测试报告

# 1.引言

## 1.1 背景

在计算机视觉领域，尤其是语义分割任务中，近年来的研究进展显著，其中包括深度学习技术的突破性应用。对抗学习作为一种新兴的技术，通过引入生成对抗网络（GANs）来增强模型对数据分布的理解，已经在图像生成、风格迁移和数据增强等方面展现出巨大潜力。此外，半监督学习方法，如自训练、多实例学习和伪标签方法，有效地结合了有限的标注数据与大量的未标注数据，以提升模型在各种复杂场景下的鲁棒性和准确性。这些方法通过不同的方式利用未标注数据，如基于聚类的特征学习或利用模型自身的预测作为额外的监督信号，已经在减少对大规模标注数据集的依赖方面取得了令人鼓舞的成果。随着这些技术的不断发展和完善，它们为解决语义分割中的挑战提供了新的视角和工具，推动了计算机视觉在自动驾驶、医疗诊断和增强现实等领域的实际应用。

## 1.2 研究动机

通过探索和实现一种基于对抗学习的半监督语义分割方法，旨在利用少量标注数据结合大量未标注数据进行有效训练，以提高模型的泛化能力和性能，同时降低对昂贵手动标注工作的依赖，进而推动计算机视觉技术在各种实际应用场景中的广泛应用。

# 2.测试方法

## 2.1 测试环境

硬件环境：单个NVIDIA TitanX GPU，具备12 GB的显存

软件框架：pytorch

## 2.2 数据集

PASCAL VOC 2012数据集是一个在计算机视觉领域广泛使用的基准数据集，它为语义分割及其他视觉识别任务提供了标准化的评估平台。该数据集包含了来自日常生活场景的图像，涵盖了包括动物、车辆、室内和室外环境等20个类别的对象。每个类别的对象都在图像中被精确标注，为研究者提供了丰富的、像素级的标注信息。PASCAL VOC 2012数据集因其平衡的类别分布、清晰的标注质量以及在学术界的普及度，成为衡量语义分割算法性能的重要标准之一。

Cityscapes数据集专注于提供城市街景图像的像素级标注，它包含了大量的城市驾驶场景，图像分辨率高，场景复杂，为语义分割算法提供了更具挑战性的测试条件。数据集中的图像被标注了19个类别，包括不同的行人、车辆、道路、建筑物等城市元素，这些细致的标注信息极大地促进了城市场景理解的研究。Cityscapes数据集因其高难度和真实性，成为评估和推动语义分割技术在复杂城市环境中应用的重要资源。通过在Cityscapes上的测试，研究者可以验证算法在处理高分辨率图像和复杂场景时的鲁棒性和准确性。

## 2.3 测试流程

* **环境配置**：首先配置适当的硬件环境和安装所需的软件框架，接着对PASCAL VOC 2012和Cityscapes数据集进行预处理和划分
* **网络训练**：了解分割网络和全卷积判别网络的架构，根据说明文档设置关键的超参数和优化器；之后实施训练过程，包括对抗训练和半监督训练，同时计算并优化由交叉熵损失、对抗损失和半监督损失组成的多任务损失函数。
* **评估指标：**训练完成后，在验证集上进行模型评估，使用平均交并比作为主要的性能指标，交并比（IoU）是一个常用的衡量图像分割精度的指标，它通过计算预测分割区域与真实标注区域之间的交集与并集的比例来评估分割结果的质量。
* **结果分析：**评估过程中，进行定量测试和对比测试结果分析，以确定模型性能并识别改进方向。

# 3.参数配置

对于分割网络，使用了带Nesterov加速的随机梯度下降（SGD）作为优化器，其中动量参数设置为0.9，权重衰减为。初始学习率设置为，并按照幂为0.9的多项式衰减降低。

对于判别网络，使用了Adam优化器，学习率设置为，动量参数设置为0.9和0.999，并同样采用幂为0.9的多项式衰减策略。

# 4.测试结果

## 4.1 定量测试结果

分别在PASCAL VOC 2012数据集和Cityscapes数据集进行测试。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Methods |  | Data | Amount |  |
|  | 1/8 | 1/4 | 1/2 | Full |
| baseline | 66.0 | 68.3 | 69.8 | 73.6 |
| baseline+ | 67.6 | 71.0 | 72.6 | 74.9 |
| baseline++ | 68.8 | 71.6 | 73.2 | N/A |
| FCN-8s (Long et al., 2015) | N/A | N/A | N/A | 67.2 |
| Dilation10 (Yu & Koltun, 2016) | N/A | N/A | N/A | 73.9 |
| DeepLab-v2 (Chen et al., 2017) | N/A | N/A | N/A | 77.7 |

**Table 1：results on the VOC 2012 validation set**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Methods |  | Data | Amount |  |
|  | 1/8 | 1/4 | 1/2 | Full |
| baseline | 52.4 | 58.3 | 62.6 | 66.4 |
| baseline+ | 53.8 | 59.1 | 63.7 | 67.7 |
| baseline++ | 54.2 | 59.7 | 64.5 | N/A |
| FCN-8s (Long et al., 2015) | N/A | N/A | N/A | 65.3 |
| Dilation10 (Yu & Koltun, 2016) | N/A | N/A | N/A | 67.1 |
| DeepLab-v2 (Chen et al., 2017) | N/A | N/A | N/A | 70.4 |

**Table 2 results on the Cityscapes validation set**

## 4.2 对比测试结果

对比不同方法在PASCAL VOC 2012数据集上进行测试。

|  |  |  |
| --- | --- | --- |
| Methods | Baseline | Adversarial |
| Luc et al.(2016) | 71.8 | 72.0 |
| ours | 73.6 | 74.9 |

**Table 3 Adversarial learning comparison with Luc et al.(2016) on VOC 2012 validation set**

|  |  |  |  |
| --- | --- | --- | --- |
| Methods | Data amount | Fully-supervised | Semi-supervised |
| Papandreou et al.(2015) | Full | 62.5 | 64.6 |
| Souly et al. (2017) | Full | 59.5 | 64.1 |
| ours | Full | 66.3 | 68.4 |
| Souly et al. (2017) | 30% | 38.9 | 42.2 |
| ours | 30% | 57.4 | 60.6 |

**Table 4 Semi-supervised learning comparisons on VOC 2012 validation set without using additional labels of SBD**

# 5.结果分析

测试结果显示，通过结合对抗损失和半监督学习策略，半监督语义分割方法在PASCAL VOC 2012和Cityscapes数据集上相比基线模型实现了性能上的提升。具体来说，对抗损失的加入带来了一定程度的准确率增加，而半监督损失的进一步引入则带来了额外的性能增益。

通过使用对抗学习框架和全卷积判别器网络，该方法能够在有限的标注数据情况下提高分割性能。具体来说，对抗损失的引入使得分割网络能够学习到更接近真实标注的高阶结构信息，而半监督损失利用判别器网络产生的置信度图作为自监督信号，进一步增强了模型对未标记数据的学习。实验结果在PASCAL VOC 2012和Cityscapes数据集上均显示出，与仅使用监督学习的方法相比，该方法在语义分割任务上取得了显著的性能提升。此外，超参数分析和模型的定量与定性结果，都验证了该方法在提高分割精度和利用未标记数据方面的潜力。