

THE USE OF DEEP LEARNING IN REMOTE SENSING FOR MAPPING IMPERVIOUS SURFACE: A REVIEW PAPER

Sulaiman Mahyoub*¹, Hassan Rhinane¹, Mehdi Mansour¹, Abdelhamid Fadil², Akensous Youness³ AL Sabri Ahmed⁴

¹ Laboratory Geosciences, Department of Geology, Faculty of Sciences, Hassan II University Casablanca, Morocco

² Hassania School of Public Works, Morocco

³ Laboratory of Applied Geology, Geomatics and Environment, Department of Geology, Faculty of sciences Ben M'sik, Hassan II University Casablanca, Morocco

⁴ Department of Physics, laboratory of engineering and materials, Faculty of sciences Ben M'sik, Hassan II University Casablanca, Morocco

aldoubai100@gmail.com*, h.rhinane@gmail.com, medimansour@gmail.com, fadil.abdelhamid@gmail.com, akensous1990@gmail.com, a.alsabri2011@gmail.com

Commission 4, WG 7

KEY WORDS: Impervious Surface, Remote Sensing, Deep Learning, CNNs, DNNs, Review.

ABSTRACT

In recent years, deep convolutional neural networks (CNNs) algorithms have demonstrated outstanding performance in a wide range of remote sensing applications, including image classification, image detection, and image segmentation. Urban development, as defined by urban expansion, mapping impervious surfaces, and built-up areas, is one of these fascinating issues. The goal of this research is to explore at and summarize the deep learning approaches used in urbanization. In addition, several of these methods are highlighted in order to provide a comprehensive overview and comprehension of them, as well as their pros and downsides.

1. INTRODUCTION

Remote sensing has been successfully applied in many fields, such as classification and change detection. Retrieving accurate information from remotely sensed data remains a challenging task. Satellite images have complex patterns that are difficult to understand due to their heterogeneity (Adam, 2014). Identification of built-up areas is essential for territorial planning, climate change studies, population relocation, etc... However, image processing for classification and change detection involves certain pre-processing procedures that depend on the method used. Therefore, the remote sensing community is engaged in the development and use of new automatic remote sensing methods and in improving the performance of certain aspects, such as pre-processing, segmentation and classification. Neural networks, which form the basis of deep learning, have been used for many years in different aspects of science. However, the remote sensing community has turned to deep learning as an excellent solution for change detection and land use classification.

In recent years, the emergence of DLs has sparked real interest in neural networks. Since 2014, DL algorithms have achieved considerable success in many image analysis tasks, including land use and land cover classification (LULC), scene classification, and object detection (Chen et al. 2014; Chen et al. 2015; Cheng et al. 2016; Kussul et al. 2017; Romero et al. 2015; Vetrivel et al. 2018; Yu et al. 2017; Zou et al. 2015). To date, most studies on DL have been either general reviews on DL algorithm development (LeCun et al. 2015; Zhang et al. 2016) or detailed thematic reviews for a few sensitive areas, e.g., speech recognition (Hinton et al. 2012) and medical image recognition (Litjens et al. 2017). Although other studies have examined applications of DL in remote sensing (Li et al. 2018; Liu et al. 2018) many important areas have not been researched.

As an example, the review by Liu et al. (2018) focused only on the application of DL for remote sensing data fusion, and ignored its application in other areas of remote sensing (Ma et al. 2019).

Deep learning algorithms have been used in remote sensing applications for their outstanding performance in change detection, land cover classification and segmentation (Everingham et al. 2010; Lee et al. 2015; Lin et al. 2014). One of the most renewed topics in remote sensing is the monitoring and assessment of urbanization. In particular, Buildings, roads and a number of other surfaces. Many studies have been dedicated to this issue using different remote sensing methods. In this work, we intend to present deep learning approaches that have been applied to urban mapping, such as built-up areas, etc.

Not so long ago, the traditional ML approaches were gained attention for detecting the built-up area and impervious surface like Support vector machine, Random forest, k-nearest neighbour and decision trees classifiers. But recently, most of researchers, scholars and practitioners shifts their focuses to Deep learning to overcome the challenge of feature extraction in traditional ML, that is not automatic and handcrafted features generation is a time-consuming stage moreover improve their results accuracies (Hashemi-Beni et al. 2020). number of studies proves that deep learning can effectively deal with the limits of handmade features classifying the objects by extracting the features directly from the input data (Lee et al. 2015).

Deep learning methods in the domain of remote sensing applications have been extensively explored in several prior review publications, such as (Zhang et al. 2016), and (Zhu et al. 2017). In (Ma et al. 2019) presents a detailed and focused overview of deep learning in remote sensing in general, rather than focusing on a single application. recently, Deep learning

algorithms for semantic segmentation of remote sensing images have been the focus of a paper (Ma et al. 2019). To the best of our knowledge, this is the first study to offer a comprehensive overview of deep learning approaches for analyzing and determining impervious surfaces using earth observation images. This study covers all peer-reviewed, Scopus indexed papers, and we want to provide a strong reading foundation to novices in this topic.

The remainder of this paper will be organized as follows. Second section Introduce the selection criteria of this survey and the statistics of the works in relation to the mean subject. Then the methodology in section 3 following that, presentation of papers works presented in section 4, in section 5 a short presentation of papers works, finally, this document ends with a summary of the subject.

2. DATA SELECTION

The methodology utilized in this research is based on a stream of inclusion and exclusion criteria for this literature evaluation, which resulted in the selection of 34 papers from the Scopus Database. The advanced research's textual content, which is the title, is selected first, followed by the inclusion and exclusion keywords, as indicated in the table below. We attempt to focus exclusively on deep learning with impervious surface articles using these criteria, in order to provide a clear and helpful global overview for novices and readers. Furthermore, all selected articles were written in English and covered the period from 2011 to the present (October 22, 2021), with the most interesting ones highlighted.

```
PUBYEAR > 2010 TITLE ( ( "deep convolution neural networks"
OR "DNNs" OR "segmentation" OR "cnns" ) AND (
"impervious surfaces" OR "built-up" OR "built-up" OR "urban
area" OR "urban expansion" ) AND NOT "building height" AND
NOT "SAR" AND NOT "SAR Image" AND NOT "lidar" AND
NOT "radar" AND NOT "fusion" AND NOT "Integrating" AND
"deep learning" AND NOT "water body" AND NOT "3D point"
AND NOT "3D" AND NOT "point cloud" AND NOT
"photogrammetric imagery" AND NOT "photogrammetric images"
AND NOT "aerial images" ) AND ( LIMIT-TO ( LANGUAGE ,
"English" ) )
```

2.1 Preparation in electronic form

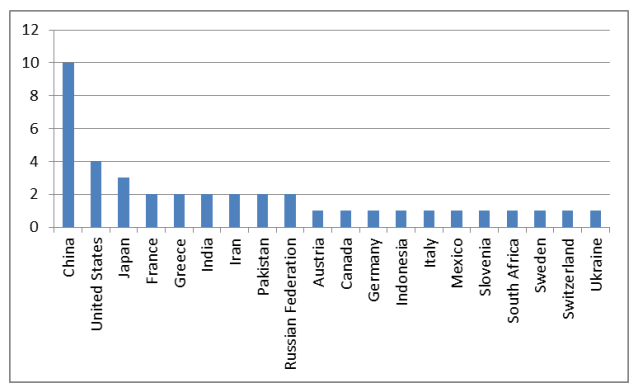


Figure 1. Number of papers by country.

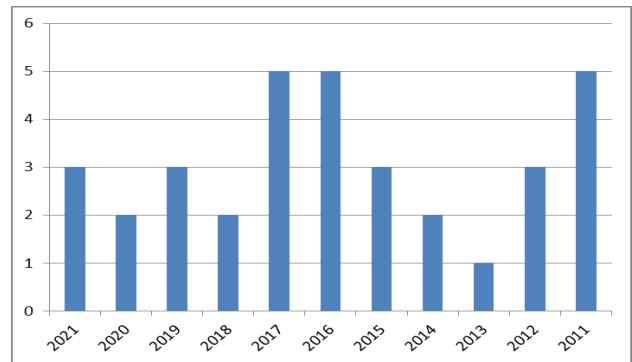


Figure 2. Number of papers by year.

The figure presented below illustrates the statistics of the works found via the search using the characters present from this section is that died that the works related to the theme follow an important increase in the years since 2017. Hence, the importance of the topic that we present in this paper.

A simple statistical analysis was performed from the above data to assess the current status and trends of DL use in the field of remote sensing. This analysis consisted of firstly calculating the number of papers published annually from 2011 to 2021, as shown in Figure 2; secondly the number of journal articles by regions 1, as shown in Figure 1. Finally, a presentation of the applied DLs for three types of classification tasks: LULC classifications, object detection, scene recognition, the applications of DL in remote sensing were mainly presented in this work. In 2021, however, the number of journal articles on this topic now exceeds the number of conference papers, which demonstrates the growing maturity of the research: most studies have focused on LULC classification, object detection and scene classification.

To carry out these analyses, the studies have mainly adopted reference image datasets. Therefore, there are a limited number of studies that have focused entirely on actual practical applications of DL for these remote sensing tasks. Although the use of DL models for segmentation, fusion, and registration has been less frequently studied, many journal articles have nevertheless pointed out that the models are relatively reliable and can be used for further analysis(Wang et al. 2018). This demonstrates that DL offers important and diverse perspectives in the field of remote sensing, which distinguishes it from other conventional classification algorithms.

In remote sensing image analysis, the CNN model has been the most commonly used, followed by the AE model (including the SAE model). (including the SAE model). RNN, DBN, and GAN models, the greatest popularity of the CNN is due to its unique characteristics that make it very suitable for processing multiband remote sensing images in which the pixels are arranged regularly(Audebert et al. 2018). As a result, there are a limited number of studies that have focused entirely on actual practical applications of DL for these remote sensing tasks. Although the use of DL models for segmentation, fusion, and registration has been less frequently studied, many journal articles have nevertheless pointed out that the models are relatively effective for these tasks (Kemker et al. 2018; Wang et al. 2018; Xing et al. 2018). This demonstrates that the CNN model offers important and diverse perspectives in the field of remote sensing and specifically for classification and

segmentation, which distinguishes it from other conventional classification algorithms in remote sensing image analysis.

3. METHODOLOGY

In this section, a CNN Architecture technique used for the classification framework for built-up area extraction is discussed.

CNN is one of the most popular computer vision algorithms that can be used for classification, due to its ability to process image data efficiently. The CNN model consists of multiple pooling and convolution operations, it is very effective in finding a more abstract and robust and efficient representation of image features in the input data (Maggiori et al. 2016). The weights are shared in the case of CNN. The weights are shared locally and connected to the same output unit form a filter In the case of CNNs, (Romero et al. 2015).

CNN architecture consists of several convolution and pooling layers. To generate convolved feature maps, subsequently kernel functions as filters are used in the convolution layers. The convolved features are generalized to higher levels by a subsampling layer, which makes the features more abstract and robust. this architecture is Similar to the structure of the primary visual cortex system where simple and complex cells are stacked in layers, the convolution and pooling layers are mixed in the CNNs. noting that the number of convolution and pooling layers can be different and usually depends on the application and the available data.

At the end the pooling layers are inserted after the convolution layers, so that to reduce the spatial size and computational complexity. In addition, the features become more robust, so that the model used is less likely to be overfitted. The most important feature of the CNN is the sharing of weights in the convolution layers, so that the same filter bank can be used for all pixels in a particular layer. Filter bank can be used for all pixels in a particular layer.

4. PRESENTATION OF PAPERS WORKS

In this section, various studies used for the classification for built-up area extraction are presented.

(Núñez 2015a)suggested an automated segmentation approach based on Cellular Neural Network (CNN) for endmember recognition to urban impervious surface, which was applied to a Very High-Resolution image (VHR) of WorldView-2 (WV-2) with the same date as Landsat-8 OLI. According to the findings, CNN segmentation was the most accurate approach, followed by classic single threshold-based segmentation.

(Bodani et al. 2017)develops a completely automated deep learning technique for detecting built-up land areas in LISS-4 satellite ortho-imagery using a convolutional neural network. Aside from that, the author goes into the design, training, and performance assessment of a deep convolutional network for this remote sensing application.

(Tan et al. 2017)comprises of a numerous structure of deep convolution neural network (CNN) to extract built-up area automatically, which can blend the information of panchromatic and multispectral remote sensing pictures. Experiments on a large number of Gaofen-2 images show that the suggested approach outperforms existing methods. The CNN structure is

implemented using the Caffe framework, and the classification accuracy reaches.

In Jakarta Province, (Irwansyah et al. 2020) suggested a semantic image segmentation deep learning module for building detection using aerial photograph images. The average training accuracy of the UNet model was 0.83, while the testing accuracy was 0.87.

(Khryashev and Ivanovsky 2019)create a convolutional neural network-based approach for automatically extracting building locations from high-resolution aerial imagery The Sorensen-Dice coefficient of similarity was utilized by the researcher to compare the outcomes of algorithms with genuine masks. Before training the created convolutional neural networks, these masks were produced automatically and split into smaller portions in conjunction with corresponding aerial images.

(Fu et al. 2019) introduced an object-based deep CNN framework that combines object-based image analysis (OBIA) and deep CNNs to extract and predict impervious surfaces from high-spatial-resolution (HSR) images in town-rural. The proposed strategy, according to the author, offers a greater effective identification accuracy. In comparison to existing methods such as OBIA, FCN-8s, and U-Net.

(Iqbal and Ali 2020) build a segmentation network (as an encoder-decoder) module for built-up environments such as built-up areas in forests and deserts, mud buildings, tin, and colorful roofing. The module achieves substantial improvements in IoU ranging from 11.6 to 52 percent over existing state-of-the-art techniques.

(McGlinchy et al. 2021) map impervious surfaces at the pixel level from WorldView-2 in a mixed urban/residential environment using fully convolutional neural networks (FCNNs), precisely U-Net with a ResNet-152 encoder. The FCNN's given set of input multispectral bands are 3, 4, and 8. According to the research, residential areas performed well, whereas more densely built regions performed poorly. Fully convolutional neural network for impervious surface segmentation in mixed urban environment

5. IMPERVIOUS SURFACE DEEP LEARNING APPLICATIONS

The majority of the Deep learning studies that applied to impervious surfaces or built-up areas were split between semantic image segmentation, image classification object-based and hybrid method; object-based deep CNN - hybrid (Fu et al. 2019),classification (Tan et al. 2017), Semantic segmentation(Bodani et al. 2017; Iqbal and Ali 2020; Irwansyah et al. 2020; Khryashev and Ivanovsky 2019; Khurshid and Khan 2013; Núñez 2015b; Zhao et al. 2014).

6. DISCUSSION

Through this short survey research we deduce that, most of the studies have been used the UNET model for extracting built-up areas, as well as studies of extracting urban areas because it depends on the segmentation level rather than pixel level.

7. CONCLUSIONS

In this study, publications related to Deep learning in almost all subfields of remote sensing especially in building cartography were systematically analysed through a meta-analysis. Subsequently, several major areas of remote sensing using DL were summarized, including scene classification, LULC, and segmentation finally, a more in-depth review to have a presentation of the most effective and usable DL method in these sub-domains.

Therefore, this study provides an overview of the various remote sensing applications in which DL is used, as well as the specific opportunities and challenges of its use. For scene classification, semantic segmentation, and LULC classification, a supervised DL model (e.g., CNN) must be based on large amounts of training samples. In practice, the cost of acquiring training samples is relatively high.

For LULC classification, DL has shown super-accurate performance compared to traditional classifiers. Nevertheless, the performance of DL in LULC classification is still inferior to that of scene classification and object detection. This can be attributed to the frequent use of reference datasets for scene classification and object detection in previous studies. Therefore, DL developments for LULC classification have also employed high-resolution images on many occasions in the preceding publications, instead of real remote sensing data, as standard hyperspectral data.

DL models have been adapted in the remote sensing domain as well to allow their implementation in non-standard image processing tasks, such as object-based image analysis and time series analysis. For object-based image classification, a patch-based strategy is a generally accepted method (Huang et al., 2018; Fu et al., 2018), which integrates CNNs with OBIA. The critical problem with this approach is determining the values of relevant parameters (e.g., patch size), as the classification accuracy is largely affected by the values of these parameters.

For the pre-processing of remote sensing image, DL models including CNN have recently enjoyed considerable success in remote sensing image fusion, and therefore, other DL models such as RNN and GAN should be introduced in this field. The important limitation of DL in image registration is the lack of publicly available training datasets of available public training datasets, which should be a future goal of the remote sensing community. In the end, in addition to experiments on reference datasets, applications of DL technologies in remote sensing images deserve more profound research in the future.

REFERENCES

Audebert, N., Le Saux, B., & Lefèvre, S. (2018). Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140, 20-32

Bodani, P., Rastogi, K., & Gupta, U.K. (2017). Semantic segmentation of liss-4 satellite imagery for mapping built-up areas using deep learning convolutional neural network. In, *38th Asian Conference on Remote Sensing - Space Applications: Touching Human Lives, ACRS 2017: Asian Association on Remote Sensing*

Chen, Y., Lin, Z., Zhao, X., Wang, G., & Gu, Y. (2014). Deep Learning-Based Classification of Hyperspectral Data. *IEEE*

Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 2094-2107

Chen, Y., Zhao, X., Jia, X.J.I.J.o.S.T.i.A.E.O., & Sensing, R. (2015). Spectral-spatial classification of hyperspectral data based on deep belief network, 8, 2381-2392

Cheng, G., Zhou, P., Han, J.J.I.T.o.G., & Sensing, R. (2016). Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images, 54, 7405-7415

Everingham, M., Van Gool, L., Williams, C.K., Winn, J., & Zisserman, A.J.I.j.o.c.v. (2010). The pascal visual object classes (voc) challenge, 88, 303-338

Fu, Y., Liu, K., Shen, Z., Deng, J., Gan, M., Liu, X., Lu, D., & Wang, K. (2019). Mapping impervious surfaces in town-rural transition belts using China's GF-2 imagery and object-based deep CNNs. *Remote Sensing*, 11

Hashemi-Beni, L., Gebrehiwot, A.J.T.I.A.o.P., Remote Sensing, & Sciences, S.I. (2020). Deep Learning for Remote Sensing Image Classification for Agriculture Applications, 51-54

Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., & Sainath, T.N.J.I.S.p.m. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups, 29, 82-97

Iqbal, J., & Ali, M. (2020). Weakly-supervised domain adaptation for built-up region segmentation in aerial and satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 263-275

Irwansyah, E., Heryadi, Y., & Santoso Gunawan, A.A. (2020). Semantic Image Segmentation for Building Detection in Urban Area with Aerial Photograph Image using U-Net Models. In, *3rd IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology, AGERS 2020* (pp. 48-51): Institute of Electrical and Electronics Engineers Inc.

Kemker, R., Salvaggio, C., Kanan, C.J.I.j.o.p., & sensing, r. (2018). Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning, 145, 60-77

Khryashev, V., & Ivanovsky, L. (2019). Urban areas analysis using satellite image segmentation and deep neural network. In D. Rudoy, & V. Murgul (Eds.), *2019 Innovative Technologies in Environmental Science and Education, ITESE 2019: EDP Sciences*

Khurshid, M.H., & Khan, M.F. (2013). Segmentation of Built up area from SPOT 5 multispectral satellite images. In, *2013 33rd IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2013* (pp. 1963-1966). Melbourne, VIC

Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A.J.I.G., & Letters, R.S. (2017). Deep learning classification of land cover and crop types using remote sensing data, 14, 778-782

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436-444

Lee, S.H., Chan, C.S., Wilkin, P., & Remagnino, P. (2015). Deep-plant: Plant identification with convolutional neural

- networks. In, *2015 IEEE international conference on image processing (ICIP)* (pp. 452-456): IEEE
- Li, Y., Zhang, H., Xue, X., Jiang, Y., Shen, Q.J.W.I.R.D.M., & Discovery, K. (2018). Deep learning for remote sensing image classification: A survey, *8*, e1264
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C.L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740-755). Cham: Springer International Publishing
- Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., & Sánchez, C.I.J.M.i.a. (2017). A survey on deep learning in medical image analysis, *42*, 60-88
- Liu, Y., Chen, X., Wang, Z., Wang, Z.J., Ward, R.K., & Wang, X.J.I.F. (2018). Deep learning for pixel-level image fusion: Recent advances and future prospects, *42*, 158-173
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166-177
- Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P.J.I.T.o.g., & Sensing, R. (2016). Convolutional neural networks for large-scale remote-sensing image classification, *55*, 645-657
- Núñez, J.M. (2015a). Segmentation of Urban Impervious Surface Using Cellular Neural Networks. In, *Iberoamerican Congress on Pattern Recognition* (pp. 509-516): Springer
- Núñez, J.M. (2015b). Segmentation of urban impervious surface using cellular neural networks. In A. Pardo, & J. Kittler (Eds.), *20th Iberoamerican Congress on on Pattern Recognition, CIARP 2015* (pp. 509-516): Springer Verlag
- Romero, A., Gatta, C., Camps-Valls, G.J.I.T.o.G., & Sensing, R. (2015). Unsupervised deep feature extraction for remote sensing image classification, *54*, 1349-1362
- Tan, Y., Ren, F., & Xiong, S. (2017). Automatic extraction of built-up area based on deep convolution neural network. In, *37th Annual IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2017* (pp. 3333-3336): Institute of Electrical and Electronics Engineers Inc.
- Vetrivel, A., Gerke, M., Kerle, N., Nex, F., Vosselman, G.J.I.j.o.p., & Sensing, R. (2018). Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning, *140*, 45-59
- Wang, S., Quan, D., Liang, X., Ning, M., Guo, Y., Jiao, L.J.I.J.o.P., & Sensing, R. (2018). A deep learning framework for remote sensing image registration, *145*, 148-164
- Xing, Y., Wang, M., Yang, S., Jiao, L.J.I.J.o.P., & Sensing, R. (2018). Pan-sharpening via deep metric learning, *145*, 165-183
- Yu, X., Wu, X., Luo, C., Ren, P.J.G., & Sensing, R. (2017). Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework, *54*, 741-758
- Zhang, L., Zhang, L., Du, B.J.I.G., & Magazine, R.S. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art, *4*, 22-40
- Zhao, S., Wu, H., Tu, L., & Huang, B. (2014). Segmentation of Urban Areas Using Vector-Based Model. In Y. Zheng, P. Thulasiraman, B.O. Apduhan, Y. Nakamoto, N. Ning, & Y. Sun (Eds.), *11th IEEE International Conference on Ubiquitous Intelligence and Computing and 11th IEEE International Conference on Autonomic and Trusted Computing and 14th IEEE International Conference on Scalable Computing and Communications and Associated Symposia/Workshops, UIC-ATC-ScalCom 2014* (pp. 412-416): Institute of Electrical and Electronics Engineers Inc.
- Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F.J.I.G., & Magazine, R.S. (2017). Deep learning in remote sensing: A comprehensive review and list of resources, *5*, 8-36
- Zou, Q., Ni, L., Zhang, T., & Wang, Q. (2015). Deep Learning Based Feature Selection for Remote Sensing Scene Classification. *IEEE Geoscience and Remote Sensing Letters*, *12*, 1-5