



四川大學
SICHUAN UNIVERSITY

基于图像大模型增强的恶劣 天气交通目标检测方法研究

汇报人：侯睿

时间：2025.4.18



目录

CONTENT

01 研究背景及意义

04 演示视频

02 相关工作

03 项目主要内容

1



PART
ONE



研究背景及意义



无人驾驶一直是我国重点支持的领域项目。



中华人民共和国中央人民政府
www.gov.cn



首页 | 简 | 繁 | EN | 登录 | 邮箱 | 无障碍

首页 > 政策 > 国务院政策文件库 > 国务院部门文件

字号: 默认 大 超大



标题: 交通运输部关于促进道路交通自动驾驶技术发展和应用的指导意见 发文机关: 交通运输部

发文字号: 交科技发〔2020〕124号 来源: 交通运输部网站

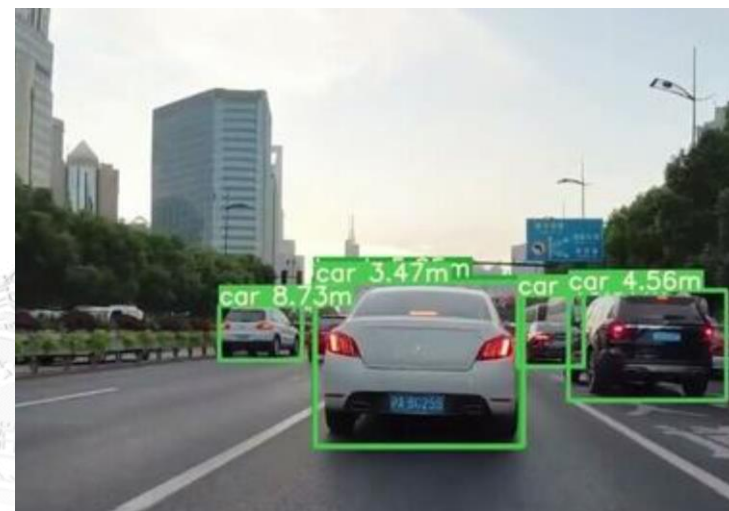
主题分类: 工业、交通\其他 公文种类: 意见

成文日期: 2020年12月20日

交通运输部关于促进道路交通自动驾驶技术发展和应用的指导意见

交科技发〔2020〕124号

目标检测方法是无人驾驶的核心
其精度直接决定了无人驾驶系统的智能化水平



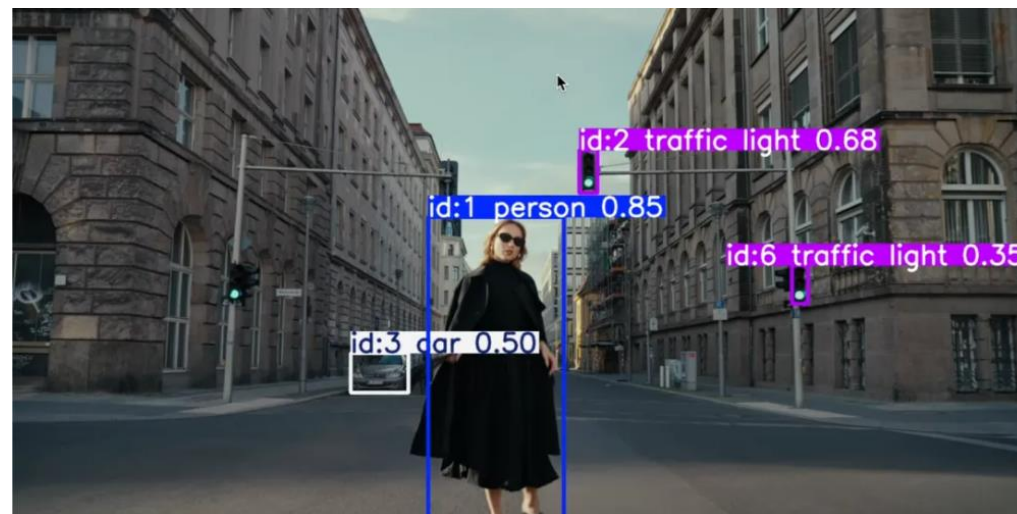


基于图像大模型增强的恶劣天气交通目标检测方法研究



四川大學
SICHUAN UNIVERSITY

当前
目标检测的性能在**晴朗天气**下已取得显著进展



但目标检测模型在**恶劣天气**条件下的表现并不佳
原因包括：能见度受限、轮廓模糊、对比度低等





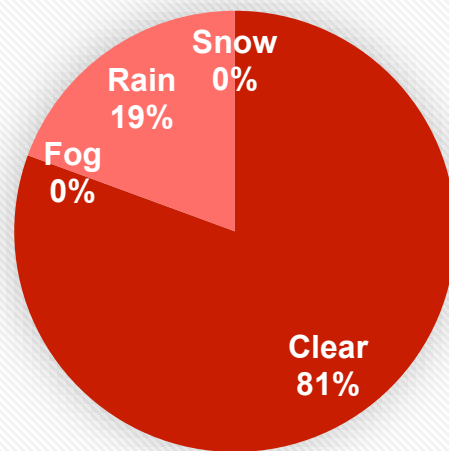
[1] Thomas Rothmeier and et.al., I Had a Bad Day: Challenges of Object Detection in Bad Visibility Conditions,

交通目标检测在恶劣天气下的性能**不佳**
核心问题在于**缺乏恶劣天气下的训练数据**^[1]

广泛应用于自动驾驶的数据集
Cityscapes、KITTI、NuScenes和BDD100K,
都是在**正常晴朗视觉条件**下捕获的图像

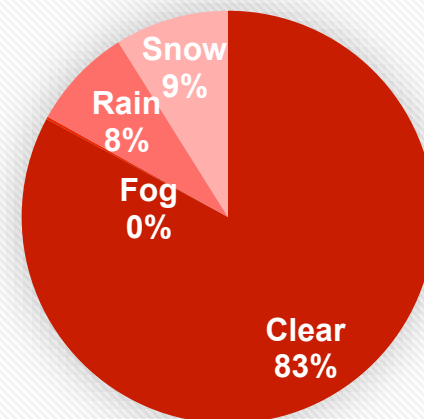
当前的交通目标检测训练集**缺乏**包含**雾、雪、雨**等**恶劣天气条件**的图像数据。

NuScenes



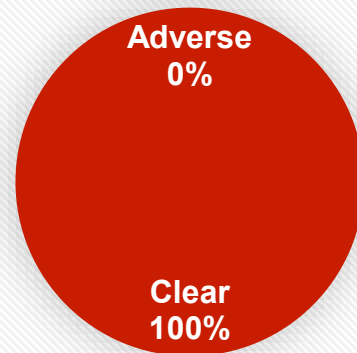
■ Clear ■ Fog ■ Rain ■ Snow

BDD100K



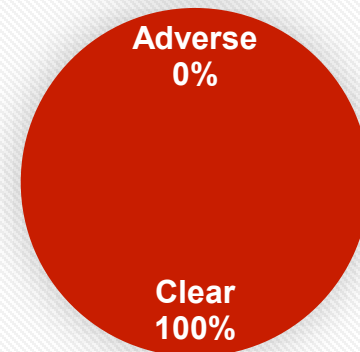
■ Clear ■ Fog ■ Rain ■ Snow

Cityscapes



■ Clear ■ Fog ■ Rain ■ Snow

KITTI



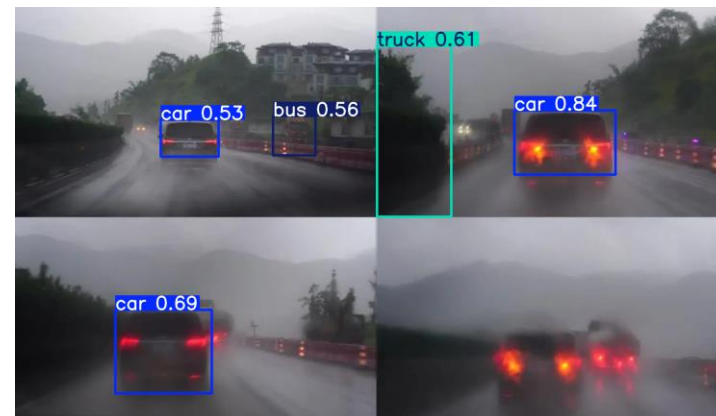
■ Clear ■ Fog ■ Rain ■ Snow



然而
恶劣天气条件罕见 & 不可控事件发生
使收集现实世界的恶劣天气驾驶数据集**极其困难**

我们提出一种基于图像大模型增强的恶劣天气交通目标检测方法
利用**图像大模型及其微调策略**
加以**物理渲染后处理**
构建符合真实恶劣天气下的驾驶场景数据
用于**增强后训练**交通目标检测模型
大幅提升其在恶劣天气下的表现

我们的研究
不仅显著提升无人驾驶系统在雨、雾等恶劣天气下的交通目标**检测精度**
也为**低成本且“边缘场景”**数据生成提供了新思路
为推动无人驾驶的**商业化落地**做出重要贡献



增强前

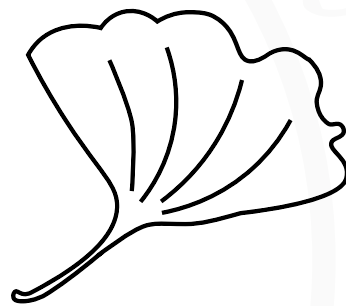


增强后

1



PART
TWO



相关工作

相关工作

一、恶劣天气下的交通目标检测



四川大学
SICHUAN UNIVERSITY

图像质量恢复:

Wenhan, Yang: 用深度学习同时完成雨纹检测与去除
Yuan-Kai, Wang: 通过多尺度深度融合实现单图去雾

恶劣天气数据集构建:

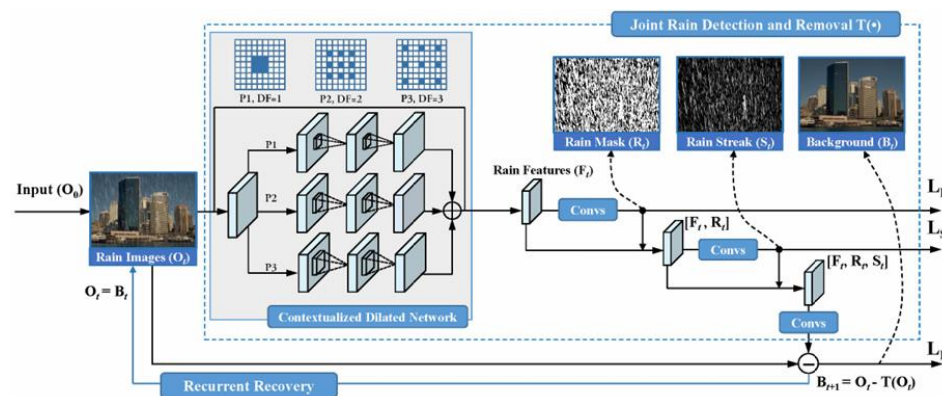
ACDC: 构建现实采集的含足够恶劣天气的数据集, 雨/雪/雾/晴各占25%
但由于天气恶劣导致收集困难, 我们认为现实采集显然不合理

Foggy Cityscapes: 物理加雾处理Cityscapes得到雾天驾驶场景

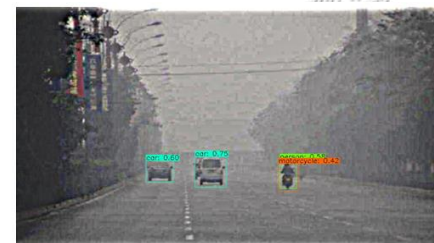
但场景多样性有限, 我们方法却能够扩展诸多场景

检测鲁棒性:

IA-YOLO等: 针对恶劣天气对YOLO进行图像自适应改造



(a) YOLO II (Baseline)



(b) IA-YOLO (Ours)

物理建模法

Alexander等：基于真实物理规律（光散射、水滴折射等）模拟雾、雨、雪的视觉效果，真实感较强



Fig. 4. The upper image is from the Cityscapes image data set, the lower figure shows that image with our fog applied. Distant objects are harder to spot than closer ones, as they are affected more by the fog.

域自适应法

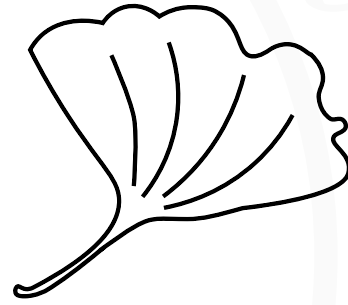
Minjun Li等：利用GAN等深度学习方法，将晴天图像的数据分布迁移到恶劣天气域



2



PART
THREE



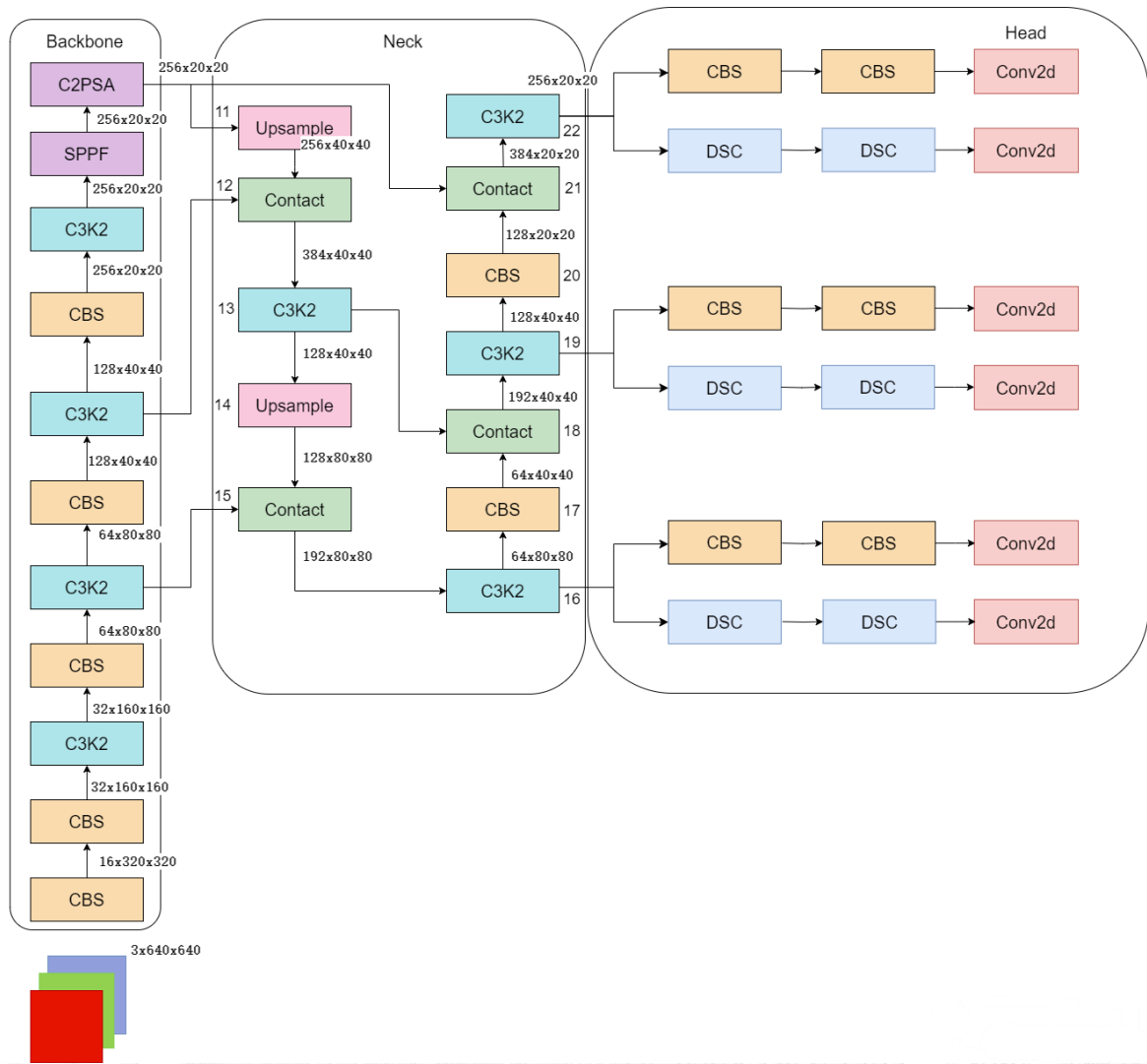
项目主要内容



基于图像大模型增强的恶劣天气交通目标检测方法研究

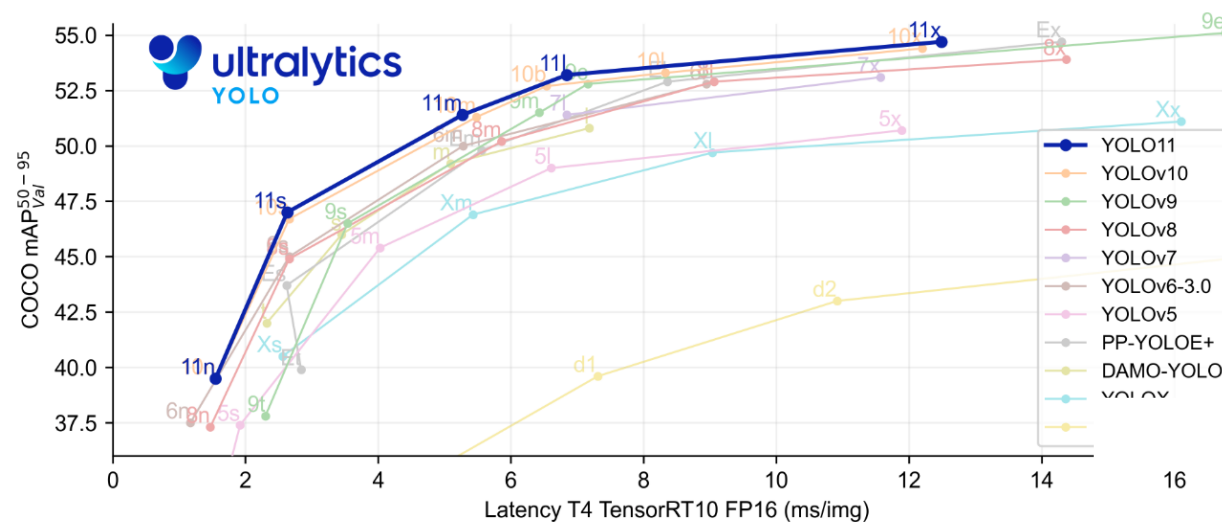


四川大学
SICHUAN UNIVERSITY



在无人驾驶场景中
YOLO11模型可做到**实时输出**目标边界框
满足**低延迟、高精度**检测需求

我们采用**YOLO11**作为我们主要研究的目标检测模型





四川大学
SICHUAN UNIVERSITY

一、目标检测模型超参数调优

我们选取“雾”作为恶劣天气研究对象

基于两个公开带标注基准数据集：

- 1. Cityscapes, 现实世界采集的**晴朗**天气城市车辆驾驶场景
- 2. FoggyCityscapes, 为Cityscapes经多重物理**雾渲染**后的版本



Cityscapes (晴)



Foggycityscapes (雾)

我们从中选取2900对构建晴/雾训练集
450对构建晴/雾验证集



| 名称 | 修改日期 | 类型 | 大小 |
|--|------------------|------|------|
| aachen_00000_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000001_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000002_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000003_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000004_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000005_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000006_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000007_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000008_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000009_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 1 KB |
| aachen_000010_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000011_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000012_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 3 KB |
| aachen_000013_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000014_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |
| aachen_000015_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 3 KB |
| aachen_000016_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 3 KB |
| aachen_000017_000019_gtFine_polygons.txt | 2024/10/28 12:52 | 文本文件 | 2 KB |

我们取**汽车**类别作为研究的交通目标
选用AP (Average Precision, 平均精度)
评估交通目标检测模型的**性能**





一、目标检测模型超参数调优

我们经多轮超参数调优实验，得出**精度和效率相对权衡最优**的超参策略：

35epoch + 60%GPU占用率计算出的batch + 预训练 + 全参数激活 + 3epoch预热训练 + 896x448 输入 + 同域验证

按照此策略

我们在 24GB RTX4090D 单卡下

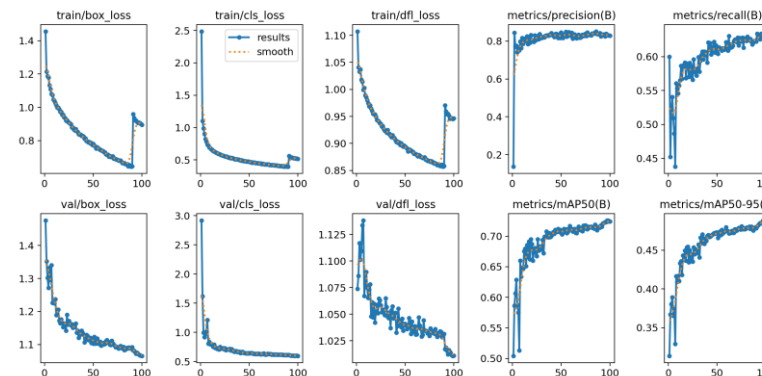
用雾天训练集

仅在8.5分钟内

训练出在晴/雾验证集上

AP分别达**0.676** 与**0.728**的突破性目标检测模型。

并以该模型作为后面实验的**基线**



| args-AP-valset | clear | fog |
|-----------------------------------|-------|-------|
| clear | 0.709 | 0.499 |
| fog | 0.579 | 0.662 |
| fog,epoch=50 | 0.588 | 0.664 |
| fog,epoch=20 | 0.535 | 0.609 |
| fog,batch=32 | 0.589 | 0.667 |
| fog,batch=8 | 0.605 | 0.649 |
| fog,batch=-1 | 0.588 | 0.666 |
| clear,pretrained=false | 0.7 | 0.449 |
| fog,freeze=23 | 0.505 | 0.494 |
| clear,freeze=11 | 0.662 | 0.483 |
| clear,pretrained=false,warmup=0.0 | 0.695 | 0.444 |
| fog,impsz=800 | 0.656 | 0.696 |
| fog,impsz=480 | 0.502 | 0.592 |
| fog,impsz=[896,448] | 0.688 | 0.719 |
| fog,valset=clear | 0.591 | 0.651 |
| fog,impsz=[896,448],batch=-1 | 0.676 | 0.728 |

Table 1: Default:epoch=35,batch=16,pretrained=True,freeze=null,warmup-epochs=3.0,impsz=640,val set is of the same category as the training set



二、构建图像大模型增强的数据集



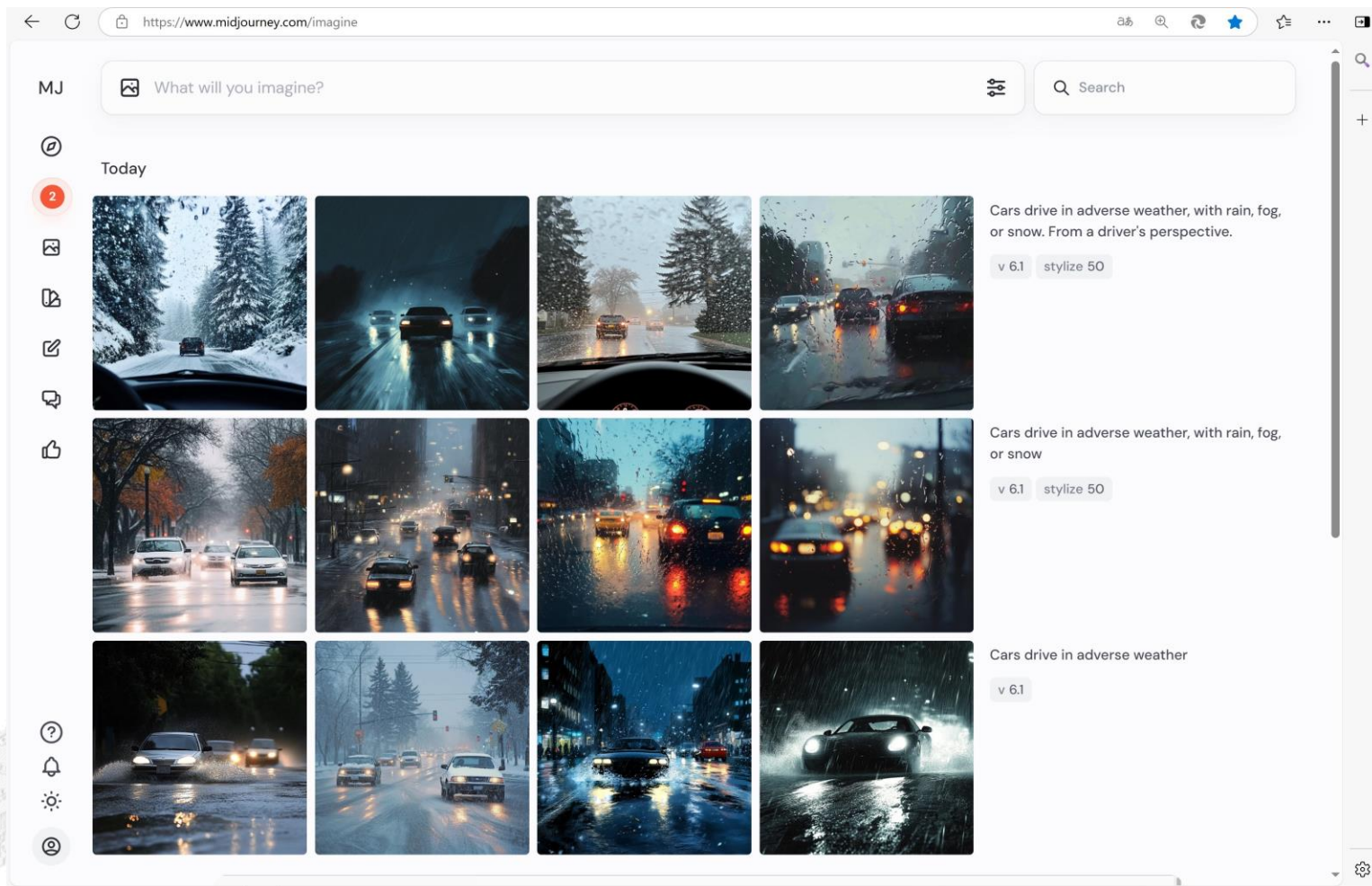
四川大學
SICHUAN UNIVERSITY

Midjourney是先进的**文本到图像的生成模型**
擅长将自然语言提示转换为符合语义的图像

我们利用大型语言模型
生成了各种**详细的文本描述**

我们采用Midjourney
生成约4000张含**雾雨雪等恶劣天气的驾
驶场景图像**

并手工挑选**548张最佳图像**
作为用于增强交通目标检测的**训练集**





二、构建图像大模型增强的数据集



Midjourney



四川大学
SICHUAN UNIVERSITY

我们对生成数据**进一步叠加雾的物理渲染**
以达到更加逼真的雾天驾驶场景

我们用Makesense工具
进行严格人工边界框标注
并由四位项目成员**反复独立校验**



渲染输入



渲染输出

我们将其与Foggy Cityscapes训练集混合
然后用于增强目标检测模型
并进行反复超参数调优以获取最佳性能

评估发现，相较于基线模型，增强后的模型：

晴天验证集
AP +5.8%

雾天验证集
AP +1.5%

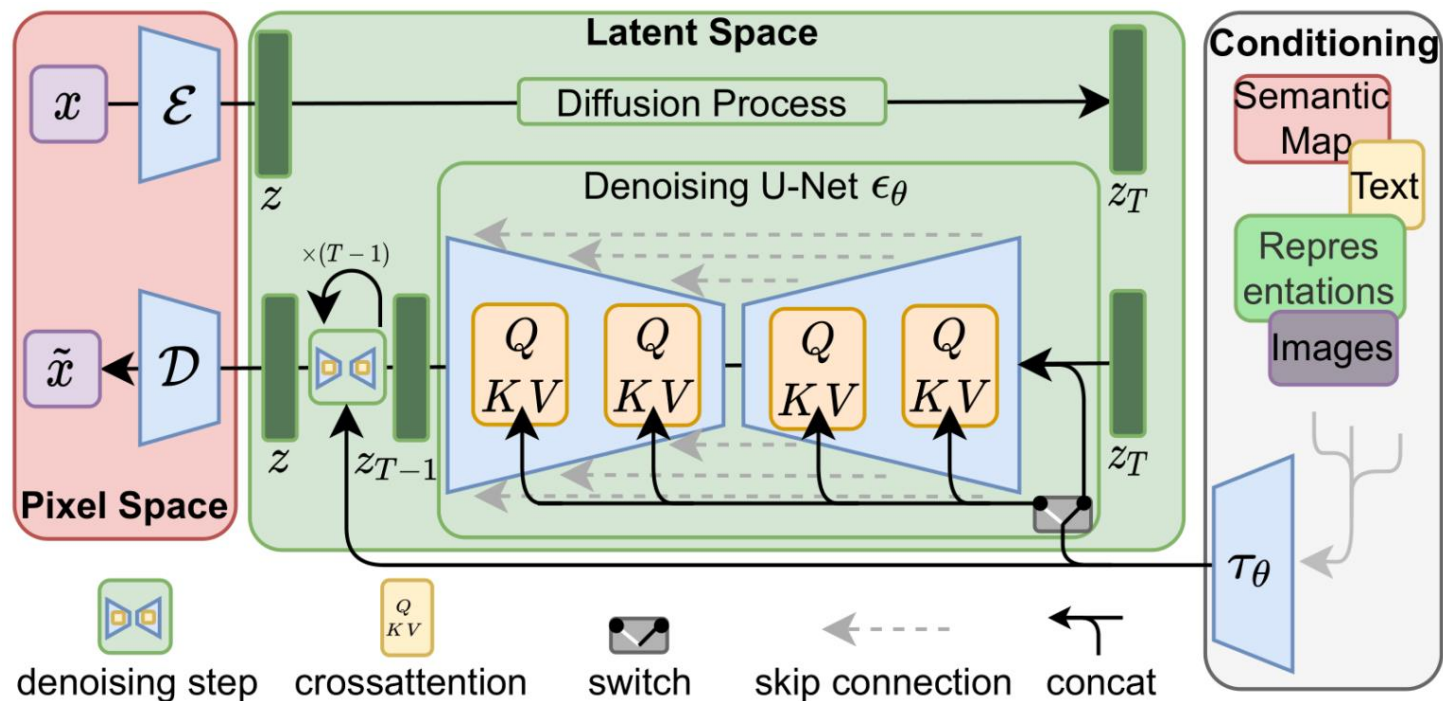




二、构建图像大模型增强的数据集

Midjourney生图存在
生成图像**过度艺术渲染化**的局限

Midjourney为闭源商用模型
参数不可改变
我们提出采用另一种开源图像大模型
Stable Diffusion



我们提出先对Stable Diffusion**进行微调**
解决生成的数据过度艺术渲染化等域偏移问题
然后再生成数据增强交通目标检测的技术路径





二、构建图像大模型增强的数据集

我们提出采用LoRA
在Foggy Cityscapes训练集上
对Stable Diffusion进行微调

LoRA能够在引入**极小参数量**的基础上
实现**高效的迁移学习**

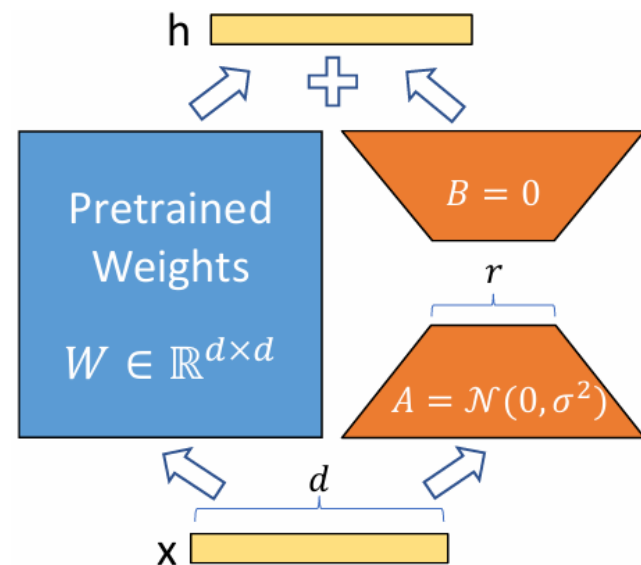
先前研究表明
仅用LoRA微调交叉注意力部分
就足以实现较好效果

我们用微调后的Stable Diffusion模型生成了约4000张合成数据
筛选出最佳的548张图像
按之前的Pipeline进行后处理与严格人工标注、反复检验

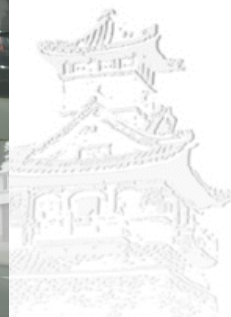
$$h = wx$$

$$w_0 + \Delta w = w_0 + BA$$

where $B \in \mathbb{R}^{d \times r}$ and $A \in \mathbb{R}^{r \times k}$, $r \ll \min(d, k)$.



Midjourney Images FoggyCityscapes StableDiffusion Images





三、恶劣天气交通目标检测增强



四川大學
SICHUAN UNIVERSITY

我们将Stable Diffusion+ LoRA生图与Foggy Cityscapes数据集混合增强交通目标检测模型，加以反复超参数调优
评估发现，相较于基线模型：

晴天验证集
AP +1.7%

雾天验证集
AP +2.3%

最终，我们将Midjourney和Stable Diffusion所生成的图像合并构建成为一个增强数据集，共1096张图像
然后用于增强交通目标检测，并进行超参数调优后，相较于基线模型，实现：

晴天验证集
AP +6.1%

雾天验证集
AP +5%





PART
FOUR



演示视频





基于图像大模型增强的恶劣天气交通目标检测方法研究



四川大學
SICHUAN UNIVERSITY



最终模型在雾天验证集上的结果展示

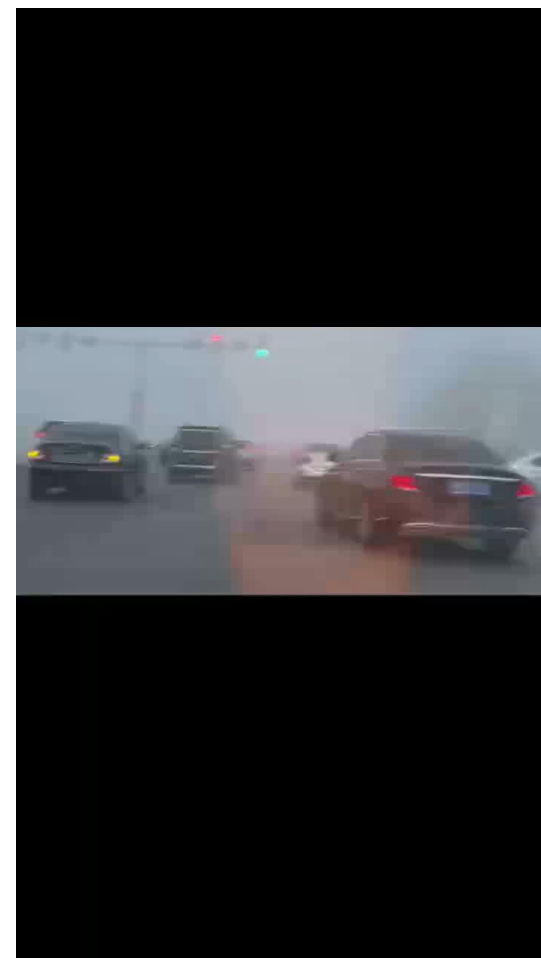
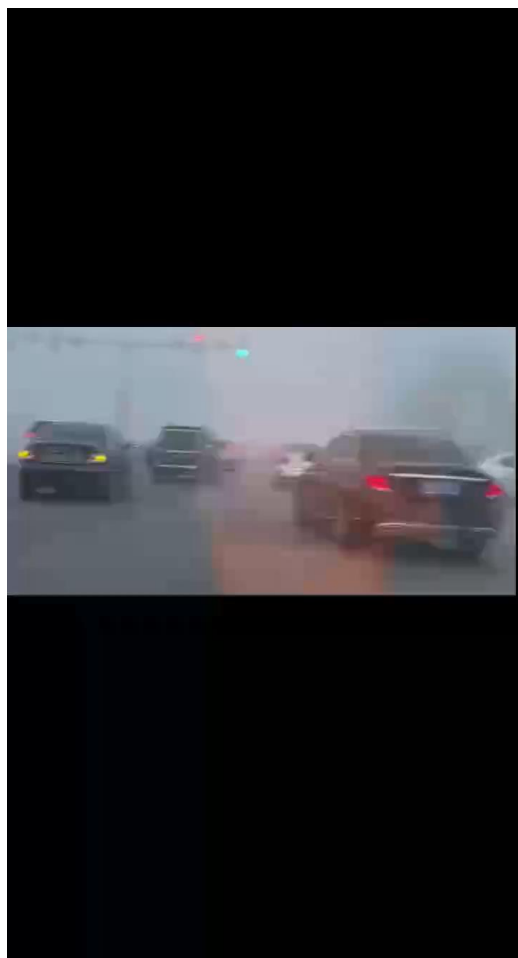




基于图像大模型增强的恶劣天气交通目标检测方法研究



四川大學
SICHUAN UNIVERSITY



真实雾天驾驶视频

Midjourney增强

Stable Diffusion增强

最终的合并增强

谢谢老师和同学们
敬请批评指正

Thanks for your listening



四川大學
SICHUAN UNIVERSITY

汇报人：侯睿

时间：2025.04.18