Zero-shot Cross-modal Retrieval of VHR EO Sketch Images

Ushasi Chaudhuri Biplab Banerjee
Avik Bhattacharya Mihai Datcu

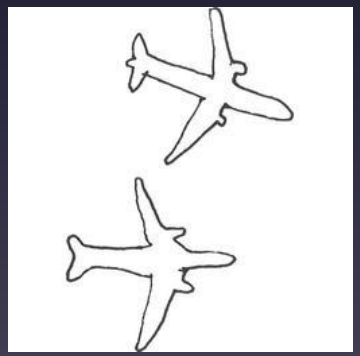
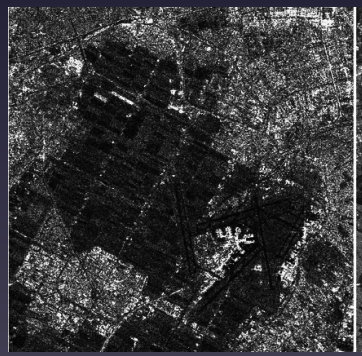
Center of Studies in Resources Engineering (CSRE)
IIT Bombay
German Aerospace Center (DLR)
Germany

Introduction

With improved storage and sensor technology:

1. Immense data collection from different satellite sensors.
2. Creates difficulty in data retrieval of **target class**.
3. Requires complicated data mining technologies.
4. Target class query searching can be difficult.

Simple solution: Use sketch-based queries!!!

Source: 1. Earth on Canvas dataset: https://ushasi.github.io/Earth-on-Canvas-dataset/
2. Meadows, P. J., and H. Laur. "Exploitation of ERS SAR imagery for land applications." EUROPEAN SPACE AGENCY-PUBLICATIONS-ESA SP 441 (1998): 43-50.
3. https://www.architecturaldigest.com/gallery/aerial-views-of-airports

Motivation:

Few more challenges:

1. Certain classes may appear directly during testing phase.
2. Insufficient or no training samples.
3. Can we can still retrieve them using just their semantic information?

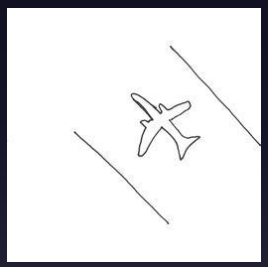
Solution: Zero-shot retrieval framework!!

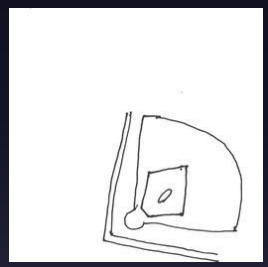
CLASSES		Image Embedding	Class Embedding
TRAINING	A	✔	✔
	B	✔	✔
	C	✔	✔
	D	✔	✔
	E	✔	✔
ZERO SHOT	F	✘	✔
	G	✘	✔
	H	✘	✔

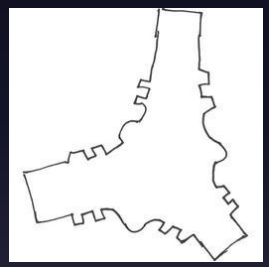
In some experiments, class embeddings of the zero-shot classes are also not used during training

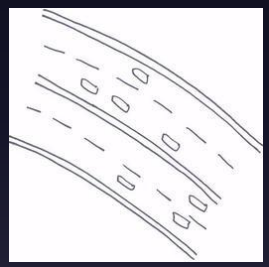
Why sketch-based image retrieval?

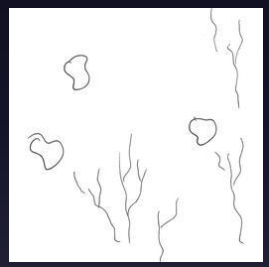
1. Minimalistic representation of data.
2. Can be quickly hand-drawn anywhere.
3. Easiest representation of the mental picture of the target query image.

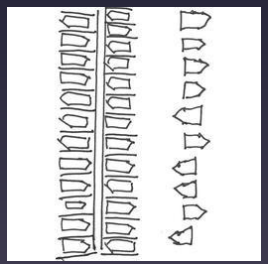

Airplane

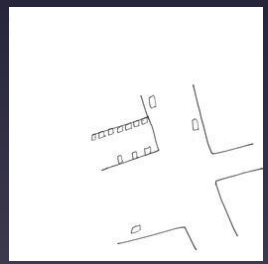

Baseball

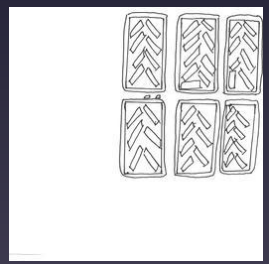

Buildings

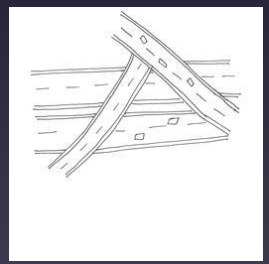

Freeway

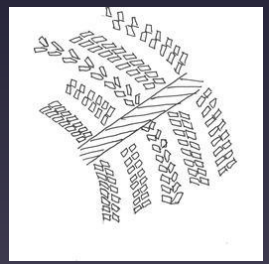

Golf course


Harbor


Intersection


Residential


Overpass


Parking

Sample sketch instances of each class.

Sketch-based image retrieval

Some examples of sketch-based image retrieval...

Sketch-based image retrieval

Idea of Framework:

1. Construct a unified space for features of both modality.
2. Features should be class-wise discriminative.
3. Domain-gap between photo and sketch should be minimized.
4. Semantic information can be simply the word2vec encoding of the class labels.

Currently tried for sketch-based aerial image retrieval.

Possible Future Extension:

1. Idea can be extended for SAR, Hyper-spectral, or multi-spectral data retrieval.
2. Preserve topology of classes by adding a semantic graph.

Relevant Literature:

[1] Fang Xu, Wen Yang, Tianbi Jiang, Shijie Lin, Hao Luo, and Gui-Song Xia. "Mental Retrieval of Remote Sensing Images via Adversarial Sketch-Image Feature Learning." *IEEE Transactions on Geoscience and Remote Sensing* (2020).

[2] Chaudhuri, Ushasi, Biplab Banerjee, Avik Bhattacharya, and Mihai Datcu. "A Zero-Shot Sketch-based Inter-Modal Object Retrieval Scheme for Remote Sensing Images." arXiv preprint arXiv:2008.05225 (2020).