

# DynamicEarth

## 全文摘要

---

这篇论文介绍了一种新的任务——开放词汇变化检测（OVCD），旨在解决现有变化检测方法依赖于预定义类别的限制，并在任何类别中检测变化。作者提出了两个训练免费的框架——M-C-I 和 I-M-C，它们利用了现有的基础模型进行 OVCD 任务。基于这两个框架，作者得到了几种方法，如 SAM-DINOv2-SegEarth-OV、Grounding-DINO-SAM2-DINO 等。在 5 个基准数据集上的广泛评估表明，他们的 OVCD 方法比现有的监督和无监督方法具有更好的泛化性和鲁棒性。为了支持进一步的研究，作者还发布了一个专门的代码库——DynamicEarth，旨在推进 OVCD 的研究和应用。

## 论文速读

---

### 论文方法

#### 方法描述

该论文提出了两种训练-free 的遥感图像变化检测框架：M-C-I 和 I-M-C。这两种框架都是基于开源模型和离线数据预处理实现的。

M-C-I 框架分为三个组件：**Mask Proposal**、**Comparator** 和 **Identifier**。首先使用 **Mask Proposal** 提取所有类别的候选区域，然后使用 **Comparator** 判断这些区域是否发生变化，并使用 **Identifier** 筛选出感兴趣的变化类别。

I-M-C 框架也包含三个组件：**Identifier**、**Mask Proposal** 和 **Comparator**。与 M-C-I 不同的是，它先通过 **Identifier** 找到感兴趣的目标实例并确定其位置，再将这些实例转换为统一的掩模格式，并使用 **Comparator** 判断它们在两个时间点的状态是否发生了变化。

#### 方法改进

M-C-I 和 I-M-C 框架都采用了 **Mask Proposal** 和 **Comparator** 组件来实现遥感图像变化检测。其中，**Mask Proposal** 使用了学习自适应分块（SAM）算法来提取感兴趣的候选区域；**Comparator** 则采用了负余弦相似度作为变化分数，并使用 DINO 或 DINOv2 提取特征向量以提高精度。

此外，I-M-C 框架还引入了 **Identifier** 组件来识别感兴趣的目标实例。它可以使用户开放词汇检测、视觉语义理解或多模态大型语言模型等技术来实现目标实例的定位。

# 解决的问题

该论文提出的 M-C-I 和 I-M-C 框架解决了传统遥感图像变化检测方法中存在的问题：

数据稀疏性和质量不佳导致难以实施完整的训练-based 检测流程；

遥感图像与自然图像之间存在差异，如视角、分辨率等因素可能影响分类器性能；

非 RGB 图像、物体尺度跨度大以及小物体等问题需要特殊处理。

因此，该论文提出的训练-free 框架能够有效地解决这些问题，从而实现更准确、更高效的遥感图像变化检测。

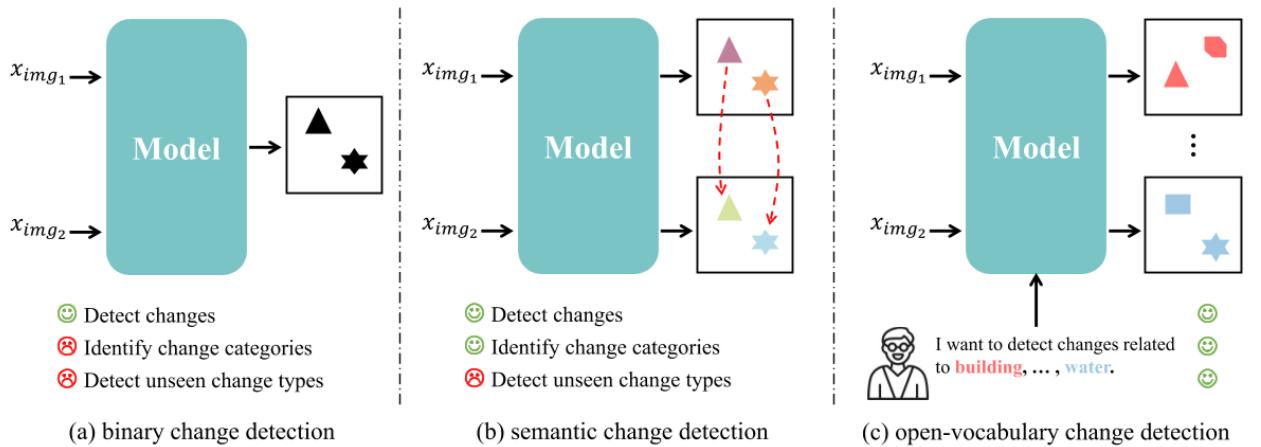


Figure 2: Different change detection tasks: (a) Binary change detection aims at discovering all (interested) changes and generating a binary mask; (b) Semantic change detection further identifies the category of changes. However, both can only be trained and evaluated on data with predefined categories; (c) Our proposed OVCD can detect changes in any category according to the user's requirements.

# 论文实验

本文主要介绍了两个框架——M-C-I 和 I-M-C 的实验结果，并进行了多个对比实验来验证其性能。其中，作者使用了四个建筑变化检测数据集和一个土地覆盖变化检测数据集来进行测试。以下是具体的实验内容：

**建筑变化检测实验：**在 LEVIR-CD、WHU-CD、S2Looking 和 BANDON 这四个数据集中，作者使用了 PCA-KM、DSFA、DCVA、GMCD、CVA、AnyChange 等方法作为对比，以及自己提出的 OVCD 方法（基于 M-C-I 和 I-M-C 框架）。实验结果显示，在简单数据中，I-M-C 框架下的方法表现更好，而在复杂数据中，M-C-I 框架下的方法表现更优。同时，对于不同的类别，识别器的性能也会影响整体的结果。

**土地覆盖变化检测实验：**在 SECOND 数据集中，作者对六种不同类别的土地覆盖变化进行了检测。实验结果显示，M-C-I 框架下的方法相对于 I-M-C 框架下的方法表现更好，因为 I-M-C 框架下的识别器很难识别一些类别，而这些类别在自然图像中通常被视为“背景”。

**OVCD 效果实验：**作者将 OVCD 方法与监督学习模型 Changer 进行了比较。实验结果显示，OVCD 方法比最好的跨数据集结果要好得多，证明了 OVCD 在现实世界中的可用性和潜力。

**多分类推理实验：**作者发现，在多分类推理时，细粒度的类别划分可能会损害 VLM 在处理遥感图像时的表现。此外，需要考虑类别信息的比较逻辑也会变得更加复杂。

总的来说，本文通过多个实验证证了 M-C-I 和 I-M-C 框架的有效性，并探讨了一些相关的问题和挑战。

Table 1: OVCD quantitative comparison on building change detection datasets. “-” denotes data missing.

Method	LEVIR-CD		WHU-CD		S2Looking		BANDON	
	$IoU^c$	$F_1^c$	$IoU^c$	$F_1^c$	$IoU^c$	$F_1^c$	$IoU^c$	$F_1^c$
PCA-KM [5]	4.8	9.1	5.4	10.2	-	-	-	-
CNN-CD [17]	7.0	13.1	4.9	9.4	-	-	-	-
DSFA [16]	4.3	8.2	4.1	7.8	-	-	-	-
DCVA [48]	7.6	14.1	10.9	19.6	-	-	-	-
GMCD [55]	6.1	11.6	10.9	19.7	-	-	-	-
CVA [2]	-	12.2	-	-	-	5.8	-	-
DINOv2+CVA [80]	-	17.3	-	-	-	4.3	-	-
AnyChange-H [80]	-	23.0	-	-	-	6.4	-	-
SCM [54]	18.8	31.7	18.6	31.3	-	-	-	-
<b>M-C-I:</b>								
SAM - DINO - SegEarth-OV	33.0	49.7	36.7	53.7	22.5	36.7	15.3	26.5
SAM - DINOv2 - SegEarth-OV	36.6	53.6	40.6	57.7	23.9	38.5	17.6	30.2
SAM2 - DINOv2 - SegEarth-OV	33.8	50.5	40.9	58.1	23.1	37.6	17.7	30.1
<b>I-M-C:</b>								
MM-Grounding-DINO - SAM2 - DINO	15.6	27.0	11.0	19.8	2.3	4.5	1.9	3.8
APE - / - DINO	53.5	69.7	56.8	72.5	10.1	18.4	7.8	14.5
APE - / - DINOv2	50.0	66.7	61.1	75.8	5.3	10.1	11.8	21.1
Grounding DINO 1.5 - SAM2 - DINOv2	-	-	-	-	-	-	-	-

## 论文总结

### 文章优点

本文提出了一种新的任务——Open-Vocabulary Change Detection (OVCD)，该任务可以实现视觉与语言之间的连接，并能够检测任何感兴趣的改变。作者

提出了两个训练免费的通用框架：**M-C-I** 和 **I-M-C**，这两个框架都充分利用了现有的基础模型。此外，作者还为地球视觉社区贡献了一个第一代的 **OVCD** 代码库——**DynamicEarth**，以促进算法开发、评估和应用。

## 方法创新点

作者提出的 **OVCD** 是一种开放词汇量变化检测的新任务，其目标是发现任何类别的变化。为此，作者设计了两个训练免费的通用框架：**M-C-I** 和 **I-M-C**。其中，**M-C-I** 框架首先使用掩模提议方法生成类无关的掩模，然后比较双时相图像或特征中的每个掩模区域来确定是否发生了变化。如果检测到变化，则将掩模区域嵌入最终的开放词汇分类器（例如 **CLIP** 或 **SegEarth-OV**）以识别变化类别。而 **I-M-C** 框架则受到后分类比较（**PCC**）方法的启发，它利用感知基础模型（如 **DINO**、**APE** 或 **Molmo**）来指导对象的识别，这些对象在格式上类似于边界框、粗略掩模或点。最后，将同一位置上的双时相图像的掩模进行比较，以确定是否有任何变化发生。

## 未来展望

尽管本文的方法在准确性方面已经取得了显著进展，但它仍然无法完全替代基于监督学习的方法。因此，未来的研究应该继续改进 **OVCD** 方法，无论是基于训练还是无训练的方式，无论是提高准确率还是效率等方面。此外，随着更多的数据集和更先进的技术的发展，我们还可以期待更多有趣的应用场景出现。