندم
۶	۷۴۷	سبزه/درختان	۱۴	۱۲۹۴	بیشه
۷	۲۶	سبزه/چمنزار کوتاه	۱۵	۳۸۰	ساختمان/سبزه
۸	۴۸۹	کاه و خاشاک	۱۶	۹۵	سنگ/برج

(بیشترین داده‌ها مربوط به کلاس ۱، ۲ و ۱۴ است).

۳- مبانی نظری

۳-۱ رویکردهای مبتنی بر یادگیری عمیق

در چند دهه اخیر با افزایش قدرت پردازنده‌های گرافیکی و در دسترس بودن مجموعه‌های حجیم از داده‌ها، استفاده از تکنیک‌های یادگیری عمیق در بسیاری از شاخه‌های مهندسی و کاربردی مطرح شده است که به موازات آن حوزه‌های پردازش تصویر نیز تحت تأثیر قرار گرفته‌اند. به طوری که در بسیاری از کاربردهای پردازش تصویر، یادگیری عمیق کارایی بالاتری را نسبت به تکنیک‌های یادگیری ماشین^۱ سنتی به دست آورده است

(Xu JL. et al., ۲۰۱۸; Li H., ۲۰۱۸). تفاوت ساختار شبکه‌های عصبی کانولوشنی که یکی از تکنیک‌های یادگیری عمیق می‌باشند، نسبت به شبکه‌های عصبی مصنوعی، محدودیت‌های اعمال شده در ارتباطات بین نرون‌ها در شبکه‌های کانولوشنی است. اگرچه این محدودیت‌ها عمومیت شبکه‌های کانولوشن را در مواجهه با داده‌های متنوع کاهش می‌دهد، لیکن ساختار این شبکه‌ها با توجه به قیود خاص آن کاملاً منطبق و مناسب با ساختار تصاویر است. همچنین در شبکه‌های عصبی کپسول با توجه به اینکه خروجی شبکه بردار می‌باشد، می‌تواند ارتباط بین ویژگی‌های استخراجی از شبکه‌های کانولوشنی را در نظر گرفت (Maggiori E. et al., ۲۰۱۶). بر این اساس، پژوهش حاضر بر روی فرایندهای تولید ویژگی‌های عمیق و طبقه‌بندی بر اساس آن متمرکز شده است که در ادامه به بیان جزئیات روش‌های بکار رفته در هر مرحله می‌پردازیم.

۳-۲ شبکه‌های عصبی کانولوشنی (CNN)

شبکه‌های عصبی کانولوشنی نوعی از شبکه‌های روبه‌جلو می‌باشند که اولین بار برای تشخیص تصویر مورد استفاده قرار گرفتند. این شبکه‌ها از مهم‌ترین و پرکاربردترین روش‌های یادگیری عمیق به طور خاص در زمینه استخراج ویژگی و طبقه‌بندی تصاویر سنجش از دور محسوب می‌گردند (LeCun Y. et al., ۲۰۱۵). شبکه‌های عصبی کانولوشنی دارای لایه‌های متعددی می‌باشند که هر یک از لایه‌ها یک سطح از ویژگی را نمایندگی می‌کنند. این سطوح متفاوت نمایندگی، توسط کنار هم قرار گرفتن ساختارهای ساده و غیرخطی به دست می‌آیند که هرکدام از ساختارها، ویژگی‌ها را به سطح انتزاع بالاتری ارتقا خواهد داد. شبکه کانولوشنی دارای چهار مشخصه اصلی: ارتباطات محلی، اشتراک وزن، لایه‌ی ادغام و مشخصه عمق است که آن را از سایر روش‌ها متمایز می‌سازد. ساختار شبکه‌های کانولوشنی بخش‌های متوالی زیر را شامل می‌شود (Bengio et al., ۲۰۱۵):

- **لایه کانولوشنی:** این لایه به منظور استخراج ویژگی‌های ضمنی از ماتریس ورودی، از فیلترهای مختلفی استفاده می‌کند.
- **لایه تجمیع:** رایج است که به صورت دوره‌ای پس از هر لایه کانولوشن برای کاهش اندازه ویژگی‌های میانی از یک لایه تجمیع استفاده شود. کاهش حجم پارامترها، کاهش محاسبات شبکه را در بر دارد و از این رو می‌توان بیش برآزش را کنترل کرد.
- **واحد اصلاح شده خطی^۲ (ReLU):** مزیت اصلی استفاده از این تابع نسبت به سایر توابع این است که هم‌زمان تمامی نرون‌ها را فعال نمی‌کند. این بدین معنی است که در یک زمان فقط تعداد کمی از نرون‌ها فعال می‌شوند و باعث می‌شود که شبکه به صورت تنک در بیاید و ویژگی‌های مهمی را یاد بگیرد.
- **لایه (Dropout):** اصطلاح 'Dropout' به واحدهای حذف شده (پنهان و آشکار) در یک شبکه عصبی اشاره دارد. این تکنیک اولین بار توسط جفری هینتون مورد استفاده قرار گرفت (Bengio et al., ۲۰۱۵). در واقع dropout

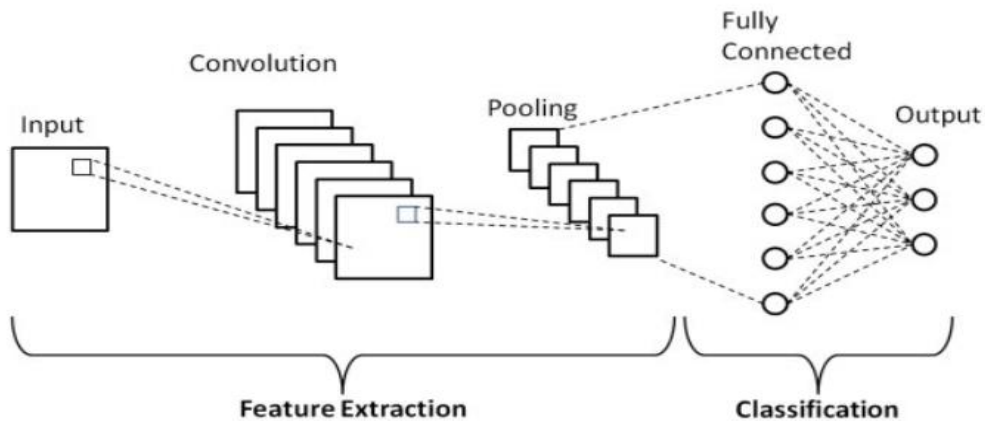
۱. Machine Learning

۲. overfitting

۳. Rectified linear unit

یک تکنیک Regularization تصادفی است که در طی هر انتشار روبه جلو واحدها در لایه‌های پنهان یا آشکار به صورت تصادفی یا موقت مطابق یک بردار از متغیرهای تصادفی برنولی بیرون انداخته می‌شوند.

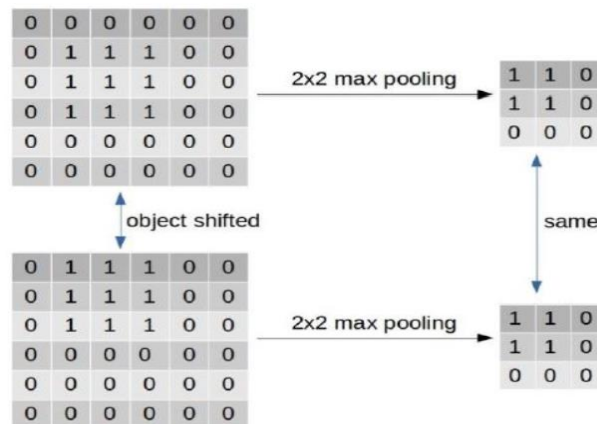
- **لایه تماماً متصل^۱**: بعد از آخرین لایه تجمیع لایه تماماً متصل وجود دارد که ویژگی‌های میانی دوبعدی را به بردارهای یک بعدی جهت ادامه فرایند یادگیری نگاشت می‌کند. از طرفی می‌توان گفت این لایه وظیفه محاسبه امتیاز دسته‌های طبقه را بر عهده دارد. لایه تماماً متصل به ما اجازه می‌دهد که خروجی شبکه را در قالب یک بردار با اندازه ثابت ارائه دهیم.



شکل ۲. معماری شبکه‌های کانولوشنی (LeCun Y. et al., ۲۰۱۵).

۳-۲-۱ مزایای شبکه کانولوشنی

مهم‌ترین مزیت استفاده از شبکه عصبی کانولوشن، توانایی آن در استخراج ویژگی‌های تصاویر به صورت خودکار با استفاده از مفهوم یادگیری عمیق است. همچنین استفاده از لایه pooling در شبکه‌های کانولوشنی باعث کاهش حجم محاسبات شده و تضمین می‌کند حتی زمانی که ویژگی‌های استخراجی، تغییرات و چرخش‌های کوچکی داشته باشند، همچنان خروجی شبکه یکسان بماند، که این ویژگی برای طبقه‌بندی نکته‌ای بسیار مهم محسوب می‌شود، همچنین استفاده از max pooling در عمل باعث می‌شود که با تغییرات اندک و جابجایی بسیار کم در تصویر، شبکه همچنان قادر به شناسایی شیء مورد نظر باشد (شکل ۳).



۴. Fully connected layer

شکل ۳. مزیت استفاده از لایه pooling در شبکه‌های کانولوشنی (Bengio et al., ۲۰۱۵).

۲-۲-۳ معایب شبکه کانولوشنی

یکی از کاربردهای شبکه‌های کانولوشنی استخراج ویژگی‌ها جهت طبقه‌بندی تصاویر می‌باشد. اما زمانی که نمونه معرفی شده به مدل آموزش دیده CNN دارای تغییراتی باشد، شبکه در تشخیص این تغییرات دچار مشکل می‌شود. دلیل این مشکل استفاده از لایه Max-pooling می‌باشد که فضای ارتباطی بین ویژگی‌ها را از بین می‌برد. در شکل ۴ تصویری از چهره در هم ریخته و غیر واقعی ارائه شده است که شبکه‌های CNN آن را به اشتباه چهره تشخیص می‌دهند چرا که ارتباط بین ویژگی‌های استخراجی و ابعاد ویژگی‌ها را در نظر نمی‌گیرند. برای این که شبکه دچار اشتباه نشود، نیاز است شبکه کانولوشنی با تصاویری از چرخش‌های متفاوت آموزش ببینند، که این امر نیازمند داده‌های آموزشی بسیار زیادی است.



شکل ۴. تصویری از چهره که نحوه قرارگیری و ابعاد اجزای آن تغییر داده شده است.

با توجه به معایب ذکر شده، استفاده از شبکه کپسول جهت رفع معایب شبکه‌های کانولوشنی پیشنهاد می‌شود.

۳-۳ شبکه کپسول

شبکه‌های عصبی کپسول روش کاملاً جدیدی از معماری یادگیری عمیق است که سعی در غلبه بر محدودیت‌ها و اشکالات شبکه کانولوشنی دارد (Xue Z. et al., ۲۰۱۵). خروجی‌های شبکه عصبی کپسول در یکسری از کپسول‌ها به صورت بردار ذخیره می‌شوند و کپسول‌ها محاسبات پیچیده‌ای را بر روی این ویژگی‌ها انجام داده و سپس نتایج حاصل از این محاسبات را به یک بردار کوچک که حاوی اطلاعات بسیار مفیدی است نگاشت می‌کنند. از این رو می‌توان ارتباط بین ویژگی‌های استخراجی شبکه را از طریق اندازه و جهت بردار حفظ کرد که هر کدام از المان‌های بردار خروجی نشان‌دهنده‌ی یک توصیف خاص از تصویر می‌باشد (Candemir S. et al., ۲۰۱۴). یکی از مهم‌ترین تفاوت‌های شبکه‌های کپسول با شبکه‌های کانولوشنی در شکل ویژگی‌های استخراج شده می‌باشد که به صورت یک بردار با خاصیت اندازه و جهت می‌باشد، در حالی که در شبکه‌های کانولوشنی و شبکه‌های تماماً متصل، خروجی‌ها به صورت عددی (اسکالر) می‌باشند. شبکه‌های کپسول مطابق شکل ۴ دارای لایه‌های `Conv1d`، `PrimaryCaps` و `DigitCaps` می‌باشد.

(Xue Z. et al., ۲۰۱۵; Candemir S. et al., ۲۰۱۴). زمانی که ویژگی‌ها در جای جای تصویر تغییر می‌کنند، با توجه به این موضوع که ویژگی‌ها به صورت بردار ذخیره شده‌اند، اندازه بردار حفظ می‌شود اما جهت آن تغییر می‌کند و در نتیجه برای شناسایی شی با چرخش‌های متفاوت، به تعداد نمونه‌های آموزشی بیشتری نیاز خواهد بود.

۱-۳-۳ اجزا شبکه کپسول پیشنهادی و محاسبه خروجی

شبکه کپسول طراحی شده در این پژوهش بسیار شبیه به شبکه کپسولی است که هینتون آن را توصیف کرده است (Candemir S. et al., ۲۰۱۴) با این تفاوت که یکسری تغییرات بر روی آن اعمال شده است. لایه‌های شبکه طراحی شده از قرار زیر می‌باشد.

- **لایه کانولوشن**

این لایه دارای ۲۵۶ فیلتر کانولوشنی به اندازه 9×9 و گام حرکتی ۲ می‌باشد که بر روی آن تابع فعال‌ساز 'ReLU' اعمال شده است. یکی از تغییراتی که در این مرحله نسبت به شبکه هینتون اعمال شده است، تغییر مقدار گام حرکتی از ۱ به ۲ تغییر داده شد. چرا که مدل اصلی بر روی مجموعه داده‌های mnist مورد استفاده قرار گرفته شده بود و تصاویر mnist به صورت 28×28 می‌باشند. با استفاده از گام حرکتی ۲ مقدار پارامترهایی که مدل در لایه بعد دریافت می‌کند بسیار کاهش می‌یابد.

- **لایه PrimeCaps**

مطابق نسخه اولیه این لایه، کپسول‌هایی به اندازه ۸، ۳۲ فیلتر 9×9 و گام حرکتی ۲ می‌باشد.

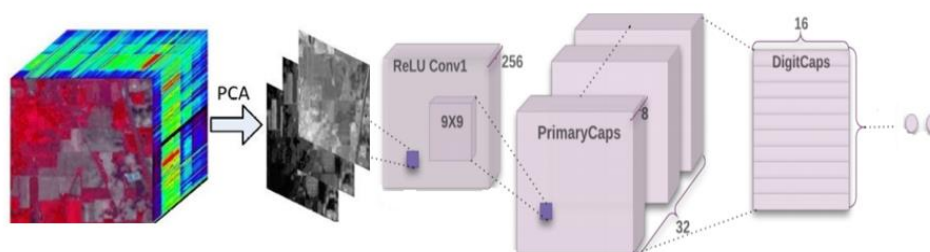
- **DigitCaps**

نسخه اولیه این لایه دارای ۱۰ کپسول ۱۶ بعدی است که بسته به مسئله این تحقیق این مقدار به ۲ کپسول ۱۶ بعدی تغییر داده شد. در این لایه مقدار الگوریتم مسیریابی نیز همانند نسخه اولیه با مقدار ۳ تنظیم شد.

- **لایه‌های تماماً متصل**

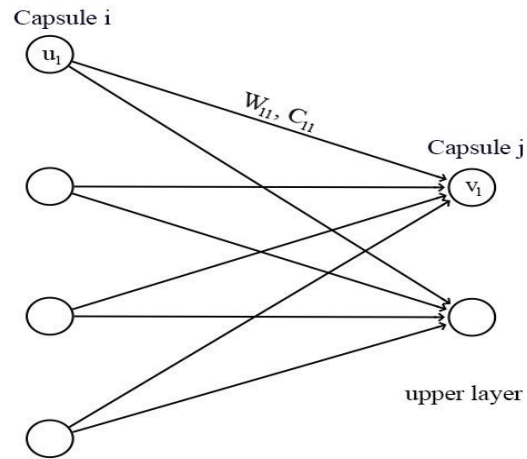
برخلاف پژوهش هینتون که در آن لایه‌های تماماً متصل در نظر گرفته نشده است، در این مطالعه لایه‌های تماماً متصل منظور شده‌اند که در آن می‌توان از چندین لایه استفاده کرد. در واقع این لایه سعی در یادگیری ویژگی‌های استخراجی از لایه DigitCaps دارد. در این تحقیق آخرین لایه تماماً متصل به اندازه‌ی طبقه‌های مدنظر این مطالعه (۱۶) است.

شکل ۵ ساختار کلی رویکرد شبکه‌های کپسول را نشان می‌دهد.



شکل ۵. شمای کلی از رویکرد شبکه‌های کپسول.

برای یک کپسول، مطابق شکل ۶ ورودی شبکه بردار u_i و خروجی بردار v_j می‌باشند که پارامترها از یک لایه به لایه دیگر منتقل می‌شوند.

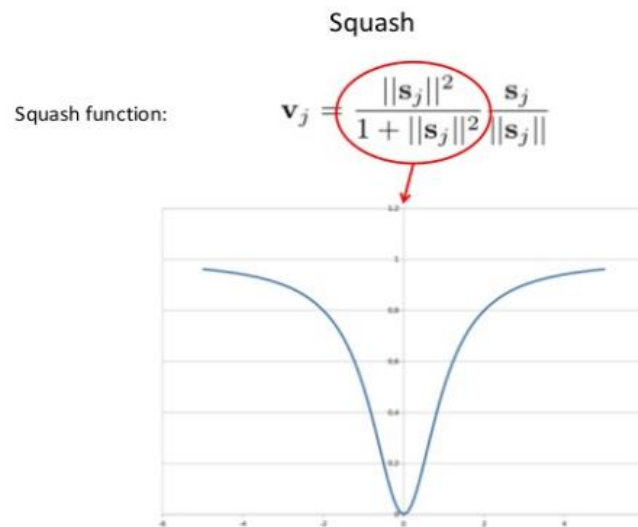


شکل ۶. ارسال پارامترها از یک لایه کپسول به کپسول‌های لایه بالاتر.

برای ارسال یک کپسول از لایه به یک لایه بعدی نیاز به فرآیند خاصی می‌باشد که در ادامه به این فرایند پرداخته شده است. برای این کار یک ماتریس تبدیل w_i را بر روی خروجی کپسول u_i لایه قبل اعمال می‌شود و سپس وزن جمعی s_j را براساس وزن‌های c_{ij} مطابق رابطه (۱) محاسبه می‌شود (Candemir S. et al., ۲۰۱۴).

$$\widehat{U}_{j|t} = w_{ij} U_j \quad \text{رابطه (۱)}$$

$$s_j = \sum c_{ij} \widehat{U}_{j|t}$$



شکل ۷. شکل تابع Squash (Candemir S. et al., ۲۰۱۴)

که در آن c_{ij} ها ضرایب اتصالی هستند که با استفاده از مسیریابی پویا محاسبه می‌شوند و $\sum c_{ij}$ برای جمع آن‌ها در نظر گرفته شده است. به طور مفهومی c_{ij} معیاری برای اندازه‌گیری میزان احتمال فعال‌سازی کپسول j توسط کپسول i می‌باشد. در شبکه های کپسول به جای اعمال ReLu، تابع squash را بر روی خروجی s_j اعمال می‌کنند. این تابع بردارهای کوچک را به ۰ و بردارهای بزرگ را به بردارهای واحدی کاهش می‌دهد و به همین دلیل خروجی یک کپسول v_j مطابق رابطه (۲) طولی بین $[-۰, ۱]$ دارد. در شکل ۷ تابع squash ارائه شده است، همچنین فرایند اعمال squash بر روی بردار ورودی s_j در زیر آورده شده است (Candemir S. et al., ۲۰۱۴):

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \cdot \frac{s_j}{\|s_j\|}$$

رابطه (۲)

$$v_j \approx \|s_j\| s_j \quad \text{for small } s_j$$

$$v_j \approx \frac{s_j}{\|s_j\|} \quad \text{for larg } s_j$$

پس از ارسال کپسول‌ها به لایه آخر تماماً متصل، این لایه نسبت به رویکردهای سنتی یادگیری ماشین مانند SVM و XGBoost دقت پایین‌تری دارند و همچنین حجم محاسبات بالایی را دارند. از این رو در رویکرد پیشنهادی مدل XGBoost به عنوان جایگزین این لایه‌ها استفاده شد.

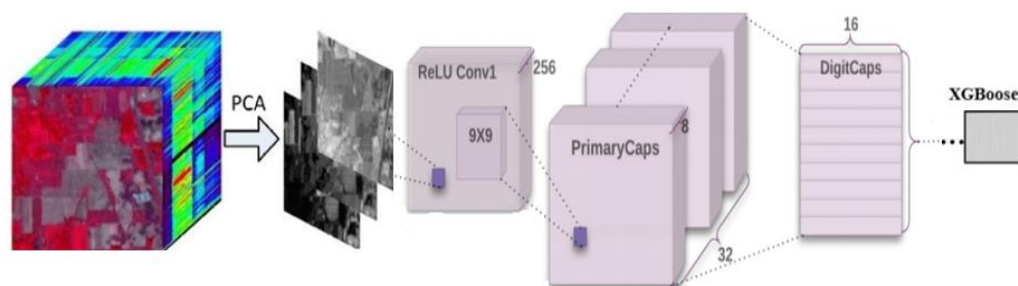
۳-۴ مدل درخت تقویتی^۲ (XGBoost)

مدل درخت تقویتی، روشی بسیار مؤثر و پرکاربرد در یادگیری ماشین است که توسط چن و گواسترین (۲۰۱۶) پیشنهاد شده است (Chen T. & Guestrin C., ۲۰۱۶). این شبکه نوعی تکنیک اجرایی کارآمد برای افزایش صحت و عملکرد ماشین افزایش گرادیان (GBM^۳) و به خصوص افزایش دقت طبقه‌بندی درخت‌های رگرسیون است. که یکی از محبوب‌ترین و کارآمدترین پیاده‌سازی‌های الگوریتم Gradient Boosted Trees می‌باشد که یک روش یادگیری نظارت شده مبتنی بر تقریب عملکرد با بهینه‌سازی توابع خاص است (Paoletti et al., ۲۰۱۹). Boosting در ML یک تکنیک یادگیری متوالی برای تبدیل یک فرضیه ضعیف به یک فرضیه قوی جهت افزایش دقت مدل است (Chen T. & Guestrin C., ۲۰۱۶). پس از حذف لایه تماماً متصل و جایگزینی آن با XGBoost ساختار مدل پیشنهادی مطابق شکل ۸ می‌باشد.

^۱ در شبکه های کپسول از مسیریابی پویایی برای محاسبه خروجی کپسول با محاسبه ضریب اتصالی به صورت تکراری استفاده می‌کنند. این فرایند مشابه الگوریتم پس انتشار در شبکه های عصبی مصنوعی می‌باشد با این تفاوت که به تعداد محدود عمل مسیریابی انجام می‌شود (به صورت پیش فرض ۳ بار)

۱. Extreme Gradient Boosting

۲. Gradient Boosting Machine



شکل ۸. شمای کلی از رویکرد پیشنهادی تلفیق شبکه‌های کپسول و درخت تصمیم تقویتی.

در این پژوهش به منظور بررسی صحت روش پیشنهادی، نتایج با سایر روش‌های ادبیات پیشین مورد بررسی قرار گرفت، که در زیر شرح مختصری از روش‌ها توضیح داده خواهد شد.

• رویکرد ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان نوعی از طبقه‌بندی کننده‌های آماری است که عملکرد کلی آن توجه بیشتری به الگوهای طبقه‌بندی و تخمین رگرسیون غیرخطی دارد. (Landgrebe DA., ۲۰۰۲) ماشین‌های بردار پشتیبان از طریق انتقال فضای ورودی به فضایی با بعد بالاتر قابلیت حل مسائل غیرخطی را دارند که این کار از طریق توابع کرنل امکان‌پذیر است. کرنل‌های متفاوت زیادی همچون کرنل گوسی، کرنل خطی، کرنل چند جمله‌ای و کرنل تانژانت هایپربولیک وجود دارند که در پژوهش انجام شده کرنل گوسی به منظور طبقه‌بندی تصاویر ابر طیفی مورد استفاده قرار گرفته است.

• رویکرد RF-۲۰۰

الگوریتم جنگل تصادفی (Landgrebe DA., ۲۰۰۲) مبتنی بر دسته‌ای از درخت‌های تصمیم‌گیری است که برچسب نهایی هر پیکسل را بر اساس برآوردی از تصمیمات درخت‌ها و غالباً بر اساس رای اکثریت تعیین می‌کند و برای بسیاری از مجموعه داده‌ها، طبقه‌بندی با صحت بالایی را تضمین می‌کند.

• شبکه‌های عصبی بازگشتی (RNN-LSTM)

LSTM^۱ یک نوع مدل یا ساختار برای داده‌های ترتیبی است که در سال ۱۹۹۵ برای توسعه شبکه‌های عصبی بازگشتی (RNN) ظهور پیدا کرد (Keshavarz A. & Ghasemiyan H., ۲۰۰۵). عبارت long term memory به وزن‌های یادگرفته شده و term memory short به حالت‌های درونی سلول‌ها اطلاق می‌شود. تغییر عمده این ساختار نسبت به RNN، جایگزین کردن لایه میانی RNN با یک بلوک است (Keshavarz A. & Ghasemiyan H., ۲۰۰۵). بزرگترین ویژگی LSTM مکان یادگیری وابستگی بلند مدت است که توسط شبکه‌های عصبی بازگشتی امکان‌پذیر نبود. برای پیش‌بینی گام

^۱ Long Short Term Memory

زمانی بعدی لازم است که مقادیر وزن‌ها در شبکه به روز رسانی شوند که این کار مستلزم حفظ اطلاعات گام‌های زمانی ابتدایی است. در تحقیق مورد بررسی شبکه‌های LSTM جهت طبقه‌بندی تصویر ابر طیفی پیاده‌سازی شد که نتایج پژوهش در جدول ۳ ارائه شده است. (Keshavarz A. & Ghasemiyani H., ۲۰۰۵)

• شبکه‌های عصبی بازگشتی (RNN-GRU)

معماری GRU^۱ (Keshavarz A. & Ghasemiyani H., ۲۰۰۵) به منظور حل کاستی‌های شبکه عصبی بازگشتی سنتی نظیر مشکل محو شدگی گرادیان و همچنین معایب موجود در معماری LSTM ارائه شده است. GRU عموماً به عنوان نسخه‌ای تغییر یافته از LSTM در نظر گرفته می‌شود چرا که هر دو این معماری‌ها از طراحی مشابهی بهره می‌برند و در بعضی از موارد بصورت یکسان نتایج عالی به دست می‌دهند. همان‌گونه که بیان شد برای حل مشکل محو شدگی گرادیان در شبکه عصبی سنتی یکی از راه حل‌ها استفاده از GRU است. این نوع معماری از مفهومی بنام دروازه به روزرسانی^۲ و دروازه بازنشانی^۳ استفاده می‌کند. این دو دروازه^۴ در اصل دو بردارند که با استفاده از آن‌ها تصمیم گرفته می‌شود چه اطلاعاتی به خروجی منتقل شده و چه اطلاعاتی منتقل نشود. نکته خاص درباره این دروازه‌ها این است که این دروازه‌ها را می‌توان طوری آموزش داد که اطلاعات مربوط به گام‌های زمانی بسیار قبل را بدون آنکه در حین گذر زمان (طی گام‌های زمانی مختلف) دستخوش تغییر شوند حفظ کند.

• رویکرد RNN-GRU-PRetanh

در این رویکرد شبکه بازگشتی با بلوک‌های GRU^۵ و همچنین تابع فعال ساز تانژانت هایپربولیک پارامتریک^۶ جهت طبقه‌بندی تصویر ابر طیفی مورد استفاده قرار گرفته است (Keshavarz A. & Ghasemiyani H., ۲۰۰۵). نتایج حاصل از این رویکرد در جدول ۳ ارائه شده است.

۳-۵ معیارهای ارزیابی

قبل از ارائه و بررسی نتایج حاصله، به بررسی معیارهای ارزیابی مورد استفاده پرداخته می‌شود. معیارهای ارزیابی مختلفی برای طبقه‌بندی دودویی وجود دارند که قبل از بررسی آن‌ها، پارامترهای مورد استفاده در این معیارها توضیح داده می‌شود. اگر یک طبقه‌بندی دودویی بر روی یک مجموعه داده تست اعمال شود، نتایج زیر مطابق شکل ۹ بدست می‌آید (Pedro Domingos, ۲۰۱۲):

۱. Gated Recurrent Unit

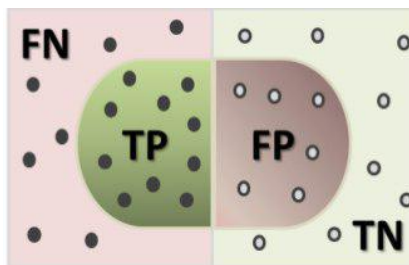
۲. Update gate

۳. Reset gate

۴. Gate

۵. Gated Recurrent Unit

۶. parametric rectified tanh (PRetanh)



شکل ۹. اعمال طبقه‌بندی دودویی بر روی یک مجموعه داده تست ۱.

در شکل ۹، چهار عبارت TP، FP، TN، و FN نمایش داده شده است. عبارت TP^۲ بیانگر تعداد نمونه‌های درست است که به درستی پیش‌بینی شده‌اند. عبارت FP^۳ نشان دهنده‌ی تعداد نمونه‌های نادرستی است که به صورت درست پیش‌بینی شده‌اند. عبارت TN^۴ بیانگر تعداد نمونه‌هایی نادرستی است که نادرست دسته‌بندی شده‌اند. عبارت FN^۵ بیانگر تعداد نمونه‌هایی درستی است که نادرست دسته‌بندی شده‌اند.

معیار Precision: در یک طبقه‌بندی دودویی، Precision نسبت بین نمونه‌های پیش‌بینی شده مثبت که دقیقاً مثبت می‌باشند، نسبت به کل پیش‌بینی‌های مثبت انجام شده می‌باشد که بصورت رابطه (۳) نشان داده می‌شود (Ghamisi P. et al., ۲۰۱۷).

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{رابطه (۳)}$$

به طور کلی عبارت Precision به معنی توانایی بالا برای پیش‌بینی درست است. معیار Recall: نسبت تعداد اسناد پیش‌بینی شده مثبت نسبت به کل اسناد موجود شامل پیش‌بینی مثبت و منفی می‌باشد که بصورت رابطه (۴) نشان داده می‌شود:

$$\text{Recall} = \frac{TP}{TP + FN} \quad \text{رابطه (۴)}$$

عبارت Recall به معنی این است که طبقه‌بندی مورد نظر توانسته است میزان زیادی از اسناد را درست طبقه‌بندی نماید. معیار F-Measure: این عبارت میانگین هارمونیک نیز نامیده شده و ترکیبی از دو معیار Precision و Recall می‌باشد. معمولاً این عبارت، F-Score متعادل نیز نامیده می‌شود و به شکل رابطه (۵) نشان داده می‌شود (Ghamisi P. et al., ۲۰۱۷).

$$F_{score} = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{رابطه (۵)}$$

معیار Accuracy: یکی دیگر از معیارهایی است که در یک طبقه‌بندی دودویی مورد استفاده قرار می‌گیرد. برای محاسبه‌ی این معیار نسبت TP و TN به تعداد کل موارد پیش‌بینی شده در نظر گرفته می‌شود بنابراین رابطه (۶) برای آن حاصل می‌شود:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad \text{رابطه (۶)}$$

۱. Apply binary classifications to a test data set

۲. True Positive

۳. False Positive

۴. True Negative

۵. False Negative

۳-۶ پارامترهای شبکه

در پژوهش پیش رو، اندازه نمونه‌های آموزشی $5 \times 5 \times 30$ است که 5×5 اشاره به اندازه دسته^۱ و ۳۰ نتیجه اعمال الگوریتم کاهش بعد PCA بر روی باندهای تصویر ابرطیفی است. در بررسی‌های انجام شده‌ای ما اندازه‌ی ۳۰ بهینه‌ترین حالت برای کاهش بعد PCA می‌باشد که در این حالت تغییرات در معیارها بسیار ناچیز می‌باشد و به طوری کلی افزایش بعد بهبودی حاصل نمی‌کند. به طور کلی اندازه فیلترهای کانولوشنی در شبکه‌های CNN یا کپسول می‌تواند ۳، ۵ و یا بیشتر باشد، که در این تحقیق از فیلتر 9×9 استفاده شده است. لایه کانولوشنی در ساختار رویکرد پیشنهادی دارای ۲۵۶ فیلتر کانولوشنی به اندازه 9×9 و گام حرکتی ۲ است که بر روی آن تابع فعال‌ساز 'ReLU' اعمال شده‌است. همچنین در ساختار رویکرد پیشنهادی در قسمت کپسول، کپسول‌هایی به اندازه ۸ با ۳۲ فیلتر 9×9 و گام حرکتی ۲ مورد استفاده قرار گرفت، همچنین نرخ Dropout در تمامی لایه‌های شبکه ۰/۵ در نظر گرفته شد. تمامی پارامترهای مدل با استفاده از optuna^۲ بدست آمده‌اند. Optuna یک کتابخانه‌ی متن باز برای جستجوی پارامترهای بهینه در مدل‌های یادگیری عمیق می‌باشد.

۸۰٪ داده‌های این تحقیق به عنوان داده‌های آموزشی جهت آموزش مدل و ۲۰٪ به عنوان داده‌های آزمون ارزیابی مدل مورد استفاده قرار گرفت.

۴- نتایج

در این بخش نتایج حاصل از شبکه کپسول پیشنهادی و شبکه CNN با استفاده از معیارهای ارزیابی Recall, Precision و F1-score (Ghamisi P. et al., ۲۰۱۷) مقایسه می‌شود. همچنین در این بخش رویکرد پیشنهادی با ۵ رویکرد مرسوم طبقه‌بندی مورد مقایسه قرار گرفت که در ادامه نتایج هر یک از این مدل‌ها ارائه می‌شود.

۴-۱ نتایج رویکرد مدل شبکه کانولوشنی (CNN)

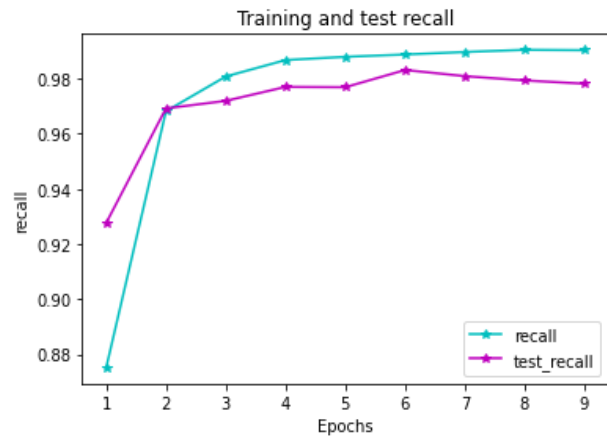
در این بخش نتایج حاصل از پیاده‌سازی شبکه CNN بر روی داده پاینس هند ارائه شده‌است. نتایج در دو حالت به ازای هر کلاس مجزا، به ازای تمامی حالت‌ها و برای مقایسه با سایر رویکردها گزارش داده شده است. برای تحلیل‌های بیشتر، نمودار مربوط به خطا و متریک‌های اندازه‌گیری در شکل ۱۰ آورده شده است. جدول ۲ نیز نتیجه حاصله از اجرای شبکه CNN بر روی مجموعه داده‌های تست بعد از ۵۰ بار تکرار مدل است که به ازای هر کلاس ارائه شده است. ستون Count بیانگر تعداد نمونه به ازای هر کلاس می‌باشد.

^۱ Patch Size

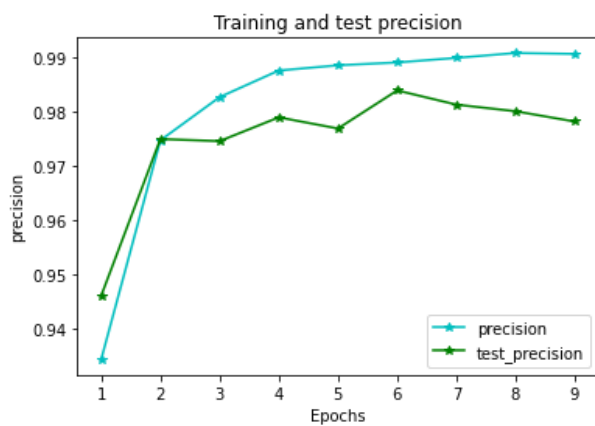
^۲ <https://optuna.org/>



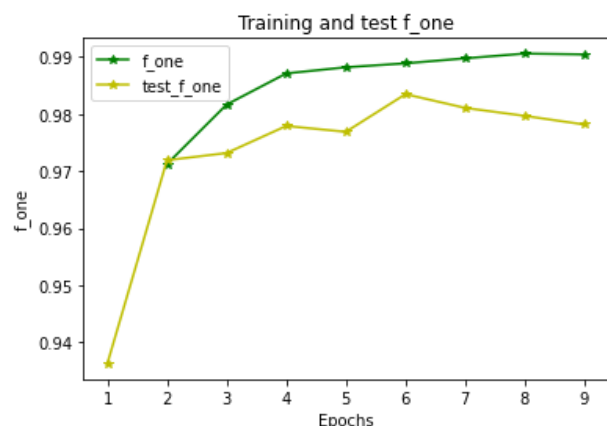
(ب)



(الف)



(ت)



(پ)

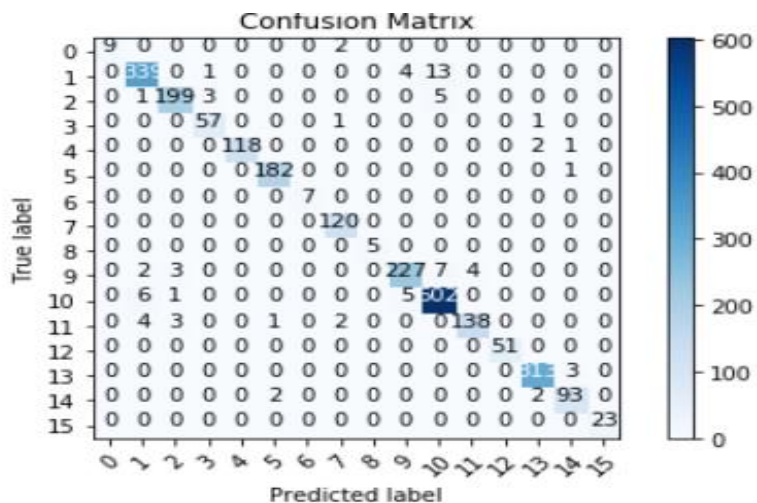
شکل ۱۰. (الف) نمودار Recall (ب) نمودار Accuracy (پ) نمودار Precision و (ت) نمودار F_1 بر روی داده‌های آموزش و تست با رویکرد پیشنهادی اول.

در این نمودارها، محور X نشان دهنده تعداد دفعات اجرای الگوریتم بر روی داده‌ها و محور Y نشان دهنده معیار بدست آمده به ازای هر بار اجرای الگوریتم است. همچنین در شکل ۱۱ ماتریس خطا مرتبط با رویکرد اول ارائه شده است. در جدول ۳ دقت

نتایج به دست آمده توسط ساختار اول و مقایسه آن با نتایج سایر پژوهش‌های انجام یافته بر روی این مجموعه از داده‌ها ارائه شده است.

جدول ۲. نتایج حاصل از ساختار پیشنهادی اول بر روی مجموعه داده‌های تست.

Count	معیارهای ارزیابی			شماره کلاس
	F ^۱ -score	Recall	Precision	
۱۱	۹۰	۸۲	۱۰۰	۱
۳۵۷	۹۶	۹۵	۹۶	۲
۲۰۸	۹۶	۹۶	۹۷	۳
۵۹	۹۵	۹۷	۹۳	۴
۱۲۱	۹۹	۹۸	۱۰۰	۵
۱۸۳	۹۹	۱۰۰	۹۹	۶
۷	۱۰۰	۱۰۰	۱۰۰	۷
۱۲۰	۹۸	۱۰۰	۹۶	۸
۵	۱۰۰	۱۰۰	۱۰۰	۹
۲۴۳	۹۵	۹۴	۹۶	۱۰
۶۱۴	۹۷	۹۸	۹۶	۱۱
۱۴۸	۹۵	۹۳	۹۷	۱۲
۵۱	۱۰۰	۱۰۰	۹۸	۱۳
۳۱۶	۹۹	۹۹	۹۸	۱۴
۹۷	۹۵	۹۶	۹۵	۱۵
۲۳	۱۰۰	۱۰۰	۱۰۰	۱۶
۲۵۶۳	۹۶,۸۶	۹۶,۸۷	۹۶,۸۸	میانگین کل



شکل ۱۱. ماتریس خطا^۱ رویکرد اول.

جدول ۳. مقایسه نتایج حاصل از ساختار پیشنهادی اول با نتایج موجود (مقادیر بر حسب درصد است).

RF-۲۰۰	SVM-RBF	RNN-LSTM	RNN-GRU-than	RNN-GRU-PRetanh	رویکرد اول	شماره کلاس
۵۸/۸۴	۶۰/۷۷	۴۶/۰۳	۶۸/۹۳	۷۰/۵۹	۹۷/۵۶	۱
۵۸/۴۲	۷۷/۶۸	۶۱/۳۷	۴۰/۹۴	۷۰/۲۸	۹۷/۵۲	۲
۸۲/۶۱	۷۹/۳۵	۸۶/۹۶	۷۸/۸۰	۸۱/۵۲	۸۵/۳۳	۳
۸۵/۹۱	۹۱/۰۵	۸۷/۰۲	۸۷/۹۲	۹۰/۱۶	۸۷/۹۲	۴
۸۰/۴۹	۸۴/۳۶	۸۶/۶۶	۸۷/۵۲	۹۱/۹۷	۸۵/۲۲	۵
۹۴/۷۶	۹۲/۰۳	۷۹/۴۹	۹۷/۲۷	۹۶/۱۳	۹۷/۴۹	۶
۷۷/۳۴	۶۹/۶۱	۵۹/۶۹	۸۲/۷۹	۸۴/۷۵	۷۴/۶۲	۷
۵۹/۴۳	۵۹/۳۱	۶۴/۸۹	۵۰/۵۸	۵۹/۶۴	۶۷/۹۹	۸
۶۳/۴۸	۷۹/۶۱	۶۰/۴۶	۷۹/۴۳	۸۶/۱۷	۸۷/۵۸	۹
۵۹/۰۶	۹۷/۵۳	۹۸/۷۷	۹۸/۷۷	۹۹/۳۸	۹۸/۷۷	۱۰
۸۸/۲۶	۸۵/۲۱	۷۵/۳۲	۸۴/۷۳	۸۴/۹۷	۸۷/۶۲	۱۱
۵۴/۸۵	۶۳/۶۴	۷۱/۸۲	۶۱/۲۱	۷۷/۵۸	۷۲/۴۲	۱۲
۹۷/۷۸	۱۰۰	۹۱/۱۱	۸۸/۸۹	۹۵/۵۶	۹۳/۳۳	۱۳
۵۸/۹۷	۸۷/۱۸	۷۹/۴۹	۷۹/۴۹	۸۴/۶۲	۷۱/۷۹	۱۴
۸۱/۸۲	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۱۵
۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۶

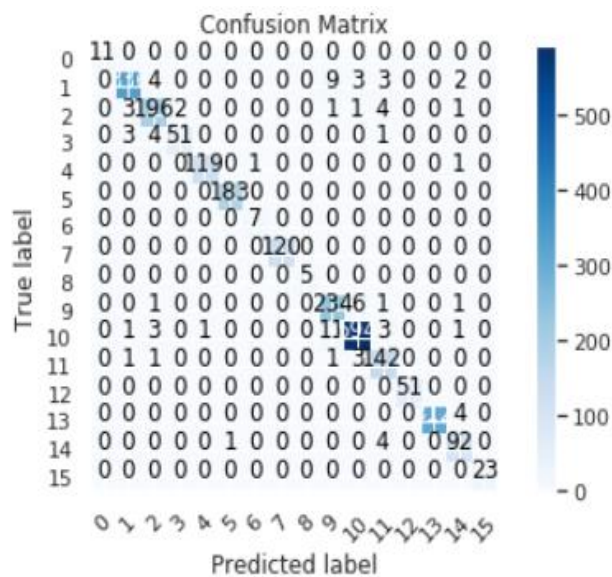
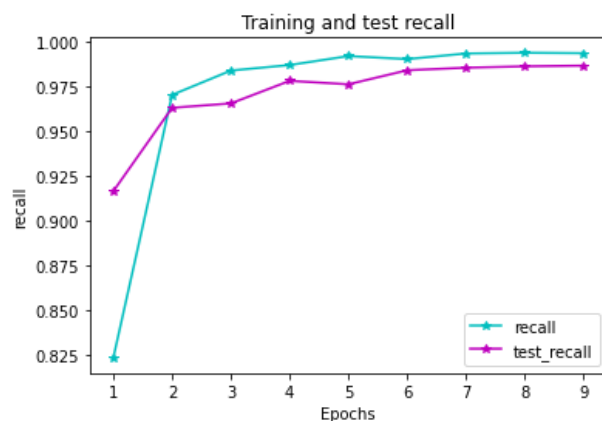
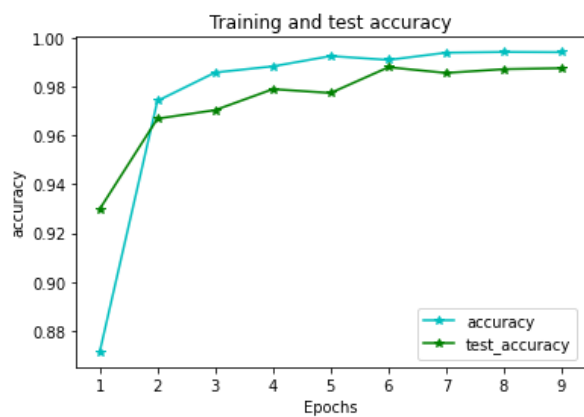
مقادیر ساختار اول در مقایسه با رویکردهای قبلی دقت قابل توجهی را به دست آورده است. این مدل نسبت به رویکرد-RNN GRU-PRetanh دقت کمتری در اکثر دامنه‌ها نشان می‌دهد. همچنین این مدل تنها توانسته است روی داده‌های کلاس Buildings-grass-trees دقت برابری را کسب کند. این در حالی است که در سایر طبقه‌ها میزان دقت نسبتاً کمتر بوده است. این مدل به دلیل داشتن عمق و تعداد فیلتر کمتر، دقت پایین‌تری در طبقه‌بندی تصاویر داشته است لذا در رویکرد پیشنهادی دوم سعی شده است با استفاده از شبکه‌های کپسول تا حدود زیادی از این مشکلات کاسته شود. در هر دو رویکرد پیشنهادی برای کنترل overffiting از early stoping استفاده شده است این سیاست در دو دور اجرایی بررسی می‌کند که اگر مقدار loss تغییری نکرده باشد یادگیری را متوقف کند. همچنین برای یادگیری پارامترهای هر دو مدل Optuna استفاده شد و بهترین پارامترها برای طبقه بندی انتخاب شد.

۴-۲ نتایج رویکرد پیشنهادی دوم شبکه‌های کپسول

در رویکرد پیشنهادی دوم نیز از ۸۰ درصد داده‌ها برای آموزش و از ۲۰ درصد آن‌ها برای تست استفاده شده است. مشابه رویکرد پیشنهادی اول، نتایج در دو حالت (به ازای هر کلاس مجزا، به ازای تمامی حالت‌ها) برای مقایسه با سایر رویکردها گزارش شده است. جدول ۴ نتایج خروجی از الگوریتم بر روی مجموعه داده‌های تست بعد از ۳۰ بار یادگیری شبکه را ارائه می‌دهد که به ازای هر کلاس گزارش داده شده است. همچنین در شکل ۱۲ ماتریس خطا مرتبط با رویکرد دوم ارائه شده است

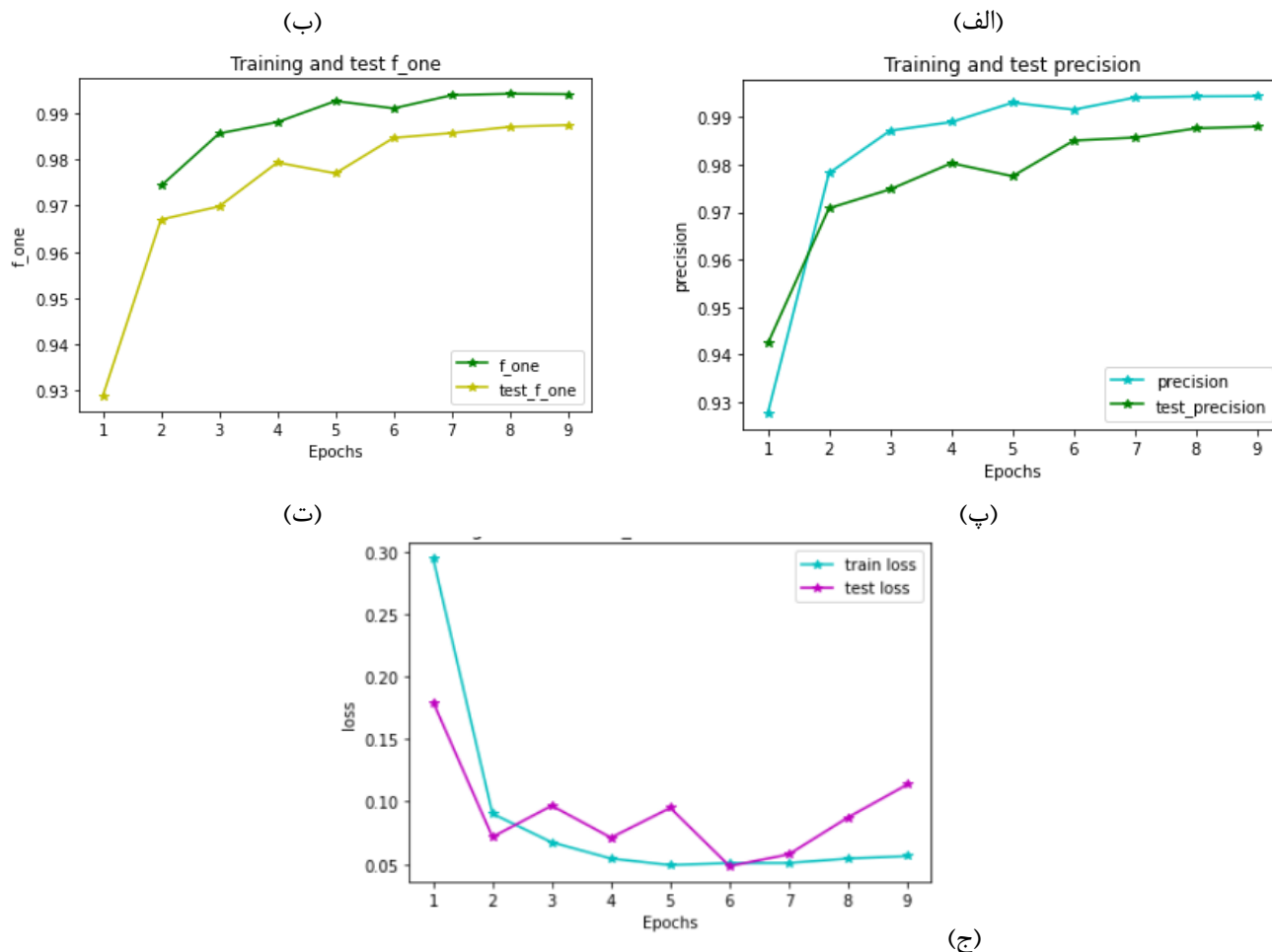
جدول ۴. نتیجه‌ی حاصله از رویکرد پیشنهادی دوم بر روی مجموعه داده‌های تست.

Count	معیارهای ارزیابی			شماره کلاس
	F1-score	Recall	Precision	
۱۱	۱۰۰	۱۰۰	۱۰۰	۱
۳۵۷	۹۶	۹۴	۹۸	۲
۲۰۸	۹۴	۹۴	۹۴	۳
۵۹	۹۱	۸۷	۹۶	۴
۱۲۱	۹۹	۹۸	۹۹	۵
۱۸۳	۱۰۰	۱۰۰	۹۹	۶
۷	۹۳	۱۰۰	۸۸	۷
۱۲۰	۱۰۰	۱۰۰	۱۰۰	۸
۵	۱۰۰	۱۰۰	۱۰۰	۹
۲۴۳	۹۴	۹۶	۹۲	۱۰
۶۱۴	۹۷	۹۷	۹۸	۱۱
۱۴۸	۹۳	۹۶	۹۰	۱۲
۵۱	۱۰۰	۱۰۰	۱۰۰	۱۳
۳۱۶	۹۹	۹۹	۱۰۰	۱۴
۹۷	۹۲	۹۵	۹۰	۱۵
۲۳	۱۰۰	۱۰۰	۱۰۰	۱۶
۲۵۶۳	۹۶,۶۲	۹۶,۶۰	۹۶,۶۸	میانگین کل



شکل ۱۲. ماتریس خطا^۱ رویکرد دوم.

نمودار مربوط با خطا و سایر متریک‌های اندازه‌گیری در شکل ۱۳ آورده شده است.



شکل ۱۳. (الف) نمودار Recall (ب) نمودار Accuracy (پ) نمودار Precision و (ت) F_1 (ج) نمودار loss بر روی داده‌های آموزش و تست با رویکرد پیشنهادی دوم.

منحنی‌های حاصل با رویکرد پیشنهادی دوم، کاملاً متفاوت از منحنی‌های حاصل از رویکرد پیشنهادی اول است. در این رویکرد، اختلاف بین دو منحنی آموزش و تست در تمامی حالات کمتر شده و این منحنی‌ها به یک مجانب افقی میل کرده‌اند. این امر نشان می‌دهد افزایش تعداد یادگیری شبکه از یک مقدار معین، تأثیری بر هیچ یک از معیارهای ارزیابی نداشته و هر چهار معیار حدوداً پس از ۷ مرحله یادگیری شبکه، بدون تغییر باقی می‌ماند. در جدول ۵ بررسی دقت‌های به‌دست‌آمده توسط رویکرد پیشنهادی دوم با نتایج حاصل از سایر پژوهش‌های انجام یافته بر روی این مجموعه از داده‌ها ارائه شده است.

جدول ۵. مقایسه نتایج حاصله از ساختار پیشنهادی دوم با نتایج سایر پژوهشگران.

شماره	رویکرد	رویکرد	RNN-GRU-	RNN-GRU-	RNN-	SVM-	RF-
-------	--------	--------	----------	----------	------	------	-----

کلاس	اول	دوم	PRetanh	than	LSTM	RBF	۲۰۰۰
۱	۹۷/۵۶	۹۵/۷۱	۷۰/۵۹	۶۸/۹۳	۶۴/۰۳	۶۰/۷۷	۵۴/۸۴
۲	۹۷/۵۲	۸۲/۷۷	۷۰/۲۸	۴۰/۹۴	۶۱/۳۷	۷۷/۶۸	۵۸/۴۲
۳	۸۵/۳۳	۸۶/۵۲	۸۱/۵۲	۷۸/۸۰	۸۶/۹۶	۷۹/۳۵	۸۲/۶۱
۴	۸۷/۹۲	۹۱/۹۳	۹۰/۱۶	۸۷/۹۲	۸۷/۰۲	۹۱/۰۵	۸۵/۹۱
۵	۸۵/۲۲	۹۲/۹۷	۹۱/۹۷	۸۷/۵۲	۸۶/۶۶	۸۴/۳۶	۸۰/۴۹
۶	۹۷/۴۹	۹۸/۱۲	۹۶/۱۳	۹۷/۲۷	۷۹/۴۹	۹۲/۰۳	۹۴/۷۶
۷	۷۴/۶۲	۸۴/۷۰	۸۴/۷۵	۸۲/۷۹	۵۹/۶۹	۶۹/۶۱	۷۷/۳۴
۸	۶۷/۹۹	۶۷/۶۴	۵۹/۶۴	۵۰/۵۸	۶۴/۸۹	۵۹/۳۱	۵۹/۴۳
۹	۸۷/۵۸	۷۶/۹۰	۸۶/۱۷	۷۹/۴۳	۶۰/۴۶	۷۹/۶۱	۶۳/۴۸
۱۰	۹۸/۷۷	۹۹/۶۰	۹۹/۳۸	۹۸/۷۷	۹۸/۷۷	۹۷/۵۳	۵۹/۰۶
۱۱	۸۷/۶۲	۸۵/۹۰	۸۴/۹۷	۸۴/۷۳	۷۵/۳۲	۸۵/۲۱	۸۸/۲۶
۱۲	۷۲/۴۲	۷۸/۵۹	۷۷/۵۸	۶۱/۲۱	۷۱/۸۲	۶۳/۶۴	۵۴/۸۵
۱۳	۹۳/۳۳	۹۶/۵۰	۹۵/۵۶	۸۸/۸۹	۹۱/۱۱	۱۰۰	۹۷/۷۸
۱۴	۷۱/۷۹	۷۹/۶۲	۸۴/۶۲	۷۹/۴۹	۷۹/۴۹	۷۸/۱۸	۵۸/۹۷
۱۵	۹۰/۹۱	۹۱/۹۱	۹۱/۹۱	۹۰/۹۱	۹۰/۹۱	۹۰/۹۱	۸۱/۸۲
۱۶	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰

رویکرد پیشنهادی دوم نسبت به رویکرد اول نتایج بهتری را در اکثر کلاسها داشته است. این ساختار توانسته است نسبت به RNN-GRU-PRetanh دقت بالاتری را بدست آورد. دلیل این امر آنست که رویکرد دوم شبکه‌های کپسول با استخراج مطلوب ویژگی‌ها توانسته است مهندسی ویژگی خودکار مناسبی را برای مدل استاندارد XGBoost ایجاد کند. همچنین با استفاده از رویکرد دوم در دامنه‌های Alfalfa, Corn-min, Grass-pasture, Grass-trees, Soybean-notill, Soybean-clean, Wheat و Buildings-grass-trees دقت بالاتری حاصل شده است. ساختار دوم در سایر کلاسها نیز دقت قابل قبولی داشته است و البته علت پایین بودن دقت این ساختار بر روی این طبقها، کم بودن تعداد نمونه‌ها در آنهاست که در نتیجه آن، مدل به خوبی نتوانسته است این داده را آموزش و یادگیری را کامل نماید.

۵- بحث و نتیجه‌گیری

طبقه‌بندی تصاویر ماهواره‌ای همواره با چالش‌های زیادی روبرو بوده است. با پیشرفت و گسترش الگوریتم‌های طبقه‌بندی و افزایش قدرت تفکیک طیفی تصاویر، با وجود افزایش کلی دقت، چالش‌های مربوط به طبقه‌بندی این تصاویر نیز تشدید شده است.

یکی از مهم‌ترین مشکلات طبقه‌بندی این تصاویر در مدل‌های یادگیری ماشین انتخاب ویژگی‌ها می‌باشد چرا که این مدل‌ها نیاز به انتخاب ویژگی به صورت دستی دارند و هر چقدر ویژگی‌های مناسبتری استخراج شوند این مدل‌ها می‌توانند طبقه‌بندی بهتری را انجام دهند. از این رو یادگیری عمیق به عنوان مدل‌های که می‌توانند انتخاب ویژگی‌ها را به صورت خودکار انجام بدهند جایگزین‌های مناسبی برای طبقه‌بندی این تصاویر می‌باشند. در این نوع یادگیری استخراج ویژگی‌ها براساس یکسری فیلتر انجام می‌شود که کاربر هیچ دخالتی در این استخراج ویژگی ندارد. برای افزایش دقت طبقه‌بندی این تصاویر، تکنیک‌های

یادگیری عمیق شبکه‌های کانولوشنی بر روی این نوع از تصاویر اعمال می‌شود. در پژوهش حاضر از روش یادگیری عمیق شبکه‌های عصبی کانولوشنی و همچنین شبکه‌های کپسول برای افزایش دقت طبقه‌بندی تصاویر ابر طیفی استفاده شد. ابعاد تصویر ورودی به شبکه از عوامل تاثیرگذار در عملکرد استخراج ویژگی توسط شبکه‌های کپسول می‌باشد؛ که در این تحقیق ابعاد $5 \times 5 \times 30$ در نظر گرفته شد. سه معیار Recall, Precision, و F1-Score برای ارزیابی نتایج استفاده و نتایج به صورت میانگین و انحراف معیار برای ۲۰ بار اجرای شبکه تهیه شد. از مجموع داده‌های در نظر گرفته شده، ۸۰ درصد برای آموزش طبقه‌بندی و ۲۰ درصد برای ارزیابی طبقه‌بندی استفاده شد. به منظور بهبود دقت طبقه‌بندی تصاویر نیز دو رویکرد متفاوت در زمینه استفاده از شبکه‌های عصبی کانولوشنی و شبکه کپسول در طبقه‌بندی تصاویر هوایی ابر طیفی به کار گرفته شد. یکی از مهم‌ترین تفاوت‌های شبکه‌های کپسول با شبکه‌های کانولوشنی در ذخیره داده‌ها در قالب یک بردار است. ویژگی بردار بودن با خود خاصیت اندازه و جهت را همراه خواهد داشت و می‌توان ارتباط بین ویژگی‌ها را نیز استخراج کرد و به دقت بالاتر رسید. در رویکرد اول از این شبکه‌ها در حالت یکپارچه برای استخراج ویژگی و در ادامه طبقه‌بندی استفاده شد و در رویکرد دوم ویژگی‌های استخراج شده توسط شبکه‌های کپسول به عنوان ورودی به مدل درخت تقویتی XGBoost برای طبقه‌بندی معرفی شد. در رویکرد اول از دولایه کانولوشنی و دولایه تماماً متصل استفاده گردید که موجب حصول دقت ۹۸ درصد بر روی داده‌های آموزش و دقت ۹۶.۳۰ درصد بر روی داده‌های تست شد. در رویکرد دوم از یک لایه کانولوشنی، یک لایه کپسول و XGboost استفاده شد؛ که منجر به دقت ۹۹ درصد بر روی داده‌های آموزش و دقت ۹۷.۵ درصد بر روی داده‌های تست شد. رویکرد پیشنهادی اول تنها در یکی از طبقه‌ها توانسته دقت برابری با رویکرد برتر در سایر پژوهش‌ها بدست آورد، در حالی که رویکرد دوم در اکثر موارد توانسته است دقت بالاتر را کسب نماید.

منابع

- Ghassemian H., Keshavarz A. & Landgrebe D., ۲۰۰۳. **Hyper-spectral image processing and analyses**. Space Magazine, ۱(۳): ۳۲-۴۱.
- Keshavarz A. & Ghasemiyan H., ۲۰۰۵. **A high-speed vector machine-based algorithm for classifying hyperspectral images using spatial correlation**. Iranian Journal of Electrical and Computer Engineering, ۳(۱): ۳۷-۴۴.
- Landgrebe DA., ۲۰۰۲. **Hyperspectral image data analysis**. IEEE Signal Processing Magazine, ۱۹(۱): ۱۷-۲۸.
- Du P., Tan K, Zhang W. & Yan Zh, ۲۰۰۸. **ANN classification of OMIS hyperspectral remotely sensed imagery: Experiments and analysis**. Congress on Image and Signal Processing, IEEE.
- Jimenez L. & Landgrebe DA., ۱۹۹۹. **Hyperspectral data analysis and feature reduction via projection pursuit**. IEEE Trans. On Geoscience and Remote Sensing, ۳۷(۶): ۲۶۵۳-۲۶۶۷.
- Lee C. & Landgrebe DA., ۱۹۹۳. **Feature extraction based on decision boundaries**. IEEE

-
- Trans. on Pattern Analysis and Machine Intelligence, ۱۰(۴): ۳۸۸-۴۰۰.
- Kaewpijit S, Moigne JL, Ghazawi TE. ۲۰۰۳. **Automatic wavelet spectral analysis for reduction of hyperspectral imagery.** IEEE Trans. on Geoscience and Remote Sensing, ۴۱(۴): ۸۶۳- ۸۷۱.
- Gualtieri JA. & Crompton RF., ۱۹۹۹. **Support vector machines for hyperspectral remote sensing classification.** ” Proceedings Volume ۳۰۸۴, ۲۷th AIPR Workshop: Advances in Computer-Assisted Recognition.
- Gualtieri JA., Chetri SR., Crompton RF. & Johnson LF., ۱۹۹۹. **Support vector machines classifiers as applied to AVIRIS data.** In summaries ۸th JPL Airborne Earth Science Workshop, JPL Pub. ۹۹-۱۷, ۲۱۷-۲۲۷.
- Gualtieri JA. & Chetri SR., ۲۰۰۰. **Support vector machines for classification of hyperspectral data.** Proc. IGARSS, Honolulu, HI, ۸۱۳-۸۱۵.
- Melgani F. & Bruzzone L., ۲۰۰۴. **Classification of Hyperspectral Remote Sensing Images with Support Vector Machines.** IEEE Transactions on Geoscience and Remote Sensing, ۴۲(۸): ۱۷۷۸ – ۱۷۹۰.
- Liyang W., Yongyi Y., Nishikawa RM., Wernick MN. & Edwards A., ۲۰۰۵. **Relevance vector machine for automatic detection of clustered micro calcifications.** IEEE Trans. Med. Imag., ۲۴(۱۰): ۱۲۷۸-۱۲۸۵.
- Fukunaga A. & Sber G., ۲۰۰۸. **Providing an optimal method based on deep learning for spectral classification Images with high resolution spatial resolution in semi-urban areas.** Journal of Geomatics Science and Technology, ۹(۲): ۱۵۱-۱۷۰.
- Bruzzone L., Chi M. & Marconconi M., ۲۰۰۶. **A novel transductive SVM for semisupervised classification of remote-sensing images.** IEEE Transactions on Geoscience and Remote Sensing, ۴۴(۱): ۳۳۶۳ – ۳۳۷۳.
- Fauvel M., Chanussot J., Benediktsson JA. & Sveinsson J., ۲۰۰۷. **Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles.** IEEE Transactions on Geoscience and Remote Sensing, ۴۵(۱۱): ۲۰۱۲-۲۰۲۰.
- Baofeng & Mahir M., 2008. **Classification of Hyperspectral Remote Sensing Images with Support Vector Machines.** IEEE Transactions on Geoscience and Remote Sensing, 42(8):1778 – 1790.
- Camps-Valls G., Gomez-Chova L., Munoz-Mari J., Vila-Frances J. & Calpe-Maravilla J., ۲۰۰۶. **Composite kernels for hyperspectral image classification.** IEEE Geoscience and Remote Sensing Letters, ۳(۱): ۹۳-۹۷.
- Camps-Valls G., Shervashidze N. & Borgwardt KM., ۲۰۱۰. **Spatio-spectral remote sensing image classification with graph kernels.** IEEE Geoscience and Remote Sensing Letters, ۷(۴): ۷۴۱-۷۴۵.
- Dehghani H., ۲۰۰۶. **Classification of distance measurement images with large dimensions and limited educational examples.** PhD Thesis, Tarbiyat Modarres University, Department of Electronic Engineering, Tehran, Iran.
- Shahhoseini R., ۲۰۰۹. **Evaluation of support vector machines in the classification of**
-

hyperspectral remote sensing data. Ms.c Thesis, University of Tehran, Tehran, Iran.

Akbari D., Safari AR. & Homayouni S., ۲۰۱۶. **Improving spectral-spatial classification of supercritical images By using spatial information to select symbols.** Scientific - Research Quarterly of Geographical Data (SEPEHR), ۲۰(۹۸): ۵-۱۴.

Ghaffari O., Voldan-Zoj MJ. & Mokhtarzadeh M., ۲۰۱۶. **Selecting the band in order to optimize the spectral separation of supercritical images.** The ۱st National Conference on Geospatial Information Technology, K.N.Toosi University of Technology, Tehran, Iran.

Ghamisi P., Plaza J., Chen Y., Li J. & Plaza AJ., ۲۰۱۷. **Advanced spectral classifiers for hyperspectral images: A review.** IEEE Geoscience and Remote Sensing Magazine, ۵(۱): ۸-۳۲.

Luo, Y., Zou J., Yao C., Li T., Bai G. & HSI-CNN: **A Novel Convolution Neural Network for Hyperspectral Image.** arXiv ۲۰۱۸, arXiv:۱۸۰۲.۱۰۴۷۸.

Wang Y., Sun A., Han J., Liu Y. & Zhu, X. **Sentiment Analysis by Capsules.** In Proceedings of the 2018 World Wide Web Conference on World Wide Web, Lyon, France, ۲۳-۲۷ April ۲۰۱۸; pp. ۱۱۶۵-۱۱۷۴.

Sabour S., Frosst N. & Hinton G.E. **Dynamic routing between capsules.** In **Proceedings of the Advances in Neural Information Processing Systems**, Long Beach, CA, USA, ۴-۹ December ۲۰۱۷; pp. ۳۸۵۹-۳۸۶۹

Hu W., Huang Y., Wei L., Zhang F. & Li H. **Deep convolutional neural networks for hyperspectral image classification.** J. Sens, ۲۰۱۵, 2015, ۲۰۸۶۱۹.

Yu Y., Gu T., Guan H., Li D. & Jin S. **Vehicle detection from high-resolution remote sensing imagery using convolutional capsule networks.** IEEE Geosci. Remote Sens. Lett, ۲۰۱۹, 16, 1894-1898.

Gao Y., Gao F., Dong J. & Li H.C. **SAR image change detection based on multiscale capsule network.** IEEE Geosci. Remote Sens. Lett, ۲۰۲۰.

Paoletti, M. E., J. M. Haut, J. Plaza & A. Plaza, ۲۰۱۸. **“A New Deep Convolutional Neural Network for Fast Hyperspectral Image Classification.”** ISPRS Journal of Photogrammetry and Remote Sensing ۱۴۵: ۱۲۰-۱۴۷. doi: ۱۰.۱۰۱۶/j.isprsjprs.۲۰۱۷.۱۱.۰۲۱.

Paoletti, M. E., J. M. Haut, R. Fernandez-Beltran, J. Plaza, A. Plaza, L. Jun & F. Pla, ۲۰۱۹. **“Capsul Networks for Hyperspectral Image Classification.”** IEEE Transactions on Geoscience and Remote Sensing ۵۷ (۴): ۲۱۴۵-۲۱۶۰. doi: ۱۰.۱۱۰۹/TGRS.۲۰۱۸.۲۸۷۱۷۸۲.

Xu Q., Wang D.Y. & Luo B. **Faster multiscale capsule network with octave convolution for hyperspectral image classification.** IEEE Geosci. Remote Sens. Lett, ۲۰۲۰.

Chen T. & Guestrin C., 2016. **Xgboost: A scalable tree boosting system.** In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining.

Xu JL., Esquerre C. & Sun DW., ۲۰۱۸. **Methods for performing dimensionality reduction in**

-
- hyperspectral image classification.** Journal of Near Infrared Spectroscopy, 26(1): 61-70.
- Li H., 2018. **Deep learning for natural language processing: advantages and challenges.** National Science Review. 0(1): 24-26.
- Maggiori E., Tarabalka Y., Charpiat G. & Alliez P., 2016. **High-Resolution Semantic Labeling with Convolutional Neural Networks.** ArXiv: 1611.01972.
- LeCun Y., Bengio Y. & Hinton G., 2010. **Deep learning.** Nature, 021: 436-444.
- Xue Z., You D., Candemir S., Jaeger S., Antani S., Long L. R. & Thoma G. R., 2010. **Chest x-ray image view classification.** In Proceedings of the 18th International Symposium on Computer-Based Medical Systems (CBMS) (pp. 66-71). Brazil: Ribeirão Preto.
- Candemir S., Jaeger S., Palaniappan K., Musco J. P., Singh R. K., Xue Z. & McDonald C. J., 2014. **Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration.** IEEE transactions on medical imaging, 33(2), 077-090.
- Bengio, Yoshua & Ian J. Goodfellow, and Aaron Courville. "Deep learning." Nature 021,7003, 2010: 436-444.
- Pedro Domingos. **A few useful things to know about machine learning.** Communications of the ACM, 2012:0
- Ghassemian H., Keshavarz A. & Landgrebe D., 2013. **Hyper-spectral image processing and analyses.** Space Magazine, 1(3): 32-41.
-