

# CNN—Based Change Detection

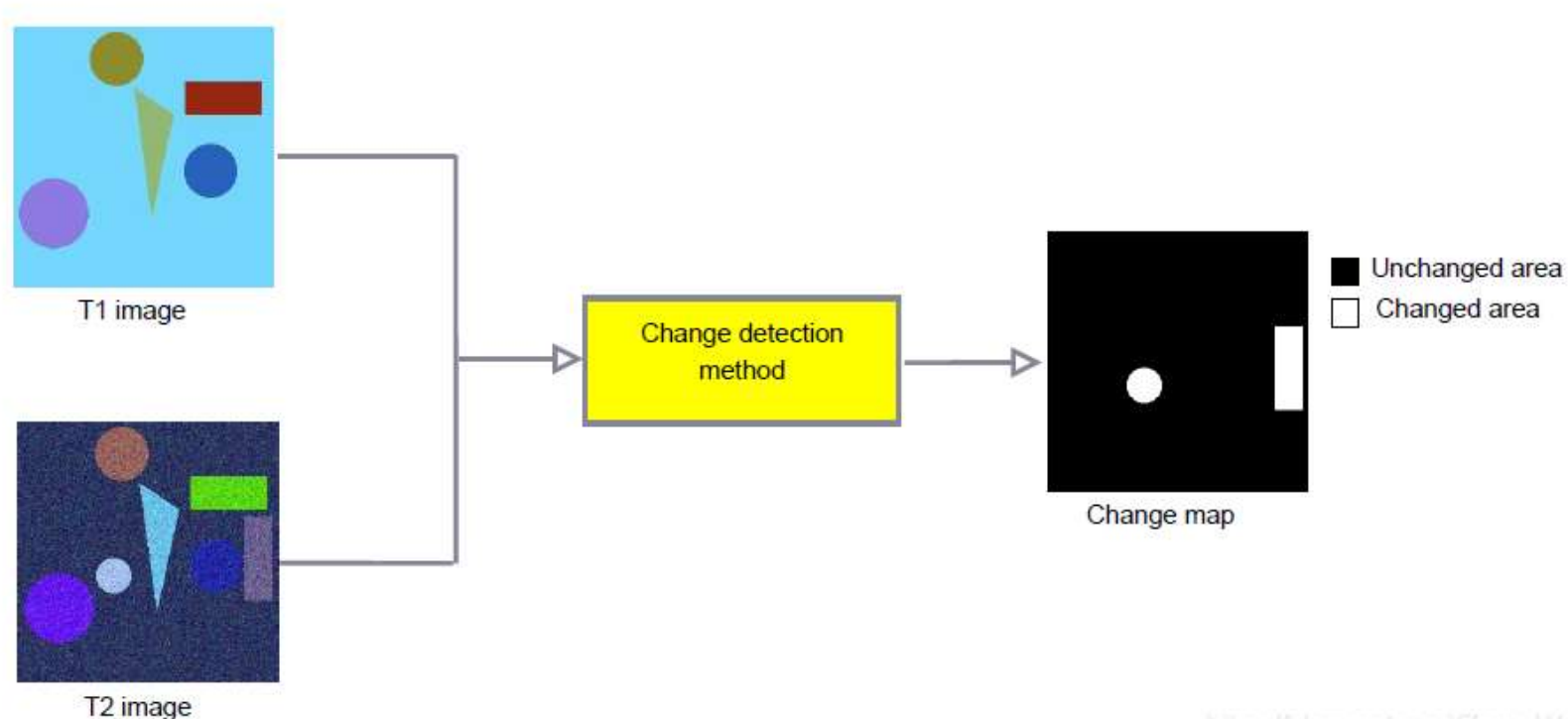
汇报人：宋雨菲

# 目录

1. Background
2. Changer: Feature Interaction is What You Need for Change Detection
3. Change is Everywhere: Single-Temporal Supervised Object Change Detection in Remote Sensing Imagery
4. Seeing Beyond the Patch: Scale-Adaptive Semantic Segmentation of High-resolution Remote Sensing Imagery based on Reinforcement Learning

# Background

变化检测任务流程： 两个不同时间的图像找变化区域，并标识变化成什么？

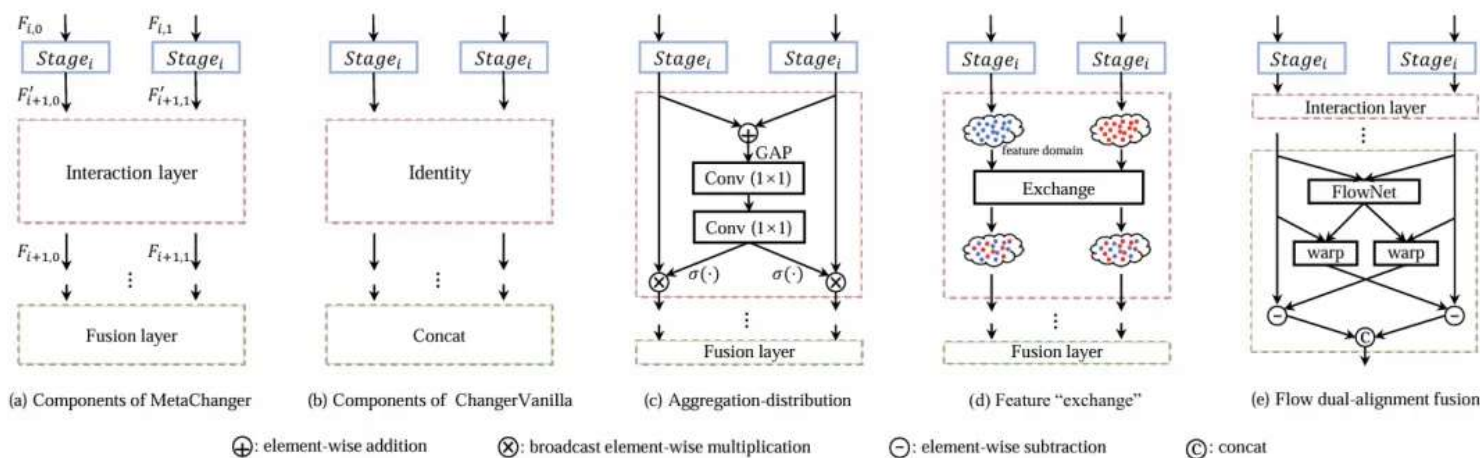


# Changer: Feature Interaction is What You Need for Change Detection

在进行变化监测时我们该注意哪些方面？

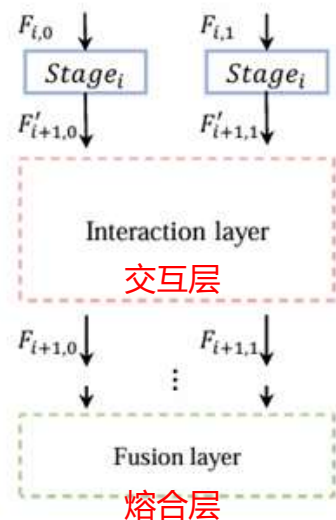
功能交互是更改检测所需注意的 **交换是有用的**

**研究目标：**旨在强调特征提取过程中**特征交互对变化检测的影响**



1. 本文提出了**MetaChanger**来探索特征交互在变化检测中的影响。为了验证特征交互的有效性，将**MetaChanger**的交互层规定为极其简单的AD和特征“交换”。
2. 研究发现，派生的**ChangerAD**和**ChangerEx**在多个变化检测数据集上均能取得竞争性的性能。广泛的消融研究表明**ChangerEx**的稳健性和可扩展性。

# MetaChanger



(a) Components of MetaChanger

可概述为:

$$Y = H(D(E_{InterAct}(X_1))D(E_{InterAct}(X_2)))$$

编码器Encoder

$$F'_{i+1,j} = \text{Stage}_i(F_{i,j}) \quad \forall j$$

$$F_{i+1,j} = \text{InterAct}(F'_{i+1,0}, F'_{i+1,1}) \quad \forall j$$

$$\text{Linear}_{(C_{in}, C_{out})}$$

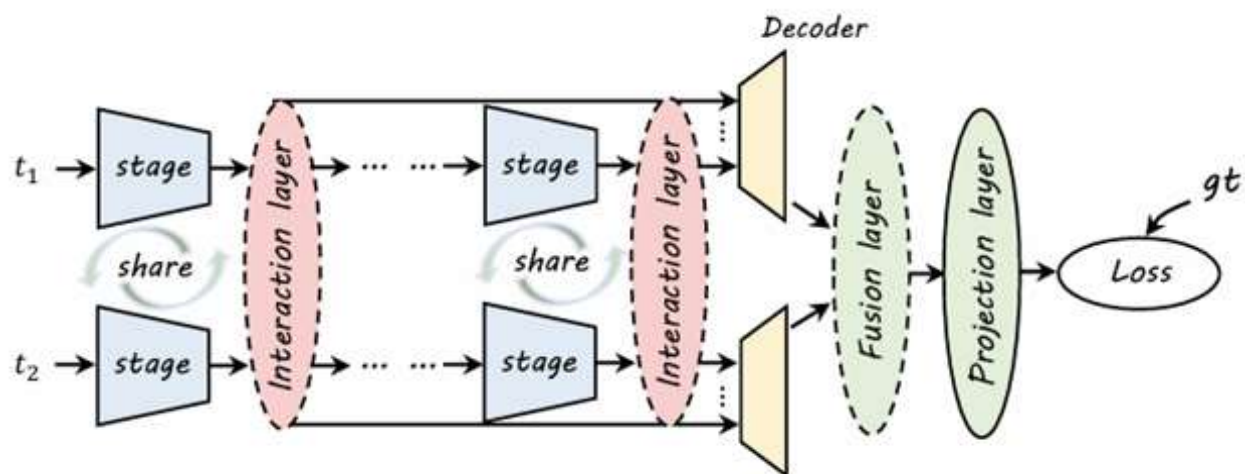
解码器Decoder

$$F_{i,j} = \text{Upsample}(\text{Linear}_{i(C_i, C)}(F_{i,j})) \quad \forall i, j$$

$$\hat{F}_j = \text{Linear}_{(4C, C)}(\text{Concat}(F_{i,j})) \quad \forall i$$

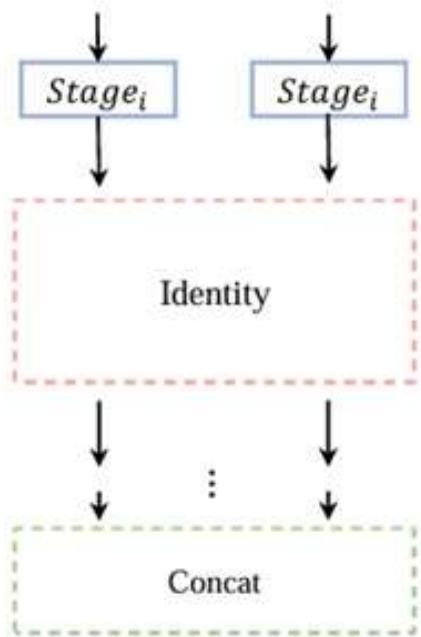
$$\hat{F} = \text{Fuse}(\hat{F}_0, \hat{F}_1)$$

$$Y = \text{Project}(\hat{F})$$



# ChangerVanilla-直接映射

为了更好地进行比较，构建了一个基准模型ChangerVanilla。ChangerVanilla没有交互层，并使用简单的连接(concat)操作作为融合层。



(b) Components of ChangerVanilla

$$InterAct_{vanilla}(x_i) = Identity(x_i)$$

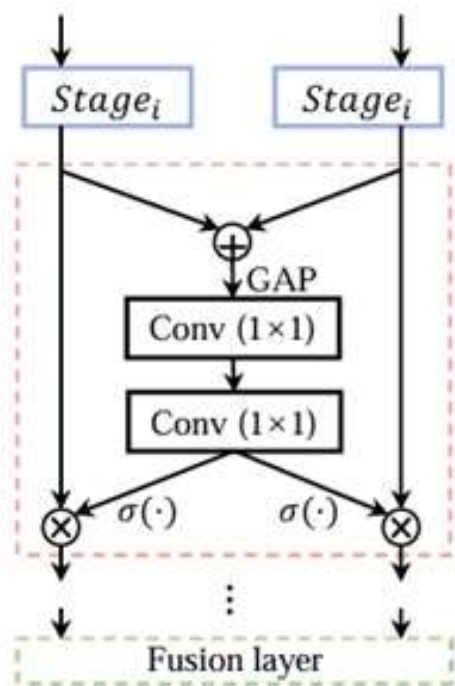
$$Fuse_{vanilla}(x_0, x_1) = Concat([x_0, x_1])$$

无参数的交互操作，也能有效提高变化检测模型的性能

# ChangerAD-特征聚合

AD特征交互层：[aggregation-distribution](#)。先进行元素加，再全局平均池化，两层MLP（第一层压缩，第二层扩张通道）。将结果进行sigmoid操作与原特征操作，相当于赋予注意力权重。

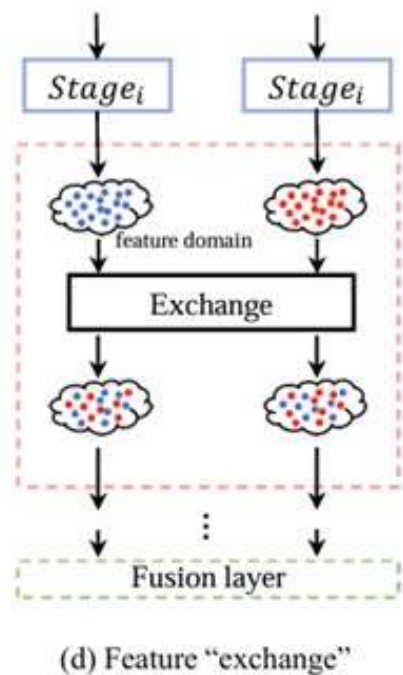
非线性激活函数



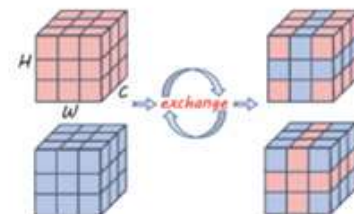
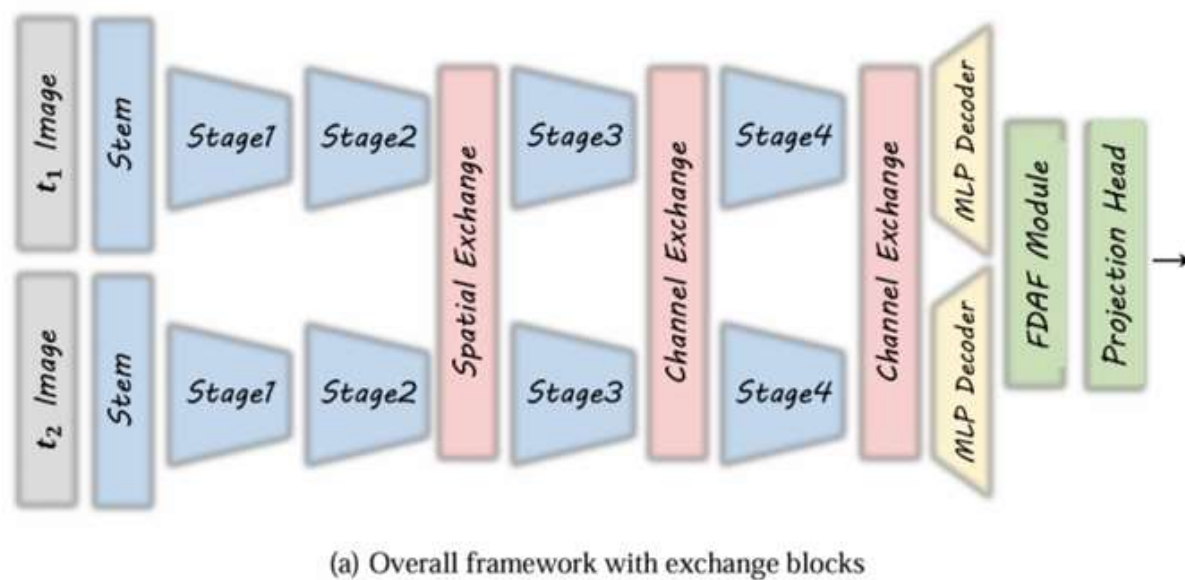
(c) Aggregation-distribution

$$InterAct_{AD}(x_i) = x_i \cdot \sigma(\hat{x}_i)$$
$$\hat{x} = MLP_{(C_i, 2C_i)}(GAP(x_0 + x_1))$$

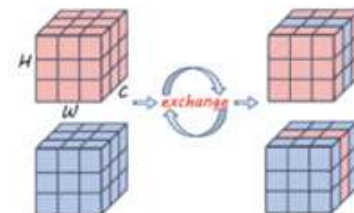
# ChangerEx-特征交换



$$\text{掩膜生成方法 } x_{0/1}(n, c, h, w) = \begin{cases} x_{0/1}(n, c, h, w), & M(n, c, h, w) = 0 \\ x_{1/0}(n, c, h, w), & M(n, c, h, w) = 1 \end{cases}$$



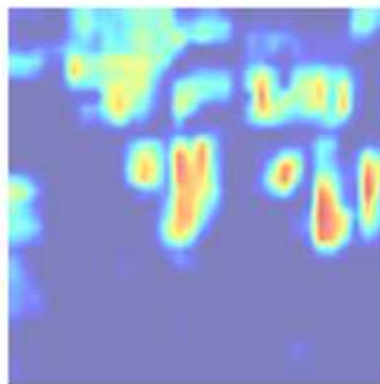
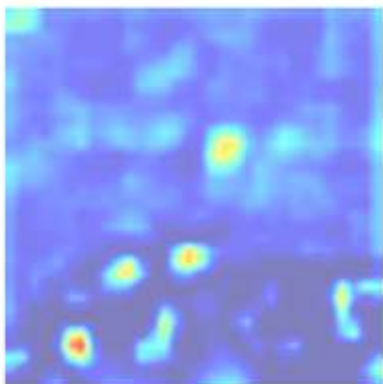
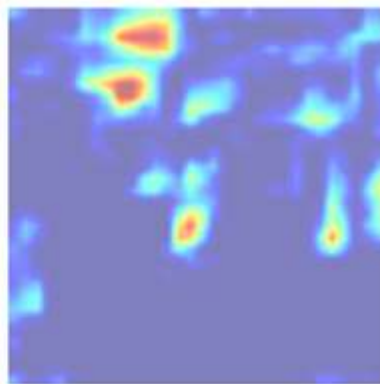
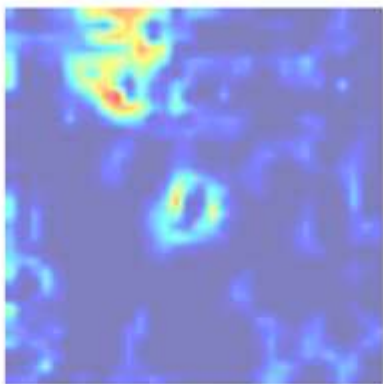
(b) Spatial exchange 空间交换



(c) Channel exchange 通道交换



# 实验热力图



(a)  $t_1/t_2$  Image

$t_2$ 时刻建筑消失

(b) w/o exchange

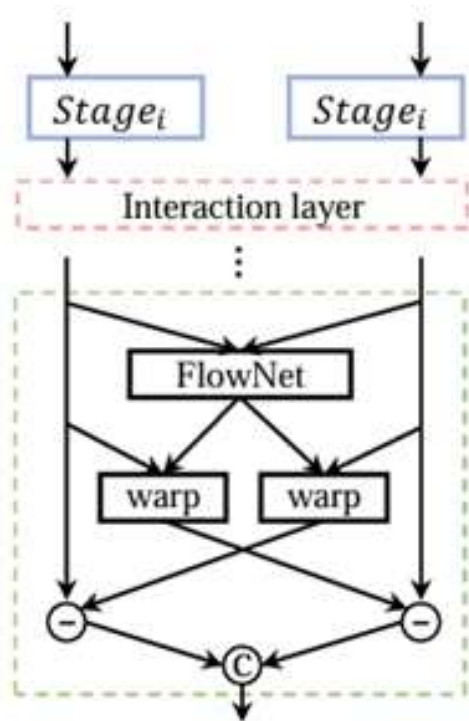
不进行特征交换

(c) w/ exchange

进行特征交换

特征交换更能建立起双时态图像之间的差异性

# FlowNet 双流对齐融合



(e) Flow dual-alignment fusion

$$x = Concat([x_0(p + \Delta p_0) - x_1, x_1(p + \Delta p_1) - x_0])$$

$$\Delta p = FlowNet(Concat([x_0, x_1]))$$

**Concat层：**经常用于将特征联合，多个卷积特征提取框架提取的特征融合或者是将输出层的信息进行融合。

**目的：**对齐双时态图像中的位置偏差

# Change is Everywhere: Single-Temporal Supervised Object Change Detection in Remote Sensing Imagery

传统双时相变化检测图像

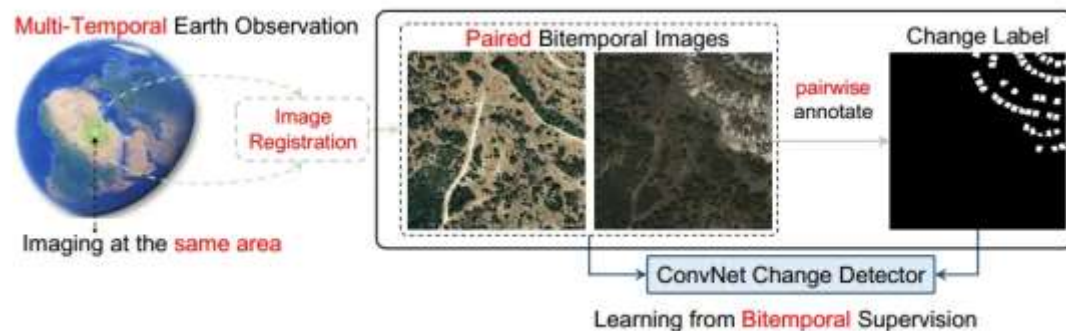
## 遥感图像中单时间监督目标变化检测

**研究目标：**旨在通过单时间监督学习的方法，利用已标记的遥感图像数据，实现在遥感图像中的目标变化检测。使用未配对的图像作为学习样本进行学习，能节约经费，学习效果会更好。

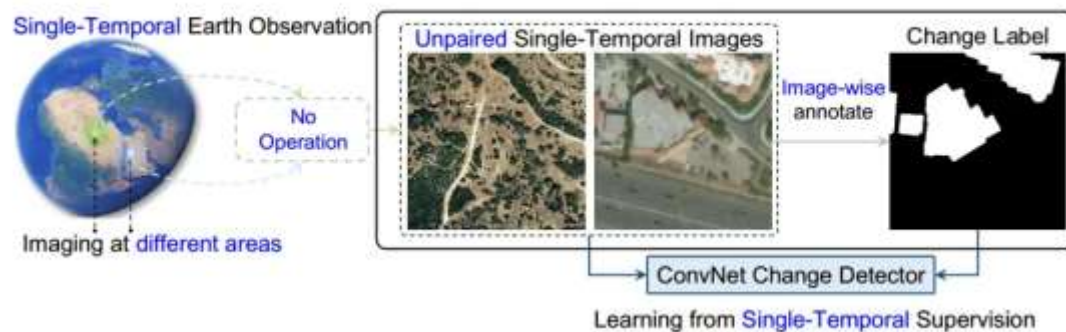
**论文创新点：**采用单时间监督学习方法，避免了需要成对图像的问题，简化了数据准备过程。

1. 针对成对双时相变化检测训练样本标注耗时、收集困难的问题，提出了一种新颖的弱监督学习算法STAR，其利用非成对单时相遥感影像构造伪双时相监督信号，以学习变化表征。

2. 同时提出了一种变化检测新架构ChangeStar，其通过核心模块ChangeMixer，可将任意语义分割模型转化为变化检测器，从而复用现有的优秀架构，揭示了语义分割与变化检测之间的内在联系。

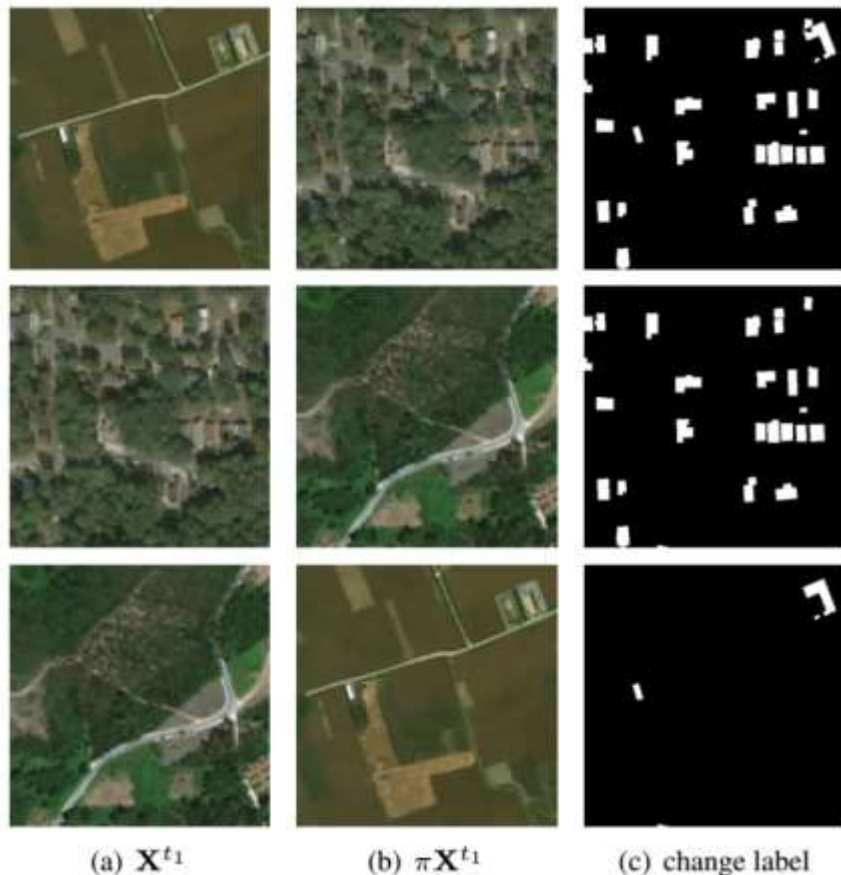


(a) Conventional Bitemporal Supervised Learning



伪变化检测图像

# 如何生成配对图像？



STAR方法:

STAR方法的关键思想是通过构造伪双时间图像对，利用单时间态数据进行变化监控信号的学习。

原始变化检测的思考

$$\min_{\theta} L(F_{\theta}(X_{t1}, X_{t2}), Y_{t1 \rightarrow t2})$$

传统双时相变化检测图像对

可以发现  $Y^{t1 \rightarrow t2}$  是唯一的监控信号来源。为了获得  $Y^{t1 \rightarrow t2}$ ，通常需要配对的语义信息来定义正样本和负样本。然而，成对的语义信息只与双时像素的语义有关，与它们的空间位置无关。相同的空间位置仅用于保证训练与推理之间的一致性。如果放宽这个条件，那么在等式中的原始问题1可以简化为以下形式，同时扩大了X的选择范围，即语义分割信息的变化与目标空间位置信息变化无关

$$\min_{\theta} L(F_{\theta}(X_i, X_j), \text{compare}(Y_i, Y_j))$$

为了提供单时间数据的变化监控信号，首先构造并利用伪双时间图像对，双时间数据在原始学习问题可以用单时态数据代替，因此学习问题可以重新表述为：

$$\min_{\theta} L(F_{\theta}(X_{t1}, \pi X_{t1}), Y_{t1} \oplus \pi Y_{t1})$$

伪变化检测图像对



# 异或 XOR

异或运算的具体逻辑为：

- 1、在伪双时态图像上有两个对象实例重叠，则将重叠区域被指定为**负样本**；
- 2、对象实例不重叠，则对象实例被设定为**正样本**；
- 3、其余的像素位置被分配为**负样本**。

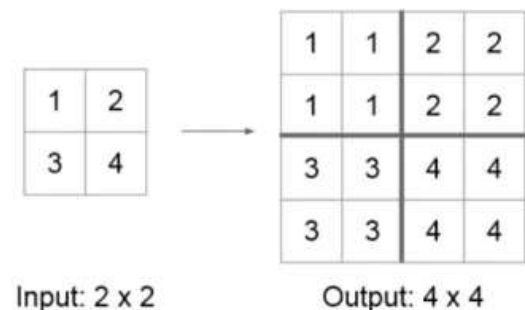
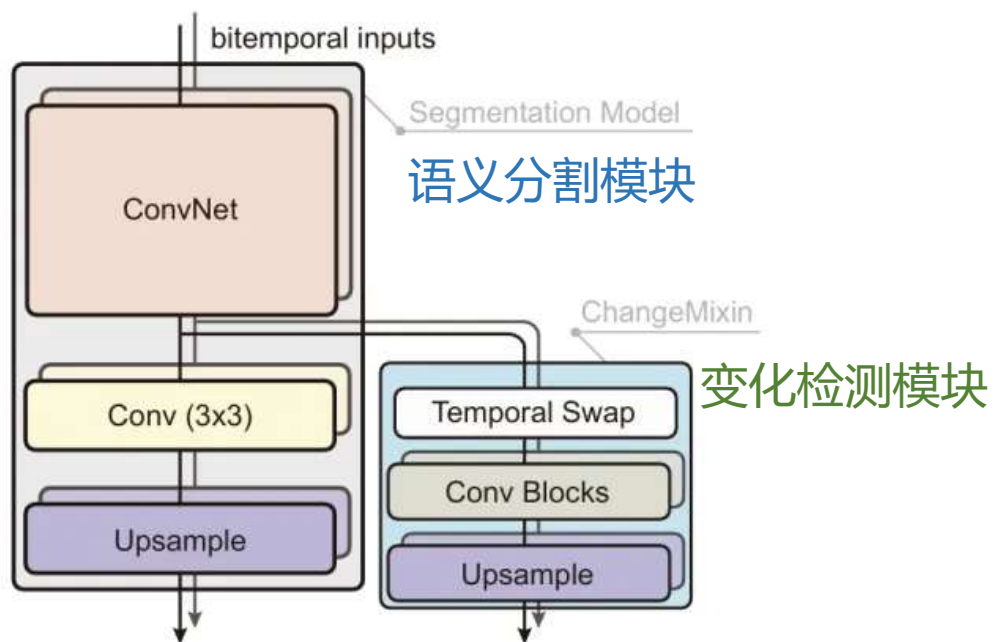
这种方法的优势在于不需要成对的标记图像，只需要利用单时间态数据，就能够获得大量训练集，来一以此训练高精度的变化检测器。



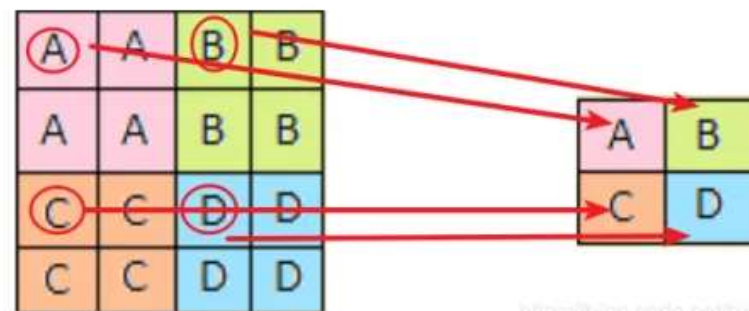
# 网络架构

网络架构：由语义分割模块和ChangeMixin模块组成，其中语义分割模块负责目标分割，ChangeMixin模块负责目标变化检测。

**ChangeMixin模块（变化检测模块）**：包含时间交换模块和多个卷积层，用于实现目标变化检测



向上采样

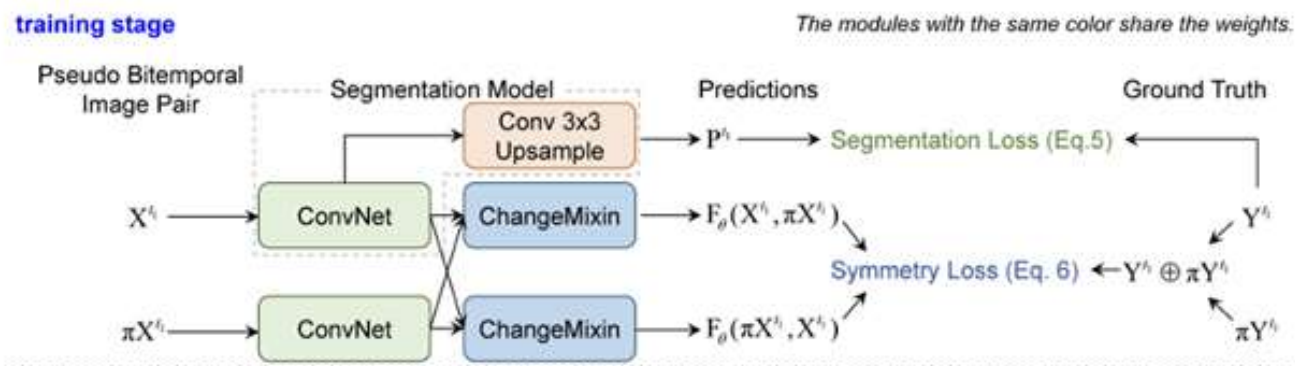


<https://blog.csdn.net/huaxi01/>

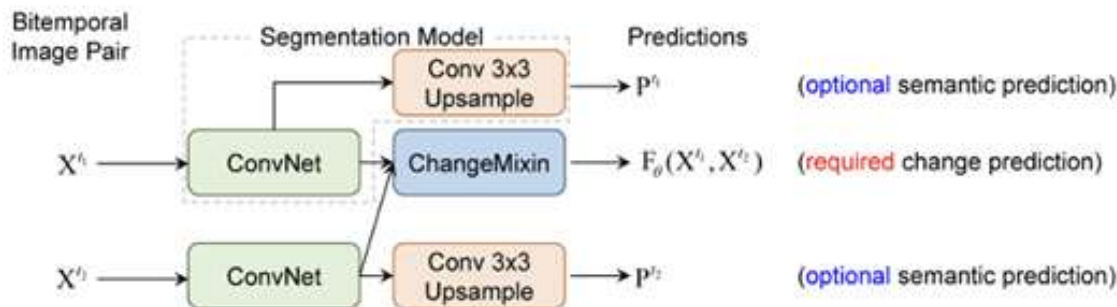
向下采样

# 训练与推导过程

## training stage



## inference stage



其中 $y$  是样本的真实标签，取值为0或1。 $p$  是模型的预测概率，表示样本属于类别1的概率。

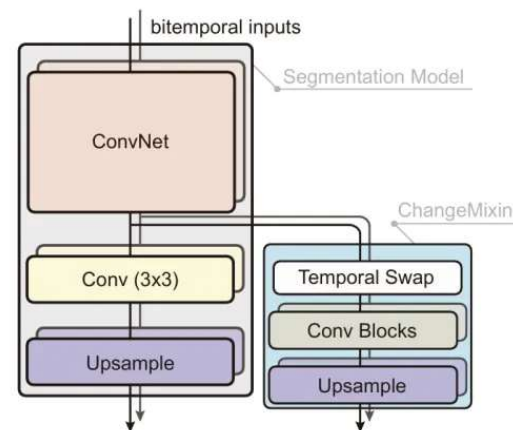
**时间交换模块 (TSM)**：负责在两个时间点的特征图之间进行交换，以提供网络架构的归纳偏差。

利用归纳偏差，即时间对称性，来稳定学习过程，缓解过拟合问题。

$TSM(T1, T2) = \text{cat}(T1, T2), \text{cat}(T2, T1)$

解决

$$F(X_1, X_2) \neq F(X_2, X_1)$$



$$\mathcal{L}_{Total} = \mathcal{L}_{Segmentation} + \mathcal{L}_{ChangeDetection}$$

$$\mathcal{L}_{Segmentation}(p, y) = -[y \log(p) + (1 - y) \log(1 - p)]$$

二元交叉熵损失函数

$$\mathcal{L}_{ChangeDetection} = \frac{[L_{Binary}(F_{\theta}(X_{t1}, \pi X_{t1}), Y_{t1} \oplus \pi Y_{t1}) + L_{Binary}(F_{\theta}(\pi X_{t1}, X_{t1}), Y_{t1} \oplus \pi Y_{t1})]}{2}$$

其中 $y$  是样本的真实标签，取值为0或1。 $p$  是模型的预测概率，表示样本属于类别1的概率。

# 实验结果

数据好坏直接影响效果

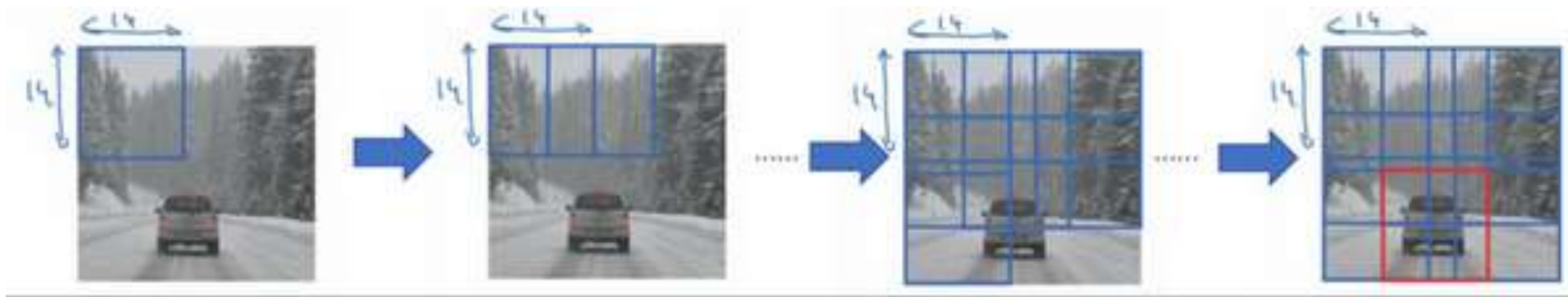
样本多，分辨率高

Method	Segmentation Model	Train on xView2 pre-disaster				Train on SpaceNet 2				$\Delta$ Params (M)	$\Delta$ MAdds (B)
		WHU		LEVIR-CD <sup>all</sup>		WHU		LEVIR-CD <sup>all</sup>			
		IoU (%)	F <sub>1</sub> (%)	IoU (%)	F <sub>1</sub> (%)	IoU (%)	F <sub>1</sub> (%)	IoU (%)	F <sub>1</sub> (%)		
PCC	PSPNet [29]	37.46	54.51	55.87	71.69	21.39	35.25	10.19	18.50	0	0
ChangeStar (ours)	+ ChangeMixin	56.44 <sub>(+18.98)</sub>	72.15 <sub>(+17.64)</sub>	61.63 <sub>(+5.76)</sub>	76.26 <sub>(+4.57)</sub>	25.56 <sub>(+4.17)</sub>	40.72 <sub>(+5.47)</sub>	15.25 <sub>(+5.06)</sub>	26.47 <sub>(+7.97)</sub>	0.16	0.63
PCC	DeepLab v3 [5]	32.46	49.01	54.77	70.78	33.08	49.72	13.78	24.23	0	0
ChangeStar (ours)	+ ChangeMixin	56.85 <sub>(+24.39)</sub>	72.49 <sub>(+23.48)</sub>	60.94 <sub>(+6.17)</sub>	75.73 <sub>(+4.95)</sub>	35.57 <sub>(+2.49)</sub>	52.48 <sub>(+2.76)</sub>	15.92 <sub>(+2.14)</sub>	27.46 <sub>(+3.23)</sub>	0.08	0.33
PCC	DeepLab v3+ [6]	35.75	52.68	55.51	71.38	23.90	38.58	9.80	17.85	0	0
ChangeStar (ours)	+ ChangeMixin	52.01 <sub>(+16.26)</sub>	68.43 <sub>(+15.75)</sub>	57.96 <sub>(+2.45)</sub>	73.38 <sub>(+2.00)</sub>	38.42 <sub>(+15.42)</sub>	55.51 <sub>(+16.93)</sub>	22.22 <sub>(+12.42)</sub>	36.36 <sub>(+18.51)</sub>	0.08	0.33
PCC	Semantic FPN [16]	38.66	55.76	56.19	71.95	27.60	43.26	7.09	13.25	0	0
ChangeStar (ours)	+ ChangeMixin	55.37 <sub>(+16.71)</sub>	71.27 <sub>(+15.51)</sub>	65.21 <sub>(+9.02)</sub>	78.94 <sub>(+6.99)</sub>	37.63 <sub>(+10.03)</sub>	54.68 <sub>(+11.42)</sub>	25.86 <sub>(+18.77)</sub>	41.10 <sub>(+27.85)</sub>	0.08	0.33
PCC	FarSeg [30]	31.66	48.09	55.09	71.04	27.69	43.37	7.97	14.77	0	0
ChangeStar (ours)	+ ChangeMixin	58.22 <sub>(+26.56)</sub>	73.59 <sub>(+25.50)</sub>	65.71 <sub>(+10.62)</sub>	79.31 <sub>(+8.27)</sub>	39.02 <sub>(+11.33)</sub>	56.14 <sub>(+12.77)</sub>	30.42 <sub>(+22.45)</sub>	46.65 <sub>(+31.88)</sub>	0.08	0.33

结果表明STAR方法在单时间监督下性能优于基线方法，并在双时间监督下取得了更好的性能。



# 传统分割网络的滑动窗口



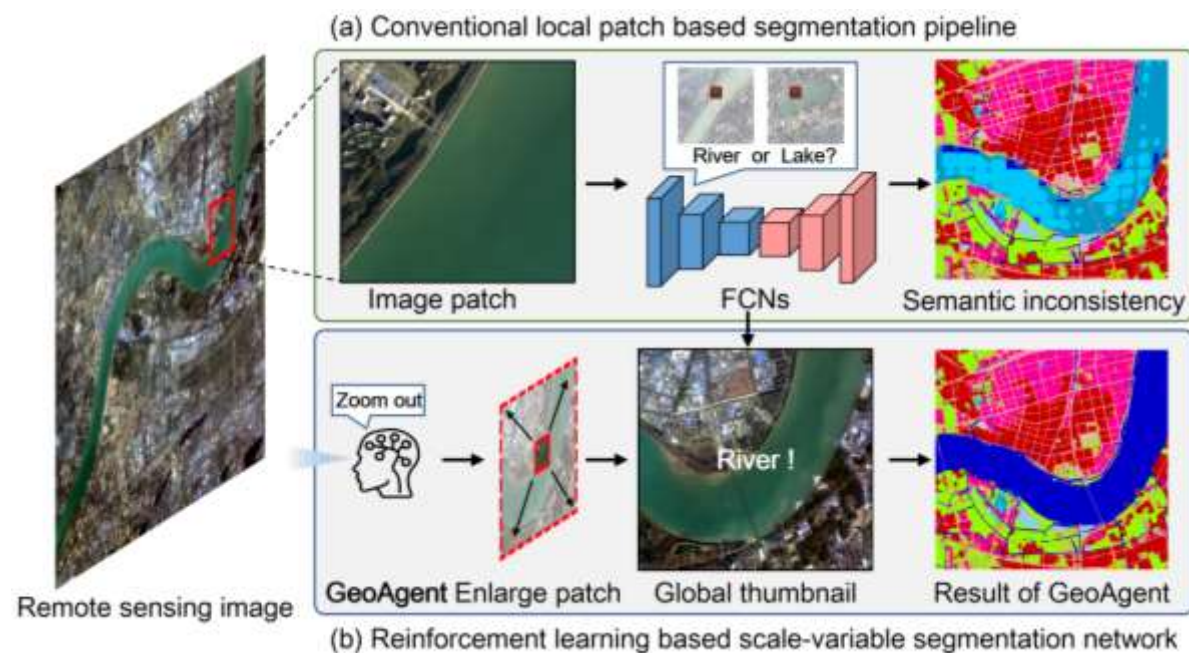
对于研究一些高分辨率遥感图像的变化检测，使用传统分割网络的滑动窗口，由于窗口的大小限制，会出现明显的**网格效应**现象。

# Seeing Beyond the Patch: Scale-Adaptive Semantic Segmentation of High-resolution Remote Sensing Imagery based on Reinforcement Learning

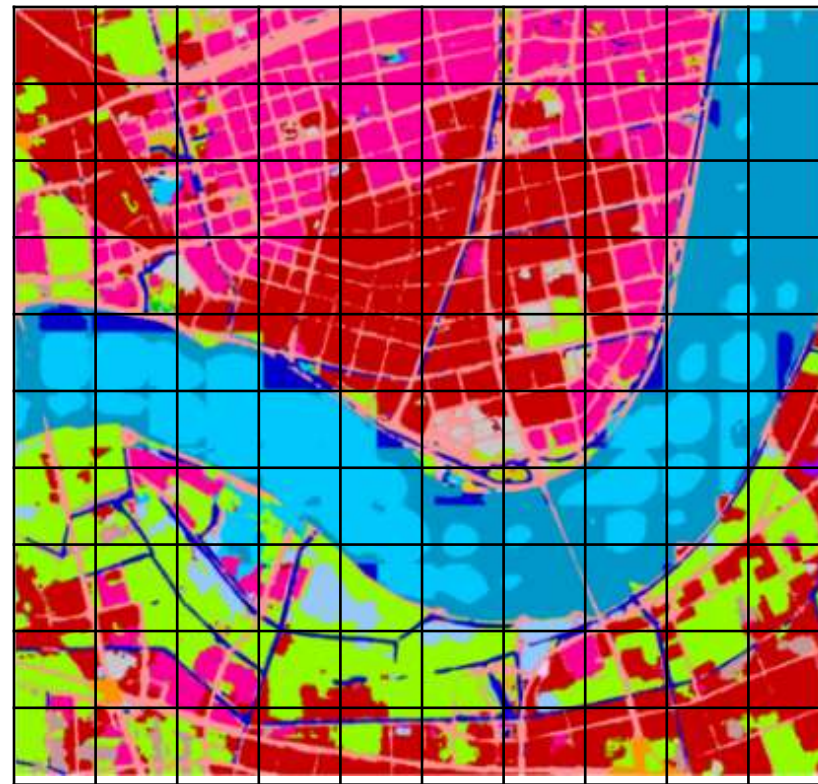
## 基于强化学习的高分辨率遥感影像尺度自适应语义分割

本文提出了一种基于深度强化学习的自适应尺度语义分割网络，称为**GeoAgent**，突破了传统分割网络的滑动窗口大小限制。在三个数据集上的实验证明，**GeoAgent**在高分辨率卫星图像的语义分割任务上取得了最先进的结果。

**研究目标：**在遥感图像分析中，基于补丁的方法在捕捉滑动窗口之外的信息方面存在限制。为了解决这一问题，本文提出了一种名为**GeoAgent**的动态尺度感知框架，该框架根据不同的地理对象自适应地捕捉图像补丁之外的适当尺度上下文信息。

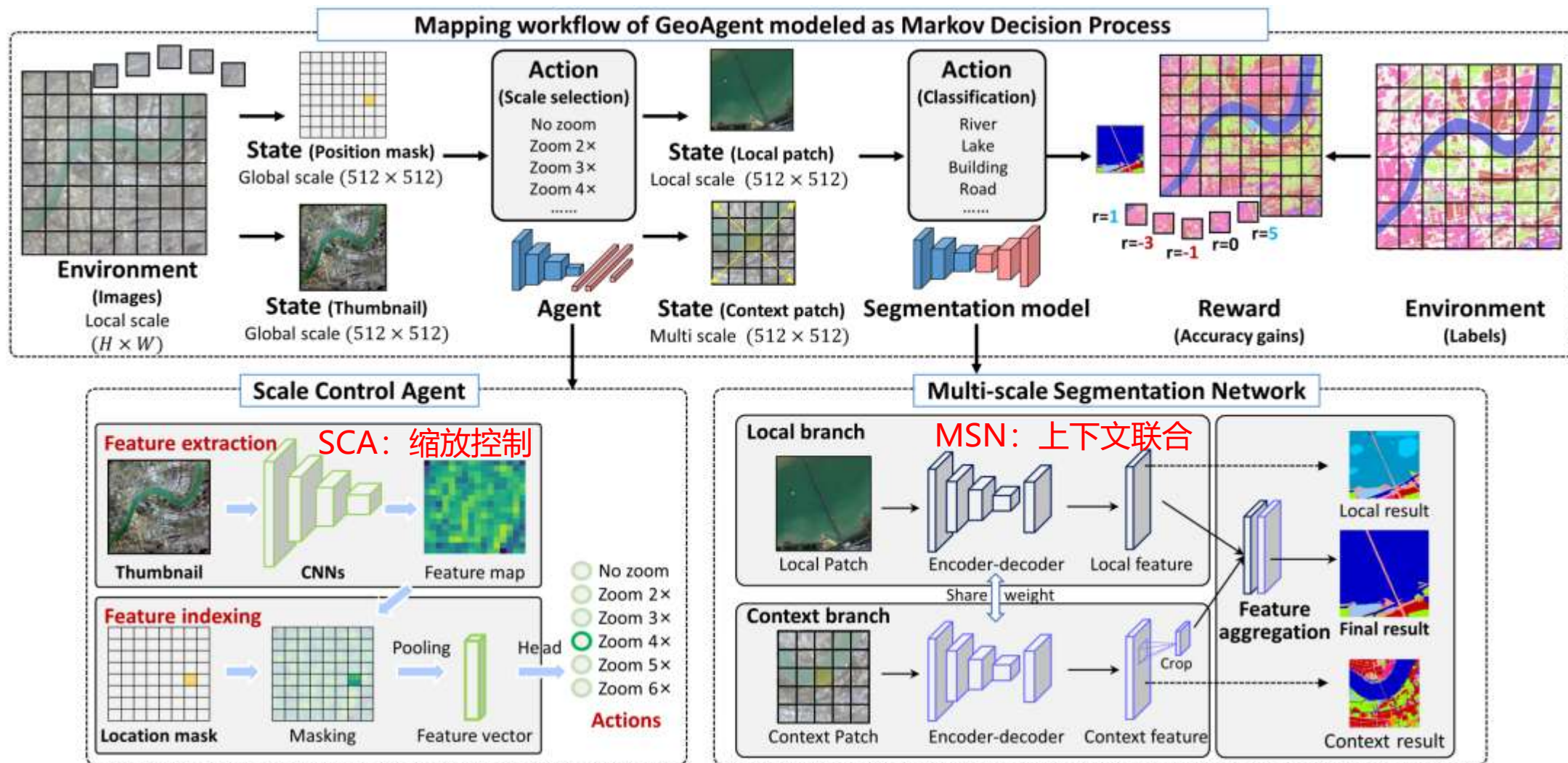


网格效应严重





# 模型总览



# 实验结果

Table 1: Segmentation accuracies of different models on three datasets.

Method	Backbone	WUSU		GID		FBP	
		IoU(%)	F <sub>1</sub> (%)	IoU(%)	F <sub>1</sub> (%)	IoU(%)	F <sub>1</sub> (%)
Local scale patch based methods							
-	UNet [33]	54.10	59.52	58.76	68.93	53.02	59.67
-	PSPNet [50]	60.37	69.29	59.04	69.34	59.69	64.27
-	FPN [19]	64.85	72.68	55.31	65.79	61.88	65.50
-	Deeplabv3+ [5]	63.39	69.97	64.89	68.65	57.08	59.46
-	UNet++ [54]	60.26	66.88	61.71	71.68	55.31	58.43
Global-local scale patch based methods							
GLNet [6]	FPN	51.71	59.09	58.69	68.79	21.86	24.34
MagNet [18]	FPN	51.47	63.46	66.98	55.94	45.05	49.09
RAZN [12]	Deeplabv3+	58.62	66.27	63.52	68.49	57.63	57.73
WiCoNet [11]	ResNet50	61.83	75.68	65.80	76.72	58.83	70.28
CascadePSP [7]	PSPNet	67.48	77.42	72.45	77.24	67.35	70.03
Scale-variable segmentation network GeoAgent							
GeoAgent	UNet	70.33 <sup>(+16.23)</sup>	77.11 <sup>(+17.59)</sup>	75.88 <sup>(+17.12)</sup>	75.23 <sup>(+6.30)</sup>	57.77 <sup>(+4.75)</sup>	61.37 <sup>(+1.70)</sup>
GeoAgent	UNet++	74.93 <sup>(+14.67)</sup>	79.17 <sup>(+12.39)</sup>	77.30 <sup>(+15.39)</sup>	78.62 <sup>(+6.94)</sup>	61.84 <sup>(+0.53)</sup>	60.01 <sup>(+1.58)</sup>
GeoAgent	Deeplabv3+	76.19 <sup>(+12.80)</sup>	80.56 <sup>(+10.59)</sup>	75.51 <sup>(+10.62)</sup>	76.07 <sup>(+7.42)</sup>	62.25 <sup>(+5.17)</sup>	63.21 <sup>(+3.75)</sup>
GeoAgent	PSPNet	75.19 <sup>(+14.82)</sup>	79.62 <sup>(+10.33)</sup>	75.57 <sup>(+10.53)</sup>	79.34 <sup>(+10.00)</sup>	69.67 <sup>(+9.98)</sup>	74.36 <sup>(+10.09)</sup>
GeoAgent	FPN	76.00 <sup>(+11.15)</sup>	82.46 <sup>(+9.78)</sup>	78.16 <sup>(+22.85)</sup>	77.56 <sup>(+11.77)</sup>	73.95 <sup>(+12.07)</sup>	74.11 <sup>(+8.61)</sup>

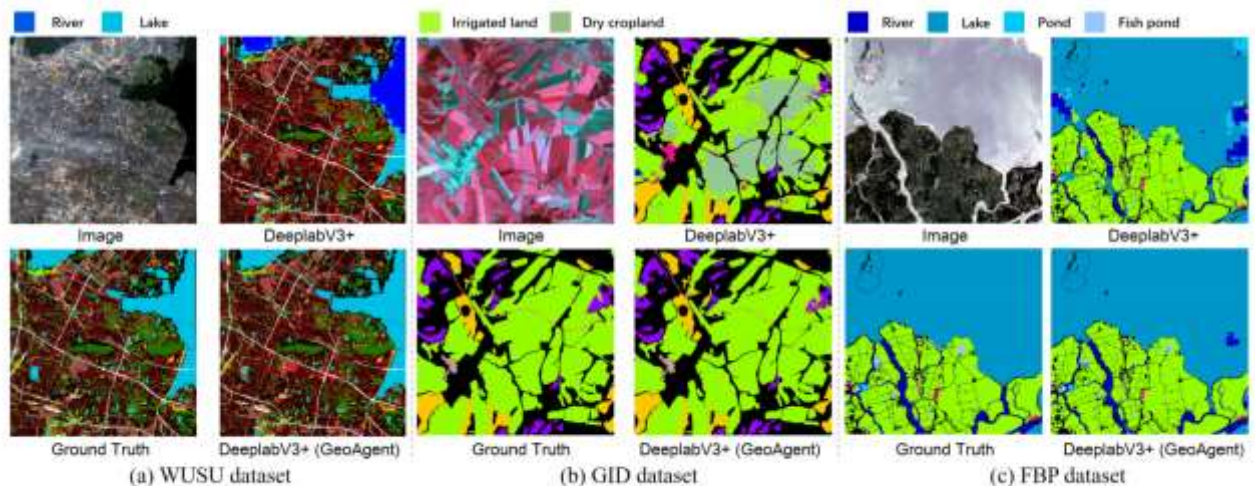
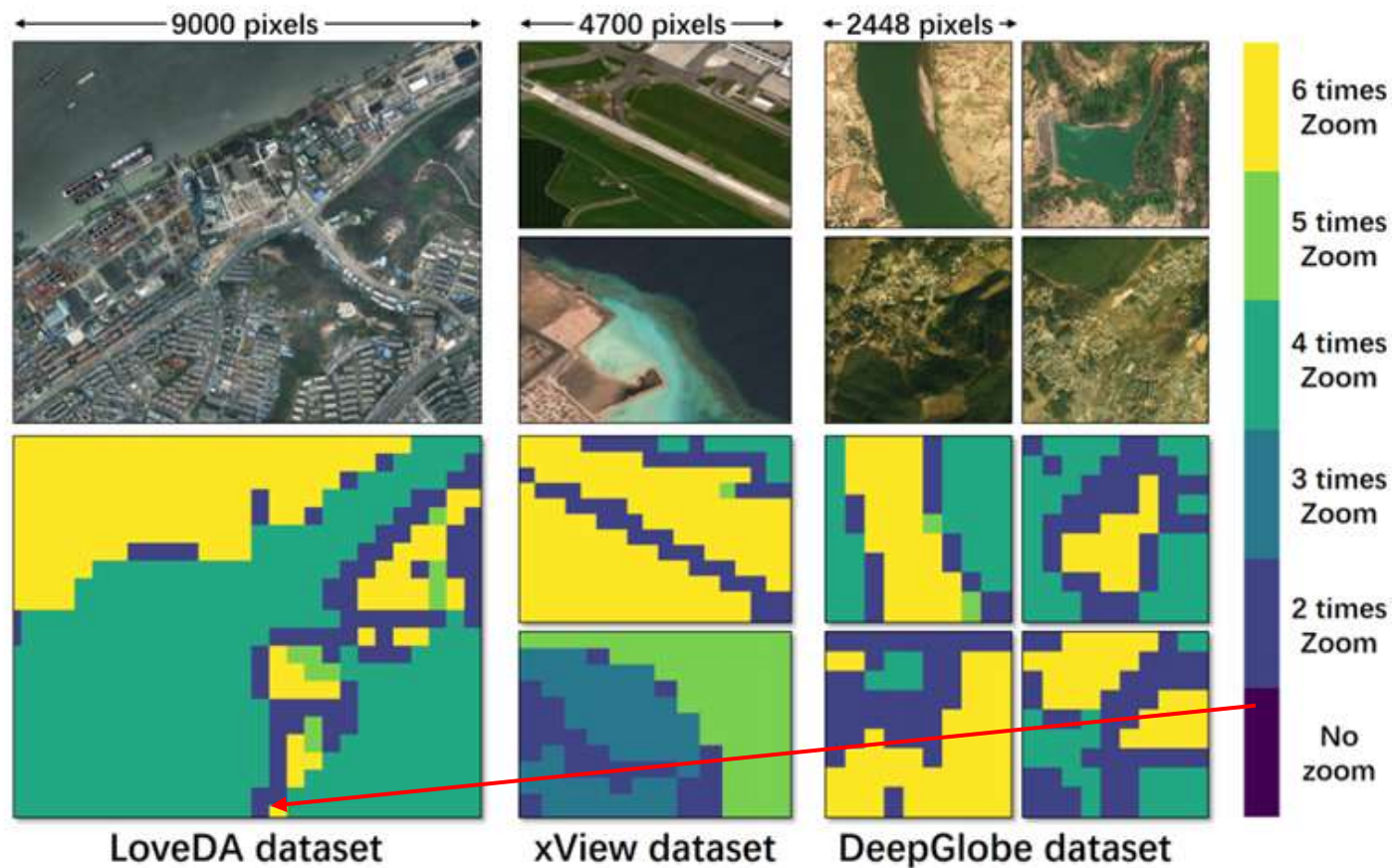


Figure 3: Segmentation results of Deeplabv3+ and modified GeoAgent models on three dataset.



# 超大分辨率遥感图像



整个实验结果表明：

GeoAgent使在高分辨率遥感图像上做的预测更加**精确**，通过“Agent”，使整个网络获得了变相的**超大感受野**。

