AI4EO

Special Topics in Applications (AIL861)
Artificial Intelligence for Earth Observation
Lecture 10

Instructor: Sudipan Saha

Some datasets

Onera Satellite Change Detection dataset: images obtained by the Sentinel-2, images obtained by the Sentinel-2, 13-band multispectral
DynamicEarthNet: 4 spectral bands (RGB + near-infrared), 75 separate areas of interest,
SECOND (SEmantic Change detectiON Dataset): 4662 pairs of aerial images, size 512 x 512, focus on 6 main land-cover classes

Supervised CD

AI4EO

Supervised CD

- ✓ Dependent on the availability of training data.
- ✓ Yields very good result if the training dataset is large.
- ✓ Yields very good result if the training dataset is strongly related to the test data.

  (Training and test data are drawn from same distribution)

↓

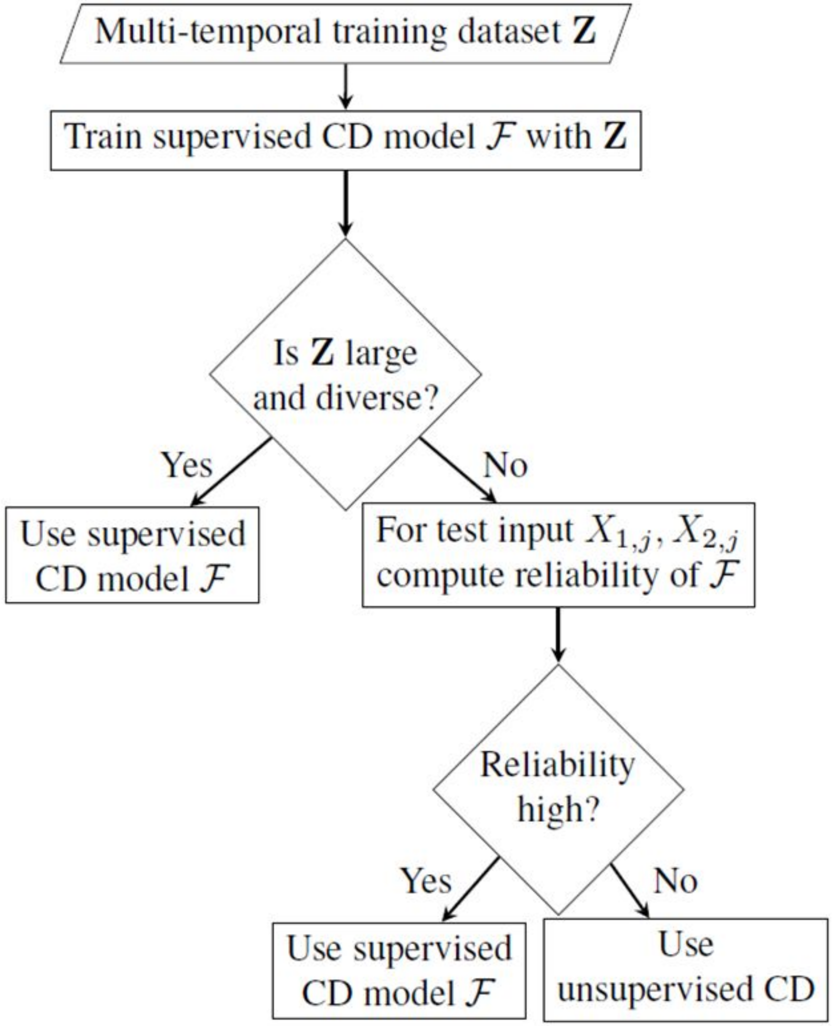
Unsupervised CD methods are preferred in the literature.

Supervised CD

- ✓ Normally during test, we inspect only the final outcome.

- ✓ May provide us with erroneous results even if the model is not confident.

Supervised or Unsupervised



Post-classification approach

Any semantic segmentation architecture can be used for supervised CD. How?

Siamese

Siamese networks are a type of neural network architecture designed to learn similarity between two inputs. The Siamese architecture consists of two identical subnetworks, which are connected to a contrastive loss function.

The idea is to feed two input images into the Siamese network, which then processes each image through the identical subnetworks, producing two feature vectors. These feature vectors are then fed into the contrastive loss function, which compares the similarity between the two vectors and generates a loss value that reflects the degree of similarity.

The contrastive loss function is designed to produce a small loss value when the two input images are similar, and a large loss value when they are dissimilar. By minimizing the loss function during training, the Siamese network learns to produce similar feature vectors for similar input images and dissimilar feature vectors for dissimilar input images.

Squeeze-Excitation-Scale

The SE block consists of two main operations: squeeze and excitation. In the squeeze operation, a global spatial average pooling operation is applied to the input feature maps, resulting in a 1D tensor with the same number of channels as the input. This operation captures the channel-wise statistics of the feature maps. In the excitation operation, two fully connected layers are applied to the 1D tensor, followed by a sigmoid activation function. This operation learns to assign different importance weights to each channel based on its contribution to the task at hand.

The output of the excitation operation is then multiplied element-wise with the original feature maps, effectively scaling the feature maps based on their importance. This recalibration of the feature maps allows the network to selectively focus on the most important features and suppress the less important ones, leading to improved performance and better generalization.

Siamese-Multitask

Siamese-Multitask is a type of neural network architecture that combines the Siamese architecture with multiple auxiliary tasks. The idea is to leverage the learned similarity metric from the Siamese network to improve performance on the auxiliary tasks, while also using the auxiliary tasks to improve the learned similarity metric.

In a Siamese-Multitask network, the identical subnetworks of the Siamese architecture are trained to perform a similarity task, such as image or text matching. At the same time, additional branches are added to the network, which perform auxiliary tasks such as object recognition, semantic segmentation, or image captioning. Each auxiliary task branch takes as input one of the images from the similarity task and produces a task-specific output.

During training, the similarity task and auxiliary tasks are optimized jointly using a combination of loss functions. The similarity task loss function is typically a contrastive loss or triplet loss that encourages the Siamese network to learn a discriminative embedding space, where similar inputs are close to each other and dissimilar inputs are far apart. The auxiliary task loss functions are designed to optimize the performance of each task separately.

The Siamese-Multitask architecture has several advantages. By using the learned similarity metric from the Siamese network to guide the training of the auxiliary tasks, the network can exploit the commonalities between the tasks, leading to better performance on all tasks. Additionally, the auxiliary tasks can help regularize the learned similarity metric, leading to better generalization and improved performance on unseen data.

Overall, the Siamese-Multitask architecture is a powerful and flexible approach for jointly learning multiple tasks in a single network, and has been successfully applied to a wide range of computer vision and natural language processing tasks.