

## 第三届IEEE地质、测绘与遥感国际学术会议

# 基于深度学习技术的遥感图像检索

报告人：康健

苏州大学 电子信息学院

2022年4月24日



## 第三届IEEE地质、测绘与遥感国际学术会议

# 报告提纲

- 1 研究背景与思路
- 2 基于深度近邻成分分析的遥感图像检索
- 3 旋转不变性约束下的遥感图像检索
- 4 跨模态遥感图像检索
- 5 研究总结与未来展望



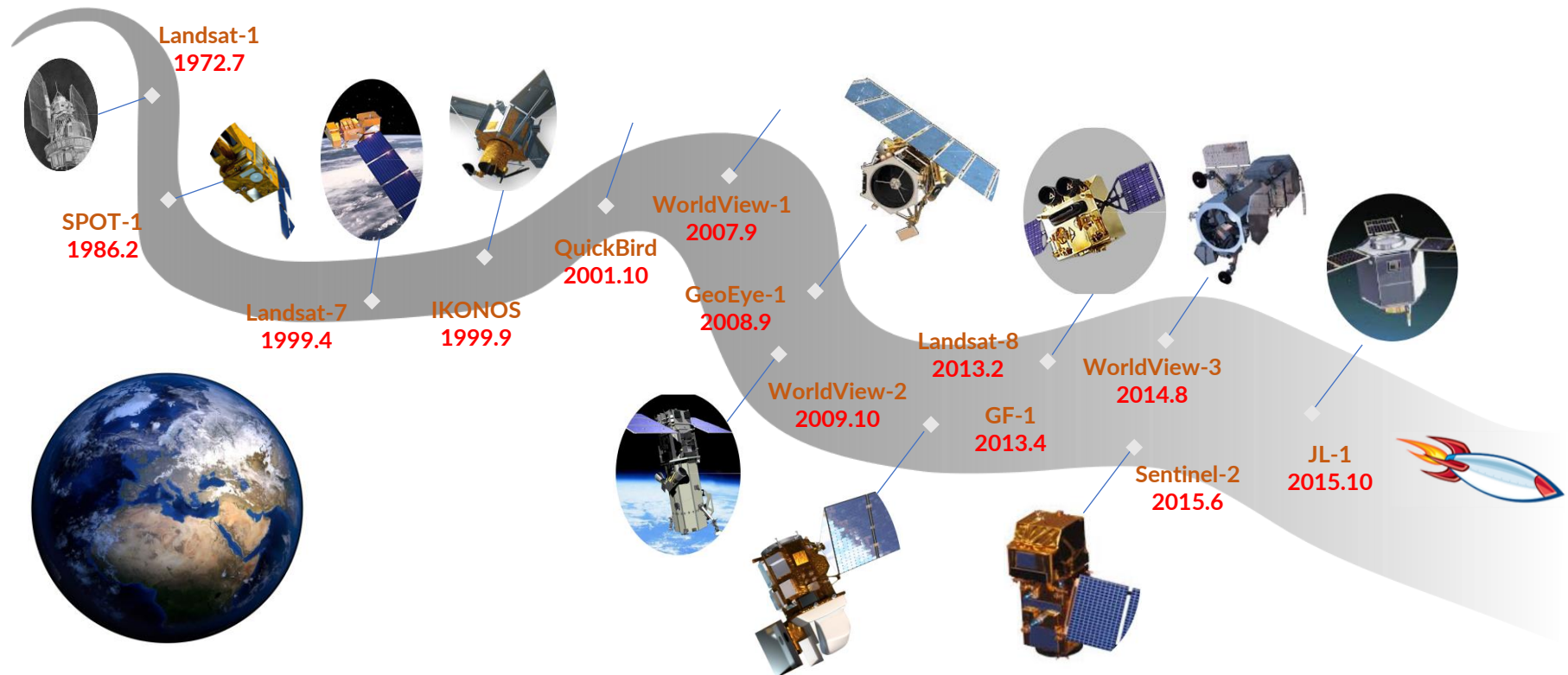
## 第三届IEEE地质、测绘与遥感国际学术会议

# 报告提纲

- 1 研究背景与思路
- 2 基于深度近邻成分分析的遥感图像检索
- 3 旋转不变性约束下的遥感图像检索
- 4 跨模态遥感图像检索
- 5 研究总结与未来展望

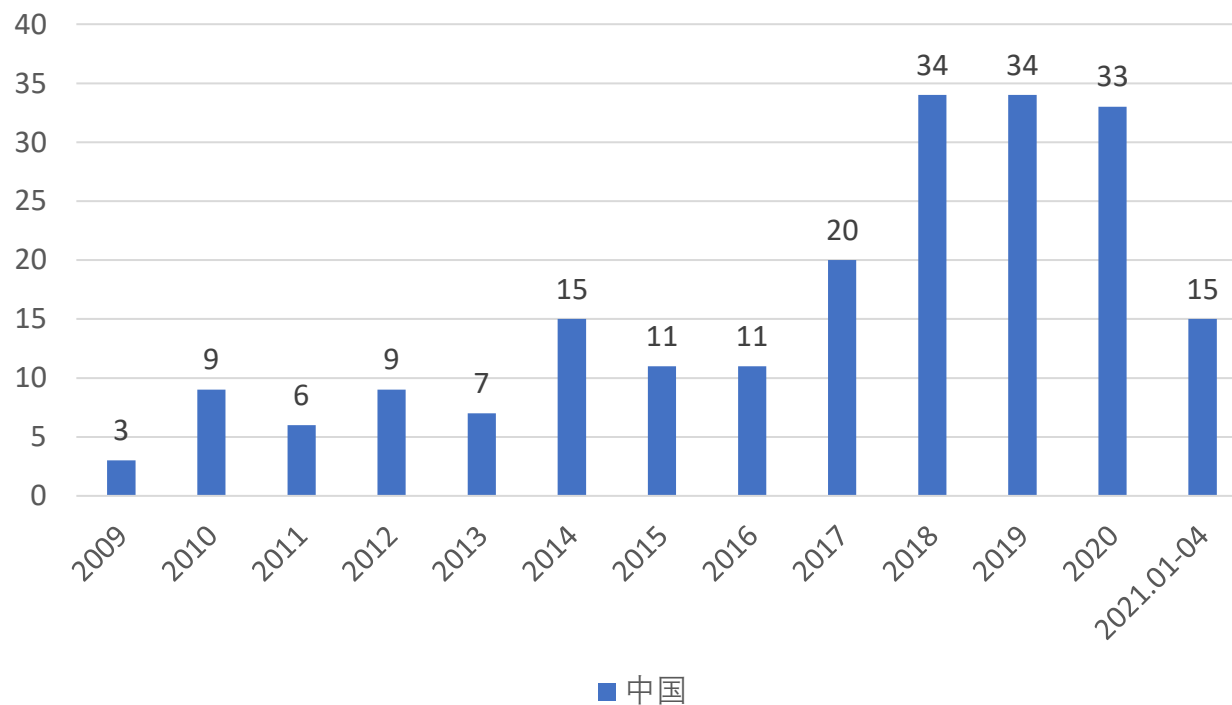


## 研究背景



## 研究背景

### 遥感卫星发射数目



遥感大数据时代加速到来

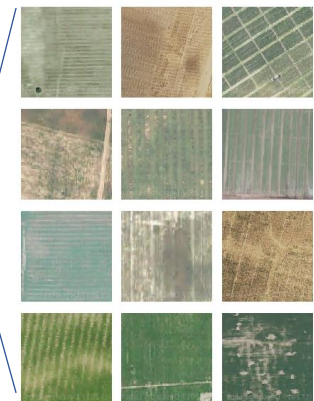


## 研究意义

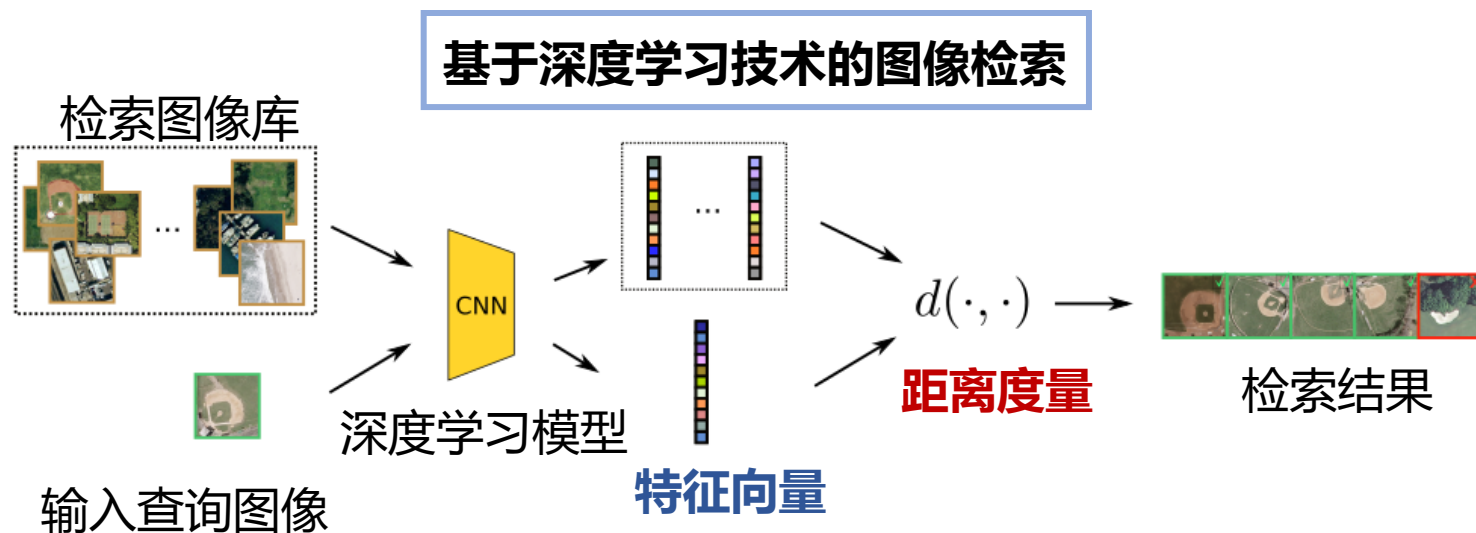
### 目标识别



### 国土资源管理



## 研究现状



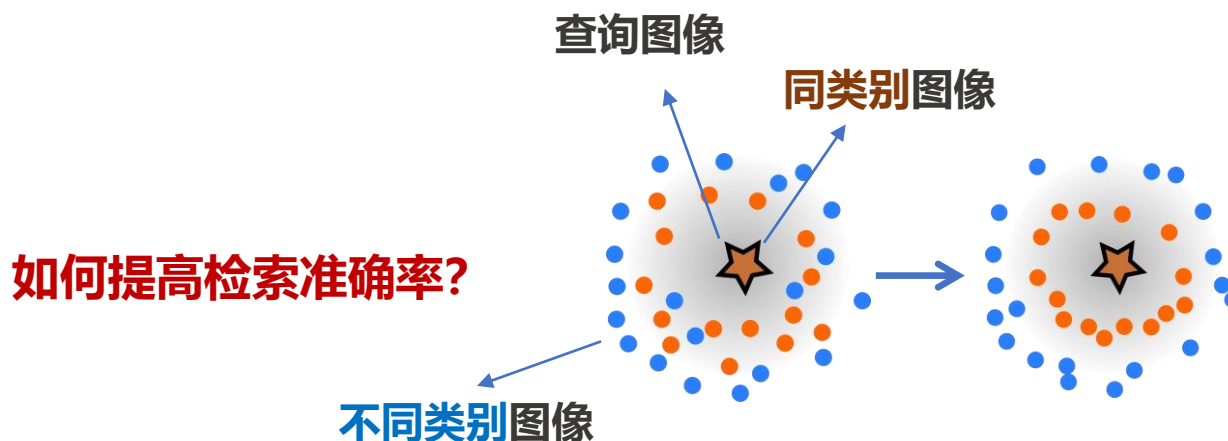
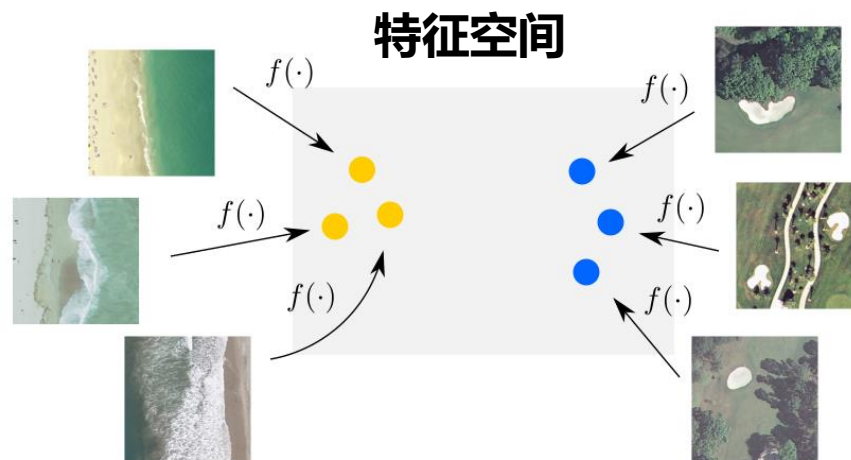
**特征向量提取：**设计新型网络结构提升图像高阶语义特征捕捉能力

**度量函数设计：**设计损失函数精确建模特征间的语义关系

## 研究现状

### 深度度量损失函数 $\mathcal{L}(\cdot)$

对于语义信息相似的图像，其特征在特征空间中距离近，若不相似，则距离远





## 研究现状

### 对比损失函数 [1]

$$\mathcal{L} = \sum_{i,j} l_{ij} \|\mathbf{f}_i - \mathbf{f}_j\|_2^2 + (1 - l_{ij})h(m - \|\mathbf{f}_i - \mathbf{f}_j\|_2)^2$$

$h(x) = \max(0, x)$  hinge loss

$l_{ij} = \begin{cases} 1, & \text{if } \mathbf{y}_i = \mathbf{y}_j, \\ 0, & \text{otherwise.} \end{cases}$  指示函数

$m$  margin 参数

**作用：**在特征空间中拉近同一类别的图像，以一定距离(margin)推开不同类别的图像。

### 三元组损失函数 [2]

$$\mathcal{L} = \sum_i h(\|\mathbf{f}_i^a - \mathbf{f}_i^p\|_2^2 - \|\mathbf{f}_i^a - \mathbf{f}_i^n\|_2^2 + m)$$

$(\mathbf{f}_i^a, \mathbf{f}_i^p, \mathbf{f}_i^n)$  基准、正例、负例图像特征三元组

$m$  margin 参数

**作用：**保证负例图像与基准图像在特征空间中的间距至少为正例图像与基准图像之间的距离加上  $m$

[1] Hadsell, R., Chopra, S., & LeCun, Y. (2006, June). Dimensionality reduction by learning an invariant mapping. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 2, pp. 1735-1742). IEEE.

[2] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

## 研究现状

### 归一化的Softmax损失函数<sup>[3]</sup>

$$\mathcal{L} = -\frac{1}{N} \sum_i \sum_c y_i^c \log\left(\frac{e^{\mathbf{w}_c^T \mathbf{f}_i}}{\sum_k e^{\mathbf{w}_k^T \mathbf{f}_i}}\right)$$

$\mathbf{w}_c$  类别基准特征(class prototype)

**作用：**在特征空间中，各个类别的所有图像特征与相应的类别基准特征(class prototype)对齐。

### 上述损失函数不足

- 传统对比损失函数通常在优化时只作用在每个批次内部中的各个图像，没有考虑到整个训练数据集中的其他图像。
- 对于三元组损失函数，需要有足够数量的图像对用来对CNN进行训练，但在大数据上很难满足这一要求。
- 归一化Softmax损失函数不能有效发掘类内图像特征之间的多样性结构。

[3] Zhai, A., & Wu, H. Y. (2018). Classification is a strong baseline for deep metric learning. arXiv preprint arXiv:1811.12649.

## 第三届IEEE地质、测绘与遥感国际学术会议

### 报告提纲

- 1 研究背景与思路
- 2 基于深度近邻成分分析的遥感图像检索
- 3 旋转不变性约束下的遥感图像检索
- 4 跨模态遥感图像检索
- 5 研究总结与未来展望



## 基于深度近邻成分分析的遥感图像检索

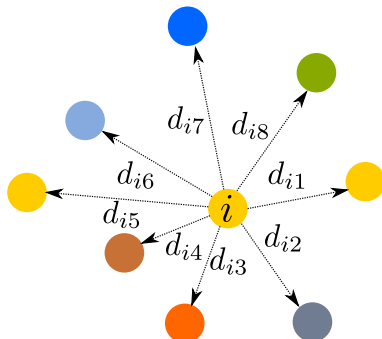
### 近邻成分分析(NCA)<sup>[4]</sup>

对于图像  $i$ , 其选择图像  $j$  作为其在特征空间的近邻的概率为:

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

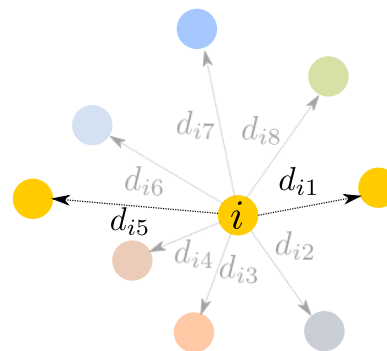
$\mathbf{A}$  特征投影线性算子

$d_{ij} = \|\mathbf{Ax}_i - \mathbf{Ax}_j\|_2^2$  特征间距



图像  $i$  被正确分类的概率为:  $p_i = \sum_{j \in \Omega_i} p_{ij}$

$j \in \Omega_i$  其他与图像  $i$  同类别的图像



近邻成分分析损失函数:

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

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

## 基于深度近邻成分分析的遥感图像检索

### 近邻成分分析(NCA)

**作用：**在特征空间中，  
最大化平均K邻近分类正确率

**优势：**善于发掘类内图像  
特征之间的多样性

### 基于深度学习的近邻成分分析(SNCA)<sup>[5]</sup>

#### 目标：

- 利用CNN模型学习非线性投影算子，增强图像语义建模能力
- 引入memory bank，用来存储所有训练图像特征，在对比学习过程中充分考虑所有图像对



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

$s_{ij}$   $\cos$ 相似性

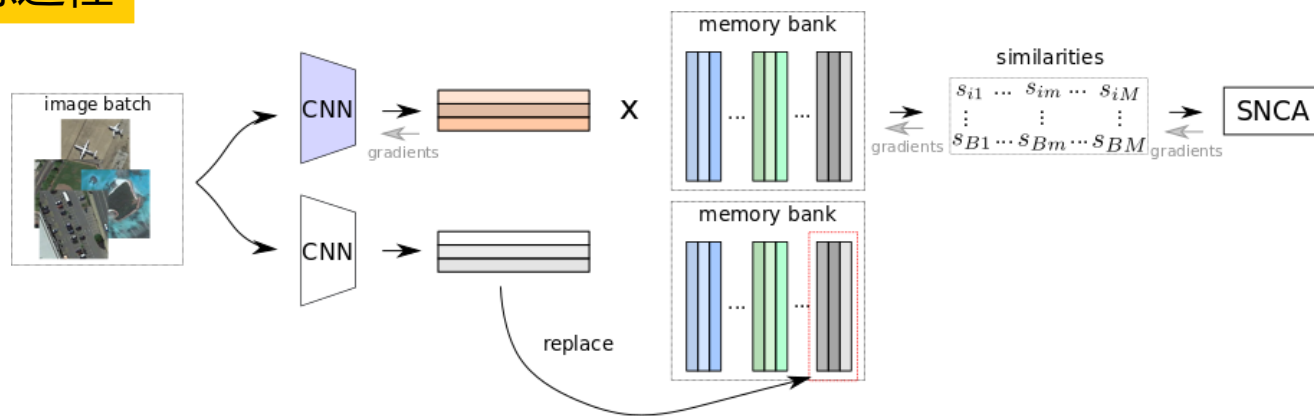
[5] Wu, Z., Efros, A. A., & Yu, S. X. (2018). Improving generalization via scalable neighborhood component analysis. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 685-701).

## 基于深度近邻成分分析的遥感图像检索

保证特征一致性：

基于动量更新机制的深度近邻成分分析<sup>[6,7]</sup>

### 训练过程



### 动量更新

$$\alpha \text{ CNN} + (1 - \alpha) \text{ CNN} \rightarrow \text{CNN}$$

[6] He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). Momentum contrast for unsupervised visual representation learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9729-9738).  
 [7] Kang, J., Fernandez-Beltran, R., Ye, Z., Tong, X., Ghamisi, P., & Plaza, A. (2020). Deep Metric Learning Based on Scalable Neighborhood Components for Remote Sensing Scene Characterization. IEEE Transactions on Geoscience and Remote Sensing.



## 基于深度近邻成分分析的遥感图像检索

提升特征分辨能力:

SNCA-CE 损失函数<sup>[7]</sup>

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \mathcal{L}_{\text{SNCA}}$$

提升类间距离

提升类内相似性度量能力

实验数据:

AID



viaduct



StorageTanks



Mountain



Commercial

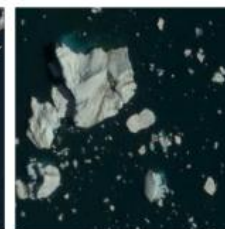
NWPU-RESISC45



wetland



ship



sea ice

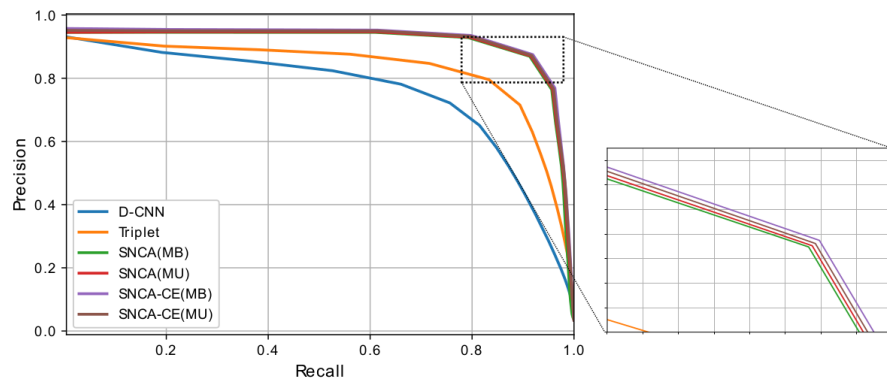


meadow

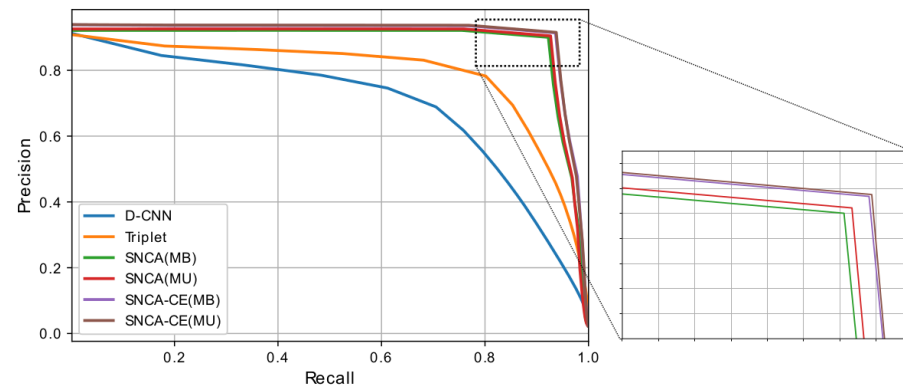
## 基于深度近邻成分分析的遥感图像检索

### 实验结果：图像检索

#### PR曲线



AID



NWPU-RESISC45

## 基于深度近邻成分分析的遥感图像检索

实验结果：图像检索



查询图像

Residential

1th

5th

10th

15th

20th

25th

30th

35th

40th



Residential

Residential

Residential

Residential

HerbaceousVegetation

Residential

Residential

Residential

Residential

对比方法

1th

5th

10th

15th

20th

25th

30th

35th

40th



Residential

Residential

Residential

Residential

Residential

Residential

Residential

Residential

Residential

所提方法

## 基于深度近邻成分分析的遥感图像检索

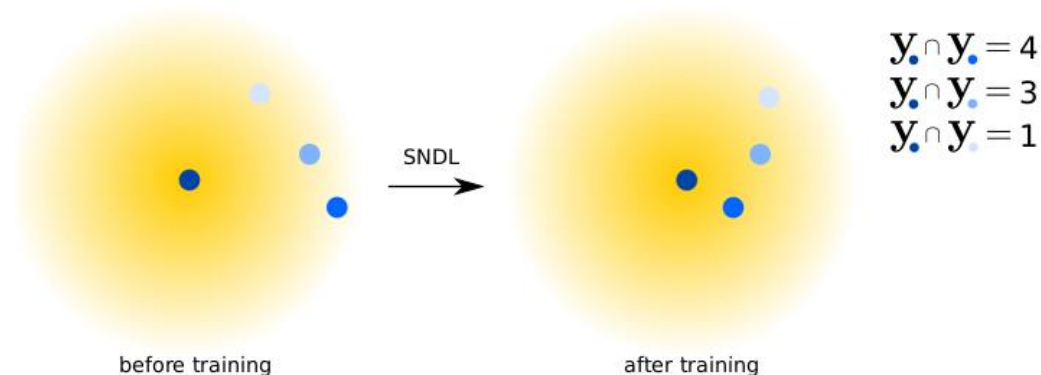
### 多标签近邻成分分析 [8]

#### 多标签遥感场景图像



裸地, 房屋, 汽车,  
树

**损失函数设计思想：**含有相同类别/标签的数量越多的两张图片，其在特征空间的距离越近。



[8] Kang, J., Fernandez-Beltran, R., Hong, D., Chanussot, J., & Plaza, A. (2020). Graph Relation Network: Modeling Relations Between Scenes for Multilabel Remote-Sensing Image Classification and Retrieval. IEEE Transactions on Geoscience and Remote Sensing.

## 基于深度近邻成分分析的遥感图像检索

### 多标签近邻成分分析损失函数<sup>[8]</sup>

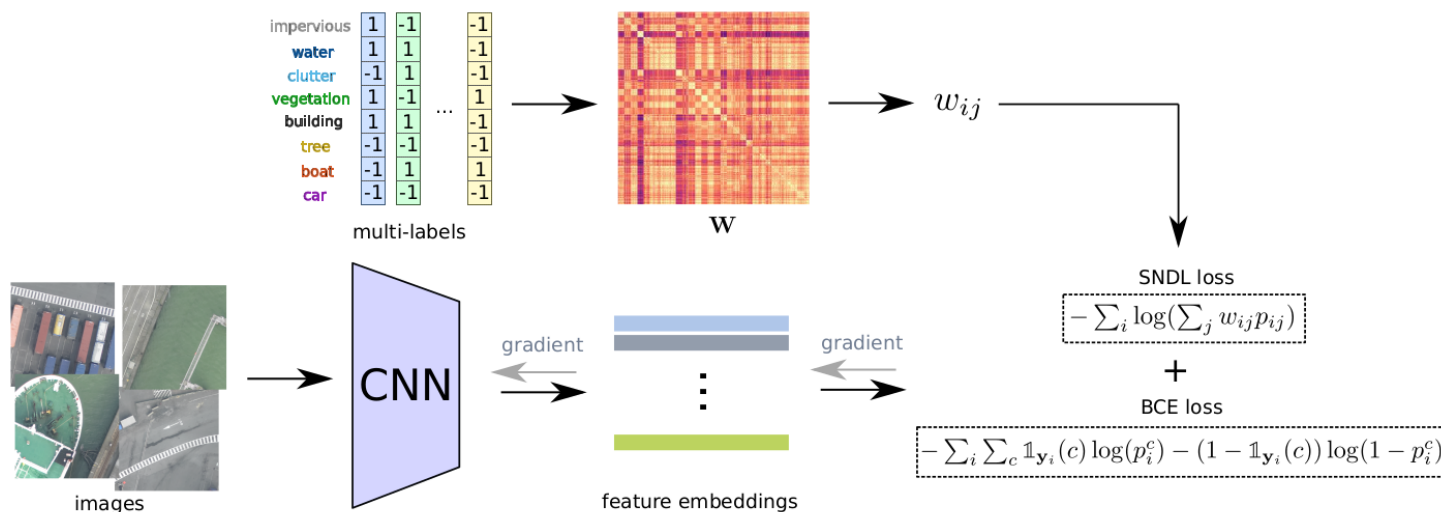
$$\mathcal{L}_{\text{SNDL}} = - \sum_i \log \left( \sum_j w_{ij} p_{ij} \right)$$

### 提升特征分辨能力:

### 联合损失函数

$$\mathcal{L} = \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{SNDL}}$$

### 方法框架:





# 基于深度近邻成分分析的遥感图像检索

## 实验结果:

		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
	GRN-SNDL	2.63	99.17	0.11	4.35	99.17	0.11	2.43	100.00	0.10
	LSEP	<b>2.75</b>	<b>99.79</b>	<b>0.09</b>	4.39	99.06	0.11	2.40	100.00	0.11
	GRN-SNDL-BCE	2.71	99.70	0.10	<b>4.47</b>	<b>99.29</b>	<b>0.09</b>	<b>2.51</b>	100.00	<b>0.07</b>
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
	GRN-SNDL	2.71	99.64	0.10	4.47	<b>99.67</b>	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
	GRN-SNDL-BCE	<b>2.80</b>	<b>99.92</b>	<b>0.08</b>	<b>4.60</b>	99.66	<b>0.07</b>	<b>2.58</b>	100.00	<b>0.06</b>
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	<b>100.00</b>	0.09
	GRN-SNDL	2.73	99.44	0.10	4.48	99.49	0.09	2.53	99.95	0.07
	LSEP	2.76	<b>99.87</b>	0.09	4.40	<b>99.75</b>	0.10	2.47	<b>100.00</b>	0.09
	GRN-SNDL-BCE	<b>2.80</b>	<b>99.87</b>	<b>0.08</b>	<b>4.59</b>	<b>99.75</b>	<b>0.07</b>	<b>2.57</b>	99.99	<b>0.06</b>



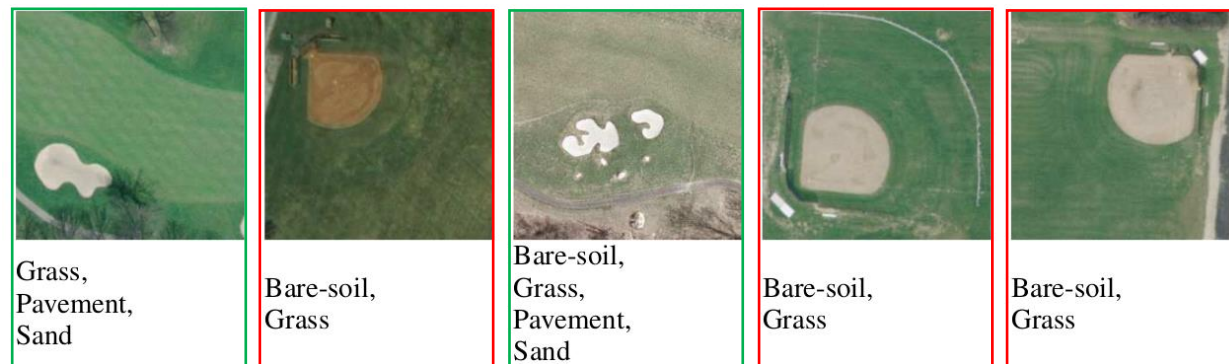
## 基于深度近邻成分分析的遥感图像检索

### 实验结果:

查询图像



LSEP[9]



所提方法



[9] Li, Y., Song, Y., & Luo, J. (2017). Improving pairwise ranking for multi-label image classification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3617-3625).

## 第三届IEEE地质、测绘与遥感国际学术会议

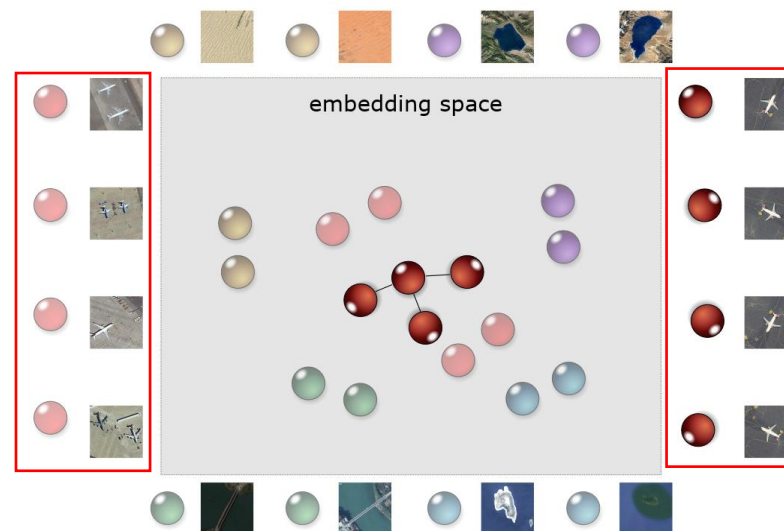
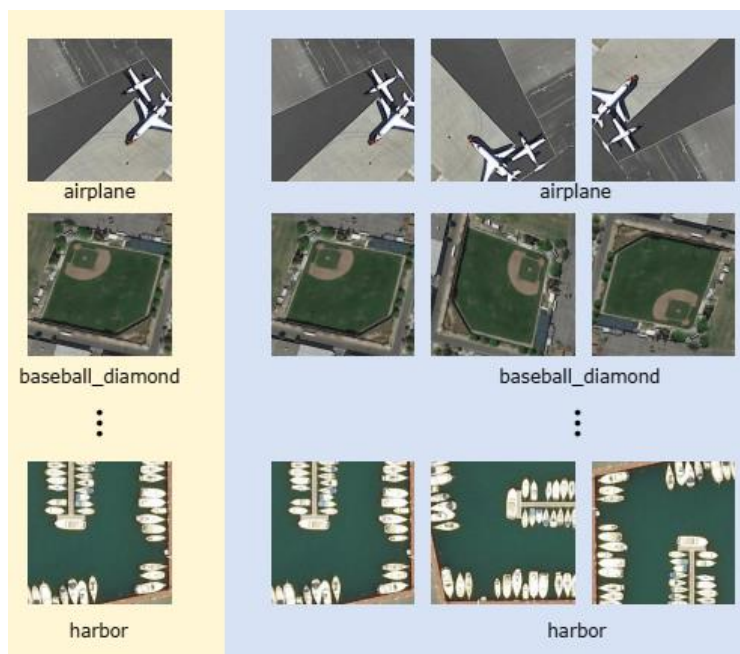
# 报告提纲

- 1 研究背景与思路
- 2 基于深度近邻成分分析的遥感图像检索
- 3 旋转不变性约束下的遥感图像检索
- 4 跨模态遥感图像检索
- 5 研究总结与未来展望



## 旋转不变性约束下的遥感图像检索

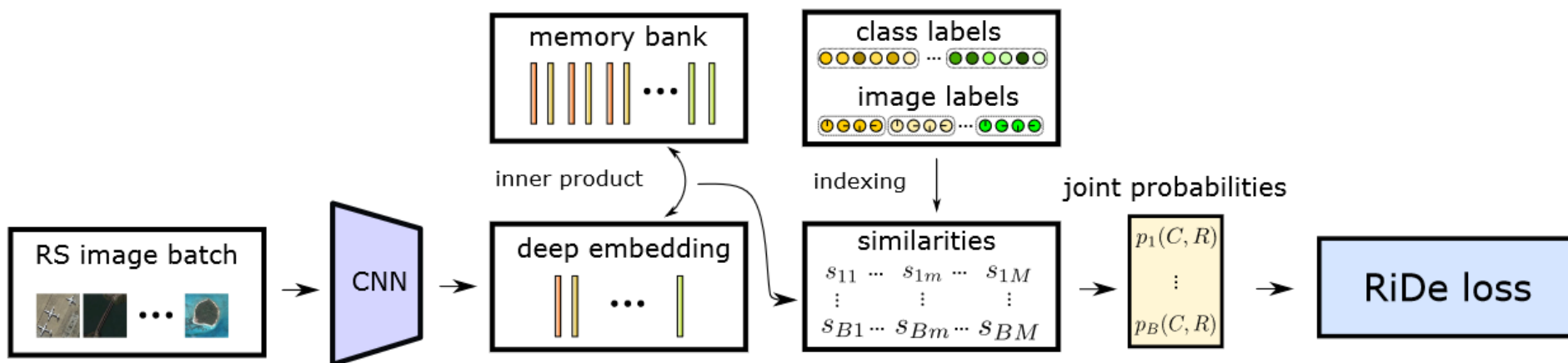
特征学习需要考虑遥感  
图像特有的**旋转不变性**



查询图片与其**不同角度**旋转图  
片的**相似性**比其他同类图片**高**

## 旋转不变性约束下的遥感图像检索

### 旋转不变的深度度量学习<sup>[10]</sup>



### 度量损失函数

$$L = - \sum_i \log(p_i^C) - \sum_i \log(p_i^R)$$

$$= - \sum_i \log \left( \sum_{j \in \Omega_i} \frac{\exp(\mathbf{f}_i^T \mathbf{f}_j / \sigma)}{\sum_{k \neq i} \exp(\mathbf{f}_i^T \mathbf{f}_k / \sigma)} \right) - \sum_i \log \left( \sum_{j \in \tilde{\mathcal{R}}_i} \frac{\exp(\mathbf{f}_i^T \mathbf{f}_j / \sigma)}{\sum_{k \neq i} \exp(\mathbf{f}_i^T \mathbf{f}_k / \sigma)} \right)$$

class discrimination      rotational invariance

类别区分项

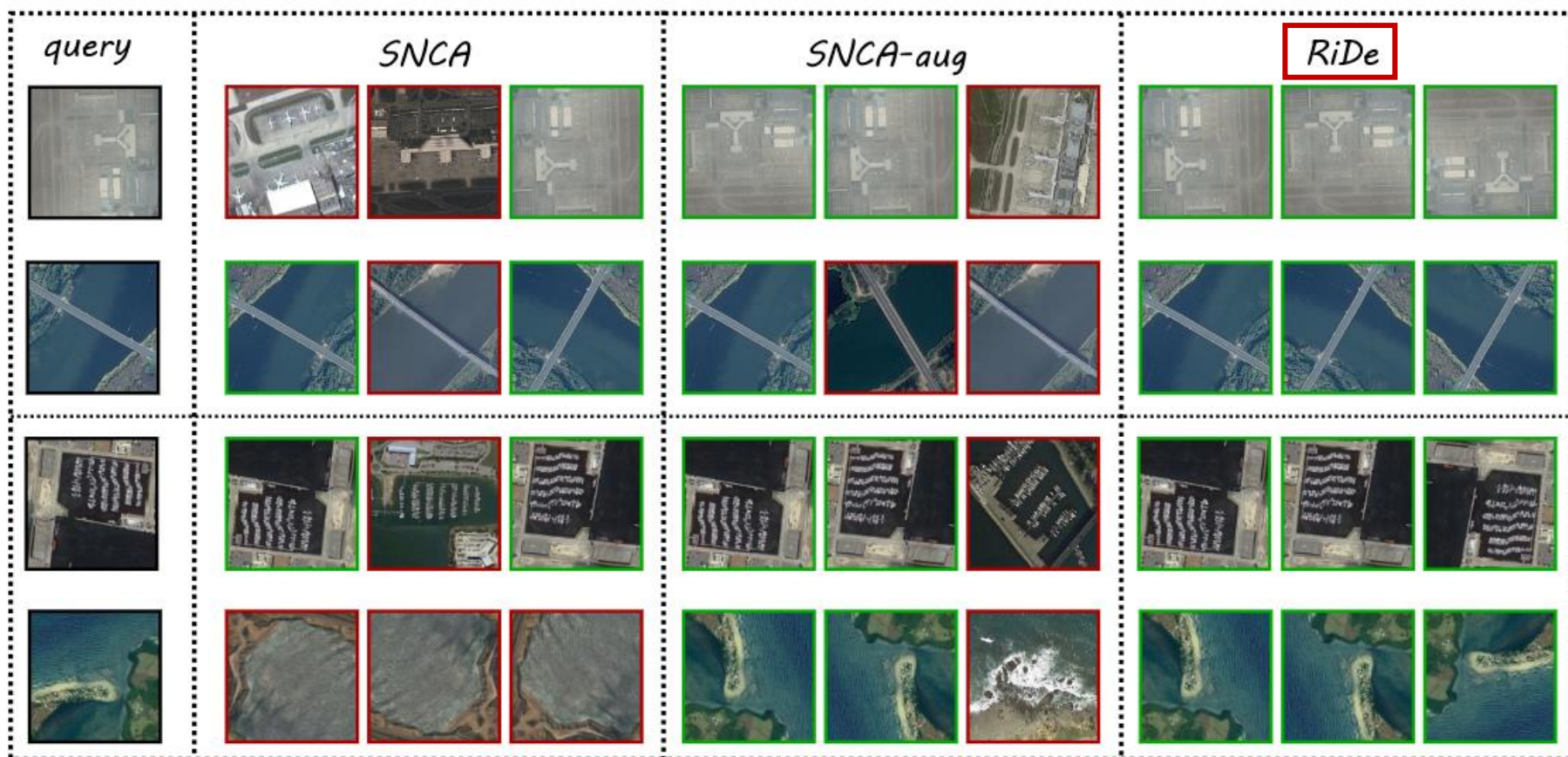
旋转不变约束

[10] **Kang, J.**, Fernandez-Beltran, R., Wang, Z., Sun, X., Ni, J., & Plaza, A. (2021). Rotation-Invariant deep embedding for remote sensing images. IEEE Transactions on Geoscience and Remote Sensing.



## 旋转不变性约束下的遥感图像检索

### 实验结果:



## 第三届IEEE地质、测绘与遥感国际学术会议

### 报告提纲

- 1 研究背景与思路
- 2 基于深度近邻成分分析的遥感图像检索
- 3 旋转不变性约束下的遥感图像检索
- 4 跨模态遥感图像检索**
- 5 研究总结与未来展望



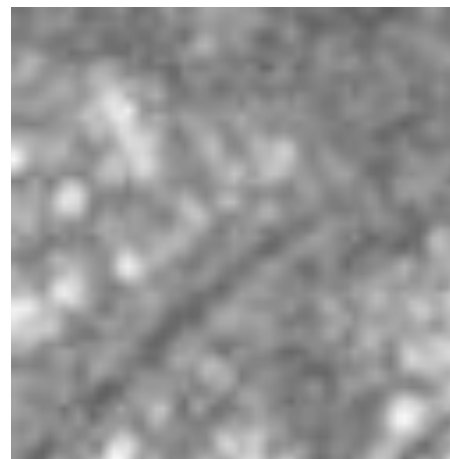


## 跨模态遥感图像检索

多模态遥感图像能更加全面地反映地物与**不同波段电磁波**相互作用的结果



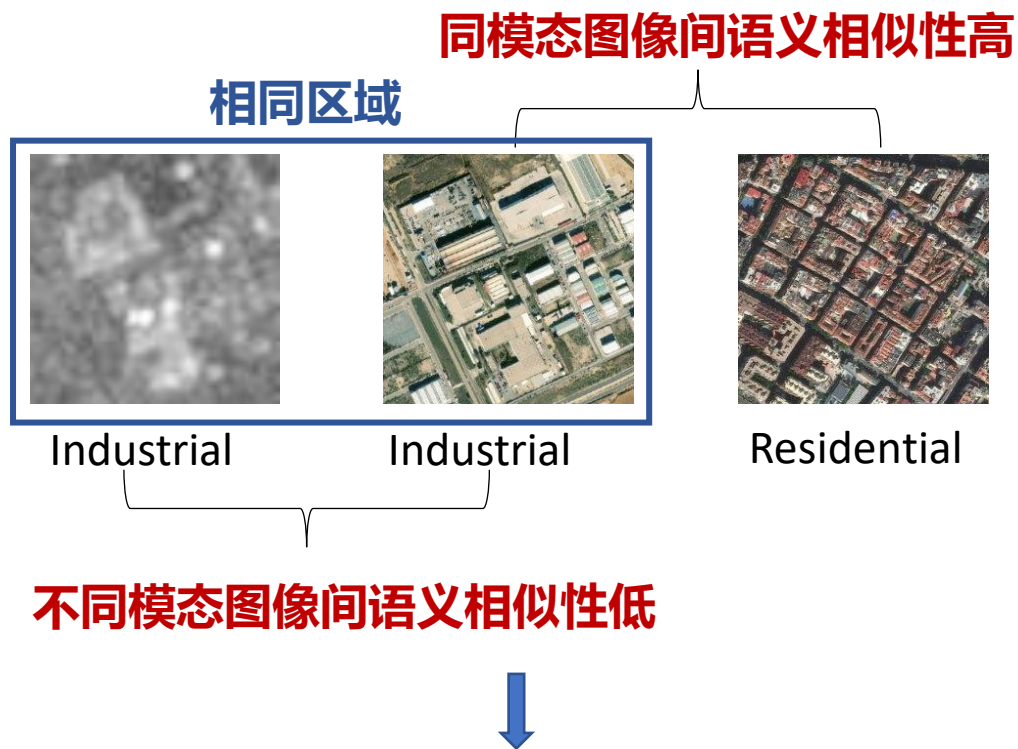
光学遥感图像



合成孔径雷达 (SAR) 图像

**图像检索方法需要满足不同模态之间的相互检索需求**

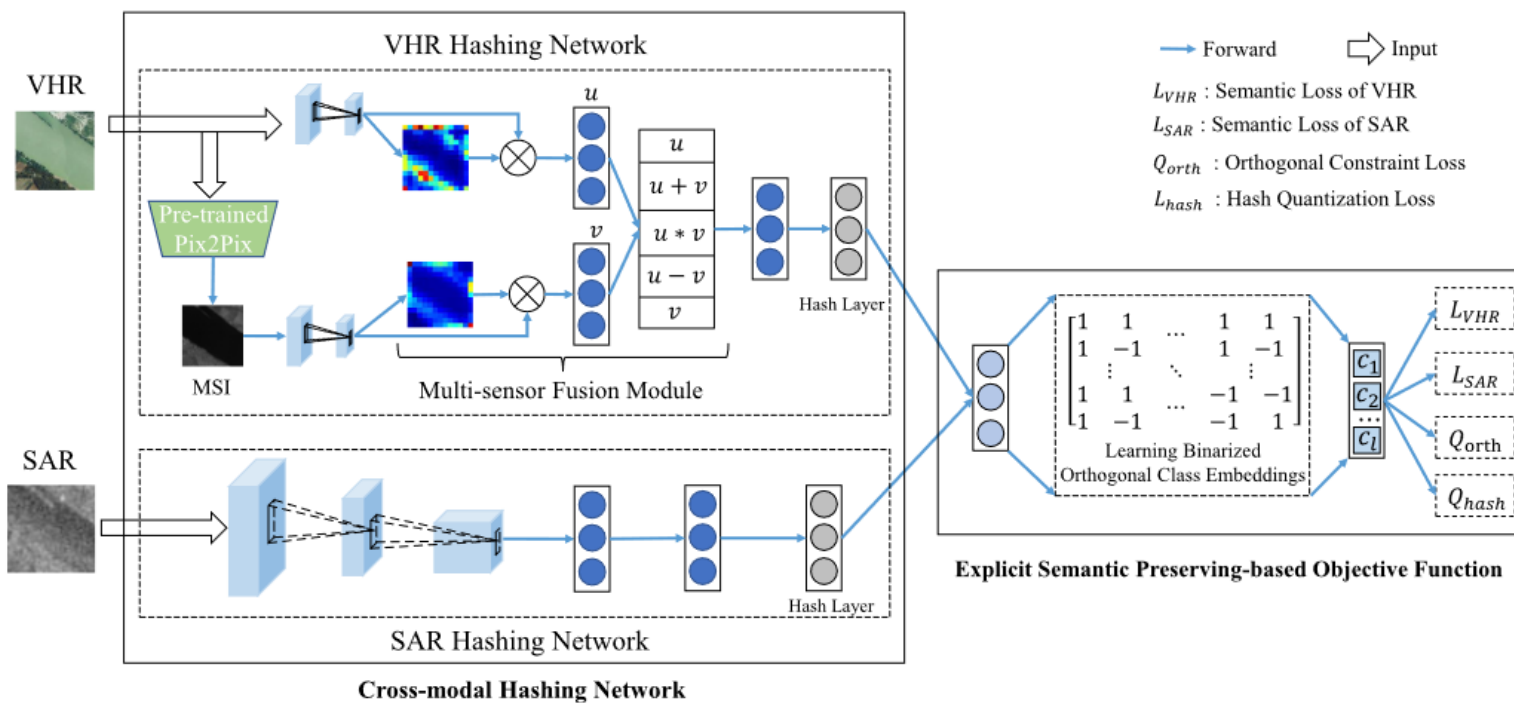
## 跨模态遥感图像检索



具有类别判别性的特征学习同时需要考虑如何减小不同模态数据之间的语义差异性

## 跨模态遥感图像检索

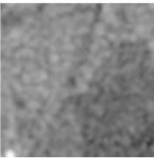






















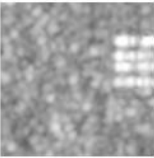




































































### 跨模态的深度哈希编码<sup>[11]</sup>



[11] Sun, Y., Feng, S., Ye, Y., Li, X., **Kang, J.**, Huang, Z., & Luo, C. (2021). Multi-sensor Fusion and Explicit Semantic Preserving-based Deep Hashing for Cross-Modal Remote Sensing Image Retrieval. IEEE Transactions on Geoscience and Remote Sensing.

# 跨模态遥感图像检索

## 实验结果:

Query	Top 10 Retrieval RS images											
SAR→VHR Highway 												Ours
												DSMHN
Residential 												Ours
												DSMHN
VHR→SAR Highway 												Ours
												DSMHN
Residential 												Ours
												DSMHN



## 第三届IEEE地质、测绘与遥感国际学术会议

### 报告提纲

- 1 研究背景与思路
- 2 基于深度近邻成分分析的遥感图像检索
- 3 旋转不变性约束下的遥感图像检索
- 4 跨模态遥感图像检索
- 5 研究总结与未来展望



## 研究总结与未来展望

### 研究总结

- 提出了**单标签到多标签**下的遥感图像检索技术，提升了检索精度。
- 考虑到遥感图像的旋转不变性，提出了**旋转不变性约束**下的检索方法。
- 解决了不同模态图像之间语义差异性大的问题，提出了**跨模态**遥感图像检索方法。

### 未来展望

- 半监督、小样本下的图像检索
- 遥感图像时间序列下的检索
- 语言描述与图像之间的相互检索



# 谢谢大家！ 敬请批评指正！



个人主页

