DOMAIN ADAPTATION OF LANDSAT-8 AND PROBA-V DATA USING GENERATIVE ADVERSARIAL NETWORKS FOR CLOUD DETECTION

Gonzalo Mateo-García, Valero Laparra, Luis Gómez-Chova

Image Processing Laboratory (IPL), University of Valencia, Spain

ABSTRACT

Training machine learning algorithms for new satellites requires collecting new data. This is a critical drawback for most remote sensing applications and specially for cloud detection. A sensible strategy to mitigate this problem is to exploit available data from a similar sensor, which involves transforming this data to resemble the new sensor data. However, even taking into account the technical characteristics of both sensors to transform the images, statistical differences between data distributions still remain. This results in a poor performance of the methods trained on one sensor and applied to the new one. In this this work, we propose to use the generative adversarial networks (GANs) framework to adapt the data from the new satellite. In particular, we use Landsat-8 images, with the corresponding ground truth, to perform cloud detection in Proba-V. Results show that the GANs adaptation significantly improves the detection accuracy.

Index Terms— Generative Adversarial Networks, Convolutional Neural Networks, Domain Adaptation, Landsat-8, Proba-V, Cloud Detection

1. INTRODUCTION

There are many Earth observation satellite missions with similar but not identical sensor characteristics. These sensor differences make that the distribution of the data acquired by them changes from one sensor to another. Therefore, derived remote sensing products from one sensor cannot be generally applicable to a different one. This is a long-standing problem that in machine learning is called *data shift* [1] and that hampers transfer learning between different satellite missions. This is particularly true in cloud detection applications since, in addition to the different sensor characteristics, spatial resolutions, and acquisition times, one has to add the extremely unpredictable cloud dynamics. Hence, neither the datasets nor the associated cloud masks used as *ground truth* can be transferred or reused from one satellite to another.

Statistical machine learning methods are even more affected by this problem since they require to gather and label new data for each new satellite mission in order to train models. This process is always time consuming and costly since the ground truth generation usually requires a manual cloud masking of a representative number of images. A possible solution is first transforming the data from another sensor to resemble data coming from the new sensor, and then training the machine learning method (e.g. a classifier) using these transformed data. However, the performance of the classifier usually decreases when applied to real data from the new sensor. The problem is that, since the transformation is not perfect, the distribution of the new sensor data differs from the distribution of the transformed data.

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness (MINECO, TEC2016-77741-R, ERDF) and the European Space Agency (ESA IDEAS+ research grant, CCN008).

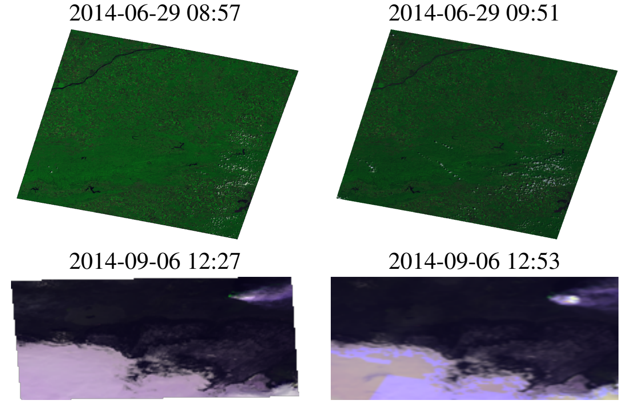

Fig. 1. Close-in-time acquisitions from Landsat-8 and Proba-V. Landsat-8 image is transformed and upscaled to resemble the optical characteristics of Proba-V. First row: Bulgaria. Second row: Vatnajökull glacier, Island.

In this work, we propose to adapt the data coming from the new sensor before applying the classifier in order to match the distribution of the transformed data. The proposal is based on Generative Adversarial Networks (GANs) [2], which have shown outstanding natural image generation capabilities. In particular we take advantage of the wealth of publicly available Landsat datasets, with the corresponding ground truth for cloud detection, to improve the cloud detection performance in Proba-V images [3, 4]. First we transform Landsat-8 to resemble Prova-V and train a cloud detection classifier. Then we use an extension of GANs to adapt the real Proba-V images to match the distribution of the transformed Landsat-8. The adaptation is unsupervised and does not require paired Landsat-8 and Proba-V images to learn the mapping, which could not be obtained due to cloud movement.

2. METHODOLOGY

Let's consider two independent datasets from different sensors: for Landsat-8 we have images with the corresponding ground truth cloud masks, $\{X_L, y_L\}$; and for Proba-V we only have images without ground truth, X_{PV}. We assume X_L and X_{PV} have similar time and space acquisitions conditions. First, we describe the employed methodology to transform (upscale) Landsat-8 to resemble Prova-V, and the cloud detection classifier. Then, we explain the proposed methodology to adapt the real Prova-V data to be similar to the transformed Landsat-8 before the classification step.

2.1. Transfer Learning from Landsat-8 to Proba-V

In a standard transfer learning approach from one sensor to another, the first step is to transform Landsat-8 images to resemble Proba-V images. First, we select the spectral bands from Landsat that overlap with Proba-V in terms of the spectral response functions. Then, the 30m resolution of Landsat-8 is *upscaled* to the 333m resolution of Proba-V. This upscaling takes into account the optical characteristics of the Proba-V sensor and the re-sampling of the 333m product described in [5]. We transform also the ground truth labels y_L at 30m using bicubic interpolation to get a Landsat upscaled dataset: $\{X_{LU}, y_{LU}\}$.

Following this approach, one can use the Landsat upscaled dataset to train a cloud detection model that, in principle, could be directly used on Proba-V images. We explored this transfer learning approach in [4] to train a fully convolutional neural network (CNN) classifier with a simplified U-Net [6] architecture, which takes as input the 4-band Proba-V images and produces a probability cloud mask.

However, after the proposed transformation, one can find still statistical differences between the Landsat-8 upscaled images and the Proba-V real data, which worsen the results of the proposed transfer learning approach. Main data shift sources are related to the different sensor spectral response functions, saturation effects, radiometric calibration, modulation transfer functions, or mixed pixels. For example, in Fig. 1, one can see significant radiometric differences between the Landsat-8 upscaled and the Proba-V real images. Proba-V contains many saturated pixels, specially in the blue channel, which is a known issue. These problems suggest that further domain adaptation can be carried out in order to improve the transfer learning results.

2.2. Generative Adversarial Domain Adaptation

Original GANs formulation finds a synthetic data generator by minimizing the Jensen-Shannon divergence between the real and the generated data distribution. More recently, conditional GANs [7] were proposed to generate samples from a conditional distribution. One application is the Generative Adversarial Domain Adaptation (GADA) [8–10] were conditional GANs formulation was modified to solve domain adaptation problems. In this work, we propose a customized version of input-level domain adaptation. Input-level domain adaptation refers to adapting the input image so that its distribution is similar to the input images of the originally trained model. This approach has the advantage that is independent of the remote sensing application, hence the developed transformation is not restricted to cloud detection problems.

Following this approach, a *generator* G is trained to adapt real Proba-V images to the upscaled Landsat-8 domain. At the same time, a *discriminator* D is trained, as an adversary of the generator, to distinguish adapted Proba-V images from upscaled Landsat-8 ones. In order to ensure consistency between the input and the output of the generator (i.e. to avoid that generated outputs do not have relation to the inputs) we add a ℓ_1 penalty to the generator loss function $\mathcal{L}(G)$. Formally, in order to tune the weights of G and D, we minimize iteratively for all the training samples the following losses:

$$\mathcal{L}(D) = \sum_i -\log(D(X^i_{LU})) - \log(1 - D(G(X^i_{PV})))$$

$$\mathcal{L}(G) = \sum_i -\log(D(G(X^i_{PV}))) + \lambda \|G(X^i_{PV}) - X^i_{PV}\|_1$$

The generator G is a 5-layer fully convolutional neural network. It consists of 2 blocks of 8 filters 3×3 separable convolution, reLU activation, and batch normalization, 2 blocks of 8 filters 3×3 separable convolution with dilation rate 2, reLU activation, batch normalization, and a layer of 1×1 convolution with 4 channels output. We used residual connections between blocks and before the final layer.

The discriminator D is also a 5-layer convolutional neural network (adapted from [7]). It consists of 4 blocks of 4×4 convolution, leakyReLU, and batch normalization. The number of filters starts in 8 for the first convolution an grows by a factor two in every layer. The convolutions are applied with a stride of 2, thus reducing by this factor the spatial size of the input. Last layer is a 1×1 convolution with 1 output channel and a sigmoid activation. Therefore, the output can be interpreted as the probability of an image to be *fake*, i.e. the probability to be a Proba-V image adapted by the generator instead of an upscaled Landsat-8 one. Once G and D are trained, we get rid of D and use the generator G to map Proba-V images to the upscaled Landsat-8 domain. Then, we can apply the CNN classifier, previously trained with the upscaled Landsat-8 datasets, to estimate the cloud mask of Proba-V images.

It is worth to note that this approach has an important benefit: it does not require simultaneous and collocated pairs of Landsat-8, X^i_{LU}, and Proba-V, X^i_{PV}, images. Otherwise simpler approaches could be used, such as a Canonical Correlation Analysis or even directly learning a generic transformation from X^i_{PV} to X^i_{LU}. However, for cloud detection problems, this is not feasible since clouds presence and location within an image highly vary even for small time differences. For instance, in Fig. 1, location of clouds has changed significantly in one hour. Moreover, having simultaneous collocated datasets from different sensors is really difficult or in some cases impossible.

3. EXPERIMENTS AND RESULTS

3.1. Experimental Setup and Data

The Landsat-8 cloud detection validation study [11] released the largest open-access manually labeled cloud mask archive for Landsat mission. It consists of two datasets for Landsat-8: the Biome dataset [12] which has 96 different acquisitions and the SPARCS dataset [13], which was collected in [14], with 80 more acquisitions. In the transfer learning approach, in order to develop the CNN using the upscaled Landsat-8 data, we split the data in a train and a test datasets. Original Landsat patches with the ground truth cloud mask from Biome and SPARCS datasets [12, 13] were upscaled from 390×390 to 32×32 size, following the transformation described in section 2.1, to resemble Proba-V:

- **Landsat-8 upscaled train dataset**: 165,601 overlapping 32×32 patches from 118 Landsat-8 products.
- **Landsat-8 upscaled test dataset**: 18,311 non overlapping 32×32 patches from 57 Landsat-8 products different than the ones in the Landsat-8 train dataset.

In addition, to validate the transfer learning approach in real Proba-V data, we use a Proba-V test set with ground truth cloud mask manually generated by the authors [3]:

- **Proba-V test dataset**: 368 non-overlapping 900×900 patches from 24 different products.

Finally, in order to train the GAN that adapts the Proba-V images we used a pseudo-simultaneous dataset where patches in the Landsat-8 upscaled train dataset are co-registered with Proba-V patches acquired the same day of the year:

- **Proba-V pseudo-simultaneous dataset**: 108,156 overlapping 32×32 patches from 134 different Proba-V Level-2A products.

It is worth noting that (*i*) we do not find collocated Proba-V patches for all the Landsat upscaled patches, and (*ii*) we do not need the ground truth cloud mask to train the GAN.

3.2. Experiment 1: Standard transfer learning

As a baseline, we show the results of the standard transfer learning approach [4] using the Landsat-8 upscaled dataset to train a cloud detection classifier based on a CNN (cf. section 2.1) without using the proposed adaptation with GANs. Table 1 shows the accuracy of this model for the Landsat upscaled datasets and for the Proba-V test dataset. We see that the model behaves well without overfitting in the Landsat upscaled domain. However, on Proba-V real images the accuracy is significantly lower.

We developed an extra experiment to evaluate how much Landsat upscaled images resemble real Prova-V images. A CNN classifier is trained to distinguish between them reaching an out-of-sample accuracy of 84.94%. These results suggest that, even though we followed a standard procedure to transform Landsat-8 images to resemble Proba-V, a data shift still exists. Therefore, there is margin for additional data statistical adaptation to improve the transfer learning results.

Table 1. Classification accuracy of the cloud detection model trained using Landsat-8 upscaled data on the different sets.

Landsat-8 train	Landsat-8 test	Proba-V test
94.41%	94.46%	88.85%

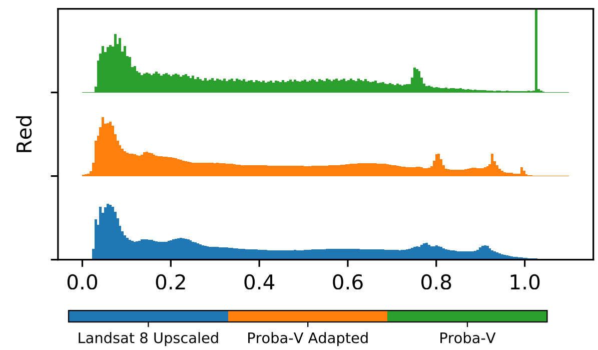

Fig. 2. Distribution of radiance values in the Red spectral channel for 10,000 patches from the Landsat-8 upscaled train dataset (blue), the Proba-V adapted data (orange), and real Proba-V data (green).

3.3. Experiment 2: Domain adaptation with GANs

Here we apply the methodology proposed in Sec.2.2 and compare with other approaches. First, we show that the generator transformation G reached its two goals: (*i*) the distribution of $G(X_{PV})$ is similar to the distribution of X_{LU}, and (*ii*) the original image information is preserved. After that we show that the classification results improve.

Figure 2 shows histograms from the input spectral channel distribution of 10,000 patches from the Landsat-8 upscaled train dataset and from the corresponding Proba-V pseudo-simultaneous dataset before and after domain adaptation. We see that the histogram of Proba-V domain adapted images is more similar to the Landsat upscaled ones. For instance, the peak of saturated values in the original Proba-V data almost disappears after domain adaptation.

The first row of Fig. 3 shows 9 patches, of 32 × 32 pixels, randomly selected from the Landsat-8 upscaled train dataset and from the Proba-V pseudo-simultaneous dataset before and after adaptation. It is easy to see that the spatial information of Proba-V images is preserved since the same patterns can be identified in both figures. However the colors and textures of adapted images are more similar to the Landsat-8 upscaled images. The second row of Fig. 3 shows predictions of the CNN cloud detection model using the above images. The snowy area in row 2 columns 1 is incorrectly identified as clouds for the non adapted Proba-V images. However, after domain adaptation, the model succeed in this difficult case.

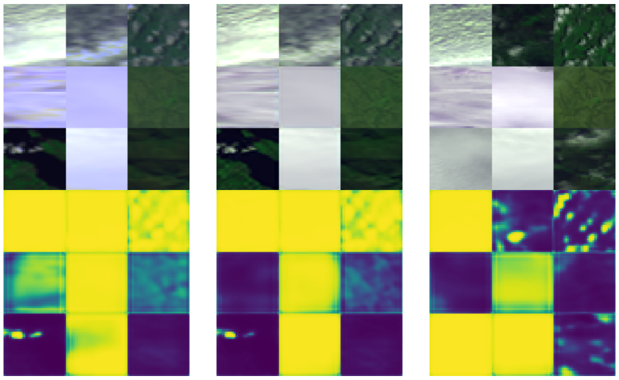

Fig. 3. First row: 32×32 patches from Proba-V pseudo-simultaneous dataset before and after adaptation and Landsat-8 upscaled patches. Second row: Cloud mask predictions.

Table 2. Cloud detection accuracy of different models on the Proba-V test dataset

Cloud detection model	Classification Accuracy
No adaptation	88.85%
Mean & std adaptation	86.97%
Adversarial domain adaptation	**90.72%**
Operational Proba-V cloud mask [15]	82.01%

Finally, Table 2 compares different approaches on the Proba-V test set. We see that the proposed adversarial domain adaptation approach increases the accuracy almost two points compared with the non adapted one. We can also see that a simple domain adaptation, removing the mean and normalizing the standard deviation, worsen the results; which indicates that more advanced domain adaptation methods are necessary. Results for the operational Proba-V cloud detection algorithm (version V101 [15]) are also given. It is worth noting that the CNN trained only with Landsat-8 data provides better detection accuracy than the operational mask.

4. CONCLUSIONS

A domain adaptation approach based on generative adversarial networks is presented to improve transfer learning from Landsat to Proba-V in the challenging application of cloud detection. Available labeled Landsat-8 datasets are spectrally and spatially transformed to resemble Proba-V characteristics, however, statistical differences remain that have to be minimized before applying models trained with one sensor data to the other. Results of the proposed adaptation are tested on real Proba-V images and demonstrate that models trained only with upscaled Landsat-8 images can provide a significantly higher cloud detection accuracy. In addition, the transformation used to adapt the Proba-V data is general and could in principle be used in any remote sensing application in which Proba-V could take advantage of Landsat-8 data.

5. REFERENCES

[1] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," in *CVPR 2011*, June 2011, pp. 1521–1528.

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27*, pp. 2672–2680. Curran Associates, Inc., 2014.

[3] L. Gómez-Chova, G. Mateo-García, J. Muñoz-Marí, and G. Camps-Valls, "Cloud detection machine learning algorithms for PROBA-V," in *IGARSS 2017*, July 2017, pp. 2251–2254.

[4] G. Mateo-García and L. Gómez-Chova, "Convolutional Neural Networks for Cloud Screening: Transfer Learning from Landsat-8 to Proba-V," in *IGARSS 2018*, July 2018, pp. 2103–2106.

[5] W. Dierckx, S. Sterckx, I. Benhadj, S. Livens, G. Duhoux, T. Van Achteren, M. Francois, K. Mellab, and G. Saint, "PROBA-V mission for global vegetation monitoring: standard products and image quality," *Int. Journal of Remote Sensing*, vol. 35, no. 7, pp. 2589–2614, 2014.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Oct. 2015, LNCS, pp. 234–241, Springer, Cham.

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *CVPR, 2017*, 2017.

[8] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial Discriminative Domain Adaptation," in *CVPR 2017*, July 2017, pp. 2962–2971.

[9] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krishnan, "Unsupervised Pixel-Level Domain Adaptation with Generative Adversarial Networks," in *CVPR 2017*, July 2017, pp. 95–104.

[10] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. Efros, and T. Darrell, "CyCADA: Cycle-Consistent Adversarial Domain Adaptation," in *ICML 2018*, July 2018, pp. 1989–1998.

[11] S. Foga, P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue, "Cloud detection algorithm comparison and validation for operational Landsat data products," *Remote Sensing of Environment*, vol. 194, no. Suppl. C, pp. 379–390, June 2017.

[12] U.S. Geological Survey, "L8 Biome Cloud Validation Masks," *U.S. Geological Survey*, , no. data release, 2016.

[13] U.S. Geological Survey, "L8 SPARCS Cloud Validation Masks," *U.S. Geological Survey*, , no. data release, 2016.

[14] M. J. Hughes and D. J. Hayes, "Automated Detection of Cloud and Cloud Shadow in Single-Date Landsat Imagery Using Neural Networks and Spatial Post-Processing," *Remote Sensing*, vol. 6, no. 6, pp. 4907–4926, May 2014.

[15] E.L.A. Wolters, E. Swinnen, I. Benhadj, and W. Dierckx, "PROBA-V cloud detection evaluation and proposed modification," Tech. Rep. Technical Note, 17/7/2015, QWG, 2015.