Improving future optical Earth Observation products using transfer learning

Peter Kettig\textsuperscript{1}, Eduardo Sanchez-Diaz\textsuperscript{1}, Simon Baillarin\textsuperscript{1}, Olivier Hagolle\textsuperscript{1}, Jean-Marc Delvit\textsuperscript{1}, Pierre Lassalle\textsuperscript{1}, and Romain Hugues\textsuperscript{3}

\textsuperscript{1}Centre national d'études spatiales (CNES), Toulouse, France
\textsuperscript{3}Institute of Technology Antoine de Saint Exupéry (IRT), Toulouse, France

Pixels covered by clouds in optical Earth Observation images are not usable for most applications. For this reason, only images delivered with reliable cloud masks are eligible for an automated or massive analysis. Current state of the art cloud detection algorithms, both physical models and machine learning models, are specific to a mission or a mission type, with limited transferability. A new model has to be developed every time a new mission is launched. Machine Learning may overcome this problem and, in turn obtain state of the art, or even better performances by training a same algorithm on datasets from different missions. However, simulating products for upcoming missions is not always possible and available actual products are not enough to create a training dataset until well after the launch. Furthermore, labelling data is time consuming. Therefore, even by the time when enough data is available, manually labelled data might not be available at all.

To solve this bottleneck, we propose a transfer learning based method using the available products of the current generation of satellites. These existing products are gathered in a database that is used to train a deep convolutional neural network (CNN) solely on those products. The trained model is applied to images from other - unseen - sensors and the outputs are evaluated. We avoid labelling manually by automatically producing the ground data with existing algorithms. Only a few semi-manually labelled images are used for qualifying the model. Even those semi-manually labelled samples need very few user inputs. This drastic reduction of user input limits subjectivity and reduce the costs.

We provide an example of such a process by training a model to detect clouds in Sentinel-2 images, using as ground-truth the masks of existing state-of-the-art processors. Then, we apply the trained network to detect clouds in previously unseen imagery of other sensors such as the SPOT family or the High-Resolution (HR) Pleiades imaging system, which provide a different feature space.
The results demonstrate that the trained model is robust to variations within the individual bands resulting from different acquisition methods and spectral responses. Furthermore, the addition of geo-located auxiliary data that is independent from the platform, such as digital elevation models (DEM$s$), as well as simple synthetic bands such as the NDVI or NDSI, further improves the results.

In the future, this approach opens up the possibility to be used on new CNES’ missions, such as Microcarb or CO3D.