International Journal of Research in Engineering, Science and Management
Volume 6, Issue 4, April 2023
https://www.ijresm.com | ISSN (Online): 2581-5792

Remote Sensing Scene Classification Using Convolutional Neural Network

V. Karthick[1], M. C. Aiswarya[2*], G. R. Jayavarshini[3], R. Sruthi[4]
[1]*Assistant Professor, Department of Electronics and Communication Engineering, Velammal College of Engineering and Technology, Madurai, India*
[2,3,4]*Department of Electronics and Communication Engineering, Velammal College of Engineering and Technology, Madurai, India*

***Abstract*: Several sectors can benefit from the use of remote sensing picture scene classification, which strives to classify remote sensing images into a number of semantic categories depending on their content. Deep learning-based remote sensing image scene categorization has generated a lot of attention and achieved significant strides as a result of these networks' strong feature learning skills. To the best of our knowledge, there hasn't been a comprehensive analysis of recent deep learning advances for scene classification in remote sensing images. This process provides a thorough evaluation of deep learning algorithms for remote sensing picture scene classification, which is crucial given the field's rapid progress. The remote sensing scene is analyzed using the deep learning algorithm from the input remote sensing photographs, and then the deep learning method is utilized. After all of the images have been trained, the deep learning technique predicts the outcome using accuracy, precision, recall, and F1-score.**

***Keywords*: Convolutional Neural Network (CNN), Remote Sensing (RS), Scene classification.**

1. Introduction

A. Introduction: Remote Sensing

Contrary to in situ or on-site observation, it is the gathering of data regarding a phenomenon or object without coming into direct touch with it. The phrase is specifically used in reference to gathering knowledge about the Earth and other planets. Geography, land surveying, and the majority of Earth science disciplines (such as hydrology, ecology, meteorology, oceanography, glaciology, and geology) are just a few of the disciplines that use remote sensing. It also has uses in the military, intelligence, commercial, economic, planning, and humanitarian sectors, among other things. The phrase "remote sensing" is currently used to describe the detection and classification of Earth-based objects using satellite- or aircraft-based sensor technologies. Based on propagating signals, it includes the surface, the atmosphere, and the oceans (e.g. electromagnetic radiation). It can be divided into "active" and "passive" remote sensing, depending on how a signal is transmitted to an object by a satellite or aircraft and detected by a sensor (when the reflection of sunlight is detected by the sensor).

B. Project Introduction

We can measure and see specific structures on the Earth's surface with the help of remote sensing images, a useful data source for earth observation. The amount of remote sensing photographs is rapidly increasing as a result of advancements in earth observation technologies. This has increased the necessity of finding the best way to utilize the growing amount of remote sensing data for insightful earth observation. So, it is crucial to comprehend vast and intricate remote sensing images. Scene classification of remote sensing images has been an important study subject since it is a crucial and difficult problem for correctly interpreting remote sensing information. The goal of remote sensing picture scene classification is to accurately tag provided remote sensing images with predetermined semantic categories. Remote sensing image classification eventually developed three parallel classification branches at various levels, including pixel-level, object-level, and scene-level classification, as the spatial resolution of remote sensing images increased. It is important to note that we refer to the classification of remote sensing images at the pixel, object, and scene levels collectively as "remote sensing image classification."

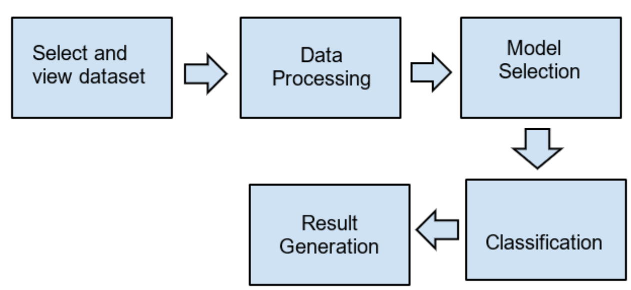

Fig. 1.

Due to the low spatial resolution of remote sensing images—the size of a pixel is comparable to the sizes of the objects of interest—researchers initially concentrated on classifying remote sensing images at the pixel level or sub-pixel level by labeling each pixel in the images with a semantic class. The classification of remote sensing images at the pixel level. In the fields of multispectral and hyperspectral remote sensing image analysis, it is still an active research topic. a guide for

*Corresponding author: aiswaryamc10@gmail.com

classifying remote sensing images. Remote sensing image classification eventually developed three parallel classification branches at various levels: classification at the pixel level, classification at the object level, and classification at the scene level. It is important to note that we use the term "remote sensing image categorization" in this context.

2. Related Work

By 2020, Xinting Yang, Song Zhang, Jintao Liu1, Qinfeng Gao, Shuanglin Dong, and Chao Zhou hoped to develop deep learning for intelligent fish aquaculture.The applications of deep learning (DL) for smart fish aquaculture were thoroughly and in-depth investigated in this process. The classification of species, behavioral studies, feeding choices, estimates of size or biomass, and predictions of water quality are the divisions. Data and algorithms, the two primary components of artificial intelligence (AI), were used to thoroughly analyze the described methods' technical details. The ability of DL to automatically extract features is by far its most significant addition, according to performance comparisons with conventional methods based on manually extracted features. A DL model can efficiently extract key fish characteristics when recognising fish. Such models outperform more conventional artificial feature extraction techniques and have demonstrated strong consistency in difficult environments like low light and high noise. Additionally, there are disadvantages to using this, such as the need for extensive data collection and DL training due to the wide variety of fish species and the body shapes and postures that vary greatly depending on the stage of development.

In 2017, Xinhuai Zou, Ming Cheng, Cheng Wang, Yan Xia, and Jonathan Li set out to implement deep learning to classify point clouds from complex forests. In this study, we presented a brand-new rasterization-based technique for classifying different tree species from TLS point clouds of intricate woodland scenes. Our approach entails individual tree extraction, noise reduction, voxel-based rasterization of tree features, and DBN model classification of tree species. Experiments reveal that both data sets attain high accuracy. A potent way to express knowledge about 3-D objects is through rasterization. We'll keep thinking about better methods to represent 3-D objects in the future.The network's initialization of its parameters does not happen at random; rather, it adjusts them beforehand in a way that makes convergence comparatively simple. The primary drawback of this endeavor is line of sight. Given the visual nature of 3D laser scanning, it is impossible to measure any surface that is not in the line of sight of the scanner.

Junwei Han, Dingwen Zhang, Gong Cheng, Lei Guo, and Jinchang Ren aimed to build object Detection in Optical Remote Sensing Images Based on Weakly Supervised Learning and High-Level Feature Learning in the year 2015. In this work, we have proposed a novel framework to tackle the problem of object detection in optical RSIs. The novelties that distinguish the proposed work from previous works lie in two major aspects. First, instead of using traditional supervised or semi supervised learning methodology, this paper developed a WSL framework that can substantially reduce the human labor of annotating training data while achieving outstanding performance. Second, we developed a deep network to learn high-level features in an unsupervised manner, which offers a more powerful descriptor to capture the structural information of objects in RSIs. It thus can improve the object detection performance further. Experiments on three different RSI data sets have demonstrated the effectiveness of the proposed work.In this work, we have proposed a novel framework to tackle the problem of object detection in optical RSIs. The novelties that distinguish the proposed work from previous works lie in two major aspects. First, instead of using traditional supervised or semi supervised learning methodology, this paper developed a WSL framework that can substantially reduce the human labor of annotating training data while achieving outstanding performance. Second, we developed a deep network to learn high-level features in an unsupervised manner, which offers a more powerful descriptor to capture the structural information of objects in RSIs. It thus can improve the object detection performance further. Experiments on three different RSI data sets have demonstrated the effectiveness of the proposed work. The Bayesian framework is used for generating accurate initial training examples and the iterative training scheme is used to gradually refine the object detector. But it also produces posterior distributions that are heavily influenced by the priors.

In 2016 saw the creation of AID: A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification by Gui-Song Xia, Jingwen Hu, Fan Hu, Baoguang Shi, Xiang Bai, Yanfei Zhong, and Liangpei Zhang. In this study, we first provide a thorough analysis of aerial scene classification by providing a concise synopsis of the available methods. We discover that the findings on the most widely used datasets are already saturated and greatly impede the advancement of aerial scene classification. The largest and most difficult dataset for the scene classification of aerial pictures, AID, is created as part of our effort to tackle the issue. The dataset serves as a benchmark resource for the research community to enhance cutting-edge aerial scene analysis techniques. The primary benefit of providing aid is that it facilitates the recovery of homes and livelihoods following a disaster. But frequently, especially in models with a lot of parameters, it has a high computational cost.

In 2016, the following individuals set out to create RSI-CB: A Large-Scale Remote Sensing Image Classification Benchmark through Crowd source Data: Haifeng Li, Xin Dou, Chao Tao, Zhixiang Hou, Jie Chen, Jian Peng, Min Deng, and Ling Zhao. Due to its many amazing characteristics, including real-time classification, quick spread speed, robust information, low cost, and large amounts of data, crowd-sourced data has become the focus of research in geographic information science internationally. The RSI-CB, which is based on data gathered from the public, presents novel viewpoints and areas of research for the development of remote sensing datasets. The number of categories and images in the RSI-CB have greatly increased when compared to previous remote sensing datasets. The RSI-CB has six categories that are based on the land-use classification standard in China, and each category has multiple subcategories. Floods and forest fires that have spread across a vast area are simpler to find, making the planning of a rescue mission simple and quick. The information produced by remote sensing data may not be accurate and may only be there temporarily, which is the main disadvantage of this approach.

In 2013, Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton set out to create a Deep Convolutional Neural Networks-based ImageNet Classification. Our findings demonstrate that a large, deep convolutional neural network is capable of breaking records on a very difficult dataset when utilizing only supervised learning. It is noteworthy that removing just one convolutional layer causes our network's performance to suffer. For instance, eliminating any one of the middle layers reduces the network's top-1 performance by around 2%. Therefore, the depth is crucial for attaining our results. Even though we anticipate that it will be helpful, especially if we are able to gain enough computational power to considerably increase the size of the network without obtaining a comparable increase in the amount of labeled data, we did not employ any unsupervised pre-training in order to simplify our experiments. As our network has grown and been trained for a longer period of time, our findings have thus far improved, but we still have a long way to go before we can match the infer temporal pathway of the human visual system. In the end, we aim to deploy very large and deep convolutional nets on video sequences where the temporal structure provides very useful information that is missing or much less visible in static images. In the absence of precise land surveying techniques, remote sensing provides a practical and somewhat inexpensive approach of rebuilding a base map. Incomplete remote sensing data collection occurs when large-scale engineering maps cannot always be produced using satellite data.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke and Andrew Rabinovich proposed Going deeper with convolutions in the year 2014. Our results seem to yield solid evidence that approximating the expected optimal sparse structure by readily available dense building blocks is a viable method for improving neural networks for computer vision. The main advantage of this method is a significant quality gain at a modest in-crease of computational requirements compared to shallower and less wide networks. Also note that our detection work was competitive despite neither utilizing context nor performing bounding box regression and this fact provides further evidence of the strength of the Inception architecture. Remotely sensed data can easily be processed and analyzed fast using a computer and the data utilized for various purposes. The main disadvantage is that powerful active remote sensing systems such as radars emit their own electromagnetic radiation that can be intrusive and also affect the phenomenon being investigated.

3. Methodology

A. Existing Methodology

To complete the RS scene categorization task in the current system, suggest a dual-branch network (ACNet). The classification model is one of four components, along with the parallel-attention model, the attention-consistent model, and the intermediate feature extraction model.

The intermediate feature extraction model acquires knowledge of the global features of the input picture pairs, which were created through spatial rotation. The local information from RS pictures is then thoroughly explored using two attentional strategies that are being used concurrently. There are a number of drawbacks connected to this practice, including It is expensive to analyze repetitive images if it is necessary to evaluate various parts of the image features, objects can be misclassified or confused, distortions may appear in an image owing to the relative motion of sensor & source

B. Proposed Methodology

To address the drawback of the current system, the proposed model is presented. By classifying the dataset of digital images from remote sensing scenes using a deep learning algorithm, this system will improve the accuracy of the neural network outcomes.

The overall categorization results perform better as a result. The accuracy is found to be more dependable when the cancer picture is predicted. Apart than eliminating the drawbacks of the current system, which include those that result from its shortcomings, this methodology has other benefits. With digital photos, CNN increases the screening accuracy, identifying remote sensing scenes takes less time, and a larger area is covered: Remote sensing permits regional surveys on a range of topics as well as the identification of very large features by allowing for the covering of very wide areas.

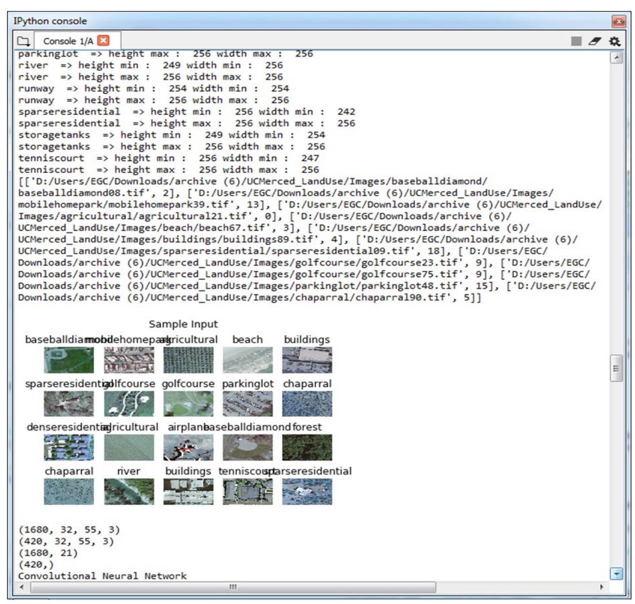

Fig. 2.

The method for selecting the data for the NWPU-RESISC45 dataset is known as data selection. In this project, the scene is located using digital images from remote sensing. Agricultural, aeroplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, motorway, golf course, harbour, intersection, medium residential, mobile home park, flyover, parking lot, river, runway, sparse residential, storage tanks and tennis court remote sensing scene images are all included in this dataset.

Obtaining rescaled data from the dataset is the process of image data pre-processing. Resize image dataset and Obtaining data are included. The size of the remote sensing scene dataset photos is scaled down to 50 in the Resize image dataset. In the process of gathering data, categorical data is referred to as variables having a limited number of rescaled values. Input and output variables for deep learning algorithms must be arrays.

Data splitting is the process of dividing a set of data into two halves, typically for cross-validation needs. One portion of the data is used to create a prediction model, while the other portion is utilized to assess how well the model is working. Analyzing image processing models requires dividing picture data into training and testing sets.Usually, the majority of the image data from a data set is used for training, and a smaller piece of the data is utilized for testing when it is divided into a training set and testing set.

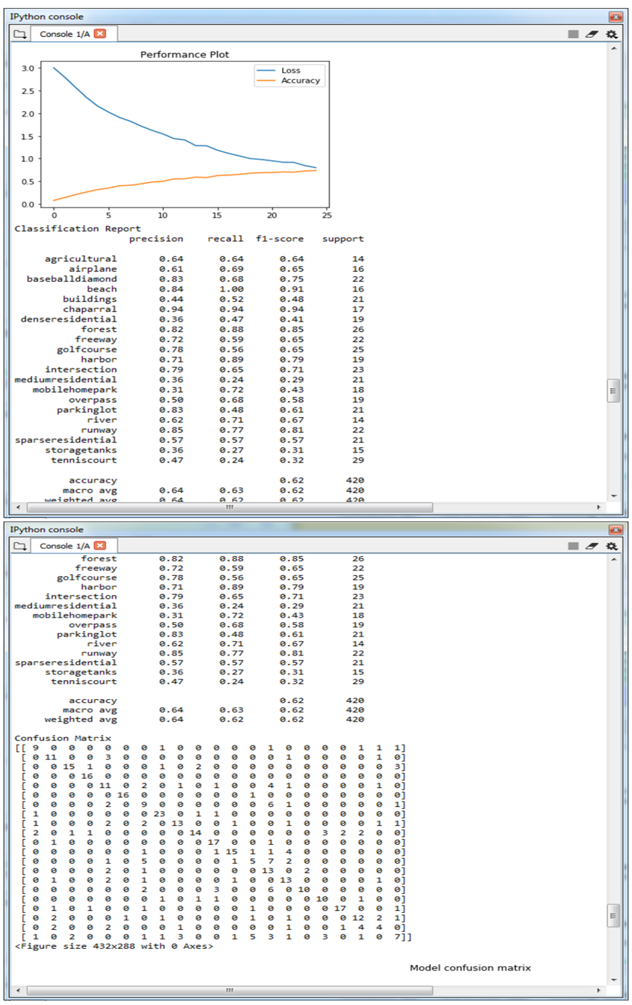

Fig. 3.

For the categorization process in this project, CNN is utilized. A class of deep neural networks known as convolutional neural networks (CNN, or ConvNet) in deep learning are most frequently used to analyze visual imagery. The fields of image and video recognition, recommender systems, image classification, medical image analysis, natural language processing, brain-computer interfaces, and financial time series all make use of them. Multilayer perceptrons are regularized variants of CNNs. Typically, multilayer perceptrons refer to completely linked networks, meaning that every neuron in one layer is coupled to every neuron in the following layer. These networks are susceptible to overfitting data since they are "completely linked." Regularization methods frequently involve adding some kind of magnitude measurement of weights to the loss function.A distinct technique to regularization is used by CNNs, which make use of the hierarchical structure of the data to patch together more complicated patterns from smaller, simpler ones. CNNs are at the lowest end of the connectivity and complexity spectrum as a result. Since the connecting arrangement between neurons mirrors how the animal visual brain is set up, convolutional networks were motivated by biological processes. The receptive field, a constrained area of the visual field, is the sole area where individual cortical neurons react to inputs. The full visual field is covered by the partial overlap of the receptive fields of several neurons.

The procedure of prediction is employed to foretell the remote sensing scene from the dataset.By improving the performance of the overall prediction results, this project will successfully predict the dataset's data.

The overall classification and forecast will be used to create the Final Outcome. Using metrics like Accuracy, Precision, Recall, and F1-measure, the performance of the suggested technique is assessed. The photos are trained to recognise the sensory scene using the unsupervised training techniques in feature. Finding the remote sensing scene from the photograph will also be improved in the graphical user interface. It will improve performance and increase accuracy while giving the result right away.

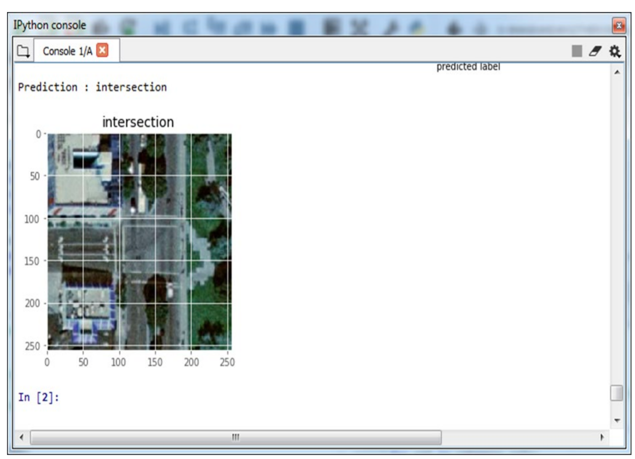

Fig. 4.

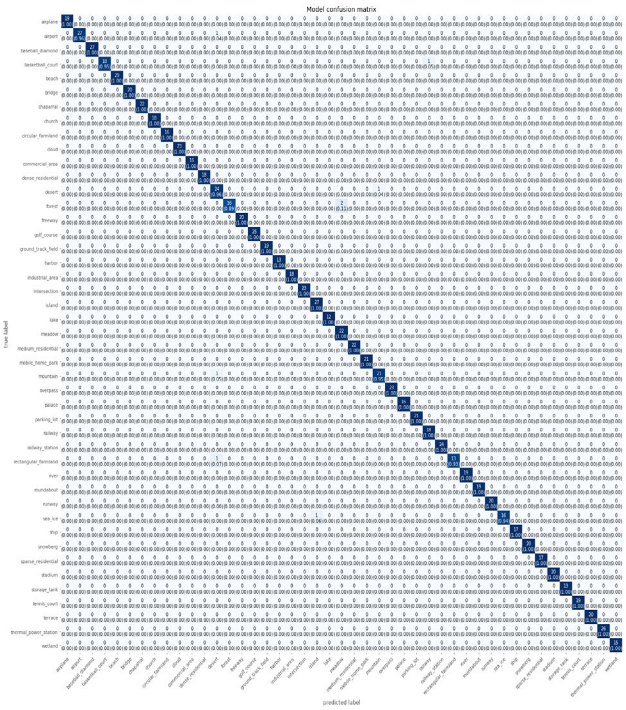

Fig. 5.

4. Conclusion

In this research, the remote sensing scene is analyzed from images by a deep learning classifier. The NWPU-RESISC45 data is used in a pre-processing technique as input data. The images are resized and turned into arrays in the pre-processing technique. It is then put through a feature selection technique, where the dataset is divided into a training dataset and a testing dataset. Then, each picture is resized and turned into an array. The remote sensing scene is finally analyzed from images using the classification technique. Implementing a deep learning method, which bases its predictions on accuracy

References

[1] Zhu et al., "Deep learning for smart agriculture: Concepts, tools, applications, and opportunities," Int. J. Agricultural Biol. Eng., vol. 11, no. 4, pp. 32–44, 2018.

[2] J. Marçais and J.-R. de Dreuzy, "Prospective interest of deep learning for hydrological inference," Ground Water, vol. 55, pp. 688–692, 2017.

[3] X. Zou, M. Cheng, C. Wang, Y. Xia, and J. Li, "Tree classification in complex forest point clouds based on deep learning," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 12, pp. 2360–2364, Dec. 2017.

[4] L. Chen, W. Yang, K. Xu, and T. Xu, "Evaluation of local features for scene classification using VHR satellite images," in Proc. IEEE Joint Urban Remote Sens. Event, 2011, pp. 385–388.

[5] G. Sheng, W. Yang, T. Xu, and H. Sun, "High-resolution satellite scene classification using a sparse coding based multiple feature combination," Int. J. Remote Sens., vol. 33, no. 8, pp. 2395–2412, 2012.

[6] L. Jiao, X. Tang, B. Hou, and S. Wang, "SAR images retrieval based on semantic classification and region-based similarity measure for earth observation," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol.8, no. 8, pp. 3876–3891, Aug. 2015.

[7] X. Tang and L. Jiao, "Fusion similarity-based reranking for SAR image retrieval," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 2, pp. 242–246, Feb. 2017.

[8] X. Tang, L. Jiao, W. J. Emery, F. Liu, and D. Zhang, "Two-stage reranking for remote sensing image retrieval," IEEE Trans. Geosci. Remote Sens., vol. 55, no. 10, pp. 5798–5817, Oct. 2017.

[9] X. Tang, L. Jiao, and W. J. Emery, "SAR image content retrieval based on fuzzy similarity and relevance feedback," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 10, no. 5, pp. 1824–1842, May 2017.

[10] Q. Zhu, Y. Zhong, L. Zhang, and D. Li, "Scene classification based on the fully sparse semantic topic model," IEEE Trans. Geosci. Remote Sens., vol. 55, no. 10, pp. 5525–5538, Oct. 2017.

[11] R. Xu, Y. Tao, Z. Lu, and Y. Zhong, "Attention-mechanism-containing neural networks for high-resolution remote sensing image classification," Remote Sens., vol. 10, no. 10, p. 1602, 2018.

[12] Q. Zhu, Y. Zhong, L. Zhang, and D. Li, "Adaptive deep sparse semantic modeling framework for high spatial resolution image scene classification," IEEE Trans. Geosci. Remote Sens., vol. 56, no. 10, pp. 6180–6195, Oct. 2018.

[13] Q. Zhu, Y. Zhong, S. Wu, L. Zhang, and D. Li, "Scene classification based on the sparse homogeneous-heterogeneous topic feature model," IEEE Trans. Geosci. Remote Sens., vol. 56, no. 5, pp. 2689–2703, May 2018.

[14] X. Tang, X. Zhang, F. Liu, and L. Jiao, "Unsupervised deep feature learning for remote sensing image retrieval," Remote Sens., vol. 10, no. 8, p. 1243, 2018.

[15] H. Sun, S. Li, X. Zheng, and X. Lu, "Remote sensing scene classification by gated bidirectional network," IEEE Trans. Geosci. Remote Sens., vol. 58, no. 1, pp. 82–96, Jan. 2020.

[16] X. Tang, C. Liu, J. Ma, X. Zhang, F. Liu, and L. Jiao, "Large-scale remote sensing image retrieval based on semi-supervised adversarial hashing," Remote Sens., vol. 11, no. 17, p. 2055, 2019.

[17] C. Liu, J. Ma, X. Tang, F. Liu, X. Zhang, and L. Jiao, "Deep hash learning for remote sensing image retrieval," IEEE Trans. Geosci. Remote Sens., to be published, doi: 10.1109/TGRS.2020.3007533.

[18] S. R. Gunn et al., "Support vector machines for classification and regression," ISIS Tech. Rep., vol. 14, no. 1, pp. 5–16, 1998.

[19] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition," IEEE Trans. Image Process., vol. 11, no. 4, pp. 467–476, Apr. 2002.

[20] X. Lu, X. Zheng, and Y. Yuan, "Remote sensing scene classification by unsupervised representation learning," IEEE Trans. Geosci. Remote Sens., vol. 55, no. 9, pp. 5148–5157, Sep. 2017.