

SPATIAL RESOLUTION SENSITIVITY ANALYSIS OF CLASSIFICATION OF SENTINEL-2 IMAGES BY PRE-TRAINED DEEP MODELS FROM BIG EARTH NET DATABASE

F. Naali¹, T. Alipour-Fard^{1*}, H. Arefi¹

¹ School of Surveying and Geospatial Engineering, University of Tehran, Tehran 1417414418, Iran – (f.naali, tayebalipour, hossein.arefi)@ut.ac.ir

KEY WORDS: Deep Learning, Convolutional Neural Network, Pre-Training, Big Earth Net, Residual Network

ABSTRACT:

Classifying and monitoring different vegetation types is important for forest management, food resources, and assessing the potential impacts of climate change. In this regard, several methods have been developed to study them using remote sensing data, and with the advent of neural networks, new methods are being proposed, especially in the field of automatic land use classification. In this research, multispectral Sentinel-2 satellite image has been used due to having spectral information and different spatial resolution for classifying plant species. Deep learning models have the ability to learn and recognize different features of images, but require a large number of training samples, so we used pre-trained ResNet networks with depths of 50, 101 and 152 layers, that trained with BigEarthNet dataset. The main purpose of this study is to evaluate the sensitivity of ResNet networks to spatial resolution. Results show that ResNet 101 was more stable than other networks, and the ResNet 50 with an overall accuracy of 76.2 has the highest accuracy at a resolution of 20 meters.

1. INTRODUCTION

Vegetation is one of the most important elements of an ecosystem. Vegetation affects living organisms, global climate and carbon cycle, so vegetation classification is important for natural resource management information on the distribution of vegetation types is a main resource for food chain planning, wildlife habitat, sustainable natural resource management, crop forests, and biodiversity conservation (Lu, Li, Moran, & Kuang, 2014). In recent years, due to increasing urbanization and natural disasters, various species of vegetation are in danger of extinction, so it is necessary to start protection and restoration programs of vegetation, which requires accurate maps of vegetation classification. Remote sensing data are known as the important sources for vegetation classification due to characteristics such as radiometric, spectral, spatial and temporal resolution. In image classification, the selection of suitable remote sensing (RS) data, evaluation and development of advanced algorithms based on neural networks to improve the performance of land cover classification have been some active research topics (Wang, Zhang, Niu, Wang, & Zhang, 2019). With the advance of digital technology and the appearance of different and new needs, it is absolutely necessary to provide modern and intelligent methods to provide an effective and compelling processing of remote sensing images. In this context, data analysis methods are essential to retrieve information from RS images, where classification is one of the main information extraction tasks providing the categorization of the observed surface at pixel level (Alipour-Fard & Arefi, 2020).

The remarkable development of deep learning models (Krizhevsky, Sutskever, & Hinton, 2012), which has been encouraged by the impressive computational and storage resources of new hardware devices and software technologies, has provided unprecedented results within RS image classification. In particular, convolutional neural network (CNN) is one of the most popular established deep learning architectures

for classification (He, Zhang, Ren, & Sun, 2016). The attractive feature of CNN is its astonishing ability to exploit the spatial correlation in data cube. Inspired by the biological visual cortex (LeCun, Bottou, Bengio, & Haffner, 1998), its architecture is based on grid-pattern receptive fields in a way that each convolution unit applies a linear function to the specific region defined by the receptive field on the input data. In this sense, the CNN builds a locally connected structure, in lieu of the standard fully connected architecture within traditional artificial networks, such as the multilayer perceptron (MLP) (Alipour-Fard, 2020). Indeed, the CNN is a stack of hierarchical n -dimensional filters, where each one comprises several weight matrices with parameter sharing mechanism, which are trained in order to learn specific patterns and features within the input data. Therefore, they are feature extractors that automatically learn hierarchies of features by adapting their weights to both the data and the conducted task, wherein bottom or lower filters (i.e. those that are more close to the input data) capture low-level features in the form of local patterns based on simple components, such as orientations of small segment of edges and outlines, whilst top/higher filters (i.e. those that are more close to the output data) extract more abstract features, which have been refined through the CNN hierarchy to extract more semantic and global information (Akbari, 2021). These final features are based on the whole input and describe its contents as a whole, and are therefore used for classification purposes.

The first CNN network was introduced by (LeCun, Bottou, Bengio, & Haffner, 1998) in 1998 with 5 continuous layers of convolution, pooling and fully connected. LeNet network is one of the first CNN networks, so it does not have the ability to reduce the calculations and the number of parameters. AlexNet Network was introduced in 2012. This network had a better potential to train due to its depth (Krizhevsky, Sutskever, & Hinton, 2012). To solve the Overfitting problem, they used the ReLU activation function and large filters with dimensions (11×11 and 5×5). VGG

* Corresponding author

network presented by (Simonyan & Zisserman, 2014). Instead of using large filters, this network uses 3×3 filters. using of small filters because it reduces the number of parameters decrease the complexity and time of computing. The depth of this network increases from 9 layers to 16 layers. Experience has shown that VGG networks have good results for classification, but the limitation for this architecture is its high computation. Although small-scale filters are used, the number of parameters used is about 140 million. ResNet presented by (He, Zhang, Ren, & Sun, 2016) Which provided an optimal way to train deeper networks. The ResNet architecture is 20 times deeper than AlexNet and 8 times deeper than VGG and has less computational complexity. In this network, in order to reduce the calculations, Residual blocks are used that provide a shortcut connection between the layers. According to research, deeper networks have more ability to extracting features. increasing the depth of the network creates problems such as Overfitting, but the presence of Residual blocks in the ResNet architecture solves this problem. In 2015, ResNet with a depth of 152 layers won the ImageNet Large Scale Visual Recognition Competition compared to other CNN and shallow networks (Alom et al., 2018). However, the successful implementation of these networks poses many challenges for researchers in the field of remote sensing. Currently one of the most important challenges in image classification of remotely sensed images is the lack of training samples, which is referred to as the small sample size problem. There are generally two basic approaches to overcome this challenge. The first approach is to collect additional training samples. This is usually achieved either by using direct data collection or annotation, which is a very costly and time-consuming task, or by generating virtual training samples using methods such as the generative adversarial network, which is also not useful in practice due to the complex nature of the relationships between classes (Alipour-Fard & Arefi, 2020). The second approach is to employ pre-training networks that are already trained on large databases. Pires de Lima et al has investigated the effect of transfer learning method on classification with CNN networks for remote sensing images. In (Pires de Lima & Marfurt, 2020), two methods were evaluated to classify images, i) VGG 19 and inception were trained directly using the scratch with small datasets. ii) Networks that were trained with ImageNet dataset were fine-tuned with remote sensing images. The results show that Using pre train networks has better performance for remote sensing classification. Most pre train networks were trained with computer vision datasets, including ImageNet and CIFAR. In this data sets, images with 3 RGB bands were used, and on the other hand, their existing features are generic, which is not suitable for remote sensing tasks. As a result, a group of researchers began to create datasets with remote sensing images that datasets are given in the table 1.

Dataset Name	Image Type	Annotation Type	Number of Images	Number of Classes
UC Merced	Aerial RGB	Single Label	2100	21
WHU-RS19	Aerial RGB	Single Label	1005	19
RSSCN7	Aerial RGB	Single Label	2800	7
AID	Aerial RGB	Single Label	10000	30
NWPU-RESISC45	Aerial RGB	Single Label	31500	45
RSI-CB	Aerial RGB	Single Label	36707	12
PatternNet	Aerial RGB	Single Label	30400	38
EuroSat	Satellite Multispectral	Single Label	27000	10

Table 1. A list of remote sensing datasets

The classes in UC Merced (Yang & Newsam, 2010), WHU-RS19 (W. Shao, Yang, & Xia, 2013), RSSCN7(Zou, Ni, Zhang, & Wang, 2015) , AID (Xia et al., 2017), NWPU-RESISC45 (Cheng, Han, & Lu, 2017), RSI-CB(Li et al., 2017), PatternNet (Zhou, Newsam, Li, & Shao, 2018), EuroSat(Helber, Bischke, Dengel, & Borth, 2019) datasets were single-label, with the class assigned to each patch being based on the dominant feature in that patch. Single-label datasets are sufficient for some remote sensing tasks such as distinguish 'building' and 'airport', but if the goal is to categorize similar classes such as 'dense forests' and 'scattered forests', CNN networks for training and extraction features need more detail, so researchers created multi-label dataset for generate accurate classification map and image retrieval (Z. Shao, Yang, & Zhou, 2018).

The Big Earth database has recently been introduced in (Sumbul, Charfuelan, Demir, & Markl, 2019), making it possible to classify images using pre-trained models on the BigEarthNet, rather than training a deep CNN from scratch. The BigEarthNet was generated using 590,326 multi-label patches of Sentinel-2 from various regions in Europe. From the transfer learning perspective, the performance of a pre-trained network in the face of changing conditions in the target image relative to the source images is an important issue that has not been addressed so far. In this research, the efficiency of the pre-trained model in the face of changing the spatial resolution of images has been investigated to produce classification maps. The main questions answered in this research are as follows:

- What is the performance of the pre-training models available on the Big Earth database in case of "changing spatial resolution"?
- Among the different variations of the ResNet as a powerful architecture, which one is more stable and less sensitive to changing the contribution of the target data into training procedure?
- If we intend to use pre-training models at an image with resolution other than the resolution of the source images, what could be the impact of fine-tuning on improving the classification results?

The rest of this article is structured as follows. Study area and dataset are explained in Section 2. Section 3 presents the detail of the proposed method. In Section 4, the experimental results are illustrated and discussed. Finally, Section 5 draws the conclusion of this article.

2. STUDY AREA AND DATA

Study area of this research is in Switzerland. Switzerland is located in $45^{\circ} 49' 2''$ N, $5^{\circ} 57' 22''$ E geographical coordinates. The predominant land use of Switzerland is forests and agricultural areas (Figure 1).

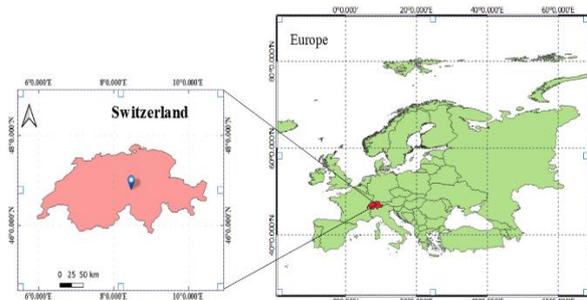


Figure 1. Location of study area

2.1 Sentinel 2 images

Sentinel-2 is developed by the European Space Agency specifically for the operational needs of the Copernicus program. The mission supports a wide range of services and applications such as agricultural monitoring, emergencies management, land cover classification or water quality. Sentinel-2 has 13 spectral bands: four bands at 10 m spatial resolution, six bands at 20 m spatial resolution, and three bands at 60 m spatial resolution (Figure 2, (Richter & Schl pfer, 2013)). The RGB display of the Sentinel-2 image is in Figure 3. This image was taken on 5/9/2020 and has a 7% cloud cover. If the cloud cover is high, the classification accuracy decreases.

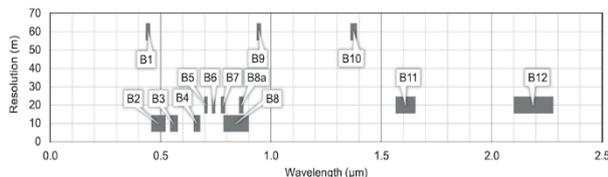


Figure 2. Sentinel-2 bands (Richter & Schl pfer, 2013)

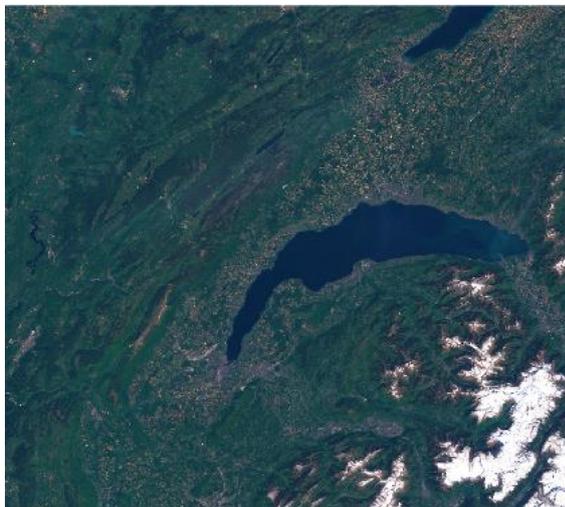


Figure 3. RGB image from Sentinel-2

2.2 BigEarthNet dataset

BigEarthNet is the first large-scale multi-label remote sensing dataset was constructed by the Remote Sensing Image Analysis (RSiM) Group and the Database Systems and Information Management (DIMA) Group at the Technische Universit t Berlin (TU Berlin). The images used in this archive are 125 Sentinel-2 tiles taken from June 2017 to May 2018. BigEarthNet is created by 125 Sentinel-2 mosaics from 10 different European countries (Austria, Belgium, Finland, Ireland, Kosovo, Lithuania, Luxembourg, Portugal, Serbia, Switzerland). One of the advantages of this dataset over other existing dataset is that each image patch is labeled by multiple land-cover classes that are provided from the CORINE Land Cover database of the year 2018 (CLC 2018) (Sumbul, Charfuelan, Demir, & Markl, 2019). BigEarthNet is much richer and more diverse than other dataset, both in terms of the number of images and the variety of classes. Atmospheric corrections were made on all tiles using Sen2Cor tool. Images with cloud and shadow coverage have been removed from the existing set of images because they are not suitable for teaching deep learning networks. The band #10 was not used in

this study due to lack of useful information about the earth's surface.

3. METHOD

In this section, we explain our proposed method in detail for satellite image classification (Figure 4). At a glance, In the first step, pre-processing is done on the images. Then we have the core of the proposed method, including convolutional neural network and how to train it, which are described in the following sections.

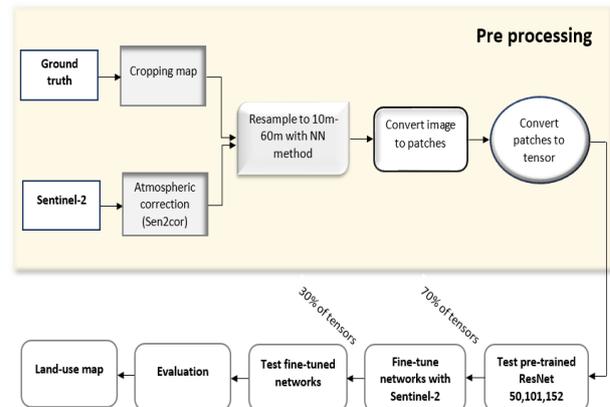


Figure 4. The research workflow

3.1 Preprocessing of satellite imagery

The Level-1C image of Sentinel-2 satellite was converted to Level-2A by Sen2cor plugin with Sentinel Application Platform (SNAP, version 7). Because Sentinel-2 imagery was georeferenced in the WGS 84 UTM 31N coordinate system, ground truth was transformed to the same coordinate system, and was clipped to the study area.

In order to investigate the sensitivity of networks to different resolutions, all bands have been resampled to a resolution of 20, 30, 40, 50 and 60 meters by the nearest neighbourhood interpolation method (Teoh, Ibrahim, & Bejo, 2008). In the next step, the image is divided into 8281 patches, each of which is part of 1) 120×120 pixels for 10m bands; 2) 60×60 pixels for 20m bands; 3) 40×40 pixels for 30m bands; 4) 30×30 pixels for 40m bands; 5) 24×24 pixels for 50m bands; 6) 20×20 pixels for 60m bands. The area of each patch is 1.44 square km ($1.2\text{km} \times 1.2\text{km}$) each image patch was annotated by multiple land-cover classes (i.e., multi-labels) according to the CORINE Land Cover database of the year 2018 (like BigEarthNet dataset), then all patches were converted to tensors to be used as CNN input in the next step.

3.2 Classification with ResNet network

Convolutional neural networks play a key role in the classification of images. Deep Convolutional Neural Networks (DCNN) is made up of a large number of convolutional layers that have the ability to extract features automatically compared to traditional image classification methods. As the number of layers increases, the network is able to recognized more complex features but deep networks are hard to train because of the vanishing gradient problem as the gradient is back-propagated to first layers, repeated multiplication may make the gradient infinitively small. As a result, as the network goes deeper, its performance gets saturated or even starts degrading rapidly. To solve this problem (He et al., 2016) suggested the Residual neural network (ResNet). The ResNet model operate skip connection, which transfers the inputs of the previous layer to the next layer

without intermediaries, and in the Back-propagation step, transfers the error of each layer to the previous layer, so the network becomes deeper and training faster. Because training deep convolutional neural network requires many training samples, in this study was used ResNet 50, ResNet 101, ResNet 152 networks that have been pre-trained with the BEN database. Two experiments were designed to evaluate pre-training models. In the first experiment, no training samples were used from the target images, and in the second experiment, to improve the performance, the networks were fine-tuned in 20 epochs with 30% of Sentinel-2 available labelled data (target image) and tested with 70% of the rest labelled data. Overall accuracy was used to evaluate the accuracy of the classification maps. In order to study the performance of the proposed network for RS classification, an implementation has been developed and tested on a hardware environment with a 6th Generation Intel® Core™ i7-6800K processor with 6M of Cache, installed over an ASUS motherboard. Also, a graphic processing unit (GPU) NVIDIA GeForce GTX 1080Ti with 8GB RAM. In order to provide an efficient implementation, the proposed model has been implemented over the TensorFlow 1.3 (an open source machine learning library developed by Google (Abadi et al., 2016)).

4. RESULTS AND DISCUSSION

Figure 5 shows the results of the first experiment. This image shows the average accuracy obtained for the four intended images. In general, it can be seen that the overall accuracy obtained in all three networks is low (less than 50%). It is expected to obtain maximum accuracy at spatial resolution of 20m, since the target and source images both have the same spatial resolution. The results of this image show that changing the resolution did not cause significant changes in the obtained results. Figure 6 shows the results of the second experiment. In this experiment, we used a limited number of training samples to fine-tune the parameters of the three networks to achieve higher accuracy.

By reducing the spatial resolution, the ResNet101 was able to maintain its efficiency, while the speed of decreasing accuracy while decreasing the spatial resolution in the ResNet152 is higher than the other two networks. This indicates that the ResNet101 is more stable and somehow invariant to spatial resolution changes. By comparing the results obtained in the first and the second experiments, we find that fine-tuning the parameters are necessary and the performance of the networks does not follow the same pattern in both cases. At a resolution of 10m, the overall accuracy is rapidly reduced due to spectral distortion.

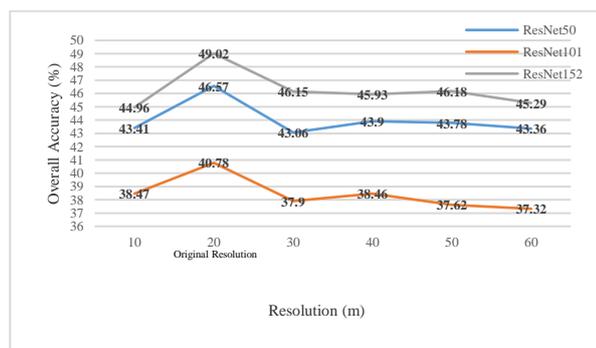


Figure 5. Average of overall accuracy obtained from the first experiment (without fine-tuning the parameters) on Sentinel-2 image.

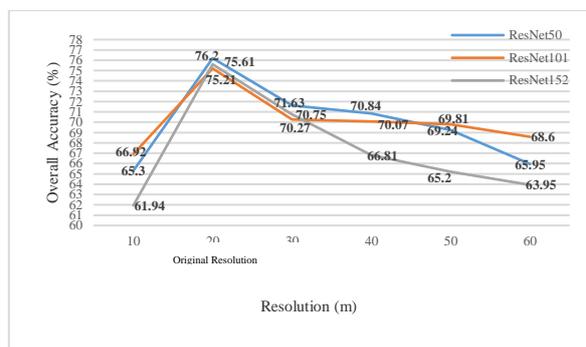


Figure 6. Average of overall accuracy obtained from the second experiment (with fine-tuning the parameters) on Sentinel-2 image.

By comparing the accuracy of each class, it can be concluded that classes that have a larger number of training samples and are integrated, such as Arable land and Mixed forest, are not greatly affected by changes in spatial resolution, but classes that are scattered, such as Land principally occupied. And the Argo-forestry class, which has few training samples, are significantly reduced by decreasing spatial resolution. Figures 7 to 9 show the graph of networks accuracy for each class.

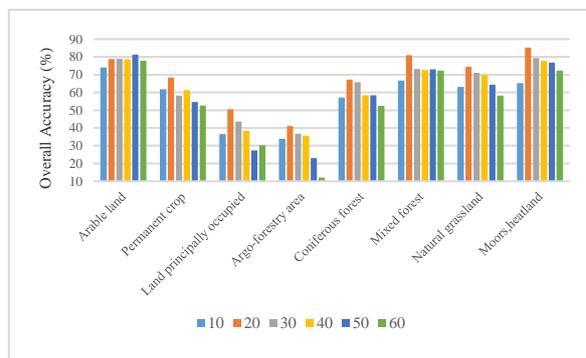


Figure 7. Average of overall accuracy obtained from the second experiment (with fine-tuning the parameters) on Sentinel-2 image for each class with ResNet 50 network.

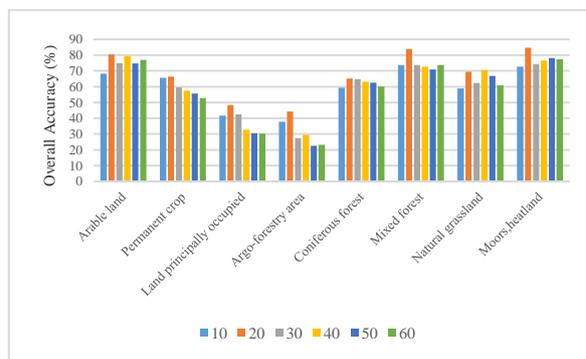


Figure 8. Average of overall accuracy obtained from the second experiment (with fine-tuning the parameters) on Sentinel-2 image for each class with ResNet 101 network.

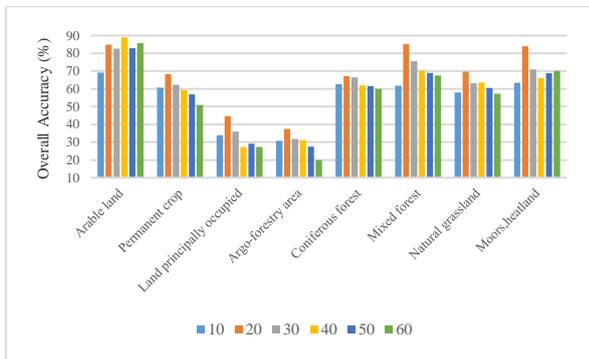


Figure 9. Average of overall accuracy obtained from the second experiment (with fine-tuning the parameters) on Sentinel-2 image for each class with ResNet 152 network.

According to the results, all networks had the highest accuracy at a resolution of 20 meters. Figures 11-13 shows classification maps at a resolution of 20 meters. According to the classification maps, we found that the performance of all three networks is relatively similar. In Compared to ground truth, the left corner of the image has different vegetation types, and all networks have mistakenly predicted the Arable land label, which is the dominant class in the region, for other classes.

Comparing the maps from all three networks, it can be seen that the ResNet 50 network performed worse in detecting the Land principally occupied by agriculture and Argo-forestry area classes marked in the red box, and the ResNet 101 network performed poorly in detecting the Natural grassland classes marked in the black box. Although the accuracy of the ResNet 50 is 76.2%, which is higher than other networks, but with the visual interpretation of the classification maps, it can be seen that the Resident 152 has performed better in detecting complex classes, As the number of hidden layers increases, the network becomes more capable of learning complex features but it needs more time and powerful processors.

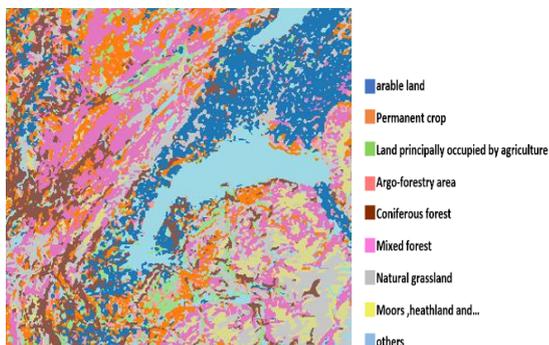


Figure 10. Ground-truth

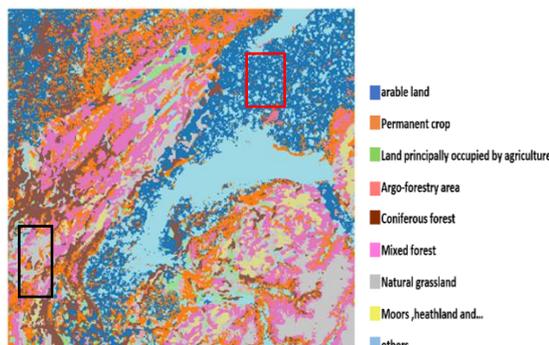


Figure 11. Classification map obtained from ResNet50 at 20 m spatial resolution with fine-tuning

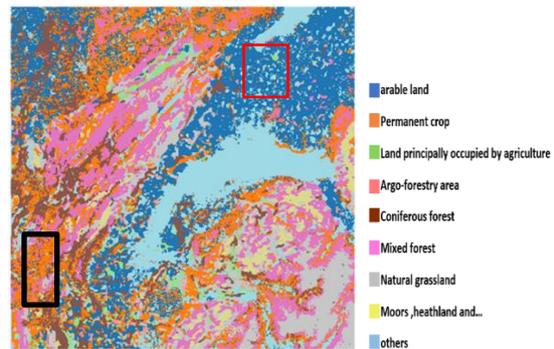


Figure 12. Classification map obtained from ResNet101 at 20 m spatial resolution with fine-tuning

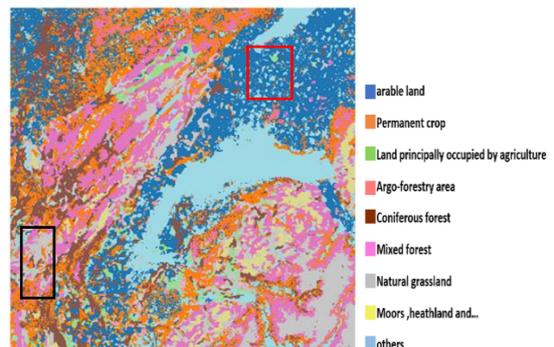


Figure 13. Classification map obtained from ResNet152 at 20 m spatial resolution with fine-tuning

5. CONCLUSIONS

The industrialization of cities has led to the extinction of various vegetation species, so high-precision classification maps are needed for the sustainability and management of natural resources. In this research, ResNet networks that have been trained with BigEarthNet dataset and the sensitivity of these networks to different resolutions of Sentinel-2 satellite image have been investigated. The overall accuracy obtained for the networks proves that The BigEarthNet dataset, due to the allocation of multiple labels for each patch and the use of multi-band satellite imagery, has the ability to highlight different vegetation features compared to other machine vision datasets. (Because only three RGB bands are used in dataset like ImageNet). This advantage can be seen for each class, however, some classes may not be well distinguishable due to the complexity and similarity of the features. Classification maps show that the ResNet 152 network has been more successful in detecting complex classes but ResNet 50 has higher overall accuracy. All networks have better performance at 20 m resolution because all bands have less spectral distortion at this resolution. The sensitivity of ResNet networks with different depth to changing spatial resolution varies. ResNet 101 network is slightly sensitive to changing spatial resolution when changing the spatial resolution of 30-50 meters.

REFERENCES

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . Devin, M. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*.

- Akbari Dotappah Sofla, R., Alipour-Fard, T., & Arefi, H., (2021) Road extraction from satellite and aerial image using SE-Unet, *Journal of Applied Remote Sensing*, 15(1) 014512, <https://doi.org/10.1117/1.JRS.15.014512>
- Alipour-Fard, T., & Arefi, H. (2020). Structure Aware Generative Adversarial Networks for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5424-5438.
- Alipour-Fard, T., Paolletti, M.E, Haut, J.M., Arefi, H, Plaza, J, & Plaza, A. (2020). Multibranch Selective Kernel Networks for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*, doi: 10.1109/LGRS.2020.2990971.
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., . . . Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*.
- Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10), 1865-1883.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7), 2217-2226.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Li, H., Dou, X., Tao, C., Hou, Z., Chen, J., Peng, J., . . . Zhao, L. (2017). RSI-CB: A large scale remote sensing image classification benchmark via crowdsourcing data. *arXiv preprint arXiv:1705.10450*.
- Lu, D., Li, G., Moran, E., & Kuang, W. (2014). A comparative analysis of approaches for successional vegetation classification in the Brazilian Amazon. *GIScience & Remote Sensing*, 51(6), 695-709.
- Pires de Lima, R., & Marfurt, K. (2020). Convolutional neural network for remote-sensing scene classification: Transfer learning analysis. *Remote Sensing*, 12(1), 86.
- Richter, R., & Schlöpfer, D. (2013). Atmospheric/topographic correction for satellite imagery (ATCOR-2/3 User Guide, Version 8.3. 1, February 2014). *ReSe Applications Schlöpfer, Langeggweg*, 3.
- Shao, W., Yang, W., & Xia, G.-S. (2013). Extreme value theory-based calibration for the fusion of multiple features in high-resolution satellite scene classification. *International Journal of Remote Sensing*, 34(23), 8588-8602.
- Shao, Z., Yang, K., & Zhou, W. (2018). Performance evaluation of single-label and multi-label remote sensing image retrieval using a dense labeling dataset. *Remote Sensing*, 10(6), 964.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019). *Bigearthnet: A large-scale benchmark archive for remote sensing image understanding*. Paper presented at the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium.
- Teoh, K. K., Ibrahim, H., & Bejo, S. K. (2008). *Investigation on several basic interpolation methods for the use in remote sensing application*. Paper presented at the 2008 IEEE Conference on Innovative Technologies in Intelligent Systems and Industrial Applications.
- Wang, M., Zhang, X., Niu, X., Wang, F., & Zhang, X. (2019). Scene classification of high-resolution remotely sensed image based on resnet. *Journal of Geovisualization and Spatial Analysis*, 3(2), 1-9.
- Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., . . . Lu, X. (2017). AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7), 3965-3981.
- Yang, Y., & Newsam, S. (2010). *Bag-of-visual-words and spatial extensions for land-use classification*. Paper presented at the Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems.
- Zhou, W., Newsam, S., Li, C., & Shao, Z. (2018). PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 197-209.
- 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(11), 2321-2325.