Problem Statement & Approach Presentation

Problem Statement

i. Study area: change detection in earth observation data is the task of recognizing changes on Earth by comparing two or more satellite or aerial images covering the same area at different times.
    i. Focus on bi-temporal change detection where changes are detected between two images
    ii. Main goal is detection of extreme changes such as natural disasters
ii. Objective: design a deep learning based solution for solving the bi-temporal change detection problem in heterogenous source images
    i. The used satellite images' spatial resolution allow for change detection in ground coverage
    ii. Not suitable for dealing small objective trees, cars and such
iii. Data: the differences between the two images are the following
    i. Domain shift between the two imaging modalities
    ii. Simple transformation
    iii. Temporal difference
    iv. Each pixel maps to a corresponding on the other modality

Problem Formulation

- Same geographical region is scanned by two sensors at two different times instants t_1 , t_2 producing the two images:
  - **X**: $X \to \mathbb{R}^{H\times W\times C_1}$
  - **Y**: $\mathrm{Y} \to \mathbb{R}^{H\times W\times C_2}$
  - Where H and W are the common height and width of the two images, and C_1 , C_2 are the respective number of channels.
- Goal is to find the areas of the common geographical region where change has occurred between t_1 and t_2. This can be formulated as the estimation of the unknown change mask:
  - $\Delta(x,y) = \begin{cases} 0 & area\ unaffected\ by\ change \\ 1 & area\ changed\ over\ time\ instances \end{cases}$
  - Where we want to estimate $\widetilde{\Delta}$ that covers most of the changed areas
- The assumptions adopted are that:
  1. A limited part of the image changed between t_1 and t_2
  2. No images containing changes are provided for training

Review of the paper: Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection

- Approach is divided in two phases
  1. Prior computation
  2. Deep learning model for final change detection trained on patches with low likelihood of change
  3. Apply trained model to all patches and compare with original patches across both modalities and use reconstruction to find change maps
- Proposed deep learning models
  1. X-Net
  2. ACE-Net

Luppino, Luigi Tommaso & Kampffmeyer, Michael & Bianchi, Filippo Maria & Moser, Gabriele & Serpico, Sebastiano & Jenssen, Robert & Anfinsen, Stian. (2020). Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection.

Prior computation

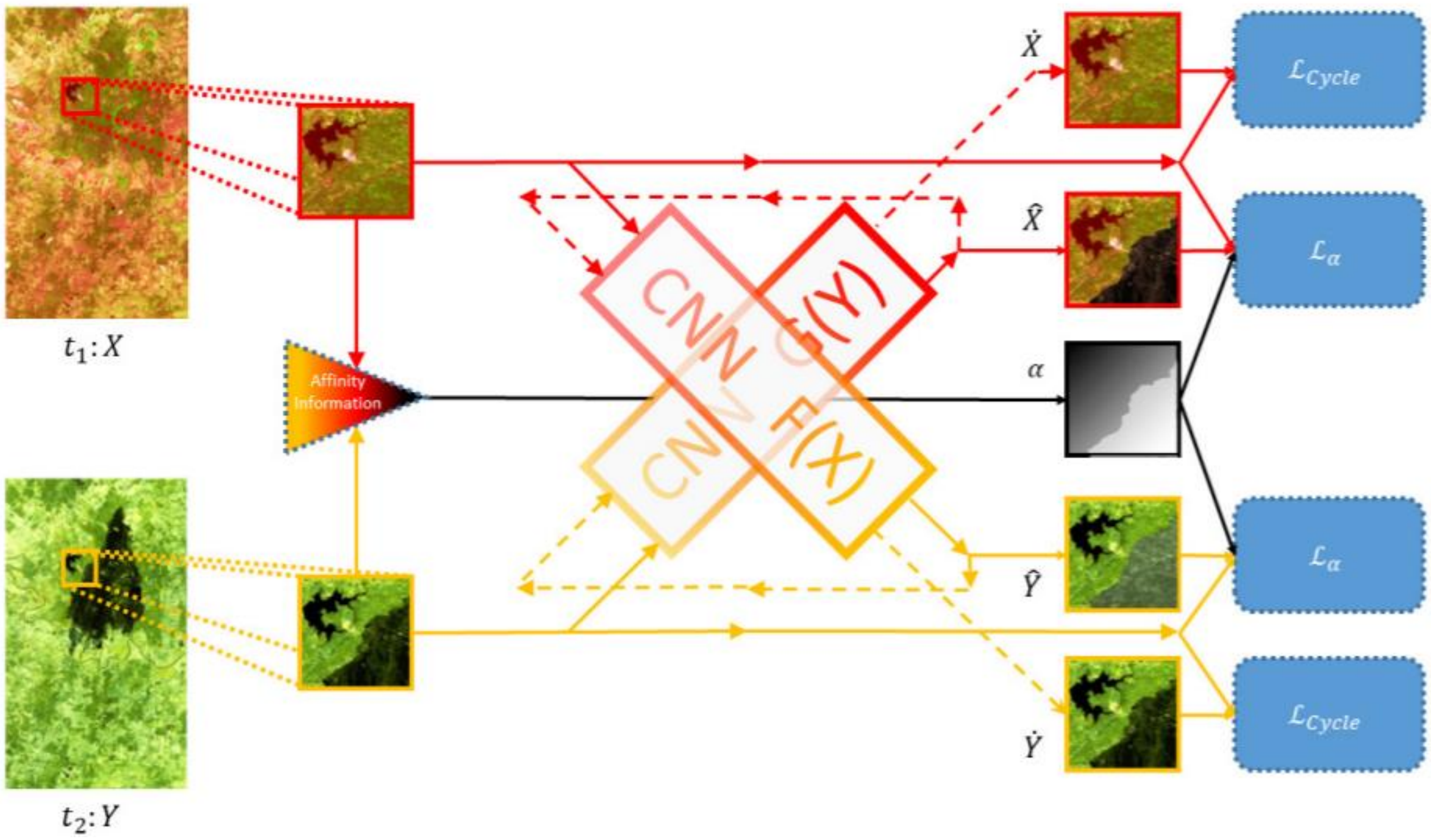
- Find a coarse detection map
- For each pair of corresponding patches compute the pairwise pixel distances and apply a gaussian kernel
- Create affinity matrices from differences in each modality
- Compute difference between the affinity matrices and threshold

Algorithm 1 Evaluation of $\boldsymbol{\alpha}$:

for all patches $p_\ell, \ell \in \{1, \ldots, |\mathcal{P}|\}$ **do**
  Compute $d^{\boldsymbol{X}}_{i,j}\ \forall i,j \in p^{\boldsymbol{X}}_\ell$ and $d^{\boldsymbol{Y}}_{i,j}\ \forall i,j \in p^{\boldsymbol{Y}}_\ell$
  Determine $h^{\boldsymbol{X}}_\ell$ and $h^{\boldsymbol{Y}}_\ell$
  Compute $A^{\boldsymbol{l}}_{i,j} = \exp\left\{-\left(\frac{d^{\boldsymbol{l}}_{i,j}}{h^{\boldsymbol{l}}_\ell}\right)^2\right\}, \boldsymbol{l} = \boldsymbol{X}, \boldsymbol{Y}$
  Compute $\alpha_{i,\ell} = \frac{1}{k^2}\sum_j |A^{\boldsymbol{X}}_{i,j} - A^{\boldsymbol{Y}}_{i,j}|\ \forall i \in p_\ell$
  Add $\alpha_{i,\ell}$ to the set $\mathcal{S}^\alpha_i\ \forall i \in p_\ell$
end for
for all pixels $i \in [1, \ldots, M]$ **do**
  Compute $\alpha_i = \frac{1}{|\mathcal{S}^\alpha_i|}\sum_{\alpha_{i,\ell} \in \mathcal{S}^\alpha_i} \alpha$
end for

Luppino, Luigi Tommaso & Kampffmeyer, Michael & Bianchi, Filippo Maria & Moser, Gabriele & Serpico, Sebastiano & Jenssen, Robert & Anfinsen, Stian. (2020). Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection.

X-Net



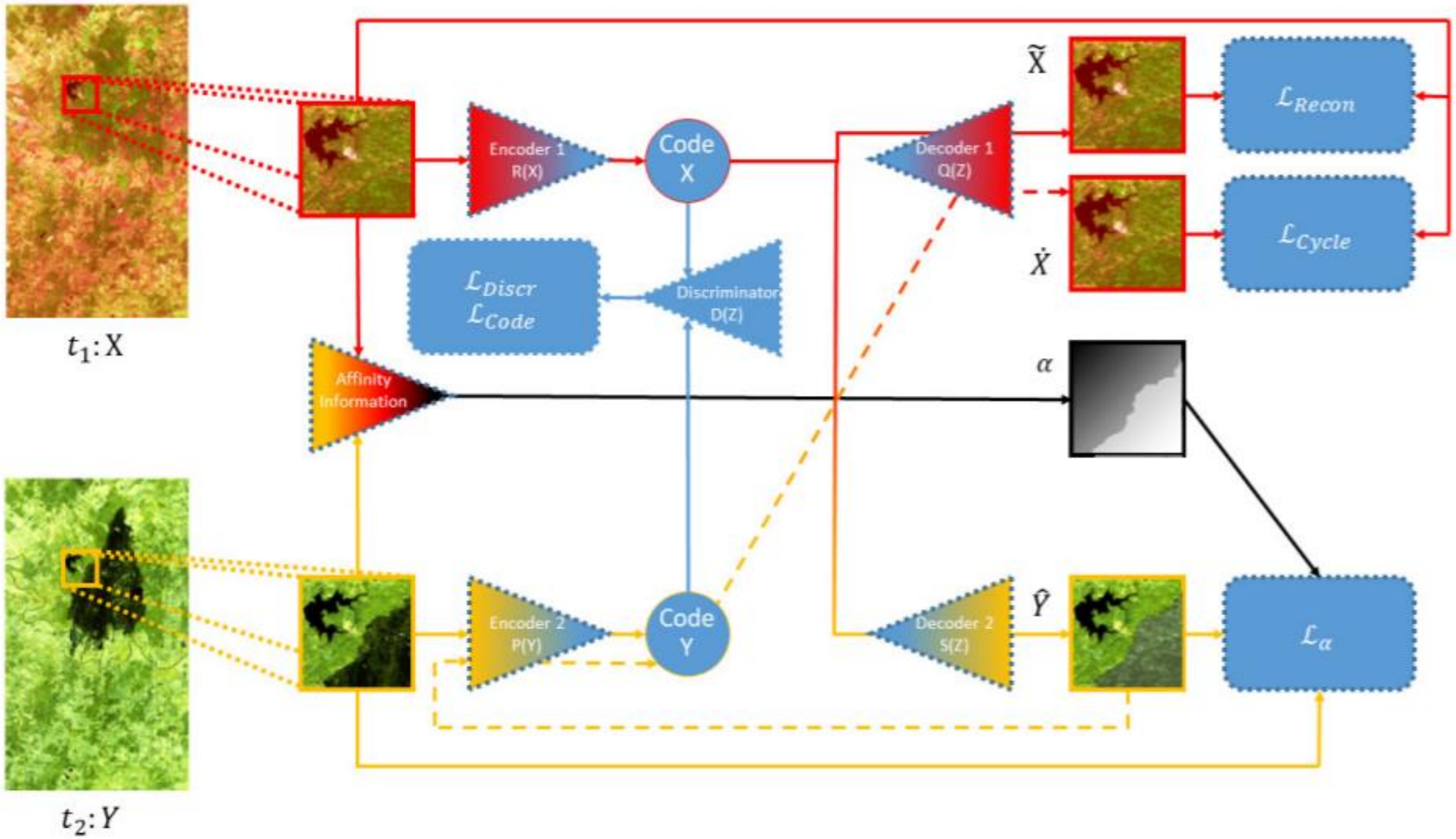
Luppino, Luigi Tommaso & Kampffmeyer, Michael & Bianchi, Filippo Maria & Moser, Gabriele & Serpico, Sebastiano & Jenssen, Robert & Anfinsen, Stian. (2020). Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection.

X-Net

- Two CNNs transform the patches from one modality to the other
  - F(x)[X->Y] G(y) [Y->X]
- The models are co-trained with the following losses
  1. Translation loss
     $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{y} \qquad \mathbf{G}(\boldsymbol{y}) = \boldsymbol{x}$
  2. Cycle consistency loss
     $\boldsymbol{F}\big(\boldsymbol{G}(\boldsymbol{y})\big) = \boldsymbol{y} \qquad \boldsymbol{G}\big(\boldsymbol{F}(\boldsymbol{x})\big) = \boldsymbol{x}$
  3. Regularization

Luppino, Luigi Tommaso & Kampffmeyer, Michael & Bianchi, Filippo Maria & Moser, Gabriele & Serpico, Sebastiano & Jenssen, Robert & Anfinsen, Stian. (2020). Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection.

ACE-Net



Luppino, Luigi Tommaso & Kampffmeyer, Michael & Bianchi, Filippo Maria & Moser, Gabriele & Serpico, Sebastiano & Jenssen, Robert & Anfinsen, Stian. (2020). Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection.

ACE-Net

- Intermediate latent space Z is introduced between X and Y
- Encoders transform patches to Z
  - F(x)[X->Z]   G(y) [Y->Z]
- Decoders transform patches from Z to X and Y respectively
  - R(z) [Z->Y]   Q(z) [Z->X]
- Discriminator trained to distinguish between encoded patches
- Adversarial training with the following losses
  1. Translation loss
     $R(F(x)) = y$ $Q(G(y)) = x$
  2. Cycle loss
     $Q\left(G\left(R(F(x))\right)\right) = x$ $R\left(F\left(Q(G(y))\right)\right) = y$
  3. Reconstruction loss
     $Q(F(x)) = x$ $R(G(y)) = y$
  4. Adversarial losses

Luppino, Luigi Tommaso & Kampffmeyer, Michael & Bianchi, Filippo Maria & Moser, Gabriele & Serpico, Sebastiano & Jenssen, Robert & Anfinsen, Stian. (2020). Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection.

Proposed solution

- A pure deep learning model
- Implementing a similar framework to the reviewed paper
- Semi-supervised approach
- Datasets
  1. Sarptical: 10,000+ pairs of no change corresponding pairs of SAR and optical images
  2. Changed image pairs for testing with ground truth
     - Gloucester flooding event
     - California flooding event
     - Texas forest fire

Architecture

1. X-net/ACE-net trained on Sarptical dataset for coarse map creation
2. Train fine change detection using low likelihood of change patches
   I. Fine tune the X-net
   II. Train a deep regression model from scratch
   III. Use a traditional machine learning regressor
3. Compute translation error for each patch and threshold to find change map