ABSTRACT

In the remote sensing community, Land Use Land Cover (LULC) classification with satellite imagery is a main focus of current research activities. Accurate and appropriate LULC classification, however, continues to be a challenging task. In this paper, we evaluate the performance of multi-temporal (monthly time series) compared to mono-temporal (single time step) satellite images for multi-label classification using supervised learning on the RapidAI4EO dataset. As a first step, we trained our CNN model on images at a single time step for multi-label classification, i.e. mono-temporal. We incorporated time-series images using a LSTM model to assess whether or not multi-temporal signals from satellites improves CLC classification. The results demonstrate an improvement of approximately 0.89% in classifying satellite imagery on 15 classes using a multi-temporal approach on monthly time series images compared to the mono-temporal approach. Using features from multi-temporal or mono-temporal images, this work is a step towards an efficient change detection and land monitoring approach.

Index Terms— Land Use, Land Cover, Remote Sensing, Satellite imagery, Sentinel-2, Planet Fusion, RapidAI4EO, CNN, LSTM.

1. INTRODUCTION

Remote sensing plays an important role in the observation and study of the earth’s surface by providing mono-temporal and multi-temporal information on land use/land cover (e.g., artificial area, forest, bare ground, agriculture, etc.). Land use is determined by how humans use it for recreation, wildlife habitats, urban development and agriculture. Whereas, land cover is the extent to which the Earth’s surface is covered by forests, wetlands, urban, and agriculture areas [1]. Land Use and Land Cover (LULC) is a method of categorizing and classifying human and natural activities in the landscape over a specific period. In the past decades, many Earth observing satellites like Sentinel-2 by European Space Agency, Landsat 7 and 8 by U.S. Geological Survey, etc have come into existence. The data from such satellites were used by researchers to analyze and track land use and land changes over time in response to human factors such as climate change, urbanization, deforestation, wildfires, etc. RapidAI4EO [2], a new satellite imagery corpus, aims to provide better monitoring of Land Use (LU), Land Cover (LC), and LULC change at a much higher level of detail and temporal frequency than is currently possible [2]. The objective of this paper is to explore the advantages of using multi-temporal (monthly) over mono-temporal (single time-step) satellite images. To achieve this, we trained our models with different land cover ontologies. Level-1: 5 classes, Level-1.5: 7 classes, Level-2: 15 classes with single time-step as well as high cadence i.e., monthly imagery from RapidAI4EO dataset.

2. RELATED WORK

2.1. LULC classification

A common application of remote sensing is creating land-use/land-cover classification maps (LULC) on satellite images. As satellite imagery becomes increasingly affordable and more precise, many government and private industries are using LULC classification maps for a wide range of applications, including land monitoring for human or natural activities, damage delineation (e.g. fire, flood forecasting), wildlife habitat preservation and planning rural and urban land use. Deep learning-based image classification models categorize an image belonging to one or more categories such as single class, multi-class and multi-label classification. The main focus of this research is to learn the multi-label patch-based distribution which can classify satellite images (e.g. 40% agri-
culture, 30% forest, 30% water bodies) using images at single timestep as well as images with monthly cadence. To achieve this, we used mono-temporal and multi-temporal approaches using a supervised classifier to extract features from multi-spectral satellite images. We contrast on the state-of-the-art Convolutional Neural Network (CNN) [3] architecture such as ResNet-50 [4] to train a patch-based multi-label classifier that recognizes multiple classes from a single imagery.

2.2. Mono-temporal approach

The mono-temporal approach uses images at a single time step (the monthly mosaic images) from the RapidAI4EO dataset, consisting of Planet Fusion (PF) and Sentinel-2 (S2) imagery as an input to the ResNet-50 model. Figure 1 represents the architecture, in which features are extracted from a single time step image using ResNet-50, and then SoftMax is applied for multi-label distribution.

2.3. Multi-temporal approach

As a way of exploring the implications of multi-temporal data, we passed the existing ResNet-50 features trained on mono-temporal images into an LSTM [5]. Figure 2 illustrates the architecture of the multi-temporal approach, in which the features are extracted from multi-temporal images using ResNet-50, concatenated and then passed to an LSTM. Later, the output is passed through two fully connected layers and then post-processed by applying softmax for patch-based multi-label distribution. The multi-temporal satellite images with monthly cadence provided us with a chance to investigate if it could further improve the multi-label LULC classification.

2.4. Training configuration

As part of the ResNet-50 training process, we used RapidAI4EO dataset, which covers 500,000 patches across the European Economic Area (EEA39), with a 70:20:10% split for training, validation, and testing respectively. These 500,000 locations comprise 1400 tiles where each tile covers an area of 8000 x 8000 meters. We further divide this 8000 x 8000 m area into 1600 small patches, on which we ran our ResNet-50 model. The ResNet-50 model is pretrained on ImageNet. Four-channel (RGB-N) images were taken as input for our model. The images were normalized using the mean and standard deviation calculated over all training images. The model is trained for 20 epochs with a learning rate of 0.0001. After every epoch, the learning rate is reduced by a factor of 0.1. We experimented with different loss functions such as Focal Loss, BCE loss and KL Divergence loss for training our model concluding KL Divergence loss to be the best. To incorporate the multi-temporal approach, we trained our LSTM model for 20 epochs with a learning rate of 0.00001 using KL Divergence loss.

3. EXPERIMENTS

In this section, we explore the new dataset for LULC classification. Furthermore, we ran experiments to show the improvement of multi-temporal times series data over mono-temporal single time step image.

3.1. Dataset

RapidAI4EO [2] will establish the basis for the next-generation Copernicus Land Monitoring Service (CLMS) under the sponsorship of the European Union’s Horizon 2020 program. The idea of such a large corpus comes from patch-based LULC EuroSAT [6] and BigEarthNet [7] corpora. RapidAI4EO aims at creating a dense spatio-temporal training corpus by combining S2 with high cadence, very high resolution and harmonized multi-spectral PF imagery at 500,000 patch locations spread over EEA39.

RapidAI4EO dataset consists of PF (3-meter spatial resolution) and S2 (10-meter spatial resolution) images for selected European locations. These selected locations are sampled across EEA39 to ensure that each country covered by Corine Land Cover 2018 (CLC 2018) is adequately repre-
3.2. Mono-temporal experiments

To begin, we trained our ResNet-50 model on RapidAI4EO dataset, i.e. 500,000 locations across EEA39 with the different land cover ontologies, Level-1 (5 classes), Level-1.5 (7 classes) and Level-2 (15 classes). This CLC nomenclature is organized into multiple levels of hierarchical classification, with 44 classes at the most detailed level and 5 classes at the least, is defined here. As the forest and semi-natural area class from Level-1 ontology contains visually very distinctive subclasses, it can be further separated into forests, shrubs, and bare areas, leading to 7 classes. The aggregated class labels are also mapped as a probability distribution like [0.4,0.5,0,0.1,0] rather than a binary one hot encoding [1,1,0,0,0]. Table 1 shows that our classification model achieves an F1-score (micro) on test set of 60.78% for 15 classes and up to 73.88% for 7 classes. When the land cover classes are aggregated into 5 classes, the F1-score (micro) rises to 81.24%.

3.3. Multi-temporal experiments

To implement the multi-temporal approach, we experiment with monthly images i.e. 15th of every month, on 3 different land cover ontologies. As observed in Table 2, the multi-temporal model achieves an F1-score of 61.67% for 15 classes and up to 74.67% for 7 classes. The overall performance increases to 81.69% when the land cover classes are aggregated into 5 classes.

3.4. Comparison of mono-temporal and multi-temporal

To demonstrate the advantage of the multi-temporal approach over the mono-temporal approach, we compared the mono-temporal approach on single time step images with the multi-temporal approach on monthly images. As can be observed from Tables 1 and 2, multi-temporal classification of monthly time-series images outperformed those based solely on mono-temporal classification. The multi-temporal model on monthly images achieves the overall improvement in F1-score of 0.89% for 15 classes, 0.79% for 7 classes and 0.45% for 5 classes when compared with mono-temporal model. As evidence from Table 3, the mono-temporal approach shows minimal advantage over multi-temporal for classes such as urban, industrial forest areas and inland waters.

4. MODEL INFERENCE AND APPLICATIONS

To evaluate the performance of our model we selected a test area that has not been used in the model training. We gen-
generated LULC maps on the test area in Sicily as shown in Figure 4. As can be observed, our model is able to classify agriculture, forest, water bodies and artificial areas precisely. These LULC maps can be used for land monitoring, planning urban land use, wildlife habitat preservation and change detection. We also showed an example of how we can use our model for the change detection part in Figure 5.

5. CONCLUSION

In this paper, we compared the LULC classification using mono-temporal i.e. single time step images and multi-temporal images on monthly cadence. For the mono-temporal approach, we train a ResNet-50 model on the RapidAI4EO dataset for multi-label classification. To incorporate multi-temporal images, we used the mono-temporal model for inference on monthly images and passed it through a LSTM. We can conclude from the evaluation that the multi-temporal approach outperforms the mono-temporal approach on F1-score (micro) by approximately 0.89%. Nevertheless, it is obvious from Table 3 that the mono-temporal approach is suitable for only a few classes while the multi-temporal approach is suitable for most of the classes. These LULC maps can be leveraged for multiple real-world Earth observation applications. A few application areas include detection of changes in land use and land cover, delineation of damaged areas, and urban planning.

6. ACKNOWLEDGEMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101004356.

7. REFERENCES

[1] C. Prakasam, “Land use and land cover change detection through remote sensing approach: a case study of kodaikanal taluk, tamil nadu,” International journal of Geomatics and Geosciences, vol. 1, pp. 150–158, 2010.

[2] Giovanni Marchisio, Patrick Helber, Benjamin Bischke, Timothy Davis, Çağlar Senaras, Daniele Zanaga, Ruben Van Der Kerkhove, and Annett Wania, “Rapidai4eo: A corpus for higher spatial and temporal reasoning,” CoRR, vol. abs/2110.01919, 2021.

[3] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroury, Bing Shuai, Ting Liu, Xingxing Wang, and Gang Wang, “Recent advances in convolutional neural networks,” CoRR, vol. abs/1512.07108, 2015.

[4] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep residual learning for image recognition,” CoRR, vol. abs/1512.03385, 2015.

[5] Sepp Hochreiter and Jürgen Schmidhuber, “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[6] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth, “Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification,” CoRR, vol. abs/1709.00029, 2017.

[7] Gencer Sumbul, Marcela Charfuelan, Begüm Demir, and Volker Markl, “Bigearthnet: A large-scale benchmark archive for remote sensing image understanding,” CoRR, vol. abs/1902.06148, 2019.