Qualifier

Retrieval of Remote Sensing Images based on Deep Learning Techniques

Ushasi Chaudhuri

Under the supervision of

Avik Bhattacharya & Biplab Banerjee

Center of Studies in Resources Engineering (CSRE)

IIT Bombay

4 January 2019

Outline

1. Introduction
   - Retrieval
   - Applications
   - Motivation
2. Objectives
   - Graph Convolutional
   - Cross-modal
3. References

Introduction

Definition (Retrieval)

The action of obtaining or consulting data stored in a computer system.

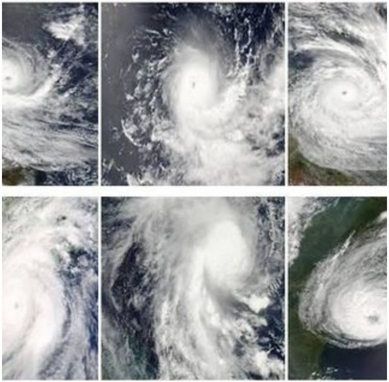
Can be used for various forms:

- Image/SAR retrieval/spotting
- Text retrieval
- cross-modal data retrieval

source:
https://www.google.com/url?sa=isource=imagescd=ved=2ahUKEwiwjfvq78nfAhVBNI8KHUFDCT8Qjhx6BAgBEAMurl=http
tutorial-on-neural-networks-for-information-
retrieval%2Fpsig=AOvVaw2Zd0FdVts4mHYkenq9esqrust=1546338999199345

Introduction - Retrieval

Finds varied applications: Cyclone detection & spotting[1,2]



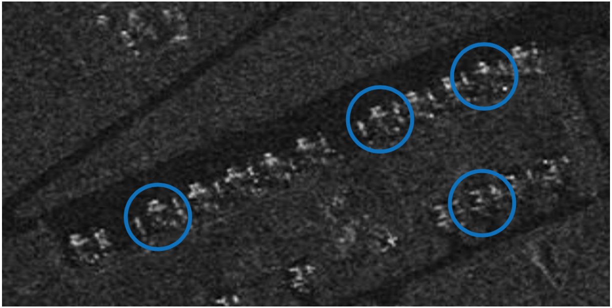
Source: https://www.researchgate.net/figure/Examples-of-retrieved-images-of-a-cyclone-query-image-using-LVP_fig3_326193252

[1] J. Liu, C. Liu, B. Wang and D. Qin, "A novel algorithm for detecting center of tropical cyclone in satellite infrared images," 2015 IEEE IGARSS, Milan, 2015.

[2] N. Jaiswal and C. M. Kishtawal, "Objective Detection of Center of Tropical Cyclone in Remotely Sensed Infrared Images," in IEEE JSTARS, vol. 6, no. 2, pp. 1031-1035, April 2013.

Introduction - Retrieval

Applications: Airplane spotting using RADARSAT-2[3][4][5]



Source: https://platform.digitalglobe.com/radarsat-2-content-gbdx/

[3] S. Bo and Y. Jing, "Region-based airplane detection in remotely sensed imagery," 2010 3rd International Congress on Image and Signal Processing, Yantai, 2010, pp. 1923-1926.

[4] Z. An and Z. Shi, "An airplane detection method for panchromatic image," 2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, 2013, pp. 189-192.

[5] X. Li, S. Wang, B. Jiang and X. Chan, "Airplane detection using convolutional neural networks in a coarse-to-fine manner," 2017 IEEE 2nd ITNEC, Chengdu, 2017, pp. 235-239.

Introduction - Retrieval

Dataset: Query	1	2	3	4	5

Query image versus the retried images from UC Merced and PatternNet. The first and the second row images are from UC Merced dataset `Airplane` and `Buildings` classes respectively, while the third and the fourth row images are from the PatternNet dataset of `ClosedRoad` and `CoastalMansion` classes.

Introduction

- CBIR performance depends on:
  1. **representation capability** of extracted features
  2. efficiency of **similarity measure**.
- If extracted features are of very high dimension, system is affected by the curse of dimensionality.
- Also, feature space should be discriminative to avoid class overlapping problem.
- Problem of exhaustive high dimensional feature matching:
  1. improve the **search criteria**
  2. **dimensionality reduction** in the feature space.

Motivation

We plan to focus our attention on two important problems:

1. **global image representation**, highlighting the interactions among the **local scene** constructs.
2. discriminative learning of an embedding space from an **irregular spatial distribution** of regions.

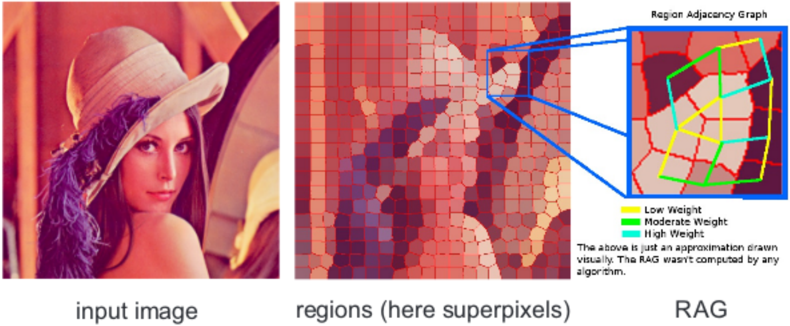


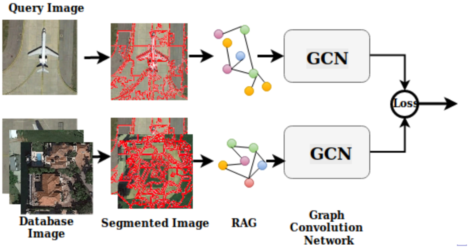
Figure: Region adjacency graph (RAG) construction.

Problem Overview:

Challenge: We have moved from an Euclidean (image) to a non-Euclidean (RAG) domain.

Possible approach to the problem:

1. Image segmentation
2. Feature extraction (shape, color, texture, etc)
3. RAG formation
4. Graph convolution and pooling
5. Image retrieval using GCN based feature embedding.



Graph Convolution and Pooling

- **Convolution:**[6]

$$\mathbf{H} = h_0\mathbf{I} + h_1\mathbf{A_1} + \ldots, \qquad \mathbf{H} \in \mathbb{R}^{N\times N} \tag{1}$$

$$\mathbf{V}_o^{(k)} = \sum_{l=1}^{l} \mathbf{H}^{(l,k)}\mathbf{V}_i^{(l)} + b \tag{2}$$

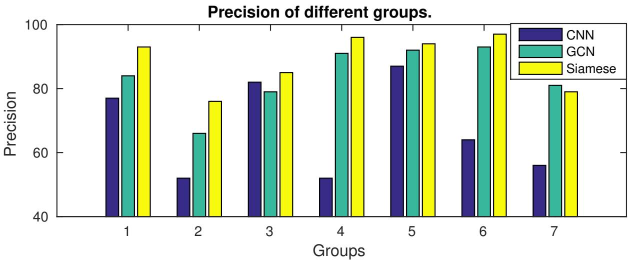
- **Contrastive loss:**

$$L^k = \sum_{n=1}^{B} \Big((1 - y_{nij})(\widehat{x}_{ni} - \widehat{x}_{nj})^2 + y_{nij}\left\{\max(0, m - (\widehat{x}_{ni} - \widehat{x}_{nj}))\right\}^2 + \alpha\left(|\widehat{x}_{ni}|_2 + |\widehat{x}_{nj}|_2\right) + \beta|\theta|_2 \Big) \tag{3}$$

[6] Input vertex features - $\mathbf{V}_i$; First-order approximation of their adjacency matrix - $\mathbf{H}$; zeroth-order and first-order adjacency matrices - h_0 and h_1

Results

Model	UC Merced ANMRR	UC Merced MAP(%)	UC Merced P@10 (%)	PatternNet ANMRR	PatternNet MAP(%)	PatternNet P@10
G-KNN	0.92	7.50	10.12	0.88	12.35	13.24
RAG-KNN	0.75	26.74	24.90	0.69	22.56	37.70
VGG-VD16	0.38	53.71	78.34	0.33	59.86	92.04
VGG-VD19	0.39	53.19	77.60	0.34	57.89	91.13
GoogLeNet	0.39	53.13	80.96	0.29	63.11	93.31
GCN	0.33	64.81	87.12	0.28	73.11	95.53
SGCN	**0.30**	**69.89**	**93.63**	**0.21**	**81.79**	**97.14**



2. Cross-modal retrieval of pan-chromatic, multi-spectral and SAR images using hashing techniques.

- Within the same dataset, retrieval is a relatively easy work. If we have different modalities of data, ex Pan images, multi-spectral images, SAR images, optical images, or texts, retrieval amongst cross domain becomes more challenging.
- A network learned for a particular modality may not give good performance on different data.

Definition (No free lunch theorem)

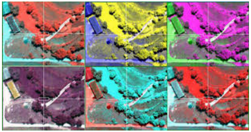
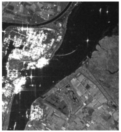
[a] If an algorithm performs well on a certain class of problems then it necessarily pays for that with degraded performance on the set of all remaining problems.

[a] Wolpert, D.H., Macready, W.G. (1997), "No Free Lunch Theorems for Optimization", IEEE Transactions on Evolutionary Computation 1, 67.

2. Cross-modal retrieval of pan-chromatic, multi-spectral and SAR images using hashing techniques.

Advantages of each data:

- **PAN images:** High spatial resolution.
- **Multi-spectral images:** high spectral resolution.
- **SAR images:** Polarization information. No cloud clutter.
- **VHR optical image:** High spatial resolution, 3 spectral channel.


1. 1m pan IKONOS image of Valparaiso, Chile
2. source: https://www.sensorsmag.com/components/hyperspectral-and-multispectral-imaging-sensors-find-hyper-flexible-applications
3. Sentinel-1 SAR dataset with C band
4. `sparseresidentialarea` from PatternNet dataset.

References

- Such, Felipe Petroski, et al. "Robust spatial filtering with graph convolutional neural networks." *IEEE Journal of Selected Topics in Signal Processing* 11.6 (2017): 884-896.
- Zhou, Weixun, et al. "Patternnet: a benchmark dataset for performance evaluation of remote sensing image retrieval." *ISPRS Journal of Photogrammetry and Remote Sensing (2018)*.
- Xinbo Gao, Lihuo Hi, et al. "Label consistent matrix factorization hashing for large scale cross modal similarity search" *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2018.
- Tinne Tuytelaars, et al., "Memory Aware Synapses: Learning what (not) to forget",*arXiv:1711.09601v4* [cs.CV] 5 Oct 2018

Thank You