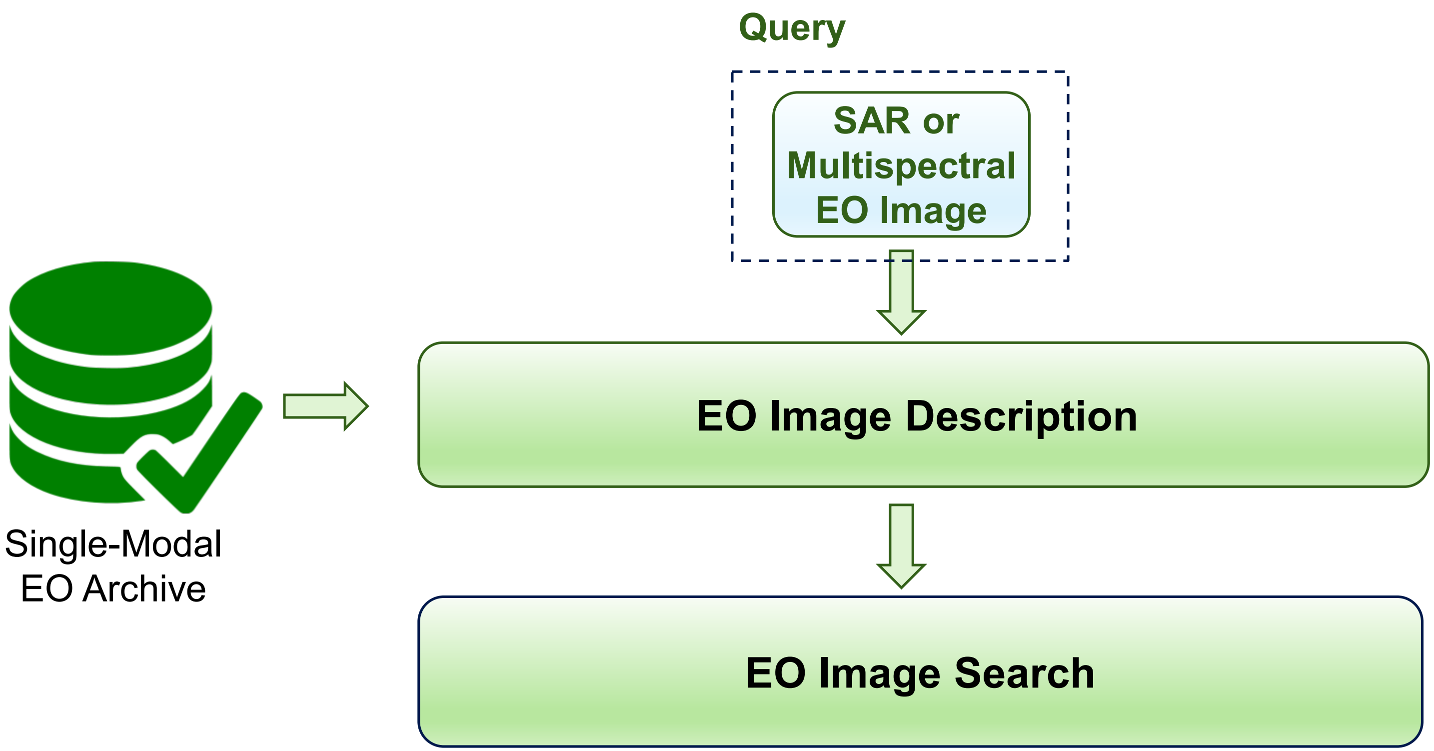
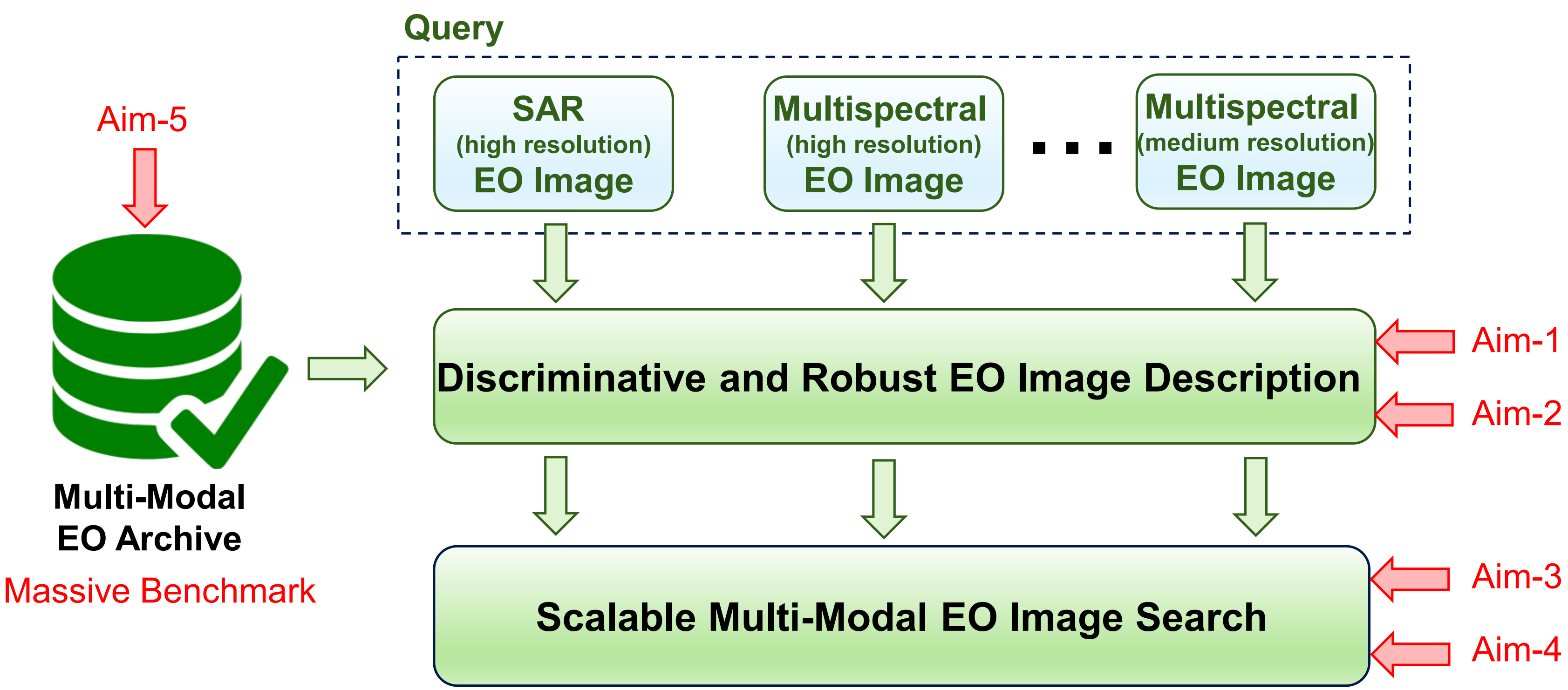
Accurate and Scalable Remote Sensing Image Search and Retrieval in Large Archives

Prof. Dr. Begüm Demir
Email: demir@tu-berlin.de
Web: https://www.rsim.tu-berlin.de & http://bigearth.eu/

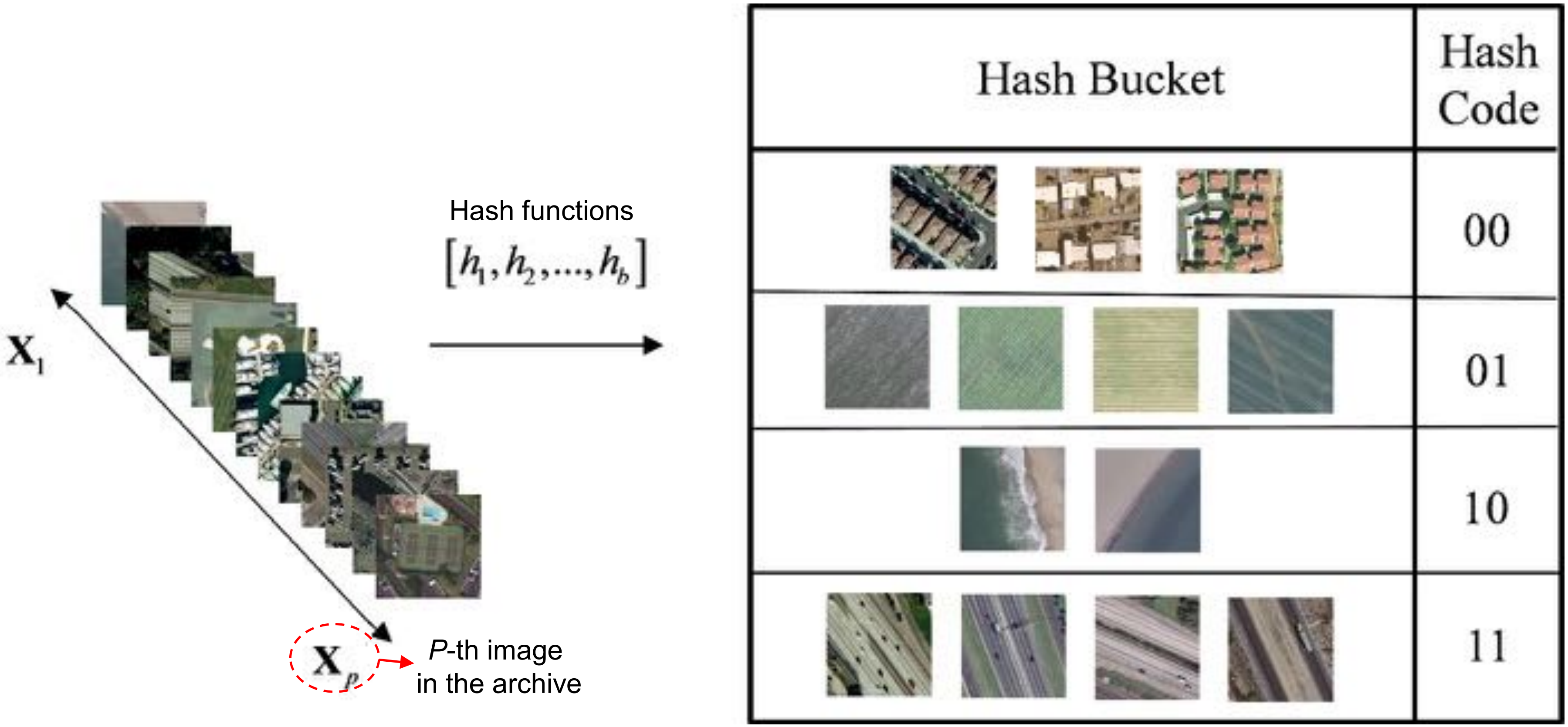
CBIR in RS


Query
SAR or Multispectral EO Image
Single-Modal EO Archive
EO Image Description
EO Image Search

BigEarth Novel Vision



Hashing Methods in Image Retrieval



B. Demir, L. Bruzzone “Hashing based scalable remote sensing image search and retrieval in large archives”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no.2, pp. 892-904, 2016.

Kernel-based Hashing Methods

✓ Two main methods that define hash functions in the kernel space:

- kernel-based unsupervised LSH hashing method (hash functions are defined by using only unlabeled images) [1].
- kernel-based supervised hashing LSH method (semantic similarity is used to define much distinctive hash functions) [2].

Kernel-based methods express the Gaussian random vector as the weighted sum of *m* images selected from the archive as:

$$v_r = \sum_{j=1}^{m} \omega_r(j) \phi(\mathbf{X}_j)$$

nonlinear mapping function

Then the hash function becomes:

$$h_r(\mathbf{X}_i) = sign\left(\sum_{j=1}^{m} \omega_r(j) \phi(\mathbf{X}_j) \phi(\mathbf{X}_i) \right) = sign\left(\sum_{j=1}^{m} \omega_r(j) K(\mathbf{X}_j, \mathbf{X}_i) \right), \; r = 1, 2, \ldots, b$$

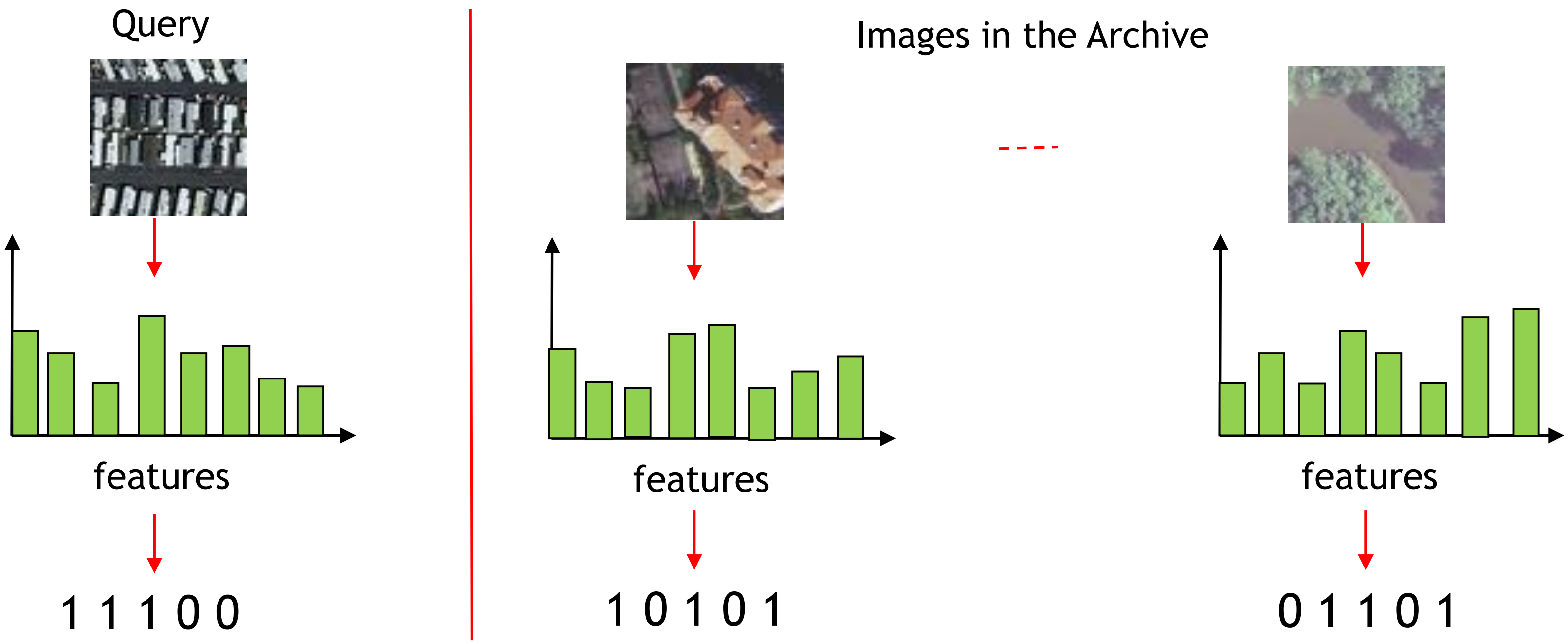
r-th hash function

kernel function

[1] B. Kulis and K. Grauman, “Kernelized locality-sensitive hashing,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 6, pp. 1092 – 1104, 2012.
[2] W. Liu, J. Wang, R. Ji, Y.-G. Jiang, and S.-F. Chang, “Supervised hashing with kernels”, Conference on Computer Vision and Pattern Recognition, Rhode Island, USA, 2012.
[3] B. Demir, L. Bruzzone, "Hashing Based Scalable Remote Sensing Image Search and Retrieval in Large Archives", IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no.2, pp. 892-904, 2016.

Pros and Cons

- ✓ These hashing methods are promising in RS for CBIR problems as they allow sub-linear time approximate similarity search with a good retrieval accuracy.



Problem: Representing a RS image with a vector of hand-crafted features, thus with a single hash code, may result in insufficient retrieval results, particularly when high-level semantic content is present in the query images.

Advances in Hashing

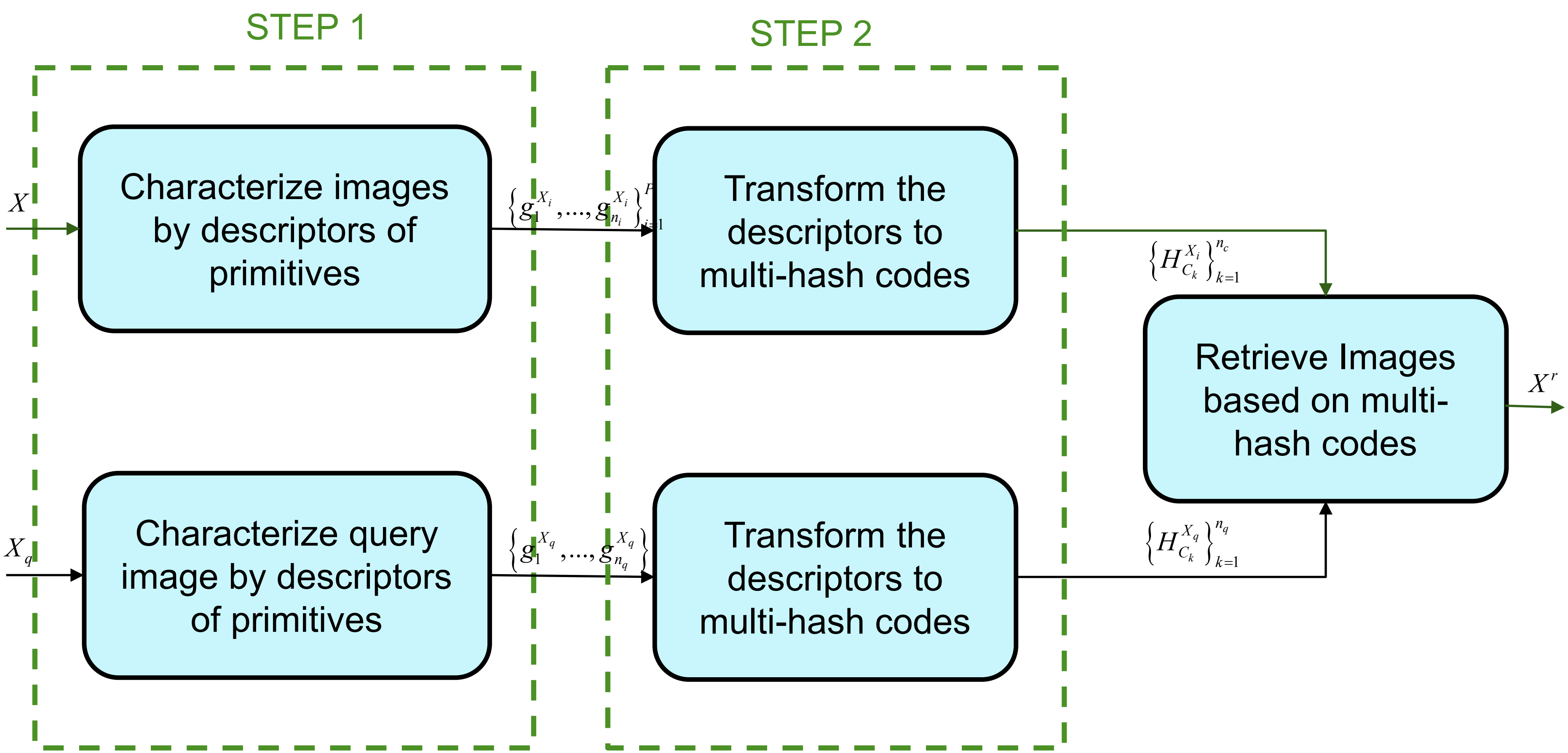
Aim: Develop hashing methods that accurately model the primitives in the definition of hashing functions.

Recent solutions: Define semantic-sensitive hashing methods:

- cluster sensitive multi-code hashing method (is unsupervised and thus does not require any annotated images).
- class sensitive multi-code hashing method (is supervised a small set of annotated images with region labels is available.
- metric learning based deep hashing network.

RSiM

Multi-Code Hashing



T. Reato, B. Demir, L. Bruzzone, "A Novel Unsupervised Multi-Code Hashing Strategy for Accurate and Scalable Remote Sensing Image Retrieval", IEEE Geoscience and Remote Sensing Letters, accepted for publication.

Unsupervised Multi-Code Hashing: Step 1

$\mathbf{X} = \left\{ \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_P \right\}$ → Extract image regions through segmentation → $\left\{ r_1^{X_i}, r_2^{X_i}, \ldots r_{n_i}^{X_i} \right\}_{i=1}^{P}$ → Compute regions' descriptors for each image

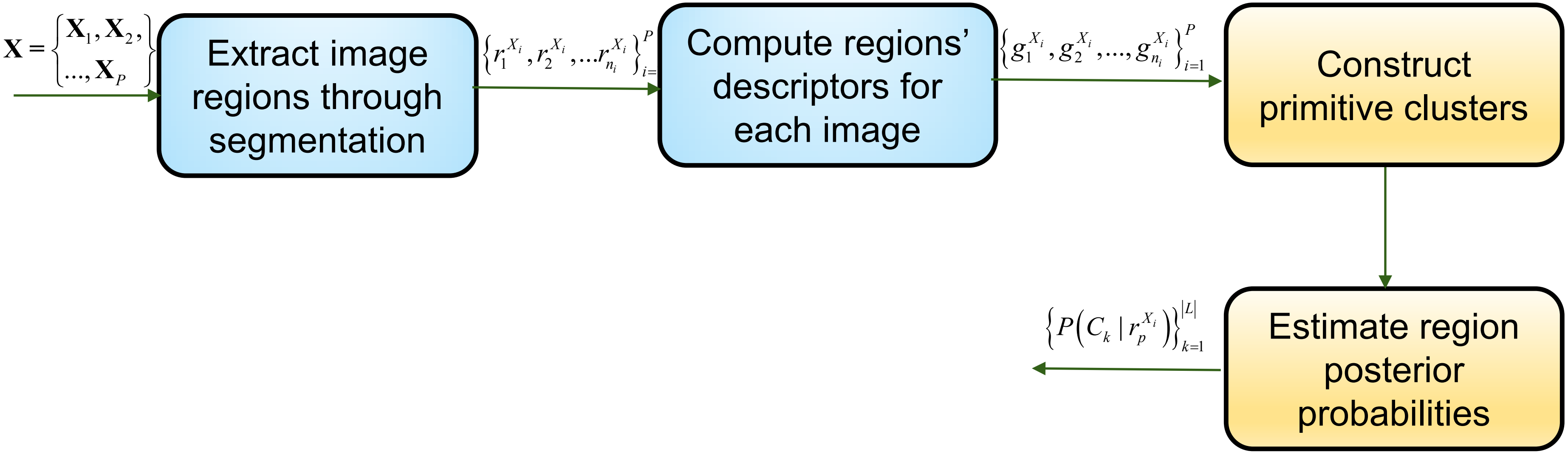
Each region is described by:

- shape features;
- texture features;
- intensity features.

✓ Any segmentation algorithm can be used, whereas in this work we have considered the parametric kernel graphs cuts algorithm.

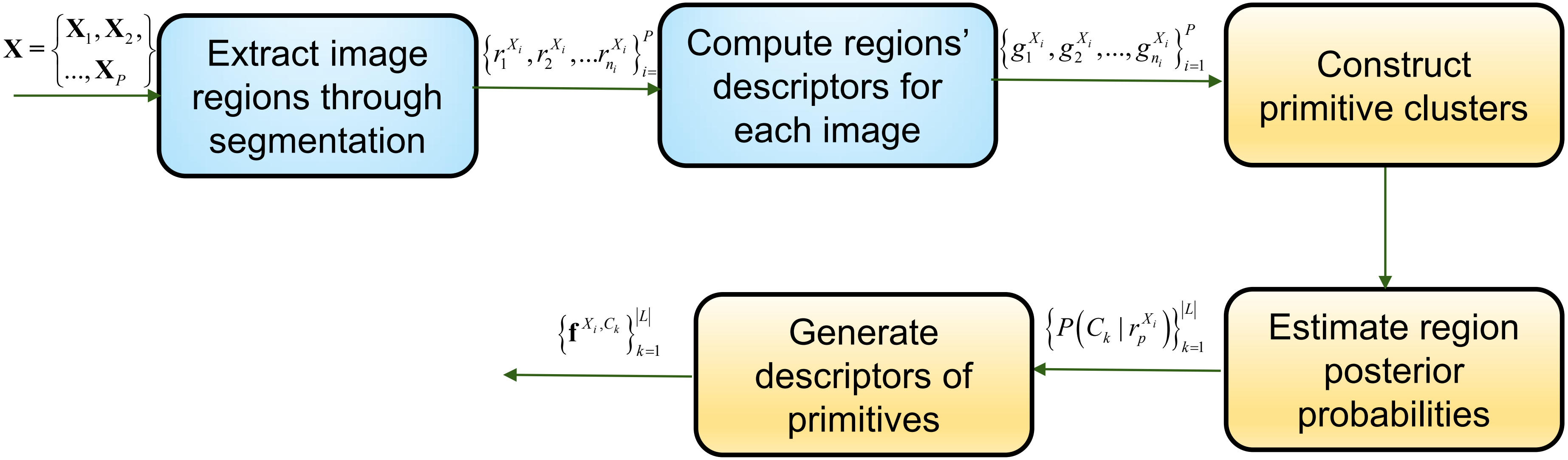
M. Ben Salah, A. Mitiche, and I. B. Ayed, "Multiregion image segmentation by parametric kernel graph cuts," IEEE Transactions on Image Processing, vol. 20, no. 2, pp. 545–557, 2011.

Unsupervised Multi-Code Hashing: Step 1



- ✓ Primitive clusters are defined clustering randomly selected regions' descriptors into n_c clusters $\{C_1, C_2, ...C_K\}$.
- ✓ This is achieved by using Gaussian mixture models, where parameters of the mixture models with n_c components are estimated by the Expectation Maximization algorithm.
- ✓ To build an accurate correspondence between the regions and the primitive clusters, $\{P(C_k \mid r_p^{X_i})\}_{k=1}^{|L|}$ are estimated from the parameters of the mixture models.

Unsupervised Multi-Code Hashing: Step 1



- ✓ Descriptors of primitives are estimated as follows:

$$
\begin{cases}
\mathbf{f}^{X_i,C_k} = \dfrac{1}{nr} \displaystyle\sum_{\forall P(C_k | r_p^{X_i}) \geq \mathrm{T}} \mathbf{g}_p^{X_i}, & \text{if } \max\limits_{p=1,2,\dots,n_i} \left\{ P(C_k | r_p^{X_i}) \right\} \geq \mathrm{T} \\
\mathbf{f}^{X_i,C_k} = \mathbf{z}, & \text{if } \max\limits_{p=1,2,\dots,n_i} \left\{ P(C_k | r_p^{X_i}) \right\} < \mathrm{T}
\end{cases}
$$

k-th class

threshold

k-th posterior probability

vector of all zero entries

Multi-Code Hashing: Step 2

- ✓ Hashing is applied to the descriptors of each primitive cluster separately from each other.

- ✓ Kernel-based unsupervised locality sensitive hashing (KULSH) is applied to the descriptors $\mathbf{f}^{\mathbf{X}_1,C_k},\mathbf{f}^{\mathbf{X}_2,C_k},\ldots,\mathbf{f}^{\mathbf{X}_P,C_k}$ of *k*-th primitive class separately from each other.

- ✓ The same process is applied for a total of *b* hash functions $[h_1^k, h_2^k, \ldots, h_b^k]$, resulting in a *b*-bits hash code $H_{C_k}^{X_i} = [h_1^k, h_2^k, \ldots, h_b^k]$ associated to each primitive class.

B. Demir, L. Bruzzone "Hashing based scalable remote sensing image search and retrieval in large archives", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no.2, pp. 892-904, 2016.

Multi-Hash-Code-Matching

- ✓ Multi-hash-code-matching scheme is used by the proposed hashing method for image retrieval.

- ✓ This scheme estimates the similarity between X_q and X_i as a sum of Hamming distances estimated between $H_{C_k}^{X_q}$ and $H_{C_k}^{X_i}, k = 1,2,\ldots,n_n$ as:

k-th hash
code of the query image

$$d^{\mathbf{X}_q,\mathbf{X}_i} = \sum_{k=1}^{n_c} H_{C_k}^{\mathbf{X}_q} \otimes H_{C_k}^{\mathbf{X}_i}, \quad \text{if } \mathbf{f}^{\mathbf{X}_q,C_k} \neq \boldsymbol{z}$$

XOR operator

k-th hash
code of the archive image

- ✓ Then, the images with the lowest distance are retrieved.

Archive Description

Data set: UCMERCED archive which consists of 2100 annotated aerial images, each of which associated with multiple labels.

dock, ship, water

buildings, cars, grass, pavement, trees

buildings, cars, grass, pavement, trees

sand, sea

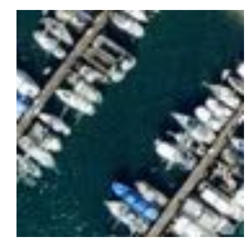
cars, pavement, trees

bare soil, buildings, grass, pavement

pavement, cars, bare soil, trees

bare soil, grass, trees

water, trees bare soil



ship, dock, water

Download the labels: **http://bigearth.eu/datasets.html**

Experimental Results

Method	Recall	Time (in seconds)	Storage Complexity
single-code hashing	58.74 %	62.7×10^{-4}	0.033 KB
multi-code hashing	**65.29 %**	62.7×10^{-4}	0.068 KB

All the experiments are implemented via MATLAB® on a standard PC with Intel® Xeon® CPU E5-1650 v2 @ 3.50GHz, 16GB RAM

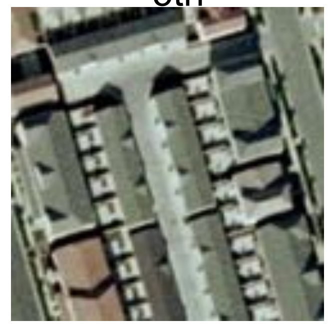
Experimental Results

Standard Single-Code Hashing method

2nd

buildings, cars, grass, pavement, trees

5th



buildings, cars, pavement

16th

pavement, sand

Query Image

building, pavement

2nd

buildings, pavement

5th

buildings, cars, grass, pavement, trees

16th

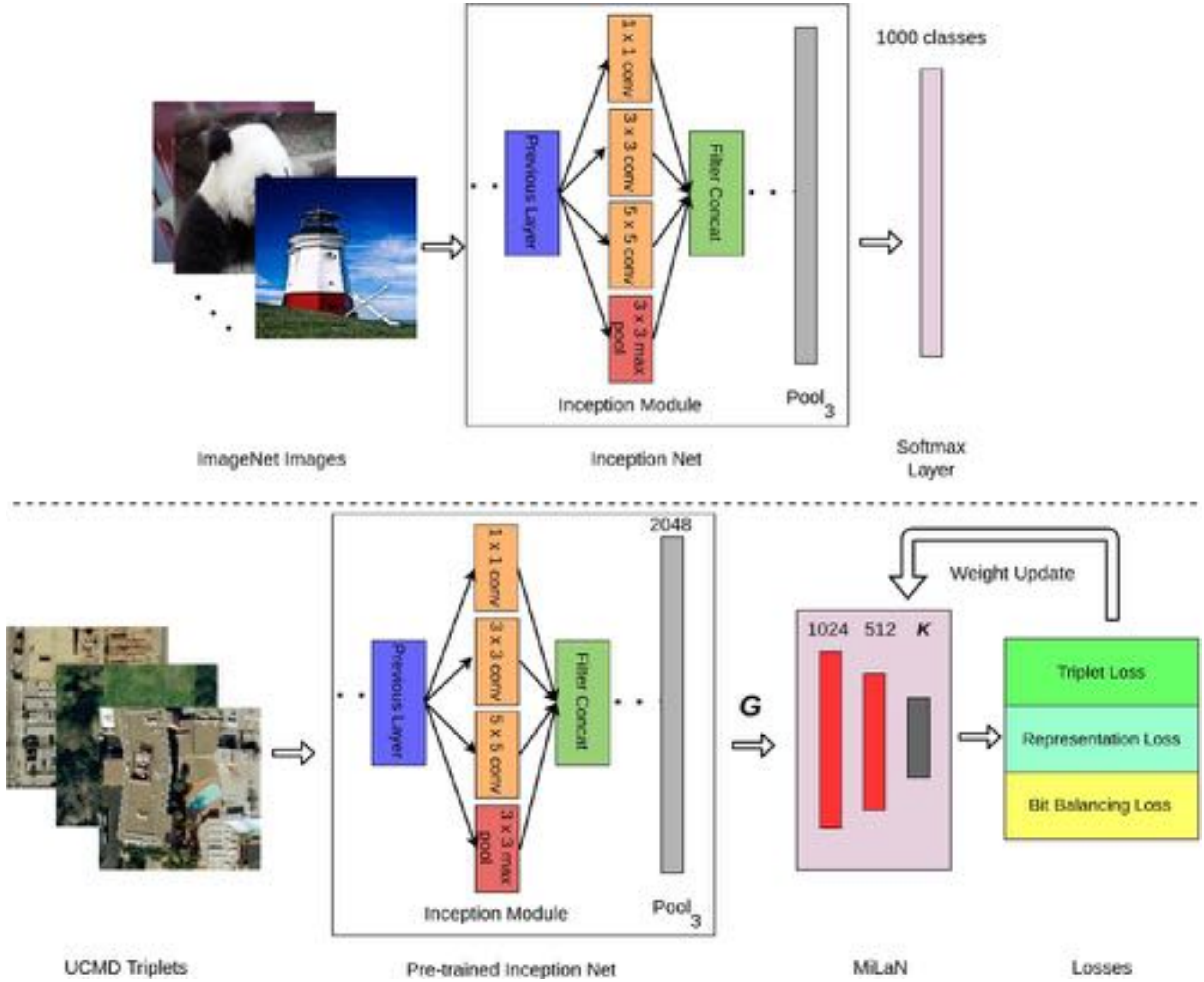
bare-soil, buildings, cars, pavement, trees

Multi-Code Hashing method

Pros and Cons

- ✓ The multi-code hashing method is promising for RS CBIR problems since it:
  - efficiently describes the complex content of RS images with multi-hash codes;
  - achieves fast and scalable image search and retrieval;
  - overcomes the limitations of the single hash codes.
- ✓ Kernel based-hashing methods in general learn hash functions in the kernel space from hand-crafted features (e.g., the bag-of-visual-words based on the scale invariant feature transform) are applied to RS CBIR problems.
- ✓ However, hand-crafted features may not accurately represent the high level semantic content of RS images. This leads to inaccurate retrieval results under complex RS image retrieval tasks.

Metric Learning based Deep Hashing



S. Roy, E. Sangineto, B. Demir and N. Sebe, "Deep Metric and hash-code learning for content-based retrieval of remote sensing images", International Conference on Geoscience and Remote Sensing Symposium, Valencia, Spain, 2018.

Experimental Results

Table: mAP and average retrieval time

Methods	Image Features		# Hash Bits K					
			K=16		K=24		K=32	
	mAP	Time (in ms)	mAP	Time (in ms)	mAP	Time (in ms)	mAP	Time (in ms)
SVM	0.556	92.3	-	-	-	-	-	-
KSLSH	-	-	0.557	25.3	0.594	25.5	0.630	25.6
Our **MHCLN**	-	-	**0.875**	25.3	**0.890**	25.5	**0.904**	25.6

S. Roy, E. Sangineto, B. Demir and N. Sebe, "Deep Metric and hash-code learning for content-based retrieval of remote sensing images", International Conference on Geoscience and Remote Sensing Symposium, Valencia, Spain, 2018.

Experimental Results

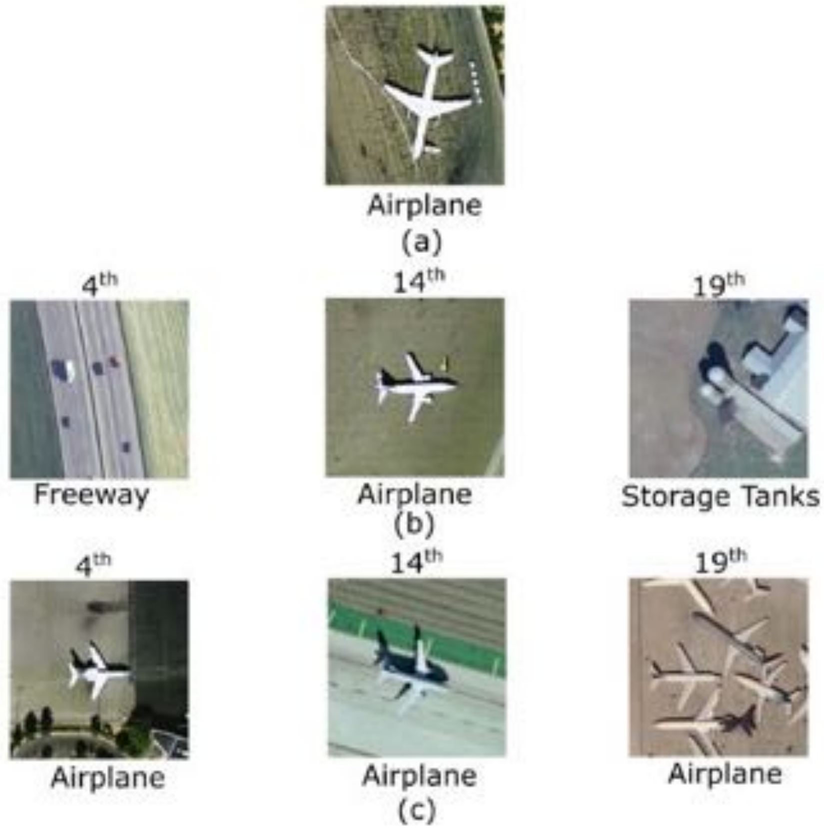
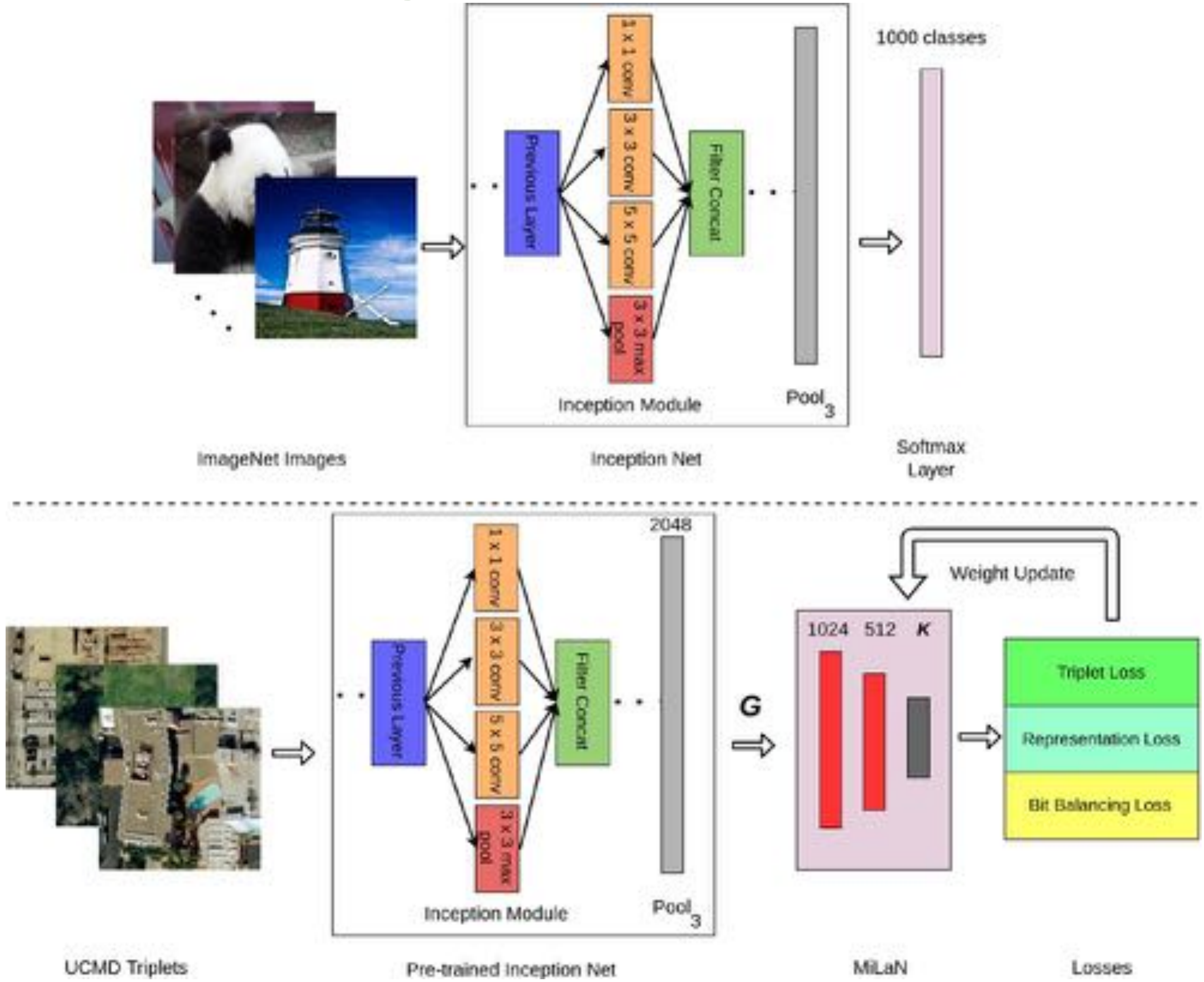


Figure: (a) Query image from UCMD, (b) images retrieved by KSLSH and (c) images retrieved by the proposed MHCLN.

Metric Learning based Deep Hashing



S. Roy, E. Sangineto, B. Demir and N. Sebe, "Deep Metric and hash-code learning for content-based retrieval of remote sensing images", International Conference on Geoscience and Remote Sensing Symposium, Valencia, Spain, 2018.

BigEarthNet- A NEW LARGE-SCALE SENTINEL-2 BENCHMARK ARCHIVE

B. Demir

BigEarthNet

✓ Contains 590,326 Sentinel-2 image patches with multiple annotations.

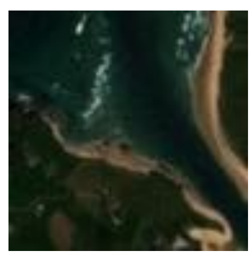
Continuous urban fabric,
Green urban areas

Non-irrigated arable land,
Fruit trees and berry plantations,
Pastures

Pastures,
Water courses,
Water bodies

Construction sites,
Non-irrigated arable land,
Pastures, Coniferous forest,
Inland marshes, Water courses

Non-irrigated arable land,
Pastures, Moors and heathland

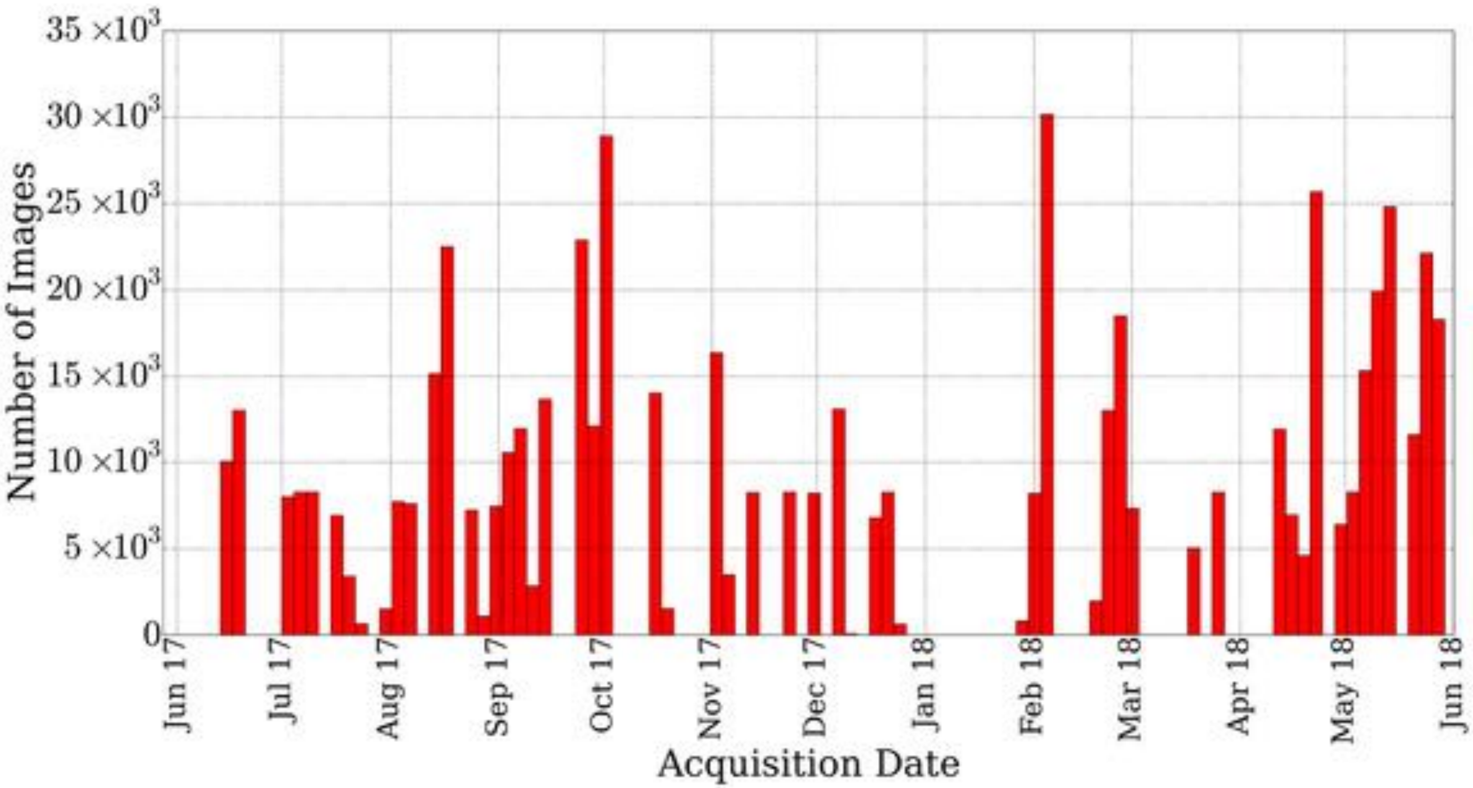


Land principally occupied by agriculture, with significant areas of natural vegetation,
Beaches, dunes, sands, Intertidal flats,
Estuaries, Sea and ocean

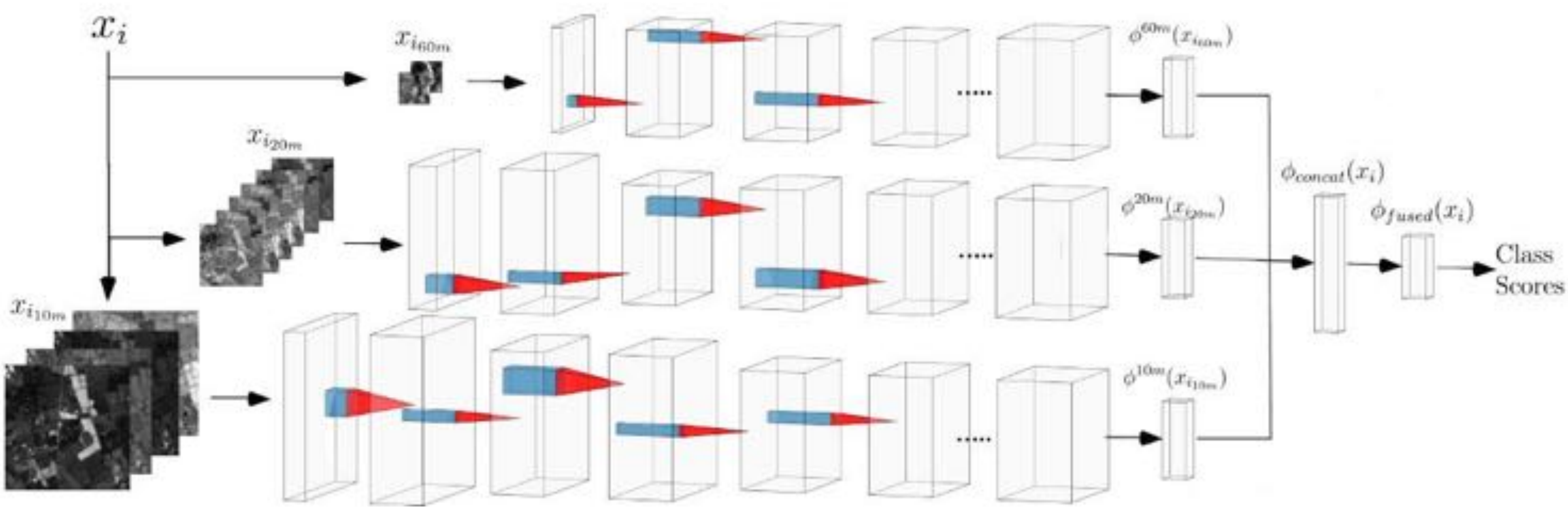
Discontinuous urban fabric,
Construction sites,
Green urban areas

BigEarthNet

- ✓ BigEarthNet has been constructed by selecting 125 Sentinel-2 tiles distributed over 10 European countries and acquired between June 2017 and May 2018.



Our Three Branch Deep Convolutional Neural Network



Single Branch CNN	Our TB-CNN
50%	67.5%

Conclusion

✓ Our hashing-based methods are promising for RS CBIR problems since they:

- efficiently describes the complex content of RS images with binary codes;
- achieves fast and scalable image search and retrieval.

✓ The BigEarthNet is 20 times larger than existing archives in RS, and thus it:

- is much more convenient to be used as a training source in the framework of deep learning;
- will make a significant advancement in terms of developments of algorithms for the analysis of large-scale RS image archives.

BigEarth

Technische Universität Berlin

Accurate and Fast Discovery of Crucial Information for Observing Earth from Big EO Archives

http://bigearth.eu/

RSiM