Conditional spatio-temporal random crop for weak labeled SAR datasets

Francesco Asaro, Gianluca Murdaca, and Claudio Maria Prati
Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Milano, Italy (francesco.asaro@polimi.it)

This work presents a methodology to improve supervised learning of segmentation tasks for convolutional architectures in unbalanced and weak labeled synthetic aperture radar (SAR) dataset scenarios, which characterize the Earth Observation (EO) domain. The presented methodology exploits multitemporality and stochasticity to regularize training by reducing overfitting and thus improving validation and test performances.

Traditional precisely annotated datasets are made of patches extracted from a set of image-label pairs, often in a deterministic fashion. Through a set of experiments, we show that this approach is sub-optimal when using weak labels since it leads to early overfitting, mainly because weak labels only mark the simplest features of the target class.

The presented methodology builds up the dataset from a multitemporal stack of images aligned with the weakly labeled ground truth and samples the patches both in time and space. The patches are selected only if a given condition of the positive class frequency is met. We show learning improvements against the traditional methodology by applying our strategy to a benchmark task, which consists of training a typical deep convolutional network, Unet (Ronneberger et al, 2015), for the segmentation of water surfaces in SAR images.

The dataset sources are Sentinel-1, calibrated sigma zero, VV-VH polarized, single-look, intensity images for the inputs, and the Copernicus’s “Water and Wetness High Resolution Layer” for the weak labels. To avoid spatial autocorrelation phenomena, the training set covers the Low Countries (Belgium, the Netherlands, and Luxembourg), while the validation and test-set span the Padana plain area (Italy). The training dataset is built up according to the methodology, while the validation and test datasets are defined in a deterministic fashion as usual.

We show the beneficial effects of multitemporality, stochasticity, and conditional selection in three different sets of experiments, as well as in a combined one. In particular, we observe performance improvements in terms of the F-1 score, which increases together with the degree of multitemporality (number of images in the stack), as well as when stochasticity and conditional rules that compensate the under-representation of the positive class are added. Furthermore, we show that in the specific framework of SAR data, the introduction of multitemporality improves the learned representation of the speckle, thus implicitly optimizing the Unet for both the filtering and segmentation tasks. We prove this by comparing the number of looks of the input patch to that of
the patch reconstructed before the classification layer.

Overall, in this framework, we show that solely using the presented training strategy, the classifier's performance improves up to 5% in terms of the F-1 score.