Generalized Scalable Neighborhood Component Analysis for Remote Sensing Image Characterization

Jian Kang[1], Ruben Fernandez Beltran[2], Antonio Plaza[3]

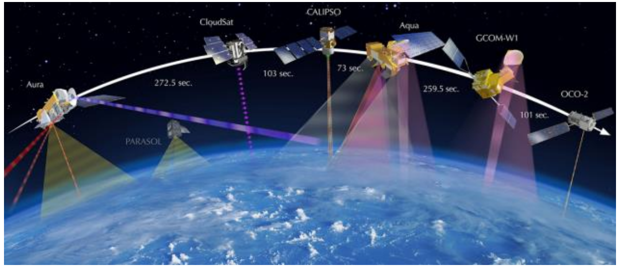
1. School of Electronic and Information Engineering, Soochow University, Suzhou 215006, China
2. Institute of New Imaging Technologies, University Jaume I, E-12071 Castellon, Spain
3. Hyperspectral Computing Laboratory, University of Extremadura, E-10003 Caceres, Spain

Outline

- Introduction
- Motivation
- Background Knowledge
- Generalized Scalable Neighborhood Component Analysis (GSNCA)
- Experiments
- Conclusion

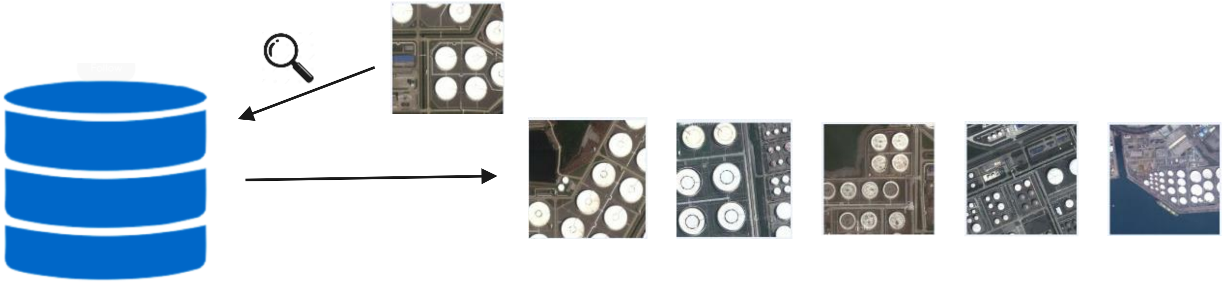
Introduction

- Remote Sensing (RS) technology development meets big Earth Observation (EO) data



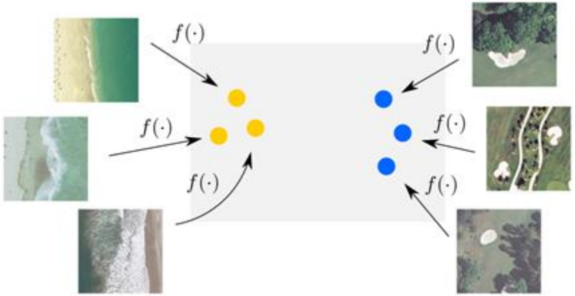
Credit: Wikipedia

- Retrieving interested contexts from big EO data is a basic task in RS



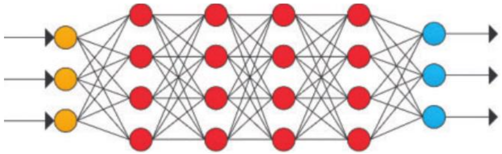
Introduction

- Characterizing the contexts of RS images with low-dimensional features is the key for achieving image retrieval



- Deep learning has been a workhorse for learning those features

$$f(\cdot) \longrightarrow$$



Introduction and Motivation

- Most state-of-the-art methods focus on designing proper CNN architectures for characterizing RS images based on deep embeddings

- Few work have focused on investigating the performance of learned embeddings and the associated metric space

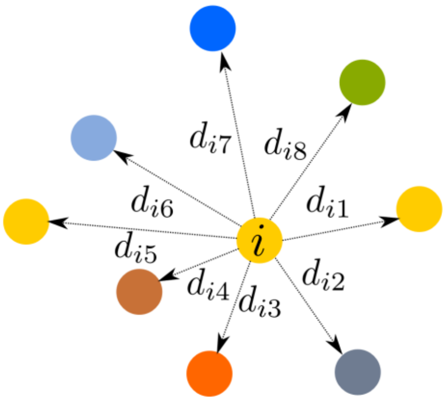
- **A unified loss function for modeling semantic relationships among RS images with both single and multiple labels**

Background Knowledge

- Neighborhood Component Analysis (NCA) [1]

$$p_{ij} = \frac{\exp(-\|\mathbf{A}\mathbf{x}_i - \mathbf{A}\mathbf{x}_j\|_2^2)}{\sum_{k \neq i} \exp(-\|\mathbf{A}\mathbf{x}_i - \mathbf{A}\mathbf{x}_k\|_2^2)},$$

  - p_{ij} the probability that the sample $\mathbf{x}_j$ is the neighbor of the sample $\mathbf{x}_i$
  - $\mathbf{A}$ the linear feature mapping
  - $d_{ij} = \|\mathbf{A}\mathbf{x}_i - \mathbf{A}\mathbf{x}_j\|_2^2$ distance of projected features in latent space



[1] J. Goldberger, G. Hinton, S. Roweis, R. Salakhutdinov. (2005) Neighbourhood Components Analysis. Advances in Neural Information Processing Systems. 17, 513-520, 2005.

Background Knowledge

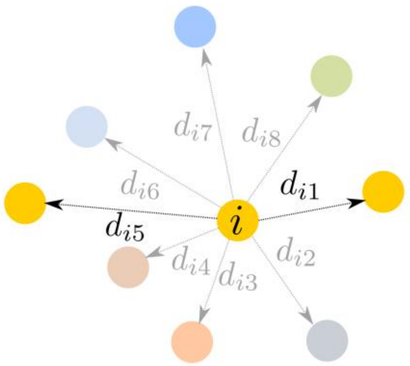
- NCA
  - The sample $\mathbf{x}_i$ can be correctly classified is with the probability

$$p_i = \sum_{j \in \Omega_i} p_{ij}$$

  - NCA loss is:

$$\mathcal{L} = -\sum_{i} \log(p_i)$$

- Objective of NCA
  - Learn the linear mapping of features which can pull the projected features sharing the same class together in the latent space



Background Knowledge

- Scalable Neighborhood Component Analysis (SNCA) [2]
  - Loss function

$\mathbf{A}$ replaced by deep learning models

$$p_{ij} = \frac{\exp(s_{ij}/\sigma)}{\sum_{k \neq i} \exp(s_{ik}/\sigma)}, \quad p_{ii} = 0$$

s_{ij} cosine similarity

$$\mathcal{L}_{\text{SNCA}} = -\sum_i \sum_{j \in \Omega_i} p_{ij}$$

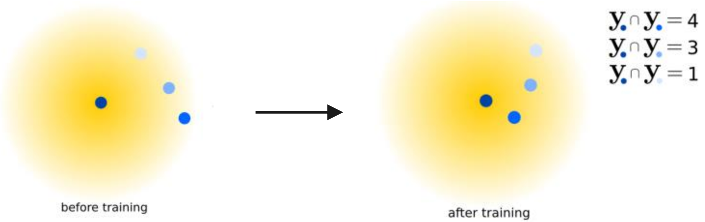
[2] Wu, Zhirong, Alexei A. Efros, and Stella X. Yu. "Improving generalization via scalable neighborhood component analysis." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.

Generalized Scalable Neighborhood Component Analysis (GSNCA)

- Generalization of SNCA to multi-label RS image characterization

Residential → [Bare-soil; Buildings; Cars; Pavement; Trees]

- **Two images sharing more labels should be closer in latent space**



Generalized Scalable Neighborhood Component Analysis (GSNCA)

- Reformulate the classification probability of $\mathbf{x}_i$ based on the indicator function

$$p_i = \sum_{j \in \Omega_i} p_{ij} \longrightarrow p_i = \sum_j \mathbb{1}_{\Omega_i}(j) p_{ij} \qquad \mathbb{1}_{\Omega_i}(j) := \begin{cases} 1 & \text{if } j \in \Omega_i, \\ 0 & \text{if } j \notin \Omega_i. \end{cases}$$

- **The indicator function can be also considered as a weighting matrix**

$$p_i = \sum_j \mathbb{1}_{\Omega_i}(j) p_{ij} \longrightarrow p_i = \sum_j w_{ij} p_{ij}$$

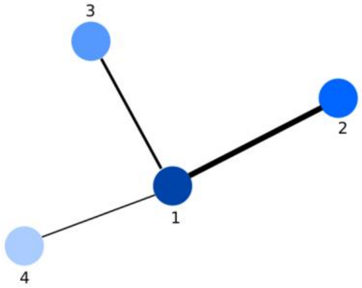
Generalized Scalable Neighborhood Component Analysis (GSNCA)

- $p_i = \sum_j w_{ij} p_{ij}$ $\quad w_{ij}$ weights measuring the level of similarities between $\mathbf{x}_i$ and $\mathbf{x}_j$

$$w_{ij} = \frac{\langle \mathbf{y}_i, \mathbf{y}_j \rangle + C}{2C}, \quad w_{ij} \in [0, 1]$$

$$\mathbf{y}_i, \mathbf{y}_j \in [+1, -1]^C$$

C the total number of classes



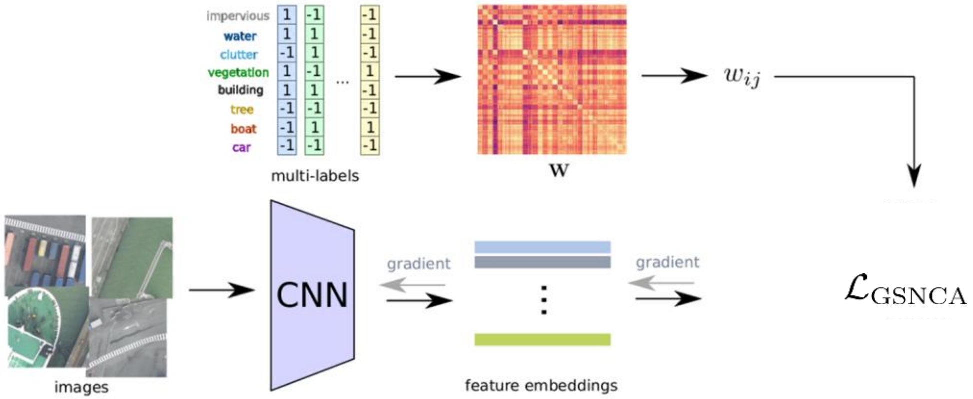
1: [Bare-soil; Buildings; Cars; Court; Grass]
2: [Bare-soil; Buildings; Cars; Court]
3: [Bare-soil; Buildings; Cars]
4: [Buildings]

$$w_{12} > w_{13} > w_{14}$$

Generalized Scalable Neighborhood Component Analysis (GSNCA)

- Loss function $$\mathcal{L}_{\text{GSNCA}} = -\sum_i \log\Big(\sum_j w_{ij} p_{ij}\Big)$$

- Framework [3]



[3] Kang, Jian, et al. "Graph relation network: Modeling relations between scenes for multilabel remote-sensing image classification and retrieval." *IEEE Transactions on Geoscience and Remote Sensing* 59.5 (2020): 4355-4369.

Experimental Setup

Dataset	UCM; AID; DFC15 (Mutli-label)
Data Splitting	Train:0.7, Val:0.1, Test:0.2
Tasks	KNN classification; Image retrieval
Metrics	WMAP; MAP; Hamming Loss; F1/2; Precision; Recall
Compared Methods	BCE; LSEP [4]

[4] Li, Yuncheng, Yale Song, and Jiebo Luo. "Improving pairwise ranking for multi-label image classification." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

Experimental Results

- KNN classification

		UCM				AID				DFC15			
		F_s^1	F_s^2	P_s	R_s	F_s^1	F_s^2	P_s	R_s	F_s^1	F_s^2	P_s	R_s
ResNet18	Contrastive	64.67	64.63	69.49	65.73	75.43	74.75	80.78	75.67	–	–	–	–
	BCE	87.76	88.23	89.19	89.19	88.31	87.42	91.77	87.25	92.74	91.88	95.38	91.55
	GSNCA	88.47	88.87	89.82	89.76	89.13	88.43	92.39	88.37	93.08	92.47	95.19	92.32
	LSEP	88.75	89.37	89.65	90.40	89.78	89.13	92.70	89.17	92.91	92.04	95.58	91.72
	GSNCA-BCE	**89.82**	**90.26**	**91.06**	**91.15**	**90.26**	**89.66**	**93.11**	**89.63**	**94.72**	**94.35**	**96.09**	**94.29**
ResNet50	Contrastive	77.02	76.78	81.50	77.49	76.31	75.88	80.42	76.88	–	–	–	–
	BCE	89.73	90.61	90.10	91.76	89.18	88.45	92.32	88.37	93.95	93.39	95.88	93.24
	GSNCA	89.68	90.11	90.71	90.91	90.43	89.88	**93.27**	89.95	94.53	94.27	95.79	94.31
	LSEP	90.36	91.26	90.57	92.43	89.43	88.65	92.74	88.58	93.52	93.03	95.49	92.96
	GSNCA-BCE	**91.31**	**91.92**	**91.98**	**92.83**	**90.95**	**90.82**	92.79	**91.08**	**95.80**	**95.78**	**96.53**	**95.95**
WideResNet50	Contrastive	74.84	74.99	78.36	76.08	81.06	80.30	85.48	80.59	–	–	–	–
	BCE	88.45	88.89	89.75	89.76	89.36	88.67	92.41	88.63	93.39	92.94	95.29	92.88
	GSNCA	90.31	90.81	91.21	91.68	90.55	89.93	93.39	89.89	94.81	94.46	96.31	94.44
	LSEP	90.22	90.81	91.13	91.79	89.40	88.56	92.87	88.44	93.68	92.92	96.11	92.67
	GSNCA-BCE	**90.81**	**91.18**	**91.97**	**91.92**	**91.02**	**90.50**	**93.66**	**90.49**	**95.96**	**95.73**	**97.13**	**95.74**

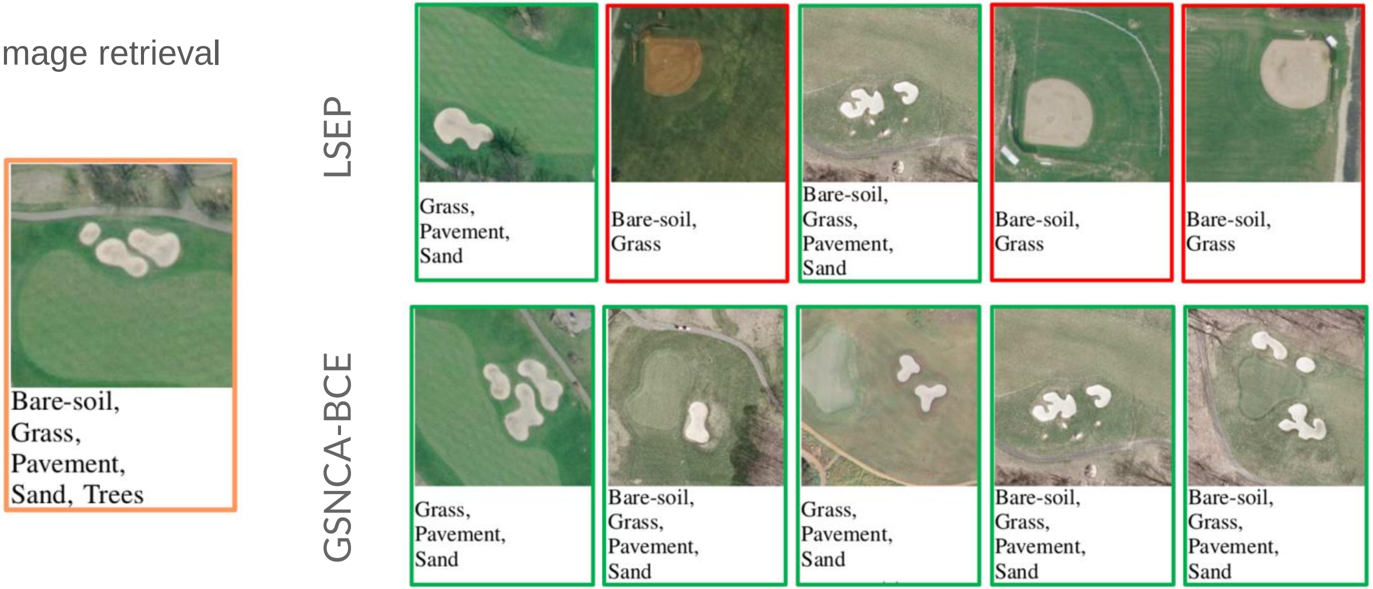
Experimental Results

- Image retrieval

		UCM			AID			DFC15		
		WMAP	MAP(%)	HL	WMAP	MAP(%)	HL	WMAP	MAP(%)	HL
ResNet18	Contrastive	1.97	86.77	0.19	3.66	93.31	0.18	–	–	–
	BCE	2.52	97.70	0.13	4.25	97.36	0.12	2.37	100.00	0.12
	GSNCA	2.63	99.17	0.11	4.35	99.17	0.11	2.43	100.00	0.10
	LSEP	**2.75**	**99.79**	**0.09**	4.39	99.06	0.11	2.40	100.00	0.11
	GSNCA-BCE	2.71	99.70	0.10	**4.47**	**99.29**	**0.09**	**2.51**	100.00	**0.07**
ResNet50	Contrastive	2.28	97.02	0.15	3.85	93.44	0.17	–	–	–
	BCE	2.64	98.99	0.11	4.33	98.31	0.11	2.45	100.00	0.09
	GSNCA	2.71	99.64	0.10	4.47	**99.67**	0.09	2.51	99.99	0.08
	LSEP	2.77	99.81	0.09	4.40	99.52	0.10	2.46	100.00	0.09
	GSNCA-BCE	**2.80**	**99.92**	**0.08**	**4.60**	99.66	**0.07**	**2.58**	100.00	**0.06**
WideResNet50	Contrastive	2.22	96.47	0.15	3.99	95.49	0.15	–	–	–
	BCE	2.62	99.37	0.11	4.39	98.93	0.10	2.45	**100.00**	0.09
	GSNCA	2.73	99.44	0.10	4.48	99.49	0.09	2.53	99.95	0.07
	LSEP	2.76	**99.87**	0.09	4.40	**99.75**	0.10	2.47	**100.00**	0.09
	GSNCA-BCE	**2.80**	**99.87**	**0.08**	**4.59**	**99.75**	**0.07**	**2.57**	99.99	**0.06**

Experimental Results

- Image retrieval



Conclusion

- Generalize the SNCA framework for multi-label RS image characterization

- Compared to the other state-of-the-art methods, the proposed GSNCA can achieve the superior performance on feature extraction for multi-label RS images

Thank you for your attention!

https://jiankang1991.github.io/