STEPWISE REFINEMENT OF LOW RESOLUTION LABELS FOR EARTH OBSERVATION DATA: PART 2

Daniele Cerra, Nina Merkle, Corentin Henry, Kevin Alonso, Pablo d'Angelo, Stefan Auer, Reza Bahmanyar, Xiangtian Yuan, Ksenia Bittner, Maximilian Langheinrich, Guichen Zhang, Miguel Pato, Jiaojiao Tian, Peter Reinartz

{FirstName.LastName}@dlr.de
Remote Sensing Technology Institute, German Aerospace Center (DLR), Oberpfaffenhofen

ABSTRACT

This paper describes the contribution of the DLR team ranking 2nd in Track 2 of the 2020 IEEE GRSS Data Fusion Contest. The semantic classification of multimodal earth observation data proposed is based on the refinement of low-resolution MODIS labels, using as auxiliary training data higher resolution labels available for a validation data set. The classification is initialized with a handcrafted decision tree integrating output from a random forest classifier, and subsequently boosted by detectors for specific classes. The results of the team ranking 3rd in Track 1 of the same contest are reported in a companion paper.

1. INTRODUCTION

The 2020 Data Fusion Contest, organized by the Image Analysis and Data Fusion Technical Committee (IADF TC) of the IEEE Geoscience and Remote Sensing Society (GRSS) and the Technical University of Munich, focused on largescale land cover mapping from globally available multimodal satellite data. The task is to train a machine learning model for global land cover mapping based on weakly annotated samples [1]. Training, validation and test data consist of triplets of Sentinel-1 and Sentinel-2 images at 10 m/px spatial resolution, and MODIS semantic labels resampled from the original 500 m/px spatial resolution [2]. The 2020 contest comprised two challenge tracks: land cover classification with low-resolution labels using the described data (Track 1), and the same problem with additional high resolution semantic labels made available for training (Track 2). This paper describes the approach developed at the German Aerospace Center (DLR) ranking 2nd in Track 2. The described results are refinements to the approach which ranked 3rd in Track 1 of the same contest, which is reported in a companion paper [3]. As a reminder, a graphical stepwise representation of Track 1 approach for a sample validation image is reported in Fig. 1.

With respect to the results obtained in Track 1, training a Random Forest (RF) classifier on higher resolution semantic

labels and refining the decision tree yields an improvement for all the semantic classes of interest.

2. METHODOLOGY

2.1. Enhanced Random Forest Classifier

In Track 2 of the IEEE Data Fusion Contest, the highresolution semantic labels associated to the validation dataset became available. These were used to train the RF classifier architecture described in [3], allowing further improvements of the best results obtained in Track 1. The parameters of the classifier based on the implementation of the Scikit-learn python library are as follows: number of trees = 500, maximum depth = 100 (longest branch in a tree), maximum leaf nodes = 100. Additionally, class imbalance within the training data was handled by setting the parameter "class_weight" to "balanced", which assigns a weight to each class inversely proportional to its frequency in the input data. We trained two models on different splits of data and different kinds of labels: the first one includes only the validation set with the high resolution labels disclosed for Track 2, while the second one contains all images from Europe and uses semantically translated labels from the 2018 Copernicus CORINE data [4] from 2018 (for details on single classes see [3]).

Visual inspection of the results suggested that the model trained on the high-resolution validation labels yielded results more consistent with respect to the training data, and made a better segmentation of natural classes such as Grasslands, Wetlands and Barren with respect to the model trained on the CORINE labels.

2.2. Urban Areas Refinement

The Urban class is initialized by the same RF used for Track 1, this time trained on the high-resolution labels for the validation dataset. The Urban mask is then expanded and refined using the same algorithm used for Track 1, employing Gabor texture parameters, NDTI (Normalized Difference Tillage Index) [5] computed from Sentinel-2, and brightness

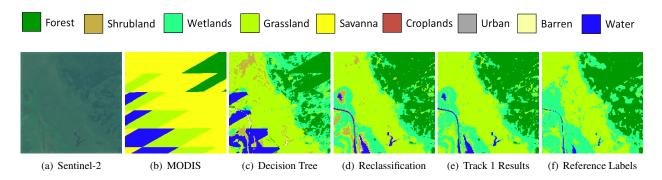


Fig. 1. Reminder for main steps of the classification strategy for Track 1. From left to right: True color combination of Sentinel-2 bands (a); coarse MODIS labelling at 500 m spatial resolution (b); output of the decision tree described in [3] (c); results after reclassification using k-means (d); final results after post-processing (e); reference labels for the selected test image (f).

of Sentinel-1. Details on this step are reported in the companion paper to this manuscript [3]. Both precision and recall for this class increase with respect to results for Track 1. False alarms are reassigned to the classes Grassland, Shrubland or Barren according to a hard thresholding of the NDVI. Although results in Table 1 show that the recall for the class has an improvement in the order of 1% only, false alarms are greatly reduced. Relevant examples are reported in Fig. 2. Similar results have been obtained by training the RF on European scenes with CORINE labels as defined in [3].

2.3. Grassland Refinement

The Grassland class as detected from the described Random Forest classifier results better defined with respect to Track 1, and is used as starting point for the final classification. The class is improved by switching to Grassland pixels originally classified as Forest having brightness value in the red band of the Sentinel-2 images above a fixed threshold set at 550. Finally, images detected as containing Wetlands but dominated by Grassland labels in MODIS are also added to the class. Grassland is the only class showing a major improvement (around 20%) with respect to Track 1 results (see Table 1). Representative results are reported in Fig. 3.

2.4. Other Classes

The other classes are refined as follows.

• The class **Forest** is expanded with the results of the Random Forest classification, and reassigned to the class **Shrubland** if values in the Short Wave Infrared (SWIR) bands of Sentinel 2 are above a fixed threshold of 710, with an NDVI below 0.7. In spite of the spatial ambiguities introduced by the lower spatial resolution of the SWIR bands (20 m), this increases the detection of Shrubland without degrading results for Forest.

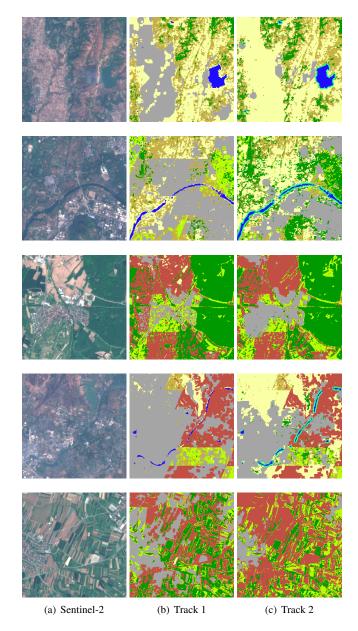


Fig. 2. True color combination of Sentinel-2 bands (a) and results for selected test images in Track 1 and Track 2 (b, c), showing improvements in urban area detection. For a legend ref. Fig. 1.

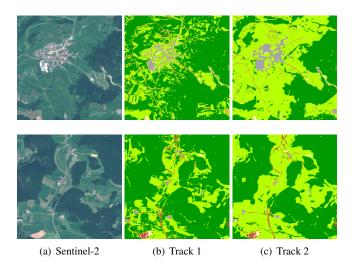


Fig. 3. True color combination of Sentinel-2 bands (a) and results for selected test images in Track 1 and Track 2 (b, c), showing improvements in the class Grassland. For a legend ref. Fig. 2.

Table 1. Classification results

Class	Track 1	Track 2
Forest	0.8012	0.8109
Shrubland	0.3538	0.3798
Grassland	0.4446	0.6299
Wetlands	0.4273	0.4388
Croplands	0.3704	0.4108
Urban	0.8140	0.8264
Barren	0.3556	0.4260
Water	0.9838	0.9860
Average Accuracy	0.5688	0.6136
Contest's Best Results	0.5749	0.6142

- The class **Wetlands** is expanded by considering images satisfying the following conditions: (1) The images contain pixels marked as Water or Wetlands in MODIS labels; (2) the variance in the red band for Sentinel-2 lies below a fixed threshold of 160, as wetlands area appear homogeneous and exhibit a smooth texture; (3) the images do not contain strong scatterers in Sentinel-1, usually indicating man-made structures. If any of the listed conditions is not met, the class is reassigned to one among the classes Barren, Brushland, Forest, or Grassland according to the NDVI.
- Barren pixels in areas dominated by crops are assigned to the Croplands class.

2.5. Evaluation of the Results

The examples reported in Figs.2-4 can be only evaluated visually, as at the time of writing the high resolution semantic labels have not been disclosed for the over 5000 images of size 256×256 contained in the test set. Results in terms of accuracy per class are reported in Table 1, along with the score for the best ranked teams. The accuracy for all semantic classes improves with respect to Track 1. Examples for the improvements in these classes are reported in Fig. 4.

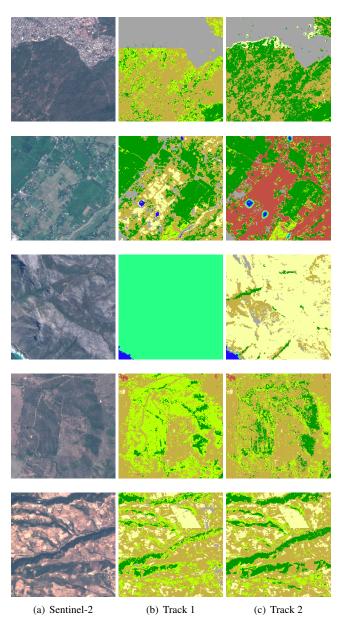


Fig. 4. True color combination of Sentinel-2 bands (a) and results for selected test images in Track 1 and Track 2 (b, c), showing improvements in classes such as Forest and Croplands, and a reduction of false alarms for classes Wetlands and Urban. For a legend ref. Fig. 2.

3. CONCLUSIONS

The 2020 IEEE GRSS Data Fusion Contest focused on the classification of Sentinel-2 and Sentinel-1 data at 10 meters spatial resolution, given MODIS-derived semantic annotations at a coarser resolution of 500 m. In Track 2 of the contest, additional semantic labels at 10 m spatial resolutions are made available for part of the available dataset. This paper describes the refinement steps taken to improve results with respect to Track 1, for which our results are described in [3]. The produced classification maps ranked 3rd and 2nd in Track 1 and Track 2, respectively.

4. ACKNOWLEDGEMENT

The authors would like to thank the research group for Signal Processing in Earth Observation at the Technical University of Munich for providing the data used in this study, and the IEEE GRSS Image Analysis and Data Fusion Technical Committee for organizing the Data Fusion Contest.

5. REFERENCES

- [1] N. Yokoya, P. Ghamisi, R. Haensch, and M. Schmitt, "2020 ieee grss data fusion contest: Global land cover mapping with weak supervision [technical committees]," *IEEE Geoscience and Remote Sensing Magazine*, vol. 8, no. 1, pp. 154–157, 2020.
- [2] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "Sen12ms a curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019.
- [3] D. Cerra, N. Merkle, C. Henry, K. Alonso, P. d'Angelo, S. Auer, R. Bahmanyar, X. Yuan, K. Bittner, M. Langheinrich, G. Zhang, M. Pato, J. Tian, and P. Reinartz, "Stepwise refinement of low resolution labels for earth observation data: Part 1," in *IGARSS* 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020.
- [4] M. Bossard, J. Feranec, and J. Otahel, "CORINE Land Cover Technical Guide Addendum 2000," *European Environmental Agency, Copenhagen*, 2000.
- [5] A. Van Deventer, A. Ward, P. Gowda, and J. Lyon, "Using thematic mapper data to identify contrasting soil plains and tillage practices," *Photogrammetric Engineering and Remote Sensing*, vol. 63, pp. 87–93, 1997.